

MIGRATION - EXPORT DE DONNÉES À PARTIR DE WINISIS

[Introduction](#)

[Deux manières d'exporter](#)

[Description de la base de démo](#)

[Export en format .iso](#)

- [Exporter la base](#)
- [Convertir avec MarcEdit](#)
- [Corriger quelques erreurs mineures](#)
- [Transformer en un vrai fichier Unimarc](#)
- [Employer MarcEdit ou le tableur ?](#)

[Export en format .txt](#)

- [Créer le format d'impression](#)
- [Imprimer la base](#)
- [Bien choisir le séparateur de champs](#)
- [Tester la validité du séparateur de champs](#)

[Import d'un fichier texte dans le tableur](#)

- [2 méthodes en fonction du séparateur](#)
- [Présence d'un zéro initial](#)
- [Préserver si nécessaire le zéro initial](#)
- [Système de numérotation et codes-barres](#)

[Et ensuite ?](#)

[Ressources complémentaires](#)



<http://amcubat.be/docpmb/>

*PMB est un SIGB : un Système Intégré de Gestion de Bibliothèque.
C'est un logiciel performant, en constante évolution, développé par la société PMB-Services.
Il est distribué sous licence libre et son code-source est ouvert.
C'est dans la même optique de partage que j'ai décidé de mettre à votre disposition un site de documentation.
Prenez le temps d'explorer ces ressources, il y a de nombreuses fonctionnalités de PMB à découvrir ...*

Anne-Marie Cubat

INTRODUCTION

Winisis est un système développé par l'Unesco – surtout entre 1998 et 2004 (la version antérieure s'appelait CDS-ISIS). Ces bases Winisis sont au départ « multifonctions », elles peuvent donc gérer aussi bien le catalogue d'une petite bibliothèque qu'un fichier de noms, adresses et numéros de téléphone. En effet, chaque utilisateur définit les champs, leur numéro, le titre, le type de contenu, le bordereau de saisie, le format d'impression, etc.

C'est pourquoi il y a une grande diversité entre les bases, même si elles concernent le même domaine (ex. bibliothèque)

- Dans Winisis, c'est l'utilisateur qui attribue le numéro et le titre aux divers champs, supposons qu'il y en ait 20.
Les numéros peuvent être 1, 2, 3, 4 ... 20 ou bien 10, 20, 30, 40 ... 200 (on va donc de 10 en 10)
- Exemple de diversité : dans une base Winisis, on a prévu un seul champ pour l'auteur (ou les auteurs), le champ n° 1, alors que dans une autre base, il y a deux champs : le champ n° 40 pour l'auteur principal, et le champ n° 50 pour les autres auteurs.

Des bases de ce genre présentent des inconvénients majeurs :

- Vu qu'elles ne sont pas conçues en soi pour gérer une bibliothèque, rien ne garantit qu'on ait respecté les normes bibliographiques fondamentales : certaines données essentielles peuvent faire défaut, ou bien être présentes mais « mélangées » (on peut par exemple retrouver dans le même champ le nom et le prénom de l'auteur)
- Vu qu'il ne s'agit pas d'une base de données relationnelle, le risque d'avoir des doublons est fort élevé (dans les notices, les noms d'auteurs, d'éditeurs, etc, voire même dans les numéros d'exemplaires)
- Il n'y a aucune uniformité entre les bases, 2 utilisateurs peuvent concevoir des bases fort différentes, donc la migration vers un « vrai » logiciel de gestion de bibliothèque sera à chaque fois un cas d'espèce, puisqu'il faudra établir dans la mesure du possible une correspondance entre les champs créés dans Winisis et les codes Unimarc
- Les bases Winisis sont assez « archaïques » dans leur architecture, et en outre Winisis ne fonctionne plus sur les PC récents (à moins d'installer une machine virtuelle), il devient urgent de récupérer les données bibliographiques qui y ont été enregistrées
- Alors que Winisis a été mis gratuitement à la disposition des utilisateurs par l'Unesco, il est paradoxal de constater que le seul export prévu se fait dans un format propriétaire (des fichiers .iso propres à Winisis). En fait, les 2 seules migrations envisagées sont de Winisis à Winisis ou de Winisis à ABCD !
- Or ABCD est le successeur de Winisis et est lui aussi développé par l'Unesco ! On ne quitte pas la maison-mère ...

DEUX MANIÈRES D'EXPORTER

Heureusement, il est quand même possible de récupérer les données qui se trouvent dans Winisis ;-)

Exporter la base en format .iso et convertir (avec MarcEdit) le fichier .iso en format Unimarc (.mrc)

- Les fichiers .iso créés par Winisis ne sont pas tout à fait illisibles. Grâce à l'excellent utilitaire MarcEdit (qui permet d'éditer le format Marc ou Unimarc), on peut convertir les fichiers .iso en un « pseudo » fichier Marc.
- Ce n'est pas un véritable fichier Marc, car les codes ne sont pas ceux de Marc 21 (ou de l'Unimarc). Au lieu de ces codes, on trouve le numéro attribué à chaque champ au moment de la création de la base Winisis.
- Mais c'est un fichier qu'on peut lire et adapter, il faudra donc remplacer les « pseudo-codes » qui s'y trouvent (les numéros des champs dans la base Winsis) par de vrais codes Unimarc si on veut importer les données
- Les caractères accentués passent en général mal, il faut faire des remplacements afin de les récupérer

Imprimer le contenu de la base dans un fichier texte et éditer ce fichier .txt dans le tableur

- Winisis permet de créer divers formats d'impression des données.
- L'astuce consiste à créer un « pseudo » format d'impression, qui va générer un fichier texte (Windows ANSI), il n'y aura donc pas de problème pour les caractères accentués.
- Chaque ligne du fichier correspond à un enregistrement, et les champs sont séparés par un caractère spécial.
- Il faut faire attention au moment de choisir le séparateur de champs. Si on peut utiliser le point-virgule comme séparateur, c'est le système le plus simple pour faciliter l'import dans Excel et les exports ultérieurs.
- Sinon quelques précautions s'imposent : il doit s'agir d'un caractère non employé dans la base (sinon le découpage se fera mal) ; en outre, il faut éviter les caractères « réservés », ceux qui ont une fonction spéciale dans Winisis.
- Transiter par le tableur offre des avantages : si on a par exemple regroupé dans le même champ le nom et le prénom de l'auteur, on peut assez facilement scinder les éléments dans Excel, alors que dans MarcEdit, il faudrait reprendre les fiches une à une. Un autre atout est qu'on peut aisément créer dans Excel des champs qui manqueraient dans la base d'origine, mais qu'il est utile de créer (ex. indexation décimale, langue de parution, fonction de l'auteur ...)
- Passer par Excel ou OpenOffice Calc permet donc d'améliorer et d'enrichir la base avant l'import.

DESCRIPTION DE LA BASE DE DÉMO

Définition de base de données - Table des champs (BIBLIO)

Etiquette	Nom	Type	Rep.	Délimiteurs
1	ISBN	Alphanumér	-	
2	Titre	Alphanumér	-	
3	Editeur	Alphanumér	-	
4	Date de parution	Alphanumér	-	
5	Nombre de pages	Alphanumér	-	
6	Auteur principal	Alphanumér	-	
7	Titre original	Alphanumér	-	
8	Langue de publication	Alphanumér	-	
9	N° exemplaire	Alphanumér	-	
10	Cote de rangement	Alphanumér	-	

Sauvegarder Vide ligne Tri ligne Efface ligne

Etape 1. Définition des champs

Définissez la structure de votre base en spécifiant les étiquettes des champs, leurs noms, types et délimiteurs.

Annuler Aide Oui

Définition de base de données - Bordereaux de saisie (BIBLIO)

Bordereau : BIBLIO

Sauvegarder Nouveau

Champs :

1	ISBN
2	Titre
3	Editeur
4	Date de parution
5	Nombre de pages
6	Auteur principal
7	Titre original
8	Langue de publication
9	N° exemplaire
10	Cote de rangement

Champs de saisie :

1	ISBN
2	Titre
3	Editeur
4	Date de parution
5	Nombre de pages
6	Auteur principal
7	Titre original
8	Langue de publication
9	N° exemplaire
10	Cote de rangement

Validation de l'enregistrement

Etape 2. Définition des Bordereaux de saisie

Annuler Aide Oui

Définition de base de données - Formats d'affichage (BIBLIO)

Nom du format : BIBLIO

Sauvegarder Nouveau

```

f2,c14,m(3048,-3048)
"ISBN (1):",tab(3048),v1/
"Titre (2):",tab(3048),v2/
"Editeur (3):",tab(3048),v3/
"Date de parution (4):",tab(3048),v4/
"Nombre de pages (5):",tab(3048),v5/
"Auteur principal (6):",tab(3048),v6/
"Titre original (7):",tab(3048),v7/
"Langue de publication (8):",tab(3048),v8/
"N° exemplaire (9):",tab(3048),v9/
"Cote de rangement (10):",tab(3048),v10/

```

Etape 3. Définition des formats d'affichage

Annuler action Annuler Aide Oui

CDS/ISIS for Windows [Version complète : FR] - [c:\winisis\data\biblio.mst - Saisie de données]

Fichier Affichage Recherche Edition Configuration Utilitaires Fenêtres Aide

NFM 1

Bordereau BIBLIO

Ajout champ : <->

978-2070654260

ISBN	978-2070654260
Titre	Encyclopédie des sciences
Editeur	Gallimard Jeunesse
Date de parution	2013
Nombre de pages	384 p.
Langue de publication	fran
N° exemplaire	1327
Cote de rangement	030 ENC SCI

CDS/ISIS for Windows [Version complète : FR] - [c:\winisis\data\biblio.mst - Saisie de données]

Fichier Affichage Recherche Edition Configuration Utilitaires Fenêtres Aide

NFM 2

Bordereau BIBLIO

Ajout champ : <->

978-2070654307

ISBN	978-2070654307
Titre	Encyclopédie de l'espace
Editeur	Gallimard Jeunesse
Date de parution	2013
Nombre de pages	256 p.
Langue de publication	fran
N° exemplaire	1354
Cote de rangement	030 ENC ESP

CDS/ISIS for Windows [Version complète : FR] - [c:\winisis\data\biblio.mst - Saisie de données]

Fichier Affichage Recherche Edition Configuration Utilitaires Fenêtres Aide

NFM 3

Bordereau BIBLIO

Ajout champ : <->

978-0349139623

ISBN	978-0349139623
Titre	Long Walk to Freedom
Editeur	Abacus
Date de parution	2013
Nombre de pages	784 p.
Auteur principal	Mandela Nelson
Langue de publication	angl
N° exemplaire	1456
Cote de rangement	910 MAN

CDS/ISIS for Windows [Version complète : FR] - [c:\winisis\data\biblio.mst - Saisie de données]

Fichier Affichage Recherche Edition Configuration Utilitaires Fenêtres Aide

NFM 4

Bordereau BIBLIO

Ajout champ : <->

978-0143117155

ISBN	978-0143117155
Titre	Invictus : Nelson Mandela and the Game that made a Nation
Editeur	Penguin Books
Date de parution	2009
Nombre de pages	288 p.
Auteur principal	Carlin John
Langue de publication	angl
N° exemplaire	980
Cote de rangement	910 CAR INV

CDS/ISIS for Windows [Version complète : FR] - [c:\winisis\data\biblio.mst - Saisie de données]

Fichier Affichage Recherche Edition Configuration Utilitaires Fenêtres Aide

NFM 5

Bordereau BIBLIO

Ajout champ : <->

978-2253140634

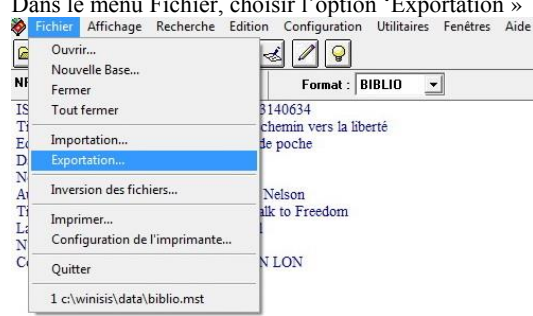
ISBN	978-2253140634
Titre	Un long chemin vers la liberté
Editeur	Le livre de poche
Date de parution	1996
Nombre de pages	767 p.
Auteur principal	Mandela Nelson
Titre original	Long Walk to Freedom
Langue de publication	fran angl
N° exemplaire	876
Cote de rangement	910 MAN LON

Comme vous le voyez, une petite base de démo qui comprend 5 enregistrements - 10 champs alphanumériques ont été définis.


EXPORT EN FORMAT .ISO

Exporter la base

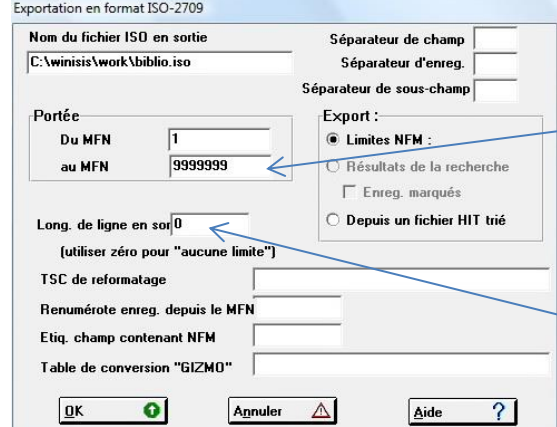
Dans le menu Fichier, choisir l'option 'Exportation »



Sélectionner le dossier WORK et donner un nom au fichier (biblio)



Exportation en format ISO-2709



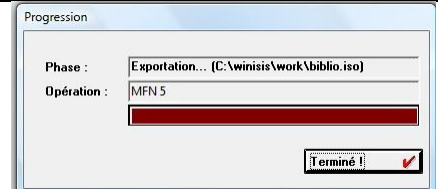
Ma base de démo est toute petite, j'ai donc exporté la totalité. Pour vos tests initiaux, réduisez éventuellement le nombre d'enregistrements à exporter. Au stade final, il faudra bien sûr tout exporter.

Attention !

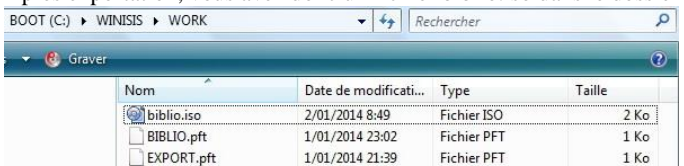
Par défaut, Winisis propose de limiter la longueur de ligne à 80 caractères, ce qui risque de provoquer des sauts de ligne intempestifs. Un souvenir des « vieilles imprimantes » !!

Pour éviter cet inconvénient, remplacer la valeur par défaut (qui est 80 caractères) par zéro.

Progression



Après exportation, vous avez donc un fichier biblio.iso dans le dossier WORK



Convertir avec MarcEdit

MarcEdit est un utilitaire gratuit, un véritable « couteau suisse » de l'édition de fichiers Marc!

Tout aussi indispensable que Notepad ++

Il permet d'éditer des fichiers Marc, et aussi de convertir des données dans ce format ou à partir de ce format.

Cerise sur le gâteau, il inclut **une fonction de conversion qui transforme les fichiers .iso de CDS-ISIS en format Marc.**

J'avais déjà travaillé quelques fois avec MarcEdit, mais vu qu'à l'époque mon but était d'importer une base Unimarc dans PMB (après avoir transformé quelques codes pour les adapter à l'Unimarc ISO 2709), et vu que je ne m'intéressais pas spécialement à Winisis à ce moment-là, je n'avais pas remarqué cette option.

Merci à Anthony Lorry (du CEDIAS), qui m'a signalé que MarcEdit offrait cette fonctionnalité.

Merci à Anthony Lorry pour l'aide efficace qu'il a régulièrement apportée à tout le monde via la liste de diffusion de PMB.

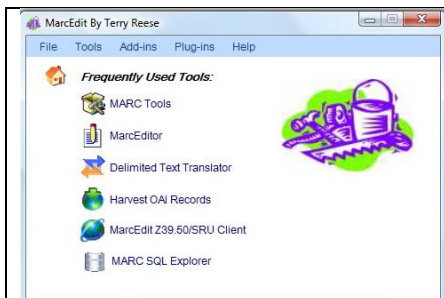
Il nous a hélas quittés beaucoup trop tôt, en janvier 2014 ...

N.B. Le catalogue en ligne du Cédias – Musée social (76.000 notices) mérite le détour : <http://cediasbibli.org/opac/>

Un bel exemple de base PMB, qui intègre des notices « classiques » et des documents numériques consultables en ligne (avec la visionneuse)

Préalable indispensable : télécharger et installer MarcEdit. Il est disponible à cette adresse-ci : <http://marcedit.reeset.net/>

On lance MarcEdit, et on choisit l'onglet « Add-ins ».



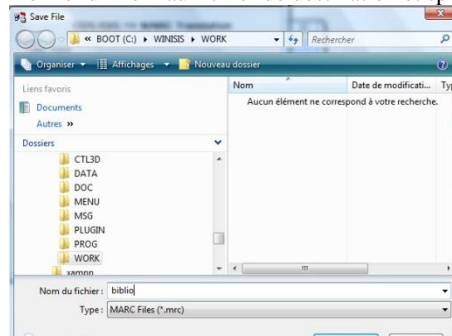
Une des options proposées dans l'onglet « Add-ins » est la traduction d'un fichier CDS-ISIS en format Marc



Choisir le fichier-source : biblio.iso



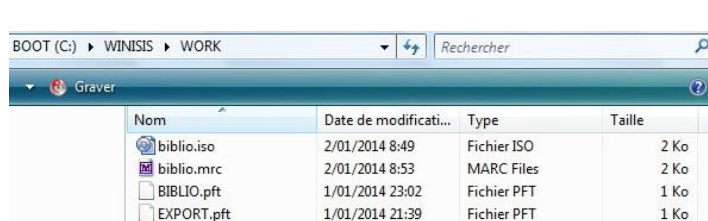
Donner un nom au fichier de destination et spécifier son format : mrc



Lancer le processus de conversion



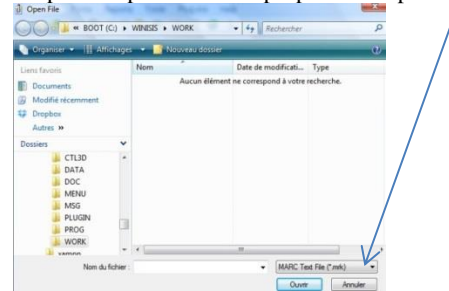
On obtient le fichier converti. : biblio.mrc



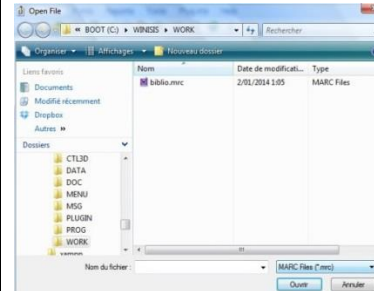
Pour visualiser le résultat obtenu, lancer MarcEditor - du moins si c'est nécessaire (vous ne l'avez peut-être pas encore fermé).



Ne pas accepter le format proposé au départ : mrc



Prendre le format .mrc et ouvrir biblio.mrc



N.B. Si le but est simplement d'éditer un fichier .mrc, mais que vous n'avez pas besoin des outils supplémentaires proposés par ce logiciel, il n'est pas nécessaire de lancer MarcEditor, il y a une autre manière plus rapide d'ouvrir un fichier .mrc. Voir illustration un peu plus loin.

Les enregistrements apparaissent.

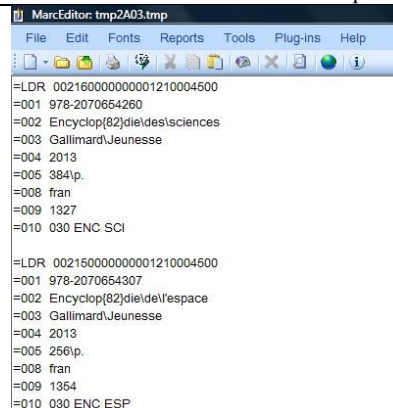
Certes, il y a encore des problèmes mineurs à régler (ex. caractères accentués), mais les données sont là.

Ce n'est pas un « vrai » fichier Unimarc, car vous n'avez pas les codes Unimarc (ex. 200\$a pour le titre du livre, 210\$c pour le nom de l'éditeur).

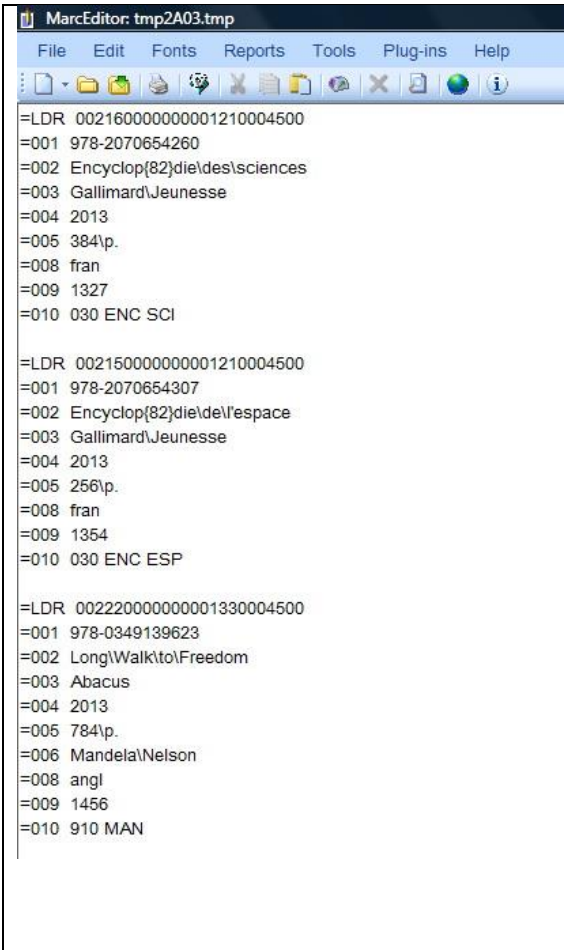
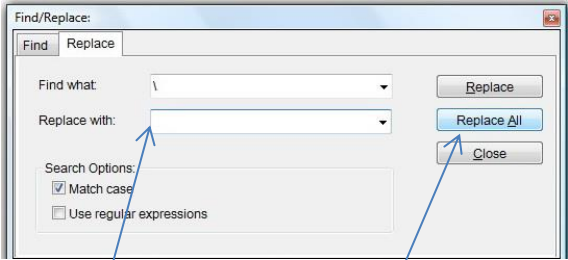
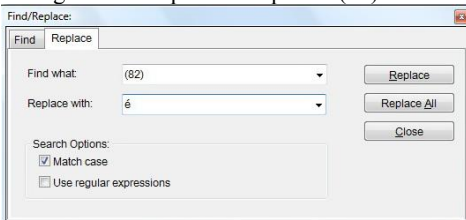
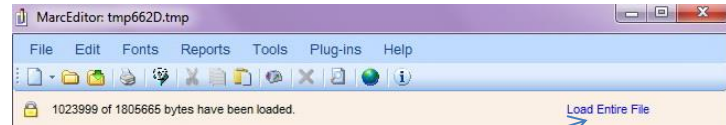
Ces codes sont remplacés par le numéro que le champ correspondant avait dans la base Winisis

Ex. champ n° 002 : titre

champ n° 003 : nom de l'éditeur



Corriger quelques erreurs mineures

 <pre> =LDR 00216000000001210004500 =001 978-2070654260 =002 Encyclop{82}die\des\sciences =003 Gallimard\Jeunesse =004 2013 =005 384\p. =008 fran =009 1327 =010 030 ENC SCI =LDR 00215000000001210004500 =001 978-2070654307 =002 Encyclop{82}die\de\l'espace =003 Gallimard\Jeunesse =004 2013 =005 256\p. =008 fran =009 1354 =010 030 ENC ESP =LDR 00222000000001330004500 =001 978-0349139623 =002 Long\Walkto\Freedom =003 Abacus =004 2013 =005 784\p. =006 Mandela\Nelson =008 angl =009 1456 =010 910 MAN </pre>	<p>Dans ce cas-ci, la 1^{re} chose à faire est de corriger les caractères accentués.</p> <p>J'ai aussi constaté une petite bizarrerie inexplicable dans l'état actuel de mes connaissances de Winisis : dans certains champs (titre, nom de l'éditeur et nombre de pages), l'espace est remplacé par une barre oblique inversée, mais par contre, il n'y a aucun problème pour la cote de rangement, les espaces y sont restés tels quels ! Je ne sais pas pourquoi ...</p> <p>Heureusement, les fonctions de remplacement de MarcEdit sont efficaces.</p>  <p>Ce champ n'est pas vide. Il contient un espace ... invisible ! En fait, j'ai demandé de remplacer partout la barre oblique inversée par un espace (un seul).</p> <p>J'ai agi de même pour la séquence (82) afin d'obtenir un « é ».</p> 
<pre> =LDR 00216000000001210004500 =001 978-2070654260 =002 Encyclopédie des sciences =003 Gallimard Jeunesse =004 2013 =005 384 p. =008 fran =009 1327 =010 030 ENC SCI =LDR 00215000000001210004500 =001 978-2070654307 =002 Encyclopédie de l'espace =003 Gallimard Jeunesse =004 2013 =005 256 p. =008 fran =009 1354 =010 030 ENC ESP </pre>	<p>Voilà le résultat obtenu après correction.</p> <p>Attention lorsque vous éditez un fichier .mrc assez long et que vous voulez faire la même modification partout (replace all) !</p> <p>Si le fichier est volumineux, au début du processus, MarcEdit ne charge en mémoire qu'une partie du fichier.</p>  <p>Ce message d'avertissement n'est pas anodin.</p> <p>Pensez à cliquer sur « Load Entire File », sinon vous risquez de croire que vous avez fait des modifications dans tout le fichier, alors qu'elles n'ont affecté qu'une partie de celui-ci.</p>

Mais cela ne suffit pas pour faire de ce fichier un véritable fichier Unimarc, il faut encore modifier d'autres éléments.

Le début des lignes ne doit pas commencer par =001, =002, =003, mais par un véritable code Unimarc.

Exemple : l'ISBN doit avoir le code 010 \$a, c'est le champ n° 1 de la base Winisis.

Donc la ligne =001 978-2070654307 doit devenir =010 \$a978-2070654307

Transformer en un vrai fichier Unimarc

Vous pouvez comparer les diverses versions		
=LDR 00285000000001450004500 =001 978-2253140634 =002 Un\long\chemin\vers\la\libert(82) =003 Le\livre\de\poche =004 1996 =005 767\p. =006 Mandela\Nelson =007 Long\Walk\to\Freedom =008 fran\angl =009 876 =010 910 MAN LON	=LDR 00285000000001450004500 =001 978-2253140634 =002 Un long chemin vers la liberté =003 Le livre de poche =004 1996 =005 767 p. =006 Mandela Nelson =007 Long Walk to Freedom =008 fran angl =009 876 =010 910 MAN LON	=LDR 00303000000001330004500 =010 \$a978-2253140634 =200 \$aUn long chemin vers la liberté =210 \$cLe livre de poche =210 \$d1996 =215 \$d767 p. =700 \$aMandela\$bNelson =200 \$dLong Walk to Freedom =101 \$afre\$sceng =995 \$f876\$k910 MAN LON\$rbook
biblio-2.mrc : le fichier "à l'état brut", tel qu'il a été converti par MarcEdit	biblio-3.mrc : le fichier légèrement corrigé espace au lieu de « \ », « é » au lieu de (82)	biblio-4.mrc : un « vrai » fichier Unimarc prêt à être importé dans PMB

Certains changements sont assez simples à réaliser.

Exemple : le nom de l'**éditeur** doit avoir le code 210 \$c, c'est le champ n° 3 de la base Winisis.

Donc la ligne =003 Gallimard Jeunesse doit devenir =210 \$cGallimard Jeunesse

Solution : un remplacement global dans tout le fichier. **Ne perdez pas de vue les espaces intermédiaires !**

=003 (attention ! c'est =003 et les 2 espaces qui suivent !) devient =210 \$c

Autres exemples de changements assez faciles de code : **date de publication, nombre de pages, titre et titre parallèle** (le titre parallèle est le titre original, dans ce cas-ci, le titre en anglais)

Mais d'autres cas sont plus complexes.

Le nom et le prénom de l'auteur doivent être scindés dans deux champs distincts.

Le code Unimarc du nom d'un auteur est 700 \$a, celui du prénom est 700\$b

Donc la ligne =006 Mandela Nelson doit devenir =700 \$aMandela\$bNelson

Remplacer =006 par =700 \$a ne pose pas de gros problème.

Le souci, c'est de mettre \$b devant le prénom (au lieu de l'espace intermédiaire entre nom et prénom).

Les langues posent un double problème.

Dans une base Winisis, on peut coder les langues comme on veut. En Unimarc, le code de l'anglais est eng, celui du français fre.

Effectuer un remplacement global n'est pas à conseiller, car vous risquez de remplacer par exemple « fran » par « fre » non seulement pour les codes des langues, mais aussi (par inadvertance) dans des mots comme France, français, franchise, safran ... !

Le 2° souci pour les langues, c'est qu'il faut les scinder, et attribuer un code à la langue de publication (101 \$a), et un autre à la langue originale (101 \$c). Au stade final, =008 fran angl doit devenir =101 \$afre\$sceng

L'exemple le plus complexe est probablement celui des **données de l'exemplaire**, qui doivent être en une seule ligne.

Les deux lignes du fichier de départ doivent donc fusionner en une seule, dont le code global est 995 :

995 \$f pour le **numéro d'exemplaire (ou n° de code-barres)**, 995 \$k pour la **cote de rangement**

=009 876

Cela devient =995 \$f876\$k910 MAN LON\$rbook

=010 910 MAN LON

N.B. J'aurais pu mettre autre chose que « book » (ex. « livre ») comme codage d'import, cela dépend du paramétrage de la base.

N.B. En outre, pour que l'import des exemplaires fonctionne correctement dans PMB, ces 2 champs (numéro et cote) ne suffisent pas, il faut aussi spécifier le **type de support** (995 \$r), c'est pourquoi j'ai ajouté ici « \$rbook », un élément ne figurait pas dans le fichier original. Explications complémentaires : <http://amcubat.be/docpmb/import-exemplaires-notices-monographies>

Les remplacements d'un code par un autre peuvent aussi vous piéger !

J'en ai fait l'expérience en préparant ce tutoriel. J'avais tout d'abord remplacé le code du champ n° 1, l'ISBN.

=001 978-2253140634 était devenu =010 \$a978-2253140634

J'ai ensuite remplacé petit à petit les autres codes, et un problème est survenu lorsque je suis arrivée au dernier champ de ma base Winisis. La version de départ était =010 910 MAN LON, et j'ai machinalement demandé le remplacement de =010 par =995 \$k

Dans un premier temps, je n'ai pas compris pourquoi la ligne de l'ISBN était devenue subitement =995 \$k978-2253140634.

Après quelques secondes de réflexion, j'ai réalisé que j'avais créé une **confusion entre le code Winisis du champ n° 10 (=010) et le code Unimarc de l'ISBN, qui est lui aussi =010 !**

Cela peut arriver pour d'autres n° : il se peut que vous ayez 11 champs Winisis, or le code Unimarc =011 \$a est celui de l'ISSN.

Si vous avez numéroté les champs de 10 et en 10 et que vous en avez au moins 20, il y aura un risque de confusion entre le 20° champ Winisis (=200) et le code Unimarc =200 \$a, qui est celui du titre de l'ouvrage.

Employer MarcEdit ou le tableur ?

MarcEdit convient parfaitement si vous voulez importer dans PMB une base Unimarc ou Marc 21, donc une base créée par un logiciel de gestion de bibliothèque respectueux des normes.

Vous pourrez faire dans MarcEdit les adaptations nécessaires, si elles ne sont pas trop complexes. J'ai par exemple remplacé le code des données d'exemplaires (qui était =HLD) par =995, et j'ai effectué d'autres changements mineurs, mais il n'y avait pas de risque de confusion (comme pour les codes – voir page précédente). En outre, je n'ai pas dû scinder des données en deux champs ou en fusionner d'autres en une seule ligne.

Au stade final, la base importée dans PMB et retravaillée dans MarcEdit ressemble fort à celle d'origine, en ce qui concerne la définition des champs et leurs relations entre eux.

Je ne vous conseillerais pas d'employer MarcEdit si la base de départ est assez ancienne et/ou peu conforme aux normes actuellement en vigueur en bibliothéconomie. Or, c'est souvent le cas pour des bases de type Winisis.

Je ne dirais sûrement pas que c'est de la faute des bibliothécaires. On ne peut pas faire des miracles avec un logiciel qui est ancien et fort limité (un système basé sur le DOS et rendu compatible Windows, rien de plus), il n'a pas été conçu au départ pour gérer le catalogue d'une bibliothèque, et son mode de fonctionnement est en soi source d'erreurs (une base de données non relationnelle génère fatalement des doublons).

Certes, MarcEdit permet de créer, à partir d'une base Winisis, un fichier Unimarc directement importable dans PMB.

Techniquement, c'est possible, mais cela ne me semble pas judicieux si la base n'est pas (ou n'est plus) conforme au niveau bibliographique. Les normes ont évolué en bibliothéconomie, les attentes des lecteurs aussi : le but d'une recherche n'est plus uniquement de vérifier si la bibliothèque a un exemplaire de tel ou tel livre, on veut savoir ce qu'elle peut proposer comme ressources sur tel ou tel sujet.

J'ai déjà analysé plusieurs bases créées avec des systèmes assez anciens, et en général, il y a des éléments à améliorer. Certaines données peuvent manquer (ex. langue de parution, indexation décimale, mots-clés), d'autres peuvent y figurer mais se trouver ensemble alors qu'elles devraient être séparées (ex. nom et prénom de l'auteur, voire même les noms et prénoms de plusieurs auteurs, tout cela dans le même champ). Pour d'autres données, c'est parfois l'inverse : j'ai déjà vu des cotes de rangement scindées en 3 champs, simplement pour faciliter l'impression des étiquettes de cote sur 3 lignes !

Je trouve préférable de compléter et de corriger tout ce qui peut l'être avant d'importer les données dans PMB.

Or certaines possibilités d'édition disponibles dans un tableur sont plus performantes ou plus faciles à mettre en œuvre que ce qu'on pourrait obtenir avec MarcEdit. Vous avez pu constater, à la lecture de la page précédente, que certains changements sont plus compliqués à opérer avec MarcEdit, et pourtant, cet éditeur est très puissant.

En outre, le passage par le tableur donne une vue d'ensemble incomparable sur la totalité de la base, et permet d'harmoniser par exemple l'indexation décimale, les cotes de rangement, les sections, de répartir des données en plusieurs champs ou au contraire les regrouper, d'ajouter des données manquantes comme les langues de publication ou la fonction de l'auteur, ou des titres de séries. On repère aussi plus vite les doublons éventuels (ex. noms d'éditeurs ou, pire encore, les numéros d'exemplaires), et on peut les corriger avant l'import. Ceci permet d'améliorer le catalogage et la recherche.

Bien sûr, il ne s'agit ici que d'un avis personnel, et certaines bases Winisis posent nettement moins de problèmes que ceux que j'ai évoqués ici.

Le choix final vous revient, de toute façon. Je me base simplement sur mon expérience de migrations antérieures, à partir de bases relativement anciennes que je pouvais récupérer en format Unimarc ou dans le format tableur, et pour lesquelles j'ai souvent préféré passer par Excel.

Même si vous employez l'autre méthode (simuler une impression, et créer ainsi un fichier texte à importer et à éditer dans le tableur), vous devrez quand même à un moment recourir à MarcEdit. Donc, il faut malgré tout passer par les deux premières étapes : exporter la base en format .iso, et convertir le fichier obtenu en format Unimarc (.mrc).

J'en reparlerai au moment opportun.

EXPORT EN FORMAT .TXT

L'export se fait via les routines d'impression. L'astuce consiste à créer un pseudo format d'impression, et à envoyer le résultat non pas vers l'imprimante, mais dans un fichier de type texte, qui sera ensuite importé et édité dans le tableur.

Je n'ai pas « inventé » cette solution, j'avais cherché du côté des formats d'impression, car je sais que c'est souvent une manière d'exporter les données, mais vu que je ne connaissais pas Winisis au départ, je ne voyais pas comment procéder.

J'ai lancé des recherches sur Internet, et comme je le fais d'habitude si les premiers résultats ne sont pas pertinents, j'ai cherché non seulement en français, mais aussi en anglais et en espagnol (en saisissant les mots « migrar de winisis »). Eureka !

J'ai trouvé une vidéo en espagnol sur YouTube : « Winisis - Migrar una base de datos Isis a un archivo de Excel ».

Un titre prometteur ... et un contenu de qualité ! Exactement ce que je cherchais.

Cette vidéo explique clairement comment créer ce format d'impression dans Winisis et comment obtenir ainsi un fichier de type texte à ouvrir ensuite dans le tableur. <http://www.youtube.com/watch?v=hksa7wNO3-s>

Je tiens ici à remercier chaleureusement BiblioConsultores, le bureau de consultants qui a réalisé cette vidéo fort utile.

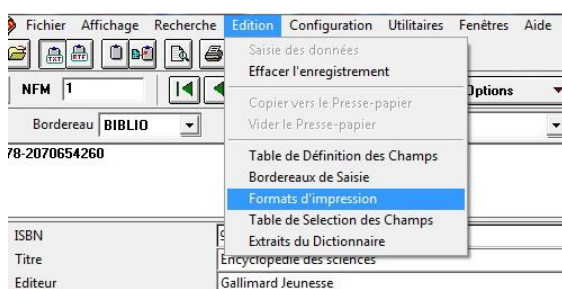
Ils sont situés à Lima, au Pérou.

<http://www.biblioconsultores.com/index.html>

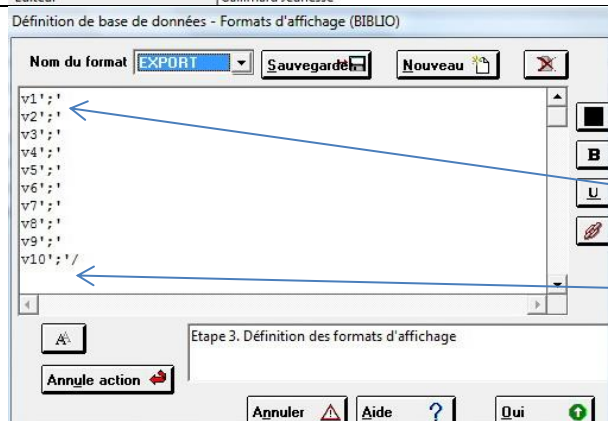
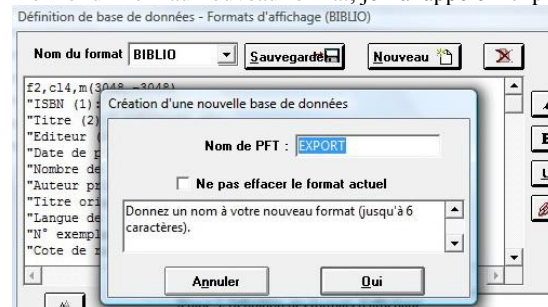
La magie d'Internet : une collaboration d'un bout à l'autre de la planète, entre des personnes qui ne se connaissent pas au départ, et qui ne se rencontreront probablement jamais, mais qui sont prêtes à partager leurs connaissances ;-)

Créer le format d'impression

Dans le menu Editions, choisir l'option « Formats d'impression »



Donner un nom au nouveau format, je l'ai appelé « export »



On y trouve la liste des champs, sous la forme v1, v2, v3 ...

Les numéros correspondent à la définition des champs dans la base Winisis (voir description de la base de démo).

v1 est donc l'ISBN, v2 le titre, v3 le nom de l'éditeur, etc.

Chaque champ est suivi d'un point-virgule inséré entre 2 apostrophes - **des apostrophes, pas des guillemets !**

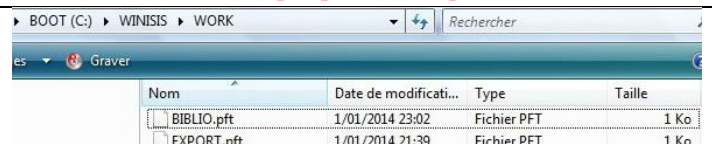
Ce point-virgule sert de séparateur de champs.

A la fin du dernier champ, comme marqueur de fin

d'enregistrement, il y a une barre oblique (elle est obligatoire).

Pas de précipitation inutile ! Ne créez pas ce format d'impression si vous n'avez pas lu le message d'avertissement ci-dessous, prenez d'abord connaissance de quelques infos importantes.

On sauvegarde ce format d'impression, il y en a désormais 2 dans le dossier WORK : le format par défaut (biblio), et celui qu'on vient de créer (export).



N.B. Il faut un séparateur de champs, mais on ne peut pas toujours employer le point-virgule, comme pour ma base de démo.

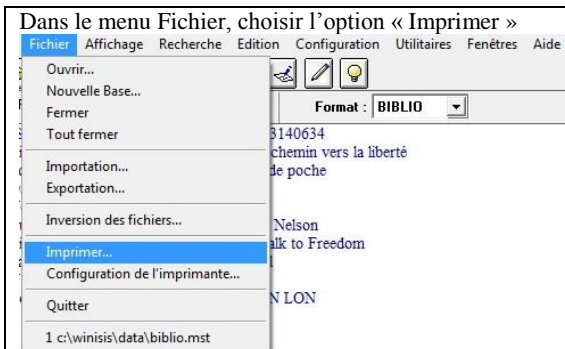
J'ai expliqué en détail quels problèmes peuvent se poser et quels critères employer pour bien choisir ce séparateur de champs.

Donc, pour éviter tout travail inutile (utiliser le point-virgule alors que cela ne conviendrait pas pour votre base), ne créez pas tout de suite votre propre format d'impression. Lisez d'abord la suite de ce chapitre, afin de prendre un « bon » séparateur !

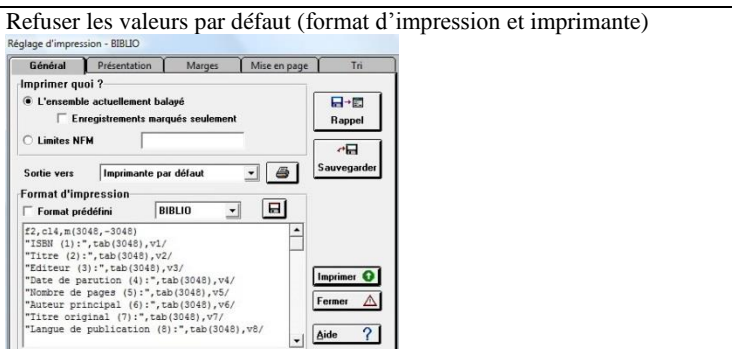
Vous reviendrez à cette page plus tard, lorsque vous serez sûr d'avoir fait le bon choix.

Imprimer la base

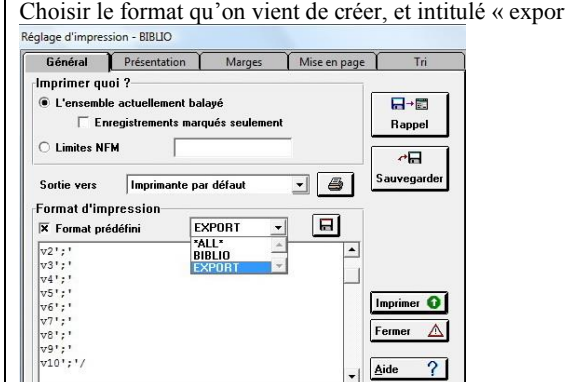
Dans le menu Fichier, choisir l'option « Imprimer »



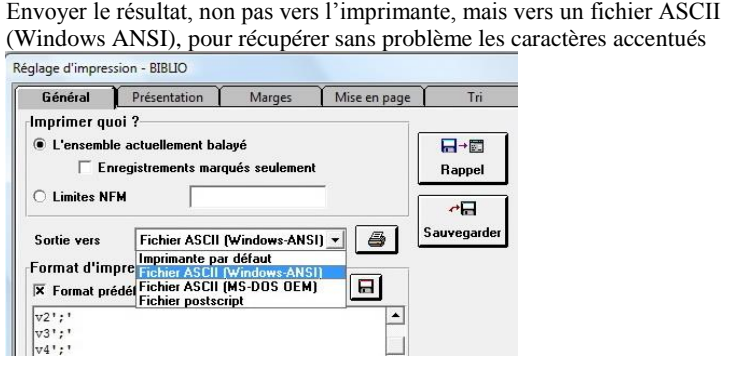
Refuser les valeurs par défaut (format d'impression et imprimante)



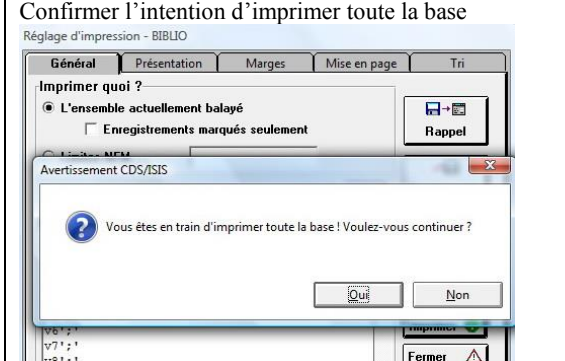
Choisir le format qu'on vient de créer, et intitulé « export »



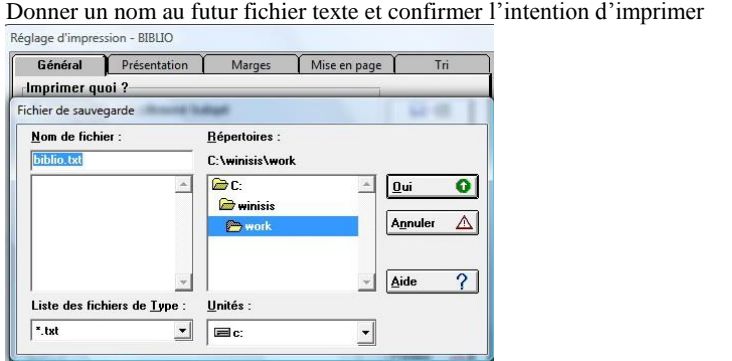
Envoyer le résultat, non pas vers l'imprimante, mais vers un fichier ASCII (Windows ANSI), pour récupérer sans problème les caractères accentués



Confirmer l'intention d'imprimer toute la base

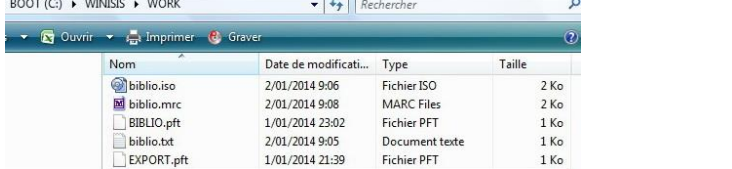


Donner un nom au futur fichier texte et confirmer l'intention d'imprimer



Il y a désormais un fichier biblio.txt dans le dossier WORK

On peut ouvrir ce fichier texte dans le bloc-notes ou dans un éditeur comme Notepad++



Chaque ligne du fichier texte correspond à un enregistrement de la base Winisis.

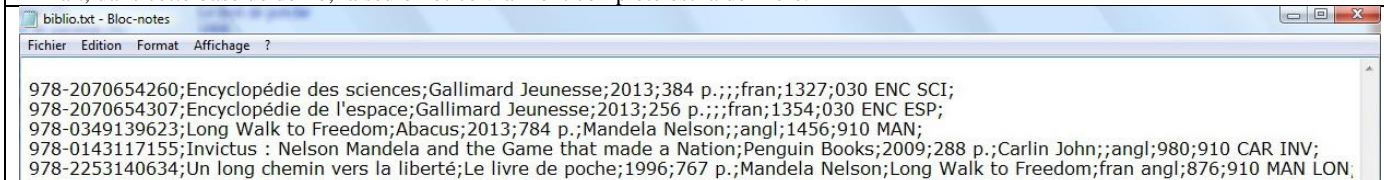
Il y a plusieurs champs par ligne, ils sont séparés par un point-virgule (c'était le caractère choisi comme séparateur de champs).

La présence de 2 ou 3 points-virgules consécutifs indique que certains champs sont vides, ils n'ont pas été complétés pour tel ou tel enregistrement (ex. à la 1^{re} ligne, entre le nombre de pages et la langue de parution).

N.B. Il est normal que certains champs restent vides pour l'une ou l'autre notice.

Exemple : un ouvrage collectif ou anonyme n'aura forcément pas de nom d'auteur ! C'est le cas des 2 premières notices (les encyclopédies).

En fait, dans cette base de démo, la seule notice vraiment complète est la dernière.



Si vous avez dû choisir comme séparateur autre chose que le point-virgule, votre fichier aura forcément un aspect un peu différent.

Bien choisir le séparateur de champs

Dans ma base de démo, le choix du séparateur de champs ne posait pas de problème particulier, car mes données ne comportaient aucun point-virgule. C'est pourquoi j'ai pris l'option la plus simple pour l'import dans le tableur, j'ai choisi le point-virgule. Mais ce n'est pas toujours possible.

Le séparateur de champs doit répondre à un 1^{er} critère vital : il ne peut pas figurer dans la base, sinon le « découpage » se fera mal, et rien ne garantit que vous le verrez tout de suite (surtout pas si le nombre de notices est élevé).

Or, la présence d'un point-virgule dans les données n'est pas si rare que cela. Voici deux exemples parmi d'autres.

- Reprenons la dernière notice de la base de démo : la version française du livre de Nelson Mandela. C'est Jean Guiloineau qui a traduit « Long Walk to Freedom ».

Le champ « auteur » aurait donc dû mentionner l'auteur principal et l'auteur secondaire (le traducteur), et il aurait pu ressembler à ceci : Mandela Nelson ; Guiloineau Jean – ou, pour respecter tout à fait les normes – Mandela, Nelson ; Guiloineau, Jean. En effet, si on mentionne plusieurs auteurs dans le même champ, on les sépare en général par un point-virgule.

- Je n'avais pas prévu de champ « mots-clés » dans la base de démo, mais si ce champ existe dans une base Winisis, il faut un séparateur entre ces mots-clés.

Il arrive souvent que des bibliothécaires emploient le point-virgule pour isoler ces mots-clés.

Même si vous pensez que le point-virgule ne figure nulle part dans vos données, il faut être vraiment sûr de son absence, sinon le « découpage » de certaines lignes du fichier posera problème, il y aura un décalage.

Au fil des années, vous avez pu oublier, et en outre vous n'êtes peut-être pas la seule personne qui catalogue ...

Cas concret - Si le champ « auteur » ressemble au départ à ceci : Mandela Nelson ; Guiloineau Jean **et** si on a choisi le point-virgule comme séparateur, ce champ « auteur » va se scinder en deux parties et la ligne (qui correspond à cette notice) comprendra 11 champs au lieu de 10

Contenu du fichier biblio.txt si la base ne contient pas de point-virgule dans les données

```
978-2070654260;Encyclopédie des sciences;Gallimard Jeunesse;2013;384 p.;;fran;1327;030 ENC SCI;
978-2070654307;Encyclopédie de l'espace;Gallimard Jeunesse;2013;256 p.;;fran;1354;030 ENC ESP;
978-0349139623;Long Walk to Freedom;Abacus;2013;784 p.;Mandela Nelson;;angl;1456;910 MAN;
978-0143117155;Invictus : Nelson Mandela and the Game that made a Nation;Penguin Books;2009;288 p.;Carlin John;;angl;980;910 CAR INV;
978-2253140634;Un long chemin vers la liberté;Le livre de poche;1996;767 p.;Mandela Nelson;Long Walk to Freedom;fran angl;876;910 MAN LON;
```

Résultat après import dans Excel si on a employé le point-virgule comme séparateur de champs

	A	B	C	D	E	F	G	H	I	J
1										
2	978-2070654260	Encyclopédie des sciences	Gallimard Jeunesse	2013	384 p.			fran	1327	030 ENC SCI
3	978-2070654307	Encyclopédie de l'espace	Gallimard Jeunesse	2013	256 p.			fran	1354	030 ENC ESP
4	978-0349139623	Long Walk to Freedom	Abacus	2013	784 p.	Mandela Nelson		angl	1456	910 MAN
5	978-0143117155	Invictus : Nelson Mandela and	Penguin Books	2009	288 p.	Carlin John		angl	980	910 CAR INV
6	978-2253140634	Un long chemin vers la liberté	Le livre de poche	1996	767 p.	Mandela Nelson	Long Walk to Freedom	fran angl	876	910 MAN LON

Aucun problème, toutes les données sont alignées correctement.

Contenu du fichier biblio.txt si la base contient déjà un point-virgule quelque part dans les données

```
978-2070654260;Encyclopédie des sciences;Gallimard Jeunesse;2013;384 p.;;fran;1327;030 ENC SCI;
978-2070654307;Encyclopédie de l'espace;Gallimard Jeunesse;2013;256 p.;;fran;1354;030 ENC ESP;
978-0349139623;Long walk to Freedom;Abacus;2013;784 p.;Mandela Nelson;;angl;1456;910 MAN;
978-0143117155;Invictus : Nelson Mandela and the Game that made a Nation;Penguin Books;2009;288 p.;Carlin John;;angl;980;910 CAR INV;
978-2253140634;Un long chemin vers la liberté;Le livre de poche;1996;767 p.;Mandela Nelson;Guiloineau Jean;Long Walk to Freedom;fran angl;876;910 MAN LON;
```

Il y a un point-virgule en trop !

Résultat après import dans Excel si on a employé le point-virgule comme séparateur de champs

	A	B	C	D	E	F	G	H	I	J	K
1											
2	978-2070654260	Encyclopédie des sciences	Gallimard Jeunesse	2013	384 p.			fran	1327	030 ENC SCI	
3	978-2070654307	Encyclopédie de l'espace	Gallimard Jeunesse	2013	256 p.			fran	1354	030 ENC ESP	
4	978-0349139623	Long Walk to Freedom	Abacus	2013	784 p.	Mandela Nelson		angl	1456	910 MAN	
5	978-0143117155	Invictus : Nelson Mandela and	Penguin Books	2009	288 p.	Carlin John		angl	980	910 CAR INV	
6	978-2253140634	Un long chemin vers la liberté	Le livre de poche	1996	767 p.	Mandela Nelson	Guiloineau Jean	Long Walk to Freedom	fran angl	876	910 MAN LON

Pas de souci pour les 4 premières notices, les données sont alignées correctement. Les ennuis commencent à la ligne 6 (la 5^e notice) !

L'ajout d'un nom d'auteur dans la colonne G de la 5^e notice (alors qu'il aurait dû être dans la colonne F) a provoqué un problème.

Il y a un décalage des autres champs de cette notice : le titre original est dans H au lieu de G, les langues dans I au lieu de H, etc.

Le séparateur de champs doit répondre à un 2^e critère tout aussi important que le 1^{er} : si on doit employer autre chose que le point-virgule, il faut que ce ne soit pas un caractère « réservé » dans Winisis, un caractère qui aurait une fonction spéciale dans ce logiciel.

Vu que je ne connais pas bien Winisis, il m'est difficile de donner une liste plus ou moins complète et exacte des caractères réservés, mais en voici déjà certains qu'on ne peut en général pas employer comme séparateur de champs :

- les symboles mathématiques + - / * ^ et d'autres symboles comme # \ \$
- l'apostrophe, les guillemets, les parenthèses et les accolades { }
- probablement aussi les symboles & @ et % (pour ces derniers, je n'en suis pas tout à fait sûre)

Évitez aussi la plupart des signes de ponctuation : je ne crois pas qu'ils aient tous une fonction spéciale dans Winisis, mais on risque d'en trouver dans les données, entre autres à cause des normes de l'ISBD.

Ex. Si les données des auteurs ou les mentions d'édition se trouvent dans le même champ, on écrit ceci

auteur1, prenom1 ; auteur2, prenom2 ; auteur3, prenom3

ville : nom de l'éditeur, date de publication → Ceci exclut la virgule, le point-virgule et les deux points comme séparateurs

Voici deux caractères que j'ai déjà testés et qui ne posent pas de problème : le risque qu'ils soient présents dans les données est quasi nul, et Winisis les accepte comme séparateurs.

- **le crochet ouvrant ou fermant []**

Attention ! Les crochets conviennent, mais pas les accolades { }

N.B. Je crois que le trait de soulignement _ convient aussi.

Attention ! A ne pas confondre avec le signe moins qui est interdit !

C'est pourquoi je conseille la prudence et une vérification grâce à MarcEdit. Il faut être sûr de l'absence, dans votre base, de point-virgule ou de tout autre caractère que vous auriez choisi comme séparateur de champs.

Tester la validité du séparateur de champs

J'avais signalé que, même si on exporte les données grâce aux routines d'impression, il fallait quand même exporter la base en format .iso et la convertir en un pseudo fichier Unimarc, et que je vous inviterais à réutiliser MarcEdit.

Le moment de retourner à MarcEdit est venu : il vous permettra de vérifier si le point-virgule (ou le caractère choisi comme séparateur de champs) est vraiment absent de votre base.

N.B. Je suppose que vous disposez déjà de votre base en 2 formats : la version .iso et le fichier.mrc obtenu grâce à MarcEdit. Relire si nécessaire le chapitre précédent : Export en format .iso

Dans ce cas-ci, le but est simplement d'éditer un fichier .mrc, et on n'a pas besoin des outils supplémentaires comme la conversion de CDS-ISIS vers Marc.

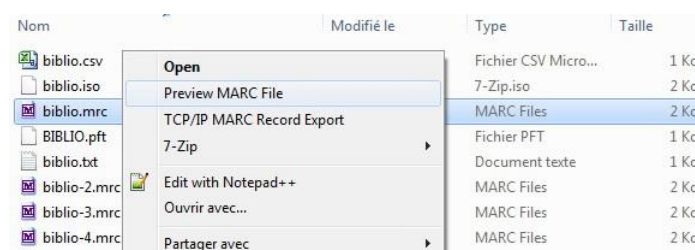
Donc il n'est pas nécessaire de lancer MarcEditor, il y a une autre manière plus rapide d'ouvrir un fichier .mrc.

On passe par le menu contextuel de l'explorateur Windows

« Parcourir » le disque dur et sélectionner le fichier biblio.mrc

Clic droit de souris pour ouvrir le menu contextuel

Choisir la 2^e option « Preview Marc File », le fichier s'ouvre immédiatement dans l'éditeur.



Attention lorsque vous éditez un fichier .mrc assez long et que vous voulez faire une vérification globale !

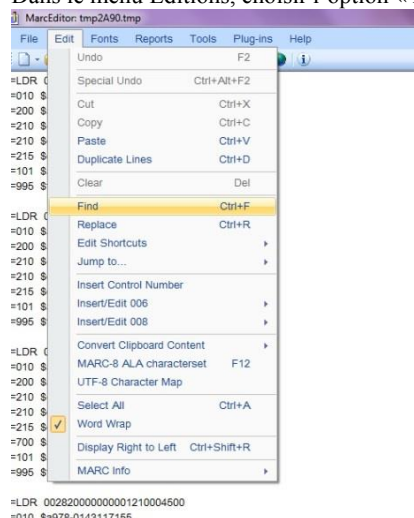
Si le fichier est volumineux, au début du processus, MarcEdit ne charge en mémoire qu'une partie du fichier.



Ce message d'avertissement n'est pas anodin.

Pensez à cliquer sur « Load Entire File », sinon vous risquez de croire que vous avez testé tout le fichier, alors que vous n'avez vérifié qu'une partie de celui-ci.

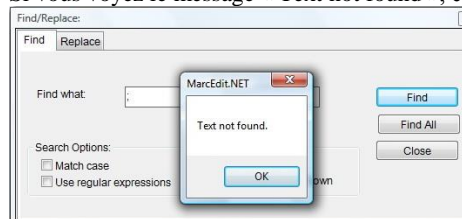
Dans le menu Editions, choisir l'option « Find »



Vérifier si le caractère choisi comme séparateur de champs est vraiment absent. Dans ce cas, j'ai vérifié s'il y avait un point-virgule.



Si vous voyez le message « Text not found », c'est bon signe ! Aucun point-virgule !



Dès que vous aurez trouvé un « bon » séparateur de champs, non employé dans votre base et sans fonction spéciale dans Winisis, vous pourrez enfin créer le format d'impression pour lequel je vous avais conseillé de patienter !!

Relire les chapitres précédents si nécessaire : créer le format d'impression et imprimer la base. Vous obtiendrez ainsi un fichier texte prêt à l'emploi.

IMPORT D'UN FICHIER TEXTE DANS LE TABLEUR

Deux cas peuvent se présenter, suivant le type de séparateur employé.

Contenu du fichier biblio-a.txt : on a employé le crochet ouvrant comme séparateur de champs

```
978-2070654260[Encyclopédie des sciences[Gallimard Jeunesse[2013[384 p.]][[fran[1327[030 ENC SCI[
978-2070654307[Encyclopédie de l'espace[Gallimard Jeunesse[2013[256 p.]][[fran[1354[030 ENC ESP[
978-0349139623[Long Walk to Freedom[Abacus[2013[784 p.][Mandela Nelson[angl[1456[910 MAN[
978-0143117155[Invictus : Nelson Mandela and the Game that made a Nation[Penguin Books[2009[288 p.][Carlin John[angl[980[910 CAR INV[
978-2253140634[Un long chemin vers la liberté[Le livre de poche[1996[767 p.][Mandela Nelson;Guiloinéau Jean[Long Walk to Freedom[fran angl[876[910 MAN LON]
```

Présence d'un point-virgule dans les données

Contenu du fichier biblio-b.txt : on a employé le point-virgule comme séparateur de champs

```
978-2070654260;Encyclopédie des sciences;Gallimard Jeunesse;2013;384 p.;;;fran;1327;030 ENC SCI;
978-2070654307;Encyclopédie de l'espace;Gallimard Jeunesse;2013;256 p.;;;fran;1354;030 ENC ESP;
978-0349139623;Long Walk to Freedom;Abacus;2013;784 p.;Mandela Nelson;;angl;1456;910 MAN;
978-0143117155;Invictus : Nelson Mandela and the Game that made a Nation;Penguin Books;2009;288 p.;Carlin John;;angl;980;910 CAR INV;
978-2253140634;Un long chemin vers la liberté;Le livre de poche;1996;767 p.;Mandela Nelson;Long Walk to Freedom;fran angl;876;910 MAN LON;
```

Pas de point-virgule dans les données

2 méthodes en fonction du séparateur

D'autres caractères que le point-virgule sont souvent employés comme séparateurs dans ce genre de fichiers de type texte, comme la tabulation ou la virgule.

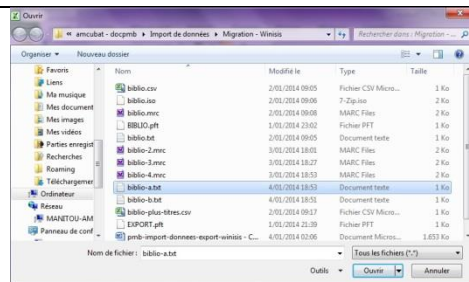
Certes, mon exemple est un peu déconcertant, il n'est sûrement pas fréquent de choisir un crochet ouvrant comme séparateur, mais dans le cas présent, je devais éviter une série de caractères réservés dans Winisis, donc les possibilités de choix étaient plus limitées que d'habitude.

La première méthode que je vais expliquer convient en fait pour tous les séparateurs de champs, le point-virgule y compris, c'est pourquoi je commence par celle-là, elle est universelle.

Le cas où le séparateur est un point-virgule est spécial : on peut parfois se permettre d'emprunter un « petit raccourci » qui accélère le processus. Mais ce n'est pas obligatoire, et ce n'est même pas toujours possible, ce qui vous oblige de temps en temps à recourir malgré tout à la 1^{re} méthode.

Ouvrir le fichier biblio-a.txt dans Excel

Il correspond au cas où le séparateur est un crochet ouvrant.

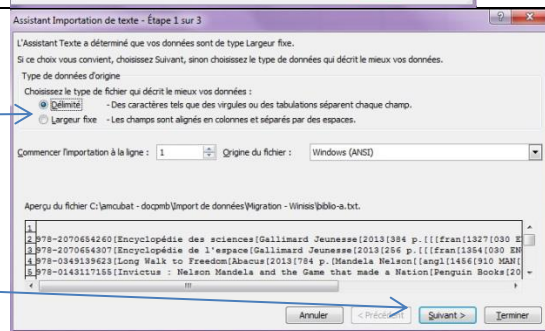


On ne choisit bien sûr pas les champs de largeur fixe, cela ne convient pas au contenu du fichier qu'on a créé.

Il faut prendre l'option « champs délimités par un caractère »

Excel cite 2 exemples de séparateurs (tabulation et virgule), mais signale aussi qu'il y en a d'autres.

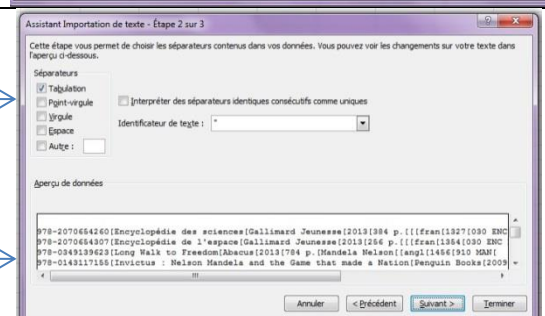
On clique sur « suivant ».



La valeur sélectionnée par défaut (tabulation) ne convient pas, ni les autres options de base (point-virgule, virgule et espace).

N.B. Par contre, si on avait ouvert le fichier biblio-b.txt (celui où le séparateur est un point-virgule), la 2^e option prédéfinie (le point-virgule) aurait convenu.

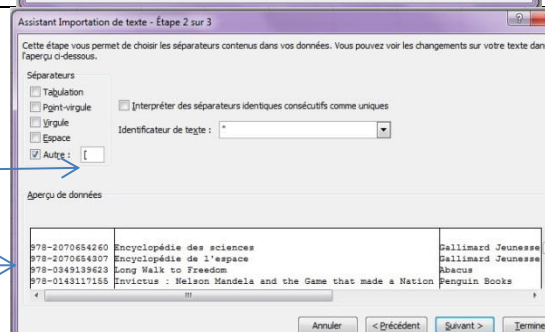
On voit bien que le contenu du fichier n'a pas encore changé.



On coche « autre », et on introduit le crochet ouvrant dans la petite zone de saisie, afin de signaler que ce sera le séparateur de champs.

L'effet est immédiat : on a un aperçu de ce que le fichier deviendra, après la répartition des données dans les colonnes.

On clique sur « suivant ».

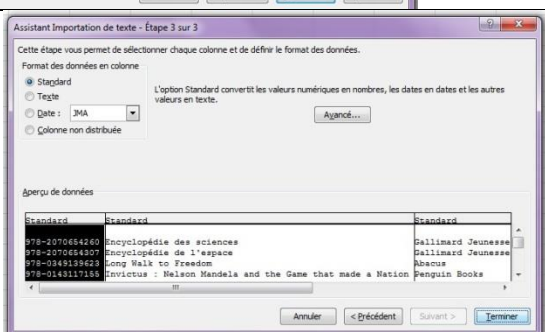


Dans ce cas-ci, on peut garder le format standard pour toutes les colonnes, inutile donc de s'attarder dans cet écran. On clique sur « Terminer ».

N.B. Il faut choisir le format texte pour certaines colonnes, si vous avez dans cette colonne des nombres comme « 00015, 00184, 01224 » (cela peut arriver pour les numéros d'exemplaires), et si vous voulez préserver le zéro initial.

Et voilà le résultat final : les données sont scindées automatiquement, sur base du séparateur choisi (le crochet ouvrant).

Notez que la présence d'un point-virgule dans la colonne F (entre les 2 noms d'auteurs) n'a posé aucun problème



	A	B	C	D	E	F	G	H	I	J
1										
2	978-2070654260	Encyclopédie des sciences	Gallimard Jeunesse	2013	384 p.			fran	1327	030 ENC SCI
3	978-2070654307	Encyclopédie de l'espace	Gallimard Jeunesse	2013	256 p.			fran	1354	030 ENC ESP
4	978-0349139623	Long Walk to Freedom	Abacus	2013	784 p.	Mandela Nelson		angl	1456	910 MAN
5	978-0143117155	Invictus : Nelson Mandela and Penguin Books	2009	288 p.	Carlin John			angl	980	910 CAR INV
6	978-2253140634	Un long chemin vers la liberté Le livre de poche	1996	767 p.	Mandela Nelson;Guiloleineau Jean	Long Walk to Freedom		fran angl	876	910 MAN LON

Vous avez pu constater que cette première méthode permet d'ouvrir le fichier texte, et de répartir facilement les données dans les diverses colonnes, quel que soit le séparateur choisi : point-virgule, virgule, tabulation, crochet, etc. Elle convient donc dans tous les cas.

Toutefois, il y a un « petit raccourci » utilisable parfois lorsque le séparateur est un point-virgule.

Il existe un format informatique appelé .CSV (Comma-Separated Values, des valeurs séparées par des virgules). C'est une variante des fichiers de type texte, qu'on peut lire dans un éditeur comme le bloc-notes ou Notepad++

Ces fichiers ont été créés au départ pour échanger simplement des données entre des logiciels différents : aucun formatage spécial des données, elles sont transférées « à l'état brut », séparées simplement par une virgule.

Ces fichiers servent souvent dans le cas de PMB et d'autres logiciels analogues, ou aussi pour les CMS (qui gèrent un site Web), car on ne peut importer des données dans des bases MySQL qu'en employant des fichiers .csv, travailler directement avec des fichiers au format tableur est impossible.

Malheureusement, ce format .csv n'est pas tout à fait normalisé !

Dans certains cas, c'est la virgule qui sert de séparateur (d'où son nom), dans d'autre cas, c'est le point-virgule.

- Si votre ordinateur est configuré en français, le séparateur de champs (appelé aussi séparateur de listes) est par défaut un point-virgule, et le séparateur entre la partie entière et la partie décimale d'un nombre est par défaut la virgule.
- Si votre ordinateur est configuré en anglais ou en américain, le séparateur de champs (appelé aussi séparateur de listes) est par défaut une virgule, et le séparateur entre la partie entière et la partie décimale d'un nombre est par défaut le point.

En fait, le séparateur de listes (ou séparateur de champs) fait partie des options régionales que vous pouvez modifier grâce au panneau de configuration de Windows : langue, pays, format des nombres, de l'heure ...

Donc, avec un ordinateur configuré en français, si vous avez gardé les valeurs par défaut,

- Vous pouvez exporter vos fichiers Excel en format .csv (enregistrer sous – choisir le format .csv). Les fichiers .csv que vous créez de cette manière auront automatiquement le point-virgule comme séparateur de champ
- Vous pouvez importer automatiquement dans Excel les fichiers .csv créés avec un point-virgule comme séparateur.

Importer automatiquement un fichier .csv dans le tableur veut dire que vous ne devez pas passer par le processus décrit à la page précédente : il ne faut pas spécifier le séparateur de champs, les données se répartissent automatiquement dans les diverses colonnes dès l'ouverture du fichier dans Excel. Il y a donc un léger gain de temps.

Mais cela ne fonctionne que si l'extension du fichier est .csv. Or, au stade actuel, le fichier que vous avez extrait de votre base Winisis est encore en format .txt. Il faut donc faire comme si vous alliez le renommer, afin de changer cette extension.

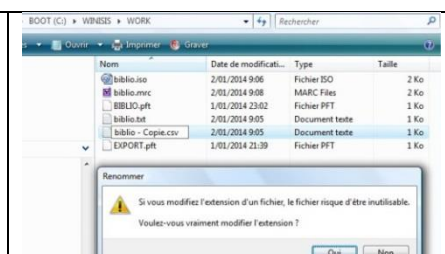
N.B. Si vous avez dans vos données des nombres comme « 00015, 00124, 01184 » (cela peut arriver pour les numéros d'exemplaires), et si vous voulez préserver ces zéros au début, ne changez pas l'extension du fichier, sinon Excel va automatiquement considérer ces valeurs comme des nombres et supprimer le zéro initial.

Voir page suivante

J'ai d'abord créé une copie de mon fichier biblio.txt, et j'ai changé l'extension de la copie.
N.B. J'aurais aussi pu renommer l'original, mais vu que je voulais garder la version d'origine en format .txt, j'ai préféré illustrer le processus avec une copie de ce fichier.

biblio – Copie.txt est donc devenu biblio – Copie.csv

Comme vous le voyez, le format .csv est un cas particulier, une variante des fichiers .txt. Changer l'extension accélère l'ouverture du fichier dans le tableur, mais cela ne change rien au format des données : fondamentalement, cela reste un fichier de type texte. D'ailleurs, malgré le message d'avertissement (« changer l'extension d'un fichier risque de le rendre inutilisable »), vous pourrez continuer à exploiter le contenu de ce fichier !



Ce « raccourci » (modifier l'extension du fichier) ne fonctionne donc pas en toute circonstance.

C'est pourquoi j'ai préféré mettre l'accent sur la 1^{re} méthode qui vous fournira le résultat escompté, quelle que soit la configuration des options régionales sur votre PC et quelles que soient les données (un zéro initial à préserver ?)

Présence d'un zéro initial

Dans certaines bases, on a parfois ajouté un ou plusieurs zéros au début des numéros d'exemplaires, afin qu'ils comportent tous le même nombre de chiffres. C'est le cas dans l'exemple ci-dessous : 001327, 001354, 001456, 000980, 000878

Contenu du fichier biblio-c.txt : on a employé le point-virgule comme séparateur de champs

```
978-2070654260;Encyclopédie des sciences;Gallimard Jeunesse;2013;384 p.;;;fran;001327;030 ENC SCI;
978-2070654307;Encyclopédie de l'espace;Gallimard Jeunesse;2013;256 p.;;;fran;001354;030 ENC ESP;
978-0349139623;Long Walk to Freedom;Abacus;2013;784 p.;Mandela Nelson;;angl;001456;910 MAN;
978-0143117155;Invictus : Nelson Mandela and the Game that made a Nation;Penguin Books;2009;288 p.;Carlin John;;angl;000980;910 CAR INV;
978-2253140634;Un long chemin vers la liberté;Le livre de poche;1996;767 p.;Mandela Nelson;Long Walk to Freedom;fran angl;000878;910 MAN LON;
```

Pas de point-virgule dans les données - Les numéros d'exemplaires comportent un zéro ou plusieurs zéros au début

La question se pose dès lors : faut-il préserver ce zéro initial ?

Dans l'absolu, je dirais « non », la présence d'un zéro initial pose toujours des problèmes, surtout pour le publipostage. C'est d'ailleurs une des règles à respecter pour le choix des numéros de lecteurs ou d'exemplaires, l'autre est d'éviter des numéros dont le nombre de chiffres est variable (pour faciliter un tri correct en toute circonstance).

<http://amcubat.be/docpmb/choisir-numeros-exemplaires-lecteurs>

En effet, pour imprimer dans PMB, on transite souvent par le tableur, or Excel a tendance à considérer tout ce qui ressemble à un nombre comme un nombre, donc si la cellule a un format « standard », Excel supprime d'office le zéro initial (qui est en effet superflu d'un point de vue mathématique).

Toutefois, si vous avez sur vos livres des étiquettes avec des numéros d'exemplaires imprimés sous forme de codes-barres, il faut préserver le zéro initial – du moins provisoirement – et changer les numéros petit à petit dans PMB

N.B. J'ai expliqué ci-dessous, dans le contexte de la solution n°2, comment changer petit à petit des numéros de codes-barres.

Si vous perdiez le zéro initial, le numéro lu avec la douchette (ex. 001327) ne correspondrait plus à ce qui est enregistré dans la base (1327), et PMB vous dirait que cet exemplaire est inconnu ! (La douchette est le lecteur de code-barres)

L'idéal, si vous n'avez pas d'étiquettes avec des codes-barres, est de profiter du changement de logiciel pour normaliser vos numéros d'exemplaires : ils auront par exemple tous 6 chiffres, et n'auront pas de zéro initial.

Donc, vous importez les numéros sans prendre de précautions particulières, et vous perdez le(s) zéro(s) du début.

2 solutions - A vous de choisir ...

N.B. La 2e solution est plus simple à mettre en œuvre

1) Vous modifiez les numéros dans Excel, vous importez dans PMB les nouveaux numéros.

- Lorsque les données sont dans le tableur, vous demandez à Excel de faire une simple addition, et d'ajouter 100.000 à tous vos numéros ! Le n° 1327 deviendra donc le 101327, le n° 876 deviendra 100876, etc. Vous importez donc les nouveaux numéros.
- **Vous pré-imprimez une série d'étiquettes** avec les nouveaux numéros de codes-barres, ex. 101327, 101456, 100876.
- **De toute façon, vous devrez probablement reprendre une partie de vos livres en main, afin de compléter le catalogage.** Donc, dès que vous en avez l'occasion, vous collez la nouvelle étiquette.
- **Rien d'autre à faire, le changement de numéro a déjà eu lieu avant l'import, donc le numéro dans PMB est correct.**
- **La concordance entre ancien et nouveau n° est évidente**, on sait tout de suite quel numéro tel ou tel exemplaire aura, mais il y a **beaucoup plus de manipulations à faire**, il faut retrouver la feuille A4 qui contient la bonne étiquette de code-barres !
- **Donc cette solution convient nettement moins bien si le nombre d'exemplaires de l'ancienne base est assez élevé et si le personnel est réduit – vous ne pouvez pas tout faire tout seul ...**

2) Vous ne changez rien dans Excel, vous importez dans PMB les anciens numéros mais sans le zéro initial.

- **Vous pré-imprimez une série d'étiquettes** avec des numéros de codes-barres corrects, ex. 100001, 100002, 100102, 100245 ..., qui n'ont donc aucun lien avec le numéro d'exemplaire précédent.
- **Vos anciens numéros sont temporairement un rien problématiques** (ils ont un nombre variable de chiffres, mais au moins, ils ont déjà perdu le zéro initial, ce qui est une bonne chose en soi), **par contre les nouveaux numéros que vous allez attribuer dans PMB ne posent aucun problème.**
- **De toute façon, vous devrez probablement reprendre une partie de vos livres en main, afin de compléter le catalogage.** Donc, dès que vous en avez l'occasion, vous collez la nouvelle étiquette (ne cachez quand même pas l'ancien numéro !)
- **Vous modifiez ensuite le numéro d'exemplaire dans la base, donc le changement de numéro se fait petit à petit dans PMB**
- **Cette solution est envisageable même si le nombre d'exemplaires de l'ancienne base est élevé, et même si on a déjà collé des étiquettes de codes-barres, mais avec d'anciens numéros inadéquats (zéro initial et/ou nombre variable de chiffres)**

Dernière possibilité que je déconseille bien entendu*Vous pouviez vous en douter ☺***Vous ne changez rien dans le tableur, vous importez dans PMB les anciens numéros – avec ou sans le zéro initial.**

- Vous ne pré-imprimez pas d'étiquettes avec des numéros de codes-barres corrects, rien ne change dans la gestion.
- Vos numéros d'exemplaires restent problématiques, ils ont un nombre variable de chiffres et/ou un zéro initial.

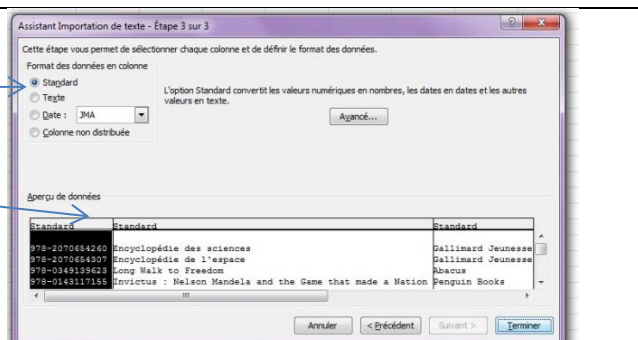
Préserver si nécessaire le zéro initial**Le fichier doit garder l'extension .txt, ne pas changer son extension en.csv, sinon le zéro va disparaître d'office !**

J'avais écrit ceci en expliquant comment ouvrir le fichier biblio-a .txt dans Excel.

Dans le cas de ma base de démo, on peut garder le format standard pour toutes les colonnes, inutile donc de s'attarder dans cet écran.

On clique sur « Terminer ».

En effet, dans cette base de démo, il n'y a aucun zéro initial dans les numéros d'exemplaires.



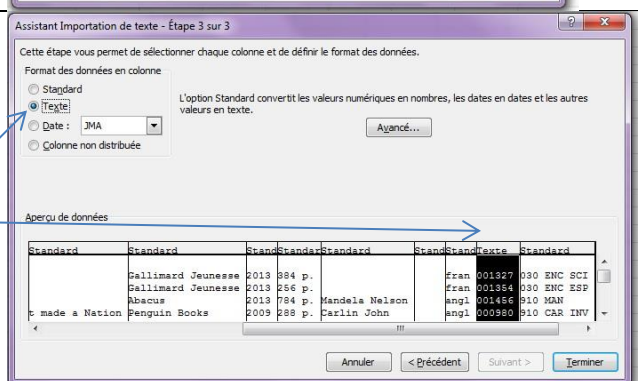
Par contre, dans le fichier biblio-c.txt, l'avant-dernière colonne comporte des zéros au début des numéros d'exemplaires.

On ne peut donc pas garder le format « standard » pour toutes les colonnes, car il « convertit les valeurs numériques en nombres ».

Il faut sélectionner la colonne qui contient les numéros d'exemplaires, et spécifier qu'elle doit avoir le format texte.

On peut alors cliquer sur « Terminer ».

Cette étape intermédiaire est indispensable, car un zéro initial est non-significatif en mathématique, il sera donc supprimé par Excel.

**Système de numérotation et codes-barres**

La douchette accélère la saisie des données et assure leur fiabilité : pas de risque d'erreurs liées aux fautes de frappe ! Même si vous n'avez pas de douchette actuellement et si vous pensez qu'un achat n'est pas possible dans l'immédiat, le changement de logiciel est le moment idéal pour imprimer et coller des étiquettes de codes-barres sur vos livres.

La migration, c'est le moment où vous allez devoir reprendre au moins une partie de votre fonds en main, pour compléter le catalogage, pour réorganiser les sections, ou simplement pour vérifier si cela vaut la peine de conserver tel ou tel livre.

La douchette viendra sûrement beaucoup plus vite que vous ne le pensez, les prix ont baissé. Si vous ne collez pas les étiquettes maintenant, à un moment où vous devez fatalement manipuler une partie de vos livres, vous allez « râler » dans quelques mois, lorsque vous devrez ressortir tous les livres des rayonnages, tout simplement parce que vous aurez reçu une douchette et qu'il faudra donc apposer les étiquettes de codes-barres que vous refusez de coller aujourd'hui ...

Par conséquent, la migration est aussi le bon moment pour réfléchir à votre système de numérotation des exemplaires. Si vous n'y pensez que dans un an ou deux, cela vous prendra plus de temps que si vous le faites maintenant, à cause des manipulations supplémentaires que cela va engendrer. Le pire, à long terme, c'est que vous allez continuer dans un système peu adéquat, les numéros des nouveaux livres que vous cataloguerez seront eux aussi problématiques.

Infos complémentaires

<http://amcubat.be/docpmb/choisir-numeros-exemplaires-lecteurs>

<http://amcubat.be/docpmb/codes-barres-concepts-base>

<http://amcubat.be/docpmb/publipostage-notions-base>

<http://amcubat.be/docpmb/avantages-logiciel-specifique-etiquettes>

<http://amcubat.be/docpmb/impression-etiquettes-codes-barres>

ET ENSUITE ?

Les données de la base d'origine sont désormais dans le tableur, mais ce n'est pas fini !

	A	B	C	D	E	F	G	H	I	J
1										
2	978-2070654260	Encyclopédie des sciences	Gallimard Jeunesse	2013	384 p.			fran	1327 030	ENC SCI
3	978-2070654307	Encyclopédie de l'espace	Gallimard Jeunesse	2013	256 p.			fran	1354 030	ENC ESP
4	978-0349139623	Long Walk to Freedom	Abacus	2013	784 p.	Mandela Nelson		angl	1456 910	MAN
5	978-0143117155	Invictus : Nelson Mandela and	Penguin Books	2009	288 p.	Carlin John		angl	980 910	CAR INV
6	978-2253140634	Un long chemin vers la liberté	Le livre de poche	1996	767 p.	Mandela Nelson	Long Walk to Freedom	fran angl	876 910	MAN LON

Il y a encore beaucoup de choses à faire, beaucoup de sujets à aborder, en voici un bref aperçu :

- Travail supplémentaire dans le tableur : réorganiser les données, les trier, les corriger, les adapter aux normes Unimarc, ajouter des éléments absents dans la base de départ mais utiles dans PMB
- Ne pas oublier de faire des sauvegardes intermédiaires !
- Export d'une feuille du tableur en format .csv : précautions à prendre
- Préparation de l'import : créer un fichier de concordance (params.xml), qui établit un lien entre chaque colonne du tableur et le code Unimarc correspondant (ex. dans la colonne 1, c'est l'ISBN, donc le code Unimarc 010 \$a)
- Import dans PMB : l'import direct n'est pas possible, car le fichier de départ n'est pas en format Unimarc ISO 2709, il est en format texte, donc on passe par un utilitaire de conversion (inclus dans PMB) qui va transformer ce fichier en un véritable fichier Unimarc qu'on pourra ensuite importer dans la base.

Je sais que je vais déclencher un phénomène de frustration, et pourtant je m'arrête à ce stade-ci des explications ☺

La raison en est toute simple.

Winisis n'est pas un cas unique, il y a d'autres catalogues qui ont été créés avec des logiciels qui ne sont pas conçus au départ pour gérer une bibliothèque, comme Access ou FileMaker, et qui impliquent aussi – en cas de migration - un transfert des données vers Excel ou OpenOffice Calc.

Même si on avait employé un logiciel spécifique de gestion de bibliothèque, les versions les plus anciennes n'offraient en général qu'une ou deux possibilités d'export : le format texte et/ou le format tableur.

En outre, de nombreux logiciels propriétaires bloquent autant que possible la migration de données vers un autre logiciel, l'export en Unimarc ou en .xml n'est par exemple pas toujours possible, ou bien c'est un Unimarc un peu différent, pas tout à fait normalisé (surtout pour les infos relatives aux exemplaires, ce qui est fort gênant).

Même Winisis, qui a pourtant été distribué gratuitement par l'Unesco, enregistre les données dans un format propriétaire et n'envisage une migration que de Winisis vers Winisis, ou de Winisis vers ABCD (le logiciel qui remplace Winisis).

Réaliser une migration en transitant par le tableur est donc assez fréquent.

En effet, même si d'autres modes d'exportation sont parfois disponibles (Unimarc), le passage par le tableur permet plus facilement d'améliorer et d'enrichir la base avant l'import dans PMB : corriger des erreurs, supprimer des doublons, scinder des données en plusieurs champs ou au contraire les regrouper en un seul, ajouter des éléments qui manquaient dans la version de départ mais qui sont utiles dans PMB, adapter la base aux normes bibliographiques actuelles, etc.

C'est pourquoi j'ai consacré des articles sur mon site à l'import de notices et d'exemplaires dans PMB à partir du format tableur, quel que soit le logiciel de départ.

Malgré les possibilités immenses offertes par le copier-coller, je n'ai quand même pas l'intention de répéter ici tout ce que j'ai déjà expliqué ailleurs ! Vous trouverez donc la suite des explications dans un autre document.

Et voici la suite : **Méthodologie - Importer des notices ou des exemplaires à partir d'un tableur**

<http://amcubat.be/docpmb/pmb-methodeologie-importer-notices-exemplaires-tableur>

N.B. Bien sûr, il y a quelques parties de chapitres identiques, je ne peux pas faire autrement !

RESSOURCES COMPLEMENTAIRES

Un zip que vous pouvez télécharger : pmb-exporter-donnees-winisis.zip

Ce zip contient tous les fichiers qui ont servi d'exemples dans ce tutoriel

- les fichiers .iso (format Winisis) et les fichiers .pft (format d'impression de Winisis) du dossier WORK de Winisis
- les fichiers .mrc (format Unimarc – à ouvrir avec MarcEdit)
- les fichiers .txt et .csv
- un fichier .xlsx (à ouvrir avec Excel ou OpenOffice Calc)

Pour éditer des fichiers en format Unimarc ou Marc 21 - Télécharger et installer MarcEdit

<http://marcedit.reeset.net/>

Le guide en ligne de PMB

L'import dans PMB se fait de 2 manières

- Si le fichier de départ est en format Unimarc ISO 2709
http://doc.sigb.net/pmb/co/admin_import_import.html
- Si le fichier de départ a un autre format (par exemple un tableur converti en format .csv, c'est-à-dire un fichier de type texte), on emploie d'abord un utilitaire (inclus dans PMB) qui convertit ce fichier en Unimarc et on l'importe ensuite dans PMB.
http://doc.sigb.net/pmb/co/administration_conversions_exports.html

Pour connaître les codes Unimarc employés dans PMB et qui correspondent aux champs de votre base

http://doc.sigb.net/pmb/co/administration_imports_references.html

Mon site personnel : la page générale sur l'import de données

<http://amcubat.be/docpmb/import-de-donnees>

Méthodologie : importer des notices ou des exemplaires à partir d'un tableur

<http://amcubat.be/docpmb/pmb-methodologie-importer-notices-exemplaires-tableur>

C'est donc la suite de ce document-ci : après l'export à partir de Winisis, l'import dans PMB.

Ce tutoriel très complet décrit toutes les étapes du processus.

L'étape suivante

<http://amcubat.be/docpmb/import-exemplaires-notices-monographies>

Ce document est surtout consacré à l'import des exemplaires et au travail « de fignotage » dans la base après l'import

- Données des exemplaires : code-barres, cote de rangement, localisation, section, code statistique, statut, support, message
- Exemple de fichier params.xml (le fichier de concordance entre les données du tableur et les codes Unimarc)
- Emploi des codages d'import afin que certaines valeurs soient fixées dans le fichier d'import (ex. section, code statistique)
- Import d'exemplaires multiples de la même notice (possible même sans numéro d'ISBN, grâce à l'astuce des ISBN fictifs)
- Traitement par lot, rôle des paniers et des requêtes SQL de sélection, de pointage et d'action afin de corriger des erreurs éventuelles ou de fignoter le catalogage
- Exemples de traitement par lot : changer facilement la localisation, la section, le code statistique, le statut ou le support - soit de tous les exemplaires qu'on vient d'importer, soit seulement de certains d'entre eux (sélectionnés par pointage)

Choix du script d'import – Les 2 variantes de func_belgique

<http://amcubat.be/docpmb/func-belgique-import-notices-exemplaires>

Si vous n'êtes pas familiarisé avec les concepts de paniers et de traitement par lot

<http://amcubat.be/docpmb/formation-pmb/paniers-traitement-par-lot-sql>

Notes techniques sur les fichiers .csv

<http://amcubat.be/docpmb/notes-techniques-format-csv>

N.B. Je rappelle que ces liens pointent vers des pages qui fournissent les explications de base.

Pour avoir tous les détails, il faut lire la version imprimable, le pdf que vous pouvez également imprimer et sauvegarder.

En outre, il y a parfois aussi un zip avec des fichiers de démo, des exemples. Pensez à le télécharger et à exploiter son contenu.

Source : <http://amcubat.be/docpmb/pmb-migration-exporter-donnees-winisis>