

# A propos des sauvegardes / restaurations du module référentiel de Moodle

Jean Fruitet - Université de Nantes - jean.fruitet@univ-nantes.fr

Je fais le point ici sur le protocole de sauvegarde / restauration du module Référentiel (version 2011101611 et ultérieures) pour les version Moodle 1.9 et Moodle 2

A propos des sauvegardes / restaurations du module référentiel de Moodle.....	1
Introduction .....	1
Architecture du module .....	1
Politique de sauvegarde .....	3
Conclusion .....	4
Bibliographie .....	4
Annexe : Quelques erreurs fréquentes lors des restaurations .....	5

## Introduction

Avec la complexification du module Référentiel les utilisateurs risquent d'être confrontés à des difficultés inattendues lors des sauvegardes / restaurations successives de cours contenant des instances du module référentiel. Cette question cruciale de cohérence des données est donc abordée ici du point de vue de l'administrateur d'un serveur Moodle qui cherche à garantir cette cohérence malgré la duplication / dissémination des données au travers du processus sauvegarde / restauration.

## Architecture du module

Il convient pour appréhender les enjeux et les difficultés de bien préciser certains points d'architecture du module référentiel.

### 1) Ne pas confondre Sauvegarde (BackUp) et Archivage

- un archivage consiste à enregistrer l'ensemble des dossiers numériques d'un utilisateur sur un support différent du serveur Moodle. Un archivage n'est pas censé permettre une restauration à l'identique dans Moodle.

- une sauvegarde consiste à enregistrer certaines ou la totalité d'un cours Moodle soit à des fins de stockage, soit à des fins de restauration ou de dissémination sur d'autres serveurs Moodle (en particulier lors de changements de version). Dans l'architecture Moodle, qui est centrée cours et non pas centrée fichier, une sauvegarde (backup) est nécessairement liée à un cours (celui qui a été sauvegardé) même si elle peut être restaurée (restore) dans un autre cours.

Le fait que la sauvegarde soit "centrée-cours" a pour effet que les informations de nature transversale (constituées par agrégation / intégration d'informations ou de contributions dans différents cours) sont délicates à adresser. Or c'est très exactement une propriété essentielle du module référentiel de permettre l'agrégation d'informations concernant une certaine occurrence de référentiel (un "référentiel, par exemple le C2i1, le B2i Collège, etc.) au travers de déclarations et de tâches évaluées dans des instances et des activités Moodle dispersées dans différents cours du même serveur.

### 2) Certaines données sont étroitement liées au cours où elles sont instanciées :

- L'instance (table referentiel) d'une activité Référentiel ;
- Les déclarations d'activité et de compétence (table referentiel\_activite) ;
- Les tâches (table referentiel\_tache) ;
- L'accompagnement (table referentiel\_accompagnement) des étudiants par des référents ;
- Les documents joints aux déclarations d'activité (referentiel\_document) ;
- Les documents joints aux tâches (referentiel\_consigne) ;

### 3) D'autres ne sont pas liées à un cours, elles sont transversales :

- L'occurrence (table referentiel\_referentiel) ;
- Les domaines (table referentiel\_domaine) ;
- Les compétences (table referentiel\_compétence) ;
- Les items de compétence (table referentiel\_item\_competence) ;

- Les certificats (table referentiel\_certificat) ;
- La liste des établissements (table referentiel\_etablissement) ;
- Les données individuelles (table referentiel\_etudiant) nécessaires à l'édition des certificats ;
- Les données de formation (table referentiel\_pedagogie) utiles à la répartition des étudiants par filières de formation.

4) Certaines données du module référentiel sont statiques (elles ne changent pas fréquemment) :

- Les rôles ;
- L'instance déclarée dans un cours ;
- L'occurrence et les Domaines / Compétences / Items du référentiel considéré ;
- Les données personnelles et de formation, les étudiants ne changeant pas - en général - de date de naissance au cours de leur scolarité :)).

5) D'autres sont dynamiques et donc constamment renouvelées ou mises à jour tout au long du processus de certification :

- Les déclarations d'activité ;
- Les tâches ;
- Les certificats.

## De la complexité du modèle de données

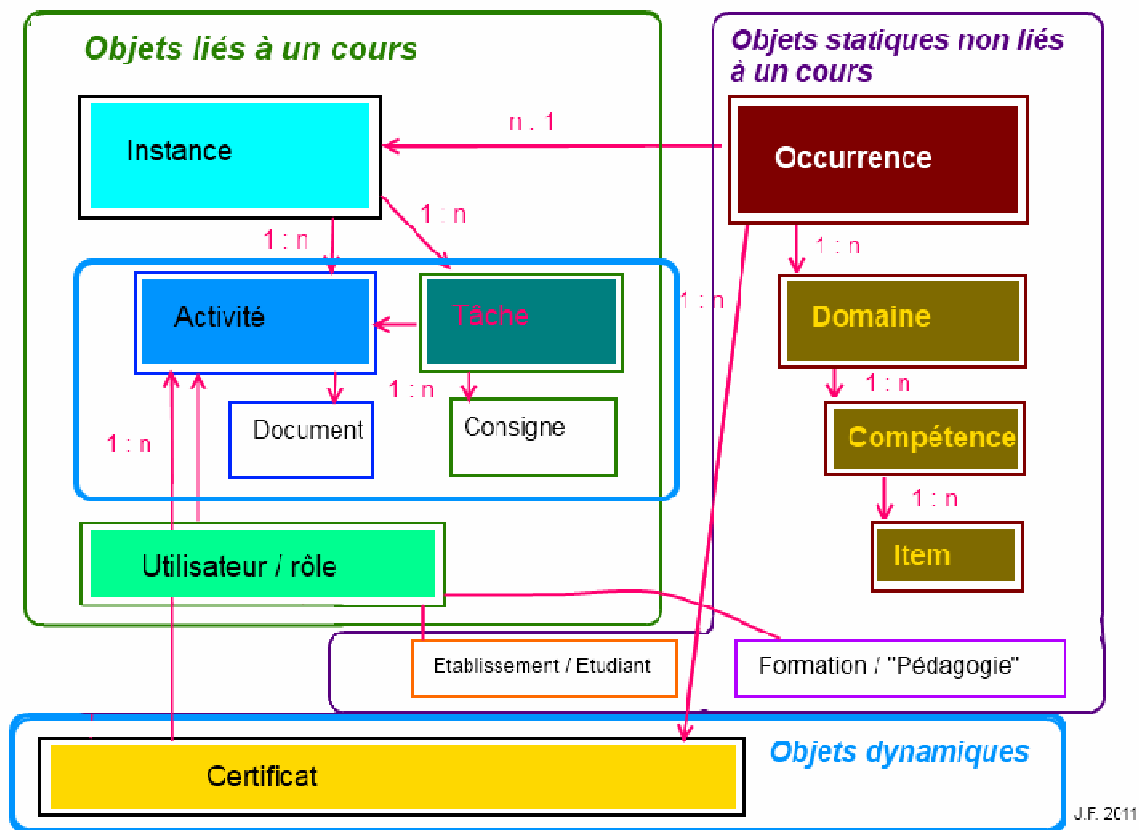
Chaque cours peut contenir une ou plusieurs instances de une ou plusieurs occurrences de référentiels.

A chaque instance est associée, pour chaque utilisateur du cours, une liste d'activités déclarées (et corollairement de compétences) et de tâches assignées.

Un certificat est la consolidation au niveau du serveur (pour une occurrence de référentiel donnée) de toutes les compétences validées pour le dit référentiel pour un utilisateur donné.

Le schéma ci-après permet de mieux percevoir la complexité de cette architecture :

### Modèle des données du module référentiel



## Politique de sauvegarde

Puisque le certificat synthétise l'état d'avancement vers la certification à un instant donné, toute restauration sur le même serveur risque de modifier cet état d'avancement, par exemple en restaurant des déclarations d'activités déclarées avant sauvegarde puis supprimées entre temps...

De même une restauration sur un autre serveur risque de remettre en cause des décisions concernant la certification de tel ou tel étudiant... par exemple en faisant sauter des verrous déposés sur leurs certificats (le rôle de ces verrous étant de geler l'état d'un certificat à un instant donné, afin de ne plus prendre en compte des déclarations de compétences ultérieures).

Cela m'a conduit à poser les règles de sauvegarde et de restauration strictes :

A) Les données liées à un cours sont sauvegardées et restaurées sans contrôle particulier, ce qui peut conduire à des données dupliquées si la restauration est faite "en place". *Il appartient donc aux utilisateurs ayant un rôle enseignants dans un cours de supprimer les tâches et les déclarations d'activité doublonnées après une restauration.*

B) Les données statiques "transversales" (Occurrence, Domaines, Compétences, Items) sont sauvegardées en l'état et restaurées si elles n'entrent pas en conflit avec des données similaires présentes sur le serveur au moment de la restauration. Cela concerne surtout l'occurrence d'un référentiel : si une occurrence identique existe sur le serveur cible, il n'en est pas créé de copie.

La vérification d'identité porte dans cet ordre sur :

- l'identité des clés des occurrences si elles sont non vides
- en cas d'échec sur l'identité du code des listes des items de compétences

Autrement dit la vérification d'identité ne porte pas sur les poids et empreintes, qui sont considérés comme des variables d'ajustement permettant de définir des protocoles de certification différents pour le même référentiel.

Par conséquent si un référentiel change de code pour une liste d'items entre deux sauvegardes, la restauration va créer un doublon de code différent qui sera considéré comme une occurrence différente. De plus les domaines, compétences et item étant dépendant hiérarchiquement d'une occurrence, si la restauration considère deux occurrences comme identiques, aucun nouveau domaine, compétence ou item ne sera enregistré (même s'il y a des différences de rédaction dans certains intitulés).

C) Les données statiques transversales liées à des utilisateurs ne sont sauvegardées que pour les utilisateurs inscrits dans le cours. Cela concerne les données de scolarité (établissement, date et lieu de naissance, affectation à une formation)

- Toute restauration de ces données vérifie que les données restaurées n'entrent pas en conflit avec les données enregistrées sur le serveur cible (étudiant ayant le même login ou numéro d'étudiant dans la sauvegarde et le serveur cible mais avec des adresses différentes, affectation à des formations / pédagogies différentes. Dans chaque cas ce sont les données du serveur cible qui sont conservées.

D) Enfin, et c'est là que les règles sont les plus contraignantes, nous devons statuer sur les données transversales hautement dynamiques tels les certificats.

Nous allons examiner différents scénarios avant de définir ces règles.

**Scénario 1 :** Les données de certification ne sont ni sauvegardées ni - a fortiori - restaurées. Cette règle stricte à quelques graves inconvénients mais aussi beaucoup d'avantages.

1) Inconvénients :

- La sauvegarde n'enregistre pas les décisions du jury ni la synthèse rédigée par l'étudiant en fin de parcours de certification.
- La sauvegarde n'enregistre pas l'état d'avancement du processus à un instant T. On ne peut pas commodément restituer un historique de certification.
- Les certificats reflétant à tout moment l'état des compétences validées dans les activités, une restauration intempestive peut tout au plus multiplier les activités doublonnées, faussant ainsi le protocole de certification si celui-ci repose sur une gestion différenciée des empreintes d'items.

2) Avantages

- Aucun risque de conflit entre des états d'avancement différents dans le processus de certification.

- Pas de risque de faire sauter des verrous posés sur les certificats lors d'une restauration intempestive.

**Scénario 2 :** La sauvegarde enregistre les données de certifications de tous les utilisateurs.

La restauration n'effectue aucun contrôle sur les données restaurées.

Ce scénario est à l'évidence un scénario catastrophe... Le risque principal est de faire perdre les informations de certification (date de certification, décision du jury, modification des verrous).

**Scénario 3 :**

- La sauvegarde n'enregistre que les données de certifications des utilisateurs du cours où s'effectue la sauvegarde **ayant effectué des déclarations d'activité dans ce cours**. Cela limite fortement l'impact d'une restauration. *Par contre il est tout à fait inutile d'envisager la sauvegarde de l'ensemble des certificats d'un référentiel en réalisant la sauvegarde d'un meta-cours où n'aurait pas été effectuées de déclarations d'activité.*

- La restauration effectue un contrôle sur les données restaurées pour garantir la cohérence de celles-ci avec les données présentes sur le serveur.

Ce scénario est à l'évidence plus complexe à mettre en oeuvre mais pour peu que les règles de restauration soient correctement rédigées - et programmées - il permet d'espérer gommer certains des inconvénients du premier scénario.

**Règle de restauration N°1 :** Pas d'écrasement du contenu des champs "synthèse", "date de certification", "décision du jury" si ces champs ont une valeur non vide sur le serveur cible.

Cela garantit que le certificat "vivant" prend le pas sur les certificats "archivés".

**Règle de restauration N°2 :** Si le verrou du certificat n'est pas posé, recalcul automatique de la liste des compétences validées et de la liste des compétences déclarées sur l'ensemble des déclarations enregistrées sur le serveur après restauration. Si le verrou est posé aucune modification de la listes de compétences validées ni des compétences déclarées...

## Conclusion

J'ai profité du portage du module référentiel vers la version Moodle 2 pour imposer rétroactivement ces règles à la version du module référentiel pour Moodle 1.9. Le code concernant les "backup / restore" du module référentiel a été totalement réécrit à partir de la version 2011101611 release '5.7.04 - 2011/10/16'.

Je recommande donc à tous les utilisateurs de Moodle 1.9 de mettre à jour leur version du module référentiel et à renouveler par la même occasion la liste de leurs sauvegardes de cours contenant une instance du module référentiel.

Ils auront ainsi la garantie que le passage de leur serveur de Moodle 1.9 vers Moodle 2.1 ou suivant récupèrera sans perte les dossiers numériques des utilisateurs.

## Bibliographie

Manuel développeur du module référentiel <sup>1</sup>.

---

<sup>1</sup> Ceci est un private joke, le manuel développeur du module référentiel consistant à vrai dire en l'ensemble des sources du programme, sources disponibles avec toute distribution du module référentiel. Si un gentil programmeur / développeur se lançait dans la rédaction de ce document, je lui en serais éternellement reconnaissant.

N.B. : A l'échelle des temps de l'artisanat du logiciel, l'éternité ne va guère au delà de deux ou trois années universelles :)) - JF

## Annexe : Quelques erreurs fréquentes lors des restaurations

### Conflits sur les utilisateurs lors de restaurations sous Moodle 1.9.7 et suivantes.

Depuis la version Moodle 1.9.7 des contrôles stricts ont été introduits pour éviter les conflits lors des restaurations.

Si vous rencontrez une erreur du type



vous devez vérifier si les utilisateurs enregistrés dans la sauvegarde ont strictement la même adresse mail que les utilisateurs de même login sur votre serveur de restauration.

Pour corriger cette erreur vous pouvez :

- soit uniformiser les adresses mail sur le serveur cible en repérant l'adresse mail enregistrée dans le fichier de sauvegarde que vous restaurez (décompressez-le et ouvrez le fichier `moodle.xml` avec un éditeur puis consultez le contenu de la rubrique `<COURSE><USERS><USER><EMAIL>...</EMAIL>` )
- soit modifier sous éditeur l'adresse mail enregistrée dans l'archive de la sauvegarde ce qui est parfois le choix le plus simple. Après modification recomprimez à nouveau l'archive et recommencez la restauration...

### Malgré une restauration réussie, les déclarations d'activité de certains utilisateurs n'apparaissent pas...

Si vous avez restauré une sauvegarde vers un nouveau cours, n'oubliez pas d'inscrire les étudiants au cours nouvellement créé. Leurs déclarations apparaîtront alors.