

Implémentation du processus Qualité Lecture Optique du SMART - Université de Liège, Belgique à la Haute École Pédagogique du Canton de Vaud, Lausanne, Suisse

Révision 1.1 du 2025-11-18.

Toute révision ultérieure rend de facto ce document obsolète.

Types de FormuLOMs

Un formuLOM est un *formulaire de Lecture Optique de Marques*, qui est conçu spécifiquement pour être lu par le lecteur optique *Axiome AXM980* à l'aide d'un programme de lecture optique (*format ADELE+, extension de fichier .AUR*) fourni par le SMART - Université de Liège.

Le format physique est une feuille de papier de grammage 90g/m² au gabarit DIN A4, dont les marques horloges (les marques noires sur le bord extérieur droit de la feuille recto (bord gauche au verso) sont situées exactement à 2.54mm de la tranche de la feuille.

Les zones de coches doivent impérativement se situer dans 38 colonnes espacées d'1/5 de pouce, correspondant aux 40 diodes du lecteur optique (*40-track 1/5"* du lecteur Axiome AXM980). L'encre rouge est invisible pour le lecteur optique, qui ne détectera que les zones cochées à l'aide d'un stylo à bille noir, bleu ou vert.

Les deux coches noires entourées de rouge placées en colonne 38 et situées sur les bords gauche supérieur et inférieur du recto (bords droit au verso) de la feuille permettent de vérifier l'équerrage de la feuille, et l'alignement des couleurs lors de l'impression.

La vérification du respect de la spécification pour les feuilles imprimées se fera à l'aide de la réglette en matière plastique **OMR forms ruler** fournie par le fabricant du lecteur optique.

HEPVaud utilise trois types différents de formuLOMs, qui sont imprimés sur du papier type *Olin Regular* à partir des fichiers suivants :

Questionnaires à choix multiples, 102 questions, recto-verso :

- **QCM5AT_102Q.pdf** : 5 propositions max. + proposition 'Aucune réponse n'est correcte' et proposition 'Toutes les réponses sont correctes' (*solutions générales implicites*)
- **QCM9_102Q.pdf** : 9 propositions max.

Ces deux types de formuLOMs (appelés QCM102) sont lus avec les programmes *VD102_A.AUR* et *VD102_V.AUR*.

Questionnaires à choix multiple avec degrés de certitude, 30 questions, recto uniquement :

- **QCMD5A_30Q.pdf**

Ce type de formuLOM (appelé QCMD30) est lu avec les programmes *VD30_A.AUR* et *VD30_V.AUR*.

Zones à cocher

Chaque formuLOM présente trois zones de données à cocher :

- Zone avec *aucune ou deux coches* en miroir côté gauche et droit pour la **forme du questionnaire** (*une lettre A, B, c ou D*),
- Zone de *5 coches* pour l'identifiant ou **matricule de l'étudiant** (*o suivie de 5 chiffres*),
- Zone de **réponses aux questions** (avec ou sans degrés de certitude) : chaque question présente une ligne principale et une ligne de repentir (*un chiffre de 1 à 5 ou 9 pour les réponses + les valeurs A et T suivant le formuLOM choisi; un chiffre de 0 à 5 pour le degré de certitude si présent*) .

Le cochage doit s'effectuer en **noircissant complètement** la case bordée de rouge, comme mentionné dans les consignes présentes sur chaque formuLOM.

En cas d'erreur, il ne faut pas raturer mais utiliser un ruban correcteur pour effacer la coche.

Lecture optique des formuLOMs

Les programmes de lecture chargés dans le lecteur optique Axiome vont générer un fichier au format texte, dit *fichier de lecture optique*, qui contiendra les données d'une feuille sur une ligne.

Chaque fichier contiendra autant de lignes que de feuilles lues par le lecteur optique.

Double lecture

Il y aura deux fichiers pour chaque lecture, qui sont appelées le *fichier de lecture optique A* (lecture de référence) et le *fichier de lecture optique V* (lecture de vérification).

- La lecture **A** est effectuée avec une intensité de **4** sur le lecteur optique et à l'aide des programmes *VD102_A.AUR* ou *VD30_A.AUR*
- La lecture **V** est effectuée avec une intensité de **3** sur le lecteur optique et à l'aide des programmes *VD102_V.AUR* ou *VD30_V.AUR*

Ces deux lectures à intensité différente permettent de détecter des erreurs de coches plus facilement par comparaison des lectures des deux fichiers A et V en générant un *fichier de données finales réconciliées*.

Afin de vérifier les données lues, l'enseignant doit joindre avec les feuilles des étudiants des feuilles appelées *feuilles de vérification* de la forme du questionnaire.

Il y aura toujours autant de feuilles de vérification qu'il y a de formes de questionnaire utilisées dans le cadre du dispositif anti-fraude lors de la passation.

Feuilles de vérification

Les en-têtes des feuilles de vérification doivent être cochés avec les valeurs suivantes :

- Feuille de vérification de la forme A :
Forme à cocher A
Matricule à cocher 99996

- Feuille de vérification de la forme B :

Forme à cocher *B*

Matricule à cocher 99997

- Feuille de vérification de la forme C :

Forme à cocher *C*

Matricule à cocher 99998

- Feuille de vérification de la forme D :

Forme à cocher *D*

Matricule à cocher 99999

Ces feuilles de vérification doivent comporter les réponses correctes cochées par l'enseignant.

Les réponses sur la feuille de vérification de la forme A doivent correspondre aux données transmises dans le *fichier de paramétrage*. Les réponses sur les feuilles de vérification des autres formes doivent être mélangées selon le mélange transmis dans le *fichier de paramétrage*.

Pour les feuilles avec degrés de certitude (QCMDC30), cette donnée ne doit pas être cochée sur les feuilles de vérification par l'enseignant.

Les feuilles des étudiants peuvent être lues dans un ordre aléatoire et différent entre la lecture A et la lecture V, notamment en raison des possibles rejets de feuilles si des données lues ne sont pas correctes (un rejet se produit par ex. si le matricule coché sur la feuille est erroné ou si la forme cochée sur la feuille est erronée, ou encore si un souci technique est apparu).

Cependant, pour des raisons logistiques, les feuilles de vérification sont généralement lues au début.

Numérotation des feuilles

Lors du lancement du programme de lecture optique, un numéro de feuille sera demandé à l'opérateur.

- Pour une nouvelle lecture, l'opérateur encodera 0000
- Pour poursuivre une lecture déjà entamée précédemment, l'opérateur encodera le numéro de feuille suivant la dernière feuille déjà lue, en format 4 chiffres (par ex, si la dernière feuille est 0411, l'opérateur encodera 0412).

Le numéro de feuille s'incrément de 1 à chaque feuille lue.

Ce numéro est également imprimé sur le formuLOM papier pour chacune des deux lectures, ce qui permet de retrouver une feuille plus aisément pour vérifier ses données.

- Le numéro de lecture A est le plus proche du *haut* du formuLOM
- Le numéro de lecture V est le plus proche du *bas* du formuLOM

Le numéro de feuille peut varier entre les deux lectures (par ex : feuilles rejetées, lecture pas dans le même ordre, etc...).

Afin de s'assurer que l'opérateur a bien lu chaque feuille une et une seule fois lors de la lecture A puis lors de la lecture V, une vérification du nombre de feuilles lues est effectuée.

Le dernier numéro de feuille lue doit être identique en lecture A et en lecture V.

Correction des rejets

Certaines feuilles mal cochées par les étudiants seront rejetées si :

- La forme du questionnaire n'est pas cochée correctement : double coche du côté gauche ou droit, coches incohérentes entre le coté gauche et droit.
- Le matricule n'est pas coché correctement : double coche ou oubli sur une ou plusieurs lignes.

En cas de rejet pour ces raisons, l'opérateur devra corriger à l'aide d'un stylo à bille vert et/ou d'un ruban correcteur les coches erronées et cocher la ou les bonnes valeurs afin que la feuille soit lue sans rejet.

Il est conseillé de consigner dans un journal les modifications effectuées sur les feuilles afin de garantir la traçabilité des actions de l'opérateur.

Description d'une ligne du fichier de lecture optique

Chaque ligne est constituée de champs séparés par un espace.

Chaque ligne comporte **6** champs (QCM5/A-T 102 Questions et QCM9 102 Questions) ou **7** champs (QCM5/A-T DC 30 Questions)

Exemple d'une ligne de la feuille de vérification de la forme A d'un QCM102 avec 60 questions :

```
1 HEPVD QCM102 099996 1 0000  
342133431122711264614112364144633274313411224361724371623144.....  
.....
```

Exemple d'une ligne de la feuille de vérification de la forme A d'un QCMD30 avec 20 questions, une seule forme :

```
1 HEPVD QCMD30 099996 . 0000 34213343112271126461.....  
54535434442533244541.....
```

Les champs sont dans cet ordre :

1. Référence de la lecture : toujours HEPVD
2. Type de feuille : QCM102 ou QCMD30
3. Matricule lu : 6 chiffres commençant toujours par un 0 (le premier chiffre, 0, est précoché sur le formuLOM)
4. Type de forme lue : 1 chiffre avec une valeur de 1 à 4 ou un point . si pas de forme cochée
5. Numéro de feuille : 4 chiffres
6. Liste des réponses aux questions lues : 102 ou 30 chiffres
7. Liste des degrés de certitude lus (*optionnel, uniquement présent si type de feuille QCMD30*) : 30 chiffres

Double lecture optique

Le numéro de feuille commence toujours à 0000 et s'incrément de 1 à chaque feuille lue.

Le fichier de lecture A et de lecture V doivent comporter les même nombre de feuilles, **le dernier numéro de feuille lue doit être identique**.

Données de lecture

Les données valides sont exclusivement numériques.

Les champs présentent soit :

- des données lues (*chiffre entre 0 et 9*)
- des données manquantes (*caractère point .*) :
une case qui n'a pas de coche alors qu'elle est attendue.
- une erreur (*caractère point d'interrogation ?*) :
par ex. double coche, rature, trait non-lu car trop fin ou mal marqué, etc.

Vérification des matricules

Chaque matricule doit comporter exactement 6 caractères et commencer par un 0.

Les matricules qui comportent des données manquantes (.) ou des erreurs (?) doivent être vérifiés et modifiés. Pour ce faire, le plus simple est de faire ressortir le numéro de feuille en lecture A et en lecture V, ce qui permet une vérification manuelle en recherchant la feuille avec le nom et le prénom inscrit par l'étudiant. Les feuilles se retrouveront dans l'ordre de la lecture V, qui est effectuée en dernier.

Il faut ensuite trier les deux fichiers A et V par matricule, et vérifier qu'ils sont concordants :
on doit retrouver la même liste de matricules dans les deux fichiers. Si ce n'était pas le cas, il faut vérifier les matricules qui diffèrent entre les deux lectures (toujours manuellement à l'aide du numéro de feuille en lecture A et en lecture V) et les corriger au besoin pour l'étape suivante.

Les matricules qui seront présents dans le fichier final **doivent impérativement faire partie du fichier signalétique fourni** (*voir plus bas*).

Les matricules des feuilles de vérification doivent être modifiés pour les valeurs suivantes :

- Feuille de vérification de la forme A :
Matricule lu **099996** à transformer en **999996**
- Feuille de vérification de la forme B :
Matricule lu **099997** à transformer en **999997**
- Feuille de vérification de la forme C :
Matricule lu **099998** à transformer en **999998**
- Feuille de vérification de la forme D :
Matricule lu **099999** à transformer en **999999**

Vérification des données

Chaque champ de données doit être vérifié par comparaison entre la ligne de lecture A et la ligne de lecture V pour un *même matricule* :

- forme du questionnaire,
- Réponses aux questions,
- (optionnel) degré de certitude (DC) associé à la réponse.

Forme du questionnaire

Pour la forme, le champ ne peut prendre que des valeurs numériques de 1 (Forme A) à 4 (Forme D). Si il y a une erreur ou une omission, il convient d'attribuer la bonne forme de questionnaire.

Les valeur de forme des feuilles de vérifications seront :

- Feuille de vérification de la forme A :
valeur **1**
- Feuille de vérification de la forme B :
valeur **2**
- Feuille de vérification de la forme C :
valeur **3**
- Feuille de vérification de la forme D :
valeur **4**

Si on n'a qu'une seule forme de questionnaire, le champ a un point comme valeur (.) dans les deux lignes, alors ce champ ne doit pas être vérifié et on peut lui attribuer la valeur **1** pour toutes les feuilles.

Réponses aux questions (et degrés de certitude)

Pour les réponses aux questions et les degrés de certitude, les champs qui comportent des données manquantes (.) ou des erreurs (?) doivent être vérifiés et corrigés.

En avançant dans le vecteur de données et en comparant question par question (*il y a 30 ou 102 valeurs suivant le type de formuLOM utilisé*) la valeur de la lecture A et la valeur de la lecture V, si la donnée est :

- **identique** : on conserve la valeur obtenue s'il s'agit d'une réponse numérique; sinon :
 - pour les données manquantes (.):
dans ce cas on considère généralement une omission et on remplace par le point par un **0**.
 - pour les erreurs de lecture (?):
dans ce cas, il convient de consulter le formuLOM papier à l'aide du numéro de feuille en lecture A et en lecture V pour cette question, et on encode la valeur finale, ou une omission (valeur **0**) si l'erreur est claire (*double coche, rature, etc...*).
- **différente** : après avoir consulté le formuLOM papier à l'aide du numéro de feuille en lecture A et en lecture V pour cette question on choisit la valeur finale à conserver (réponse à encoder), ou on encode une omission (valeur 0) si le choix est ambigu (mauvaise coche, etc...)

On obtient alors **un vecteur de données unique qui réconcilie les deux lectures A et V** pour les questions

(et un second vecteur pour les degrés de certitude, si présent). Ces vecteurs ont la même longueur que les vecteurs d'origine (*30 ou 102 valeurs suivant le type de formuLOM utilisé*).

Fichier de données finales réconciliées

Le fichier final (format **TXT**) à transmettre au SMART se présente sous la forme suivante :

Chaque ligne est constituée de champs séparés par un espace () 0x20, avec des caractères numériques ASCII uniquement.

Chaque ligne est terminée par un caractère Line Feed (*LF*) 0x0A (Format Unix).

Chaque ligne comporte **4** champs (QCM5/A-T 102 Questions et QCM9 102 Questions) ou **5** champs (QCM5/A-T DC 30 Questions)

Il ne peut y avoir que des **caractères numériques entre 0 et 9**.

Le vecteur de données des questions présente autant de chiffres que de questions possibles sur le formuLOM :

102 ou 30 suivant le modèle utilisé.

Le vecteur de données des degrés de certitude (DC) n'est présent que si on a une épreuve avec degrés de certitude, il présente alors 30 chiffres.

Exemple de fichier de données à fournir pour un QCM102 :

Fichier de paramétrage

En plus du fichier de données réconciliées, un fichier de paramétrage électronique sera demandé. Il remplace la lecture physique de la feuille de paramétrage qui ne sera plus utilisée.

Il aura une ligne d'en-tête comme ceci (caractères ASCII uniquement) :

1 Forme;Nom;Cours;Groupe_etudiants;Date;NB_questions;Nb_max_CP;Nb_max_Chap;Nb_groupe;SGI
;RC;NSP;Poids;CP;Chapitres;Formes;A1_B;A1_C;A1_D

Il devra comporter les informations suivantes :

- Forme : **vaut toujours A**
 - Nom de l'enseignant, en capitales, sans accents (**caractères ASCII uniquement**) (*max 25 caractères*)
 - Nom du cours, **caractères ASCII uniquement** (*max 40 caractères*)
 - Nom du groupe d'étudiants concerné (**caractères ASCII uniquement**) (*max 40 caractères*)
 - Date de passation de l'épreuve au format **JJ/MM/AAAA**
 - Nombre de questions, qui doit être supérieur à 1 et inférieur ou égal au nombre de questions du type de formuLOM utilisé pour administrer l'épreuve (**max 30 ou 102** suivant le type de formuLOM utilisé)

- Nombre maximal de Catégories de Performance : **vaut toujours 9**
- Nombre maximal de Chapitres : **vaut toujours 9**
- Groupe : **vaut toujours 1**
- Les Solutions Générales implicites utilisées dans toute l'épreuve : **un chiffre de 0 à 4**
 - **0** = pas de SGI utilisées
 - **1** = Aucune
 - **2** = Aucune et Toutes
- Un vecteur avec la liste des valeurs des réponses correctes pour chaque question, de longueur identique au nombre de questions de l'épreuve (**chiffres de 0 à 9**)
- Un vecteur avec la liste des nombres de solutions proposées pour chaque question, de longueur identique au nombre de questions de l'épreuve (**chiffres de 0 à 9**)
- Un vecteur avec la liste des pondérations pour chaque question, de longueur identique au nombre de questions (**la pondération par défaut est de 1**) (**chiffres de 0 à 9**).
- Un vecteur avec la liste des numéros de Catégories de Performance pour chaque question (de 0 à 9), de longueur identique au nombre de questions. La valeur **0** n'associe la question à aucune CP. (**chiffres de 0 à 9**)
- Un vecteur avec la liste des numéros de Chapitres pour chaque question (de 0 à 9), de longueur identique au nombre de questions. La valeur 0 n'associe la question à aucun chapitre. (**chiffres de 0 à 9**)
- Le nombre de forme(s) supplémentaire(s) : **un chiffre de 0 à 3**
 - **0** = forme A uniquement (*les trois champs suivants sont obligatoirement à 0*)
 - **1** = formes A et B
 - **2** = Formes A, B et C
 - **3** = Formes A, B, C et D
- La position de la question A1 dans la forme B du questionnaire (**0** si non-existent) (**un chiffre de 0 à 30 ou 102** suivant le type de formuLOM)
- La position de la question A1 dans la forme C du questionnaire (0 si non-existent) (**un chiffre de 0 à 30 ou 102** suivant le type de formuLOM)
- La position de la question A1 dans la forme D du questionnaire (**0** si non-existent) (**un chiffre de 0 à 30 ou 102** suivant le type de formuLOM)

Le fichier (format **CSV**) est constitué de la ligne d'en-tête terminée par un caractère Line Feed (*LF*) 0x0A (Format Unix), puis d'une seule ligne de données au format ASCII reprenant ces champs séparés par un **point-virgule** (;). La ligne est terminée par un caractère Line Feed (*LF*) 0x0A (Format Unix).

Exemple de fichier de paramétrage électronique :

Fichier Signalétique

Un fichier signalétique reprenant pour chaque matricule évalué, le nom, prénom, section et année académique des étudiants évalués sera fourni.

Il aura une ligne d'en-tête comme ceci (caractères ASCII uniquement) :

1 | matricule;nom;prenom;annee acad;ae code;ae lib

Ensuite, les quatre lignes des feuilles de vérification sous ce format :

```
1 999996;VERIF;A;20252026;SMART;SMART  
2 999997;VERIF;B;20252026;SMART;SMART  
3 999998;VERIF;C;20252026;SMART;SMART  
4 999999;VERIF;D;20252026;SMART;SMART
```

Il devra comporter les informations suivantes :

- Matricule : **6 caractères numériques** (commence par un **0**)
 - Nom : **caractères UTF-8** (*max 30 caractères*)
 - Prénom : **caractères UTF-8** (*max 25 caractères*)
 - Année académique : **8 chiffres** (*format 20252026*), avec une valeur identique pour toute la section.
(On change d'année académique au 15 septembre.)
 - Code de section : **toujours PEDAVAUD**
 - Libellé de section : **toujours HEPV**

Les matricules des étudiants qui seront repris dans le fichier final de données de lecture devront également se trouver dans ce fichier signalétique.

Le fichier (format **CSV**) est constitué de la ligne d'en-tête terminée par un caractère Line Feed (*LF*) 0x0A (Format Unix), puis des lignes de données encodées en UTF-8 reprenant ces champs séparés par un **point-virgule** (;).

Chaque ligne est terminée par un caractère Line Feed (*LF*) 0x0A (Format Unix).

Exemple de fichier signalétique :

```
1 matricule;nom;prenom;annee_acad;ae_code;ae_lib
2 0999996;VERIF;A;20252026;SMART;SMART
3 0999997;VERIF;B;20252026;SMART;SMART
4 0999998;VERIF;C;20252026;SMART;SMART
5 0999999;VERIF;D;20252026;SMART;SMART
6 012345;Nom de Famille;Étudiant;20252026;PEDAVAUD;HEPV
```

Fichier Adresses e-mail

Un fichier signalétique (format **CSV**) encodé en UTF-8 reprenant pour chaque matricule évalué une adresse e-mail valide sera fourni pour la publication des feedbacks.

Il n'y a pas de ligne d'en-tête.

Il devra comporter les informations suivantes :

- Matricule : **6 caractères numériques** (commence par un **0**)
- Adresse e-mail valide (*idéalement, une adresse e-mail de l'établissement*)

Les matricules des étudiants qui seront repris dans le fichier final de données de lecture devront également se trouver dans ce fichier associés à une adresse e-mail valide.

Le fichier (format **CSV**) est constitué des lignes de données encodées en UTF-8 reprenant ces champs séparés par un **point-virgule** (:).

Chaque ligne est terminée par un caractère Line Feed (LF) 0x0A (Format Unix).

Exemple de fichier d'adresses e-mail :

```
1 | 012345,etudiant@hepvd.ch
```

Recommandations

Pour le fichier signalétique et le fichier d'adresses e-mail, il est préférable de fournir un fichier signalétique unique et complet pour l'ensemble des étudiants concernés par les évaluations de la part d'un seul interlocuteur et pas des fichiers séparés reprenant les seuls étudiants des cours de la part des enseignants. Ces fichiers seront envoyés par mail à l'adresse smart@uliege.be au plus tard le jour du premier examen à corriger.

Données à fournir pour la correction d'une épreuve

Après avoir envoyé le fichier signalétique et le fichier d'adresses e-mail, les données des épreuves pourront être communiquées.

Afin de corriger une épreuve, il faudra envoyer un mail à l'adresse smart@uliege.be avec en pièces jointes :

- Le Fichier de données finales réconciliées
- Le Fichier de paramétrage

Le corps du mail comprendra les informations suivantes :

- Nom et prénom de l'enseignant
- Intitulé de l'épreuve
- Code-cours de l'épreuve
- Date de passation de l'épreuve
- Adresse e-mail de l'enseignant à qui envoyer les résultats
- Éventuellement, deux noms et adresses e-mails supplémentaires de personnes à qui envoyer les résultats.
- Barème de correction de l'épreuve avec la valeur de points à compter en cas de réponse correcte (**RC**), de réponse incorrecte (**RI**) et d'omission (**OM**):
 - Soit **Degrés de Certitude**
 - Les valeurs sont d'office calculées suivant le tableau (voir site web du SMART)
 - Soit **For Guessing**, faisant intervenir le *nombre de solutions proposées* (NSP)
 - Les valeurs sont : $\text{RC} = +1$ $\text{RI} = -1/(\text{NSP}-1)$ $\text{OM} = 0$
 - Soit **Personnalisé**
 - La valeur **RC** est toujours +1, les valeurs de **RI** et **OM** sont à discréption de l'enseignant entre 0 et -1 point avec des fractions à 0,2; 0,25; 0.33; 0.5; 0.66.
- Le ou les **Degré(s) de sévérité** à retenir (*plusieurs choix possibles entre 16, 17, 18, 19 et 20*). 20 est la valeur par défaut.
- Le nombre de décimales auquel arrondir le score final (*un choix entre 0 et 4*)
- La date limite pour la remise des notes
- Si l'enseignant utilise :
 - des pondérations particulières (*mentionner le poids sous forme d'un nombre entier uniquement*),
 - des chapitres (*mentionner les numéros de 1 à 9 et les intitulés correspondants*)
 - ou des catégories de performance (*mentionner les numéros de 1 à 9 et les intitulés correspondants*)

Pour chaque item, mentionner le ou les numéros de question qui sont concernées.

- De même il faudra mentionner si une question :
 - doit être supprimée (toutes les réponses sont considérées comme une omission)
 - doit être validée pour tous (toutes les réponses sont considérées comme correctes)
 - doit comporter une réponse supplémentaire à considérer comme valide