



**MAKINA
CORPUS**

Projet libre - csvreplicata

Manuel d'utilisation
Version 1.0 - 25/09/2008

Réf. MC-CSVREPLICATA-FR
Makina Corpus
<http://www.makina-corp.us.com>
<mailto:contact@makina-corp.us.com>



TABLE DES MATIÈRES

1 Présentation	4
2 Mise en œuvre	4
[2.1] Installation	4
[2.2] Paramétrage global	5
3 Utilisation	7
[3.1] Accès aux fonctionnalités	7
[3.2] Exportation	8
[3.3] Importation	9
[3.4] Format du fichier	11
[3.5] Particularités	12
[3.5.1] Dossiers automatiques	12
[3.5.2] Fichiers joints	12
[3.5.3] Références vers des objets	13
[3.5.4] Conflits	13
4 Paramétrage avancé	13
[4.1] Interface de marquage	13
[4.2] Handlers	14

1 PRÉSENTATION

Le produit csvreplicata va au-delà du simple système d'import/export CSV tel que l'offrait auparavant le produit ArcheCSV (i.e. choix d'un type de contenu cible + mise en correspondance entre les champs Archetypes et les colonnes du CSV).

Il permet de faire vivre sous CSV un simili Plone off-line : le fichier CSV pourra contenir des types de contenus différents, et reproduira leur arborescence. Les contenus exportés peuvent être modifiés dans le fichier CSV et de nouveaux contenus peuvent y être créés. Le fichier CSV peut ensuite être ré-importé dans le site pour mettre à jour ces contenus.

Il s'agit d'un produit opensource publié sous licence GPL par Makina Corpus.

Ce produit a été réalisé pour le compte de l'Institut Royal de Sciences Naturelles de Belgique dans le cadre du projet MARS.

2 MISE EN ŒUVRE

[2.1] Installation

Le produit peut être téléchargé depuis pypi.python.org, il suffit de décompresser l'archive et de placer le répertoire csvreplicata dans le répertoire products de Plone.

Ou bien, on peut utiliser un egg, il suffit d'ajouter Products.csvreplicata dans le buildout:

```
eggs =  
    ...  
    Products.csvreplicata
```

Ensuite, il faut installer le produit depuis l'interface Plone (Site setup – Add-on products).

[2.2] Paramétrage global

Une fois installé, le produit csvreplicata dispose d'un écran de configuration accessible depuis la page de configuration de Plone.

[▲ Up to Site Setup](#)

This is CSV Replicator settings section, you modify the default behaviour in the lists below.

CSV Replicator Settings

Encoding
Select the encoding of your file
UTF-8

Delimiter
Select the delimiter of your file
;

String delimiter
Select the string delimiter of your file
"

save

Fields and classes of Fields not handled by the CSV Import/Export

Classes of fields
Select the classes of fields handled by the CSV Import/Export

Fields
Select the fields handled by the CSV Import/Export
id
locallyAllowedTypes
TypesMode
immediatelyAddableTypes

save

Types handled by the CSV Import/Export

Select the types handled by the CSV Import/Export
ATSortCriterion
Analysis
Analysis Absolute Dating
Analysis Relative Dating
Artefact

save

Types handled by the CSV Import/Export and their respective schemas

Select the schemas handled by the CSV Import/Export for each selected types

Artefact default
description
chronology
taphonomy
technology

Artefact Assemblage default
description
chronology
taphonomy
technology

Cet écran permet à l'administrateur de gérer des paramètres qui s'appliqueront à l'ensemble des opérations d'import/export, sachant que certains d'entre eux peuvent être modifiés par l'utilisateur directement au moment de l'importation ou de l'exportation.

Encoding :

Il s'agit de l'encodage utilisé pour produire et lire les fichiers CSV. La valeur par défaut est UTF-8, mais on peut utiliser n'importe quel autre encodage (ISO-8859-1 par exemple).

Delimiter :

C'est le caractère utilisé pour séparer les colonnes du fichier CSV. La valeur par défaut est le point-virgule ';'.

Note : si on souhaite utiliser le caractère tabulation, il faut le copier/coller depuis un éditeur de texte (par exemple), car dans l'interface web, la touche tabulation fait simplement passer au champ suivant et ne saisit rien.

String delimiter :

C'est le caractère utilisé pour encadrer les valeurs de type string. Par défaut, on utilise les guillemets, c'est le caractère généralement utilisé pour le format CSV.

Fields and classes of fields not handled by the CSV Import/Export :

Il s'agit de 2 black-lists qui permettent de systématiquement ignorer certains champs ou certaines classes de champs afin de ne pas surcharger les fichiers CSV avec des informations inutiles ou inexploitable pour les utilisateurs.

Dans la liste des classes, il faut donner le nom des classes de champs Archetypes et les séparer par des retours à la ligne.

Exemple :

```
IntegerField  
FileField
```

Dans cet exemple, tous les champs de type entier ou fichier joint des différents types de contenus ne seront jamais exportés.

Dans la liste des champs, il faut donner les identifiants des champs (tels qu'ils sont définis dans le schéma du type de contenu Archetypes correspondant), et les séparer par des retours à la ligne.

Exemple :

```
id  
locallyAllowedTypes  
TypesMode  
immediatelyAddableTypes
```

Il s'agit de la liste par défaut. Ce sont des champs « système » d'Archetypes qui ne présentent aucun intérêt à priori pour les utilisateurs.

Note : le champ id doit impérativement y figurer car il est déjà exporté par csvreplicata.

Types handled by the CSV Import/Export :

Il s'agit de la liste des types de contenus qui seront pris en charge par csvreplicata.

Note : il faut appuyer sur la touche Ctrl pour sélectionner plusieurs types.

Chaque type sélectionné dans cette liste va apparaître dans la liste suivante (voir ci-dessous).

Types handled by the CSV Import/Export and their respective schematas :

Pour chaque type sélectionné, cette section permet de choisir les schematas à exporter.

Pour rappel, pour un Archetypes, les schematas sont différentes catégories dans lesquelles figurent les champs du schéma. Dans les types de bases comme Document par exemple, on trouve les schematas suivants: default, categorization, dates, ownership, settings.

Et ce sont ces schematas que l'on retrouve dans les onglets de l'écran de modification des contenus Plone:



Seuls les champs des schematas sélectionnés seront pris en charge par csvreplicata.

Cela permet de ne pas surcharger les fichiers CSV avec des informations non pertinentes. Par exemple, le schemata Settings ne concerne que le paramétrage Plone, il est fort probable qu'il ne soit pas très utile (sauf dans le cas d'une migration par le biais de csvreplicata).

3 UTILISATION

[3.1] Accès aux fonctionnalités

[3.1.1]

Les fonctionnalités de csvreplicata sont accessibles depuis l'onglet «csv import/export » présent dans tous les dossiers et les dossiers automatiques du site Plone.



Note : on peut activer cet onglet pour d'autres types si nécessaire (voir Paramétrage avancé).

La page « csv import/export » contient deux sections : la section Import et la section Export.

[3.2] Exportation

Paramètres	
Workflow states	Permet de limiter l'exportation aux contenus ayant des états workflows donnés. Le choix « Any » prend tous les états.
Depth	Indique la profondeur maximum jusqu'à laquelle aller dans l'arborescence. Si on choisit 0, il n'y a pas de limite dans la profondeur, et l'objet courant est également exporté.
Encoding	Permet de choisir un autre encodage que celui choisi dans la configuration csvreplicata du serveur.

Delimiter	Permet de choisir un autre séparateur de colonne que celui choisi dans la configuration csvreplicata du serveur.
String delimiter	Permet de choisir un délimiteur de texte que celui choisi dans la configuration csvreplicata du serveur.
Date/time format	Permet de choisir le format de date à utiliser (en fonction du logiciel qui servira à lire le CSV obtenu).
Export attached files	Indique si l'on souhaite ou pas exporter les fichiers joints. Note: si tel est le cas, l'utilisateur n'obtiendra pas un simple CSV, mais une archive ZIP contenant le CSV et tous les fichiers joints.
Types handled	Permet de limiter les types de contenus à exporter (par rapport à la liste des types sélectionnés dans la configuration csvreplicata du serveur).
Export vocabulary value	Pour tous les champs à vocabulaire, l'utilisateur peut choisir d'exporter la valeur réelle du champ ou bien le libellé correspondant. Note : Le libellé sera probablement plus explicite, en revanche il peut ne pas être unique et donc poser des problèmes de cohérence, et il peut aussi être plus facilement une source d'erreurs de saisies.

[3.3] Importation

Import

CSV File
Pick the CSV File to be imported

Encoding
Select the encoding of your file

Delimiter
Select the delimiter of your file

String delimiter
Select the string delimiter of your file

Date/time format
Indicate the date / time format of your file

Conflicts management
Choose the rule to apply if some conflicts are encountered

Server wins OR File wins (WARNING: it might produce unwanted data overwriting)

Workflow transition to new objects
Choose the workflow transition for the new objects

Import files
Submitted file must a zip archive, attached files will be updated only if they have been renamed

Yes No

Import vocabulary value
Select "Yes" to provide the value of the vocabulary instead of the key.

Yes No

Quand l'utilisateur lance l'importation, il obtient un rapport précisant le nombres de contenus mis à jour, le nombre de contenus créés, et la liste des éventuelles erreurs.

Info 1 object(s) created, 1 object(s) modified.

Error Conflict on line 13: Overlapping content modified on the server after exportation

Paramètres	
CSV	Le fichier CSV à importer (ou éventuellement le fichier ZIP si on souhaite importer des fichiers joints).
Encoding	Permet de choisir un autre encodage que celui choisi dans la configuration csvreplicata du serveur.
Delimiter	Permet de choisir un autre séparateur de colonne que celui choisi dans la configuration csvreplicata du serveur.
String delimiter	Permet de choisir un délimiteur de texte que celui choisi dans la configuration csvreplicata du serveur.

Date/time format	Permet de choisir le format de date à utiliser.
Conflicts management	En cas de conflits de mise à jour entre les contenus du serveur et ceux du fichiers CSV, l'utilisateur peut décider quelle règle de priorité appliquer (à propos des conflits, voir plus bas).
Workflow transition for new objetcs	Permet d'appliquer une action workflow sur tous les contenus nouveaux. On peut ainsi automatiquement publier tous les contenus résultant de l'importation. Note : les contenus existants restent dans le même état, même s'ils sont modifiés par l'importation CSV.
Import files	Indique si l'on souhaite ou pas importer des fichiers joints.
Import vocabulary value	Pour tous les champs à vocabulaire, l'utilisateur peut indiquer si le fichier CSV utilise la valeur réelle des champs ou bien les libellés correspondants.

[3.4] Format du fichier

La première ligne du fichier contient le chemin du dossier où a été effectué l'export (l'import devant être ensuite effectué dans le même dossier a priori, sauf si on souhaite dupliquer une arborescence), et la date de l'export au format YYYYMMDDhhmm.

Les lignes suivantes permettent de décrire les champs des types de contenus et leurs valeurs. Ainsi à chaque fois qu'on change de type de contenu :

- une ligne donne les identifiants des champs de ce type de contenu,
- une seconde ligne avec les libellés des champs (pour une meilleure compréhension pour l'utilisateur),
- une série de ligne donne les valeurs pour chacun de ces champs (une ligne par objet)

La première colonne est le chemin (i.e. path) du conteneur de l'objet (si cette colonne est vide, cela signifie que le conteneur est le dossier courant spécifié à la première ligne du fichier).

La seconde colonne est l'identifiant (i.e. id) de l'objet.

La troisième colonne est l'identifiant du type de contenu.

Les colonnes suivantes sont propres au type de contenu.

Exemple:

/test2/dossier1	20080925145955							
parent	id	type	title	description				text
Parent folder	Identifiant	Content type	Title	label_description				label_body_text
	presentation	Document	Présentation	Présentation du projet				
parent	id	type	title	description				
Parent folder	Identifiant	Content type	Title	label_description				
	les-equipes	Folder	Les équipes					
parent	id	type	title	description				text
Parent folder	Identifiant	Content type	Title	label_description				label_body_text
	les-equipes	Document	Equipe 2	Equipe de test				
parent	id	type	title	description				text image imageCaption
Parent folder	Identifiant	Content type	Title	label_description				label_body_text label_news_image label_image_caption
	reunion	News Item	Réunion	Une réunion se tiendra à l'occasion de la fin du projet. Toutes les équipes sont conviées.				
parent	id	type	title	description				
Parent folder	Identifiant	Content type	Title	label_description				
/test2	dossier1	Folder	dossier1					
parent	id	type	title	description				text
Parent folder	Identifiant	Content type	Title	label_description				label_body_text
	les-equipes	Document	Equipe 1	Equipe de développement				
parent	id	type	synonyms	description				text igYear remainType
Parent folder	Identifiant	Content type	Alternate Name	label_description				label_body_text Year of Inventorisation
	artefact1	Artefact	Individsqsq					1974 acquisition-protocol

[3.5] Particularités

[3.5.1] Dossiers automatiques

Lorsqu'on effectue l'exportation dans un dossier automatique, elle porte sur les contenus sélectionnés dans le dossier automatique.

On peut ainsi effectuer des filtrages complexes multi-critères pour exporter certains contenus en particuliers.

Le chemin des contenus (1ère colonne du fichier CSV) est le chemin complet à partir de la racine du site.

Note : on ne peut pas ré-importer un CSV dans un dossier automatique, en revanche du fait que les chemins sont complets, on peut ré-importer à la racine du site Plone.

[3.5.2] Fichiers joints

Lorsque les champs correspondant à des fichiers joints (FileField, ImageField, ...) sont exportés, l'exportation produit un fichier ZIP contenant le fichier CSV et tous les fichiers joints.

Dans le fichier CSV, la valeur du champs est le nom du fichier.

Il faut veiller à ce que les fichiers n'aient pas le même nom pour éviter les écrasements.

[3.5.3] Références vers des objets

Lorsqu'on exporte des champs de type ReferenceField, la valeur dans le fichier CSV correspond au chemin vers le contenu référencé.

Au moment de l'importation, si le chemin ne correspond à aucun contenu existant (soit parce qu'il a été supprimé, soit parce que la valeur saisie dans le CSV est incorrecte), l'utilisateur est averti que la référence est cassée.

[3.5.4] Conflits

Si un contenu est modifié dans le fichier CSV et qu'il est également modifié sur le site Plone après la date d'exportation, cela constitue un conflit.

Au moment d'importer ses données, l'utilisateur choisira le mode de résolution des conflits :

- « Server wins » : on garde la version du serveur,
- « Local wins » : on remplace la version du serveur par celle du fichier mais uniquement dans la mesure où l'utilisateur est Owner des objets concernés.

Les conflits sont signalés à l'utilisateur dans le rapport d'importation.

L'identification des conflits se base sur la date du fichier CSV (qui figure sur la première ligne).

Il est fortement recommandé de ne pas conserver les fichiers CSV une fois qu'ils ont été ré-importés, mais plutôt de les détruire, et d'en exporter un nouveau à chaque fois que nécessaire pour éviter les conflits.

4 PARAMÉTRAGE AVANCÉ

[4.1] Interface de marquage

Par défaut, l'onglet « csv import/export » est disponible sur les contenus de type Folder (Dossier) et de type Smart Folder (Dossier automatique).

Si on souhaite que cet onglet apparaisse pour d'autres types de contenu (dérivant de Folder à priori), il suffit de les marquer avec l'interface ICSVReplicable.

Pour cela il suffit d'ajouter le tag suivant dans un configure.zcml :

```
<five:implements
    class="Products.YourProduct.yourtype.YourType"
    interface="Products.csvreplicata.interfaces.ICSVReplicable"
/>
```

[4.2] Handlers

Si on souhaite que csvreplicata soit capable de traiter des types de champs autres que les champs de base d'Archetypes, il faut écrire un handler adapté.

Pour cela, il faut implémenter votre classe handler de la façon suivante :

```
from Products.csvreplicata.exceptions import *
from Products.csvreplicata.handlers.base import CSVdefault

class SpecialField(CSVdefault):
    """
    """

    def get(self, obj, field, context=None):
        """
        v = obj.Schema().getField(field).getRaw(obj)
        # v is the value returned by the field accessor
        # Make the appropriate processing to return
        # the value you want to display in the CSV
        # Note: it has to be a string
        return result

    def set(self, obj, field, value, context=None):
        """
```

```
"""  
if value=='':  
    value = None  
# value is the value read in the CSV file  
# Make the appropriate processing to compute  
# the value you will store in the field  
# then, call the store method  
self.store(field, obj, value)
```

Cette classe doit respecter les contraintes suivantes :

- hériter de CSVdefault,
- contenir une méthode get qui doit retourner la valeur à placer dans le CSV lors de l'exportation, il s'agit donc nécessairement d'une string, et l'objectif est de la rendre la plus compréhensible possible pour l'utilisateur,
- contenir une méthode set qui doit calculer la valeur à stocker dans l'objet en fonction de la valeur lue dans le CSV lors de l'importation, cette méthode doit toujours appeler la méthode store de CSVdefault pour effectuer le stockage.

Ensuite cette classe handler doit être enregistrée auprès de csvreplicata. Pour cela, on utilise la méthode setHandler de CSVReplicataTool :

```
self.portal.portal_csvreplicatatool.setHandler('Myproduct.Field.SpecialField', {'handler_class': SpecialField(),  
'file': False})
```

Le premier paramètre indique la classe field AT pour laquelle on fournit un handler, le second paramètre est un dictionnaire qui fournit la classe handler qu'on vient d'implémenter, et un booléen pour indiquer s'il s'agit d'un champ contenant des fichiers joints ou pas.



Réf. MC-CSVREPLICATA-FR
Makina Corpus
<http://www.makina-corp.us.com>
<mailto:contact@makina-corp.us.com>