

StarSafe Forms

StarSafe est un outil de gestion et d'archivage des données dont l'objectif est de pérenniser l'information avec la possibilité de la restituer de manière complètement identique à l'origine.

Il permet d'archiver les documents, et cela, quelque soit leur format, c'est à dire qu'il assure le stockage, la recherche et la consultation en ligne de ces documents.

StarSafe Forms est un module permettant de configurer les champs d'indexation pour le tri et l'archivage des documents issus d'une source SPL (fichier spool), XML, PDF ou, office 2007 (docx, xlsx).

Ce module **créé des fichiers de paramètres** enregistrés avec l'extension .XML et contenant, outre la définition des champs d'archivage, d'autres informations concernant notamment le formatage des données ainsi que les filtres à utiliser pour décoder les fichiers .SPL.

CONCEPTS

TYPES DE FICHIERS

Le module permet traiter plusieurs types de fichiers : SPL (fichier spool), XML, PDF et office 2007 (docx et xlsx).

La Fenêtre *Affichage des données* permet de visualiser le fichier SPL en mode texte et, les fichiers XML, PDF et office 2007 sous forme d'une structure XML, c'est à dire, un enchaînement de nœuds.

L'affichage du fichier PDF se distingue par le fait que les méta données du document sont, en outre, affichées sous forme d'un tableau.

PAGE DE DONNEES

En **mode ASCII** (cas de fichiers de données SPL), la page de données est l'entité de traitement résultant de la rupture du fichier spool. Les trois facteurs de rupture sont :

1. le code d'éjection,
2. le nombre de lignes par page,
3. la présence d'une chaîne de rupture.

PAGE XML

Le niveau de rupture du fichier XML définit le découpage du flux en **pages XML**.

1. : le nœud racine définit une page XML.
2. : chaque nœud élément de niveau 2 représente une page XML.

...

n : chaque nœud élément de niveau n représente une page XML.

FONCTION XPATH

Afin de faciliter l'accès aux données XML ainsi que leur extraction, nous avons intégré dans ce module, à l'instar de StarPage ou StarJet Design, une partie du langage d'interrogation de structure XML, **XPath**. Ce langage, recommandé par le W3C (World Wide Web Consortium), permet d'adresser différentes parties d'un fichier XML.

XPath, fonction d'extraction de données dans un **fichier XML**, utilise des **chemins d'accès** pour sélectionner les nœuds, ou ensembles de nœuds. Ces chemins d'accès sont d'ailleurs similaires à ceux employés dans un explorateur de fichiers.

NOEUDS

La fonction XPath traite un **fichier XML** comme une arborescence de nœuds. Il y a différents types de nœuds :

1. Nœud document ou nœud racine ()
2. Nœud élément ()
3. Nœud attribut ()
4. Nœud texte ()

• Exemple :

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<message id="203149" date="2004-07-19">
  <header>
    <author>
      <firstname>Mike</firstname>
      <lastname>Dupond</lastname>
    </author>
    <addressee type="TO">
      <firstname>Pierre</firstname>
      <lastname>Durand</lastname>
    </addressee>
    <addressee type="CC">
      <firstname>Paul</firstname>
      <lastname>Martin</lastname>
    </addressee>
    <addressee type="CC">
      <firstname>Jacques</firstname>
      <lastname>Le Guen</lastname>
    </addressee>
  </header>
</message>
```

PREDICATS

Les prédicats sont utilisés dans les fonctions XPath pour affiner une sélection, sélectionner un nœud spécifique ou un nœud contenant une valeur spécifique dans un fichier XML.

Ce sont des sous expressions placées entre crochets constituées d'une expression booléenne ou, plus simplement, d'un nombre. Dans ce dernier cas, le prédicat est interprété comme un indice de sélection.

- Exemple :

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<message id="203149" date="2004-07-19">
  <header>
    <author>
      <firstname>Mike</firstname>
      <lastname>Dupond</lastname>
    </author>
    <addressee type="TO">
      <firstname>Pierre</firstname>
      <lastname>Durand</lastname>
    </addressee>
    <addressee type="CC">
      <firstname>Paul</firstname>
      <lastname>Martin</lastname>
    </addressee>
    <addressee type="CC">
      <firstname>Jacques</firstname>
      <lastname>Le Guen</lastname>
    </addressee>
  </header>
</message>
```

/message/header/addressee/firstname[lastname='Martin'] **renvoie**

<firstname>Paul</firstname>

/message/header/addressee/firstname[1] **renvoie**

<firstname>Pierre</firstname>

CHAMPS D'INDEXATION

Les champs d'indexation définissent la façon de retrouver les documents. Après archivage des documents, les tables d'index sont conservées dans une base de données.

Il existe 2 types de champs :

1. Un **champ simple** ne peut prendre qu'une seule valeur sur un document (=ensemble de pages) ou une page XML, et présente un nombre maximum de

paramètres de définition.

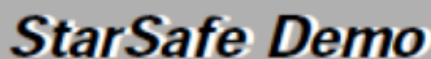
Chaque champ simple est 1) obligatoire **et/ou** 2) recherché de façon optimisée dans la base de données.

Un champ est **obligatoire** lorsque, lors de l'archivage, tous les documents ou pages XML ne le contenant pas sont rejetés de la base de données.

Un champ est **recherché de façon optimisée dans la base de données** lorsque, pour chaque document ou page XML, sa valeur est stockée dans la table d'index. Ceci augmente la vitesse de recherche donc, conséquemment, accélère le temps de réponse mais, augmente la taille de la base.

1. Un **champ multiple** peut prendre plusieurs valeurs sur un document ou une page XML et, est facultatif.

- Exemple de la facture ci-dessous (mode ASCII) :



Saleman : Margaret Peacock		TOTAL AMONT 1 889,40 F
-------------------------------	--	---------------------------

Nom du champ simple	Format	Longueur	Type
Numéro de facture (1)	Caractère	8	Recherche optimisée + Obligatoire
Identifiant du client (2)	Caractère	5	Recherche optimisée
Date de commande (3)	Date	8	Recherche optimisée
Numéro de commande (4)	Entier	5	Recherche optimisée
Date de facture (5)	Date	8	Recherche optimisée
Frais de port (6)	Réel	8	Pas de Recherche optimisée

Montant TTC (7) Réel 10 Pas de Recherche optimisée

Nom du champ multiple	Format	Longueur	Type
Référence (8)	Entier	3	Multiple
Description (9)	Caractère	30	Multiple

- Exemple en mode XML :

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<message id="203149" date="2004-07-19">
  <header>
    <author>
      <firstname>Mike</firstname>
      <lastname>Dupond</lastname>
    </author>
    <addressee type="TO">
      <firstname>Pierre</firstname>
      <lastname>Durand</lastname>
    </addressee>
    <addressee type="CC">
      <firstname>Paul</firstname>
      <lastname>Martin</lastname>
    </addressee>
    <addressee type="CC">
      <firstname>Jacques</firstname>
      <lastname>Le Guen</lastname>
    </addressee>
  </header>
</message>
```

Dans un message électronique, le numéro d'identifiant du message <message id>, la date du message <message date>, et l'auteur <author> sont uniques. Ils peuvent donc être définis en tant que champ simple.

Par contre, les destinataires <addressee> sont multiples. Ils peuvent être définis en tant que champ multiple.

CHAMP D'IMPORTATION

Un champ d'importation, **exclusivement utilisé dans le cas d'un chargement effectué par le loader externe**, est un champ pré-défini, IMPORT_FILE, faisant référence, par le biais d'une balise Path_file dans le fichier de données, au chemin du (ou des) fichier(s) à importer.

CONTRAINTES ET PRE-REQUIS

Java Runtime Environment (JRE version 5.0 ou supérieure), machine virtuelle Java permettant d'exécuter des applications et des applets Java, doit être installé sur le poste utilisateur.

UNICODE

Les caractères standards utilisés en Europe de l'Ouest et en Amérique (ISO-8859-1 ou ASCII étendu) sont stockés sur un octet. Mais, cet ensemble de caractères est insuffisant pour coder les caractères d'autres pays comme les caractères arabes ou chinois.

Unicode est un standard recensant tous les caractères existant dans le monde et définissant des codes pour ces caractères, pour une utilisation internationale.

La liste de ces codes de caractères est en évolution. Elle est organisée en blocs de caractères UCS ou, UCS (Universal Character Set, ISO 10646, Unicode) character blocks, chaque bloc étant attaché à un ensemble de langues.

Les caractères vont actuellement de 0x00 à 0x10FFFF (soit de 0 à 1 114 111) et sont codés au maximum sur 21 bits (6 octets).

Parmi les caractères Unicode, il existe des caractères de combinaison qui s'utilisent pour ajouter un accent sur le caractère précédent.

Plusieurs encodages sont possibles pour représenter des caractères Unicode (environ une quinzaine).

Le codage **UTF-8** permet d'utiliser de 1 à 6 octets par caractère Unicode. Un caractère de code inférieur à 128 est codé sur un octet. Puis, selon des plages de code, on utilise des octets supplémentaires (nécessité d'utilisation d'un algorithme de codage/décodage). L'avantage est qu'un texte en US-ASCII est codé de la même façon qu'en UTF-8.

Le codage UCS-2, ou **UTF-16**, permet de représenter un caractère UNICODE sur 2 octets (de 0 à 65535, les autres caractères pouvant être codés avec un système nommé "surrogate pairs"). Les octets de poids fort et de poids faible peuvent être inversés (big endian, **UTF-16BE**, octet de poids fort en tête, ou little endian, **UTF-16LE**, octet de poids faible en tête) en fonction des systèmes où les données sont générées.

Une séquence d'octets placée en début de fichier (BOM : byte order mark) peut permettre de connaître le mode endian des caractères de ce fichier:

FF FE = little endian FE FF = big endian

Le Bloc-note Windows permet de sauvegarder des données dans un fichier, en codage ANSI (Latin-1, ISO-8859-1), Unicode (UTF-16LE), Unicode inversé (UTF-16BE), ou UTF-8.

UTILISATION DE L'INTERFACE

CAS DE FICHIERS VOLUMINEUX

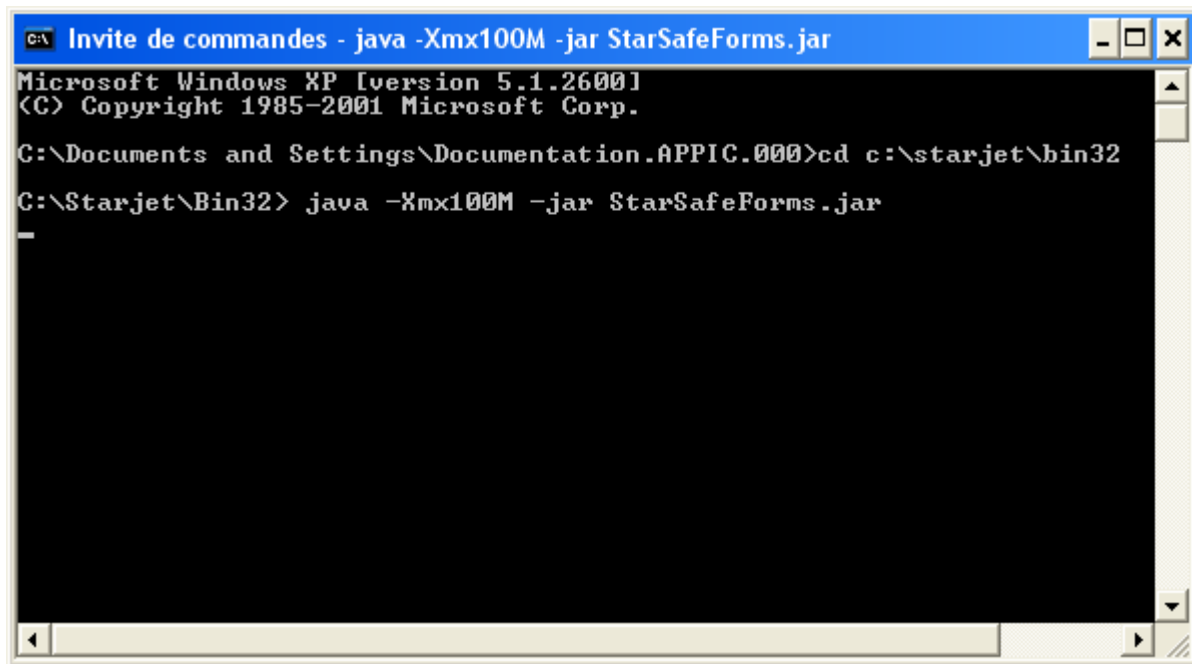
Si vous êtes amenés à effectuer le chargement de fichiers XML volumineux, il est nécessaire d'augmenter la taille maximale de la mémoire allouée à la pile

de la machine virtuelle Java. Pour ce faire, ajouter le paramètre `-Xmx` dans la ligne de commande comme suit.

```
\> java -Xmx<taille mémoire maximale> -jar StarSafeForms.jar
```

- Exemple : pour fixer à 100 Méga octets la taille maximale de la mémoire allouée à la pile de la machine virtuelle Java

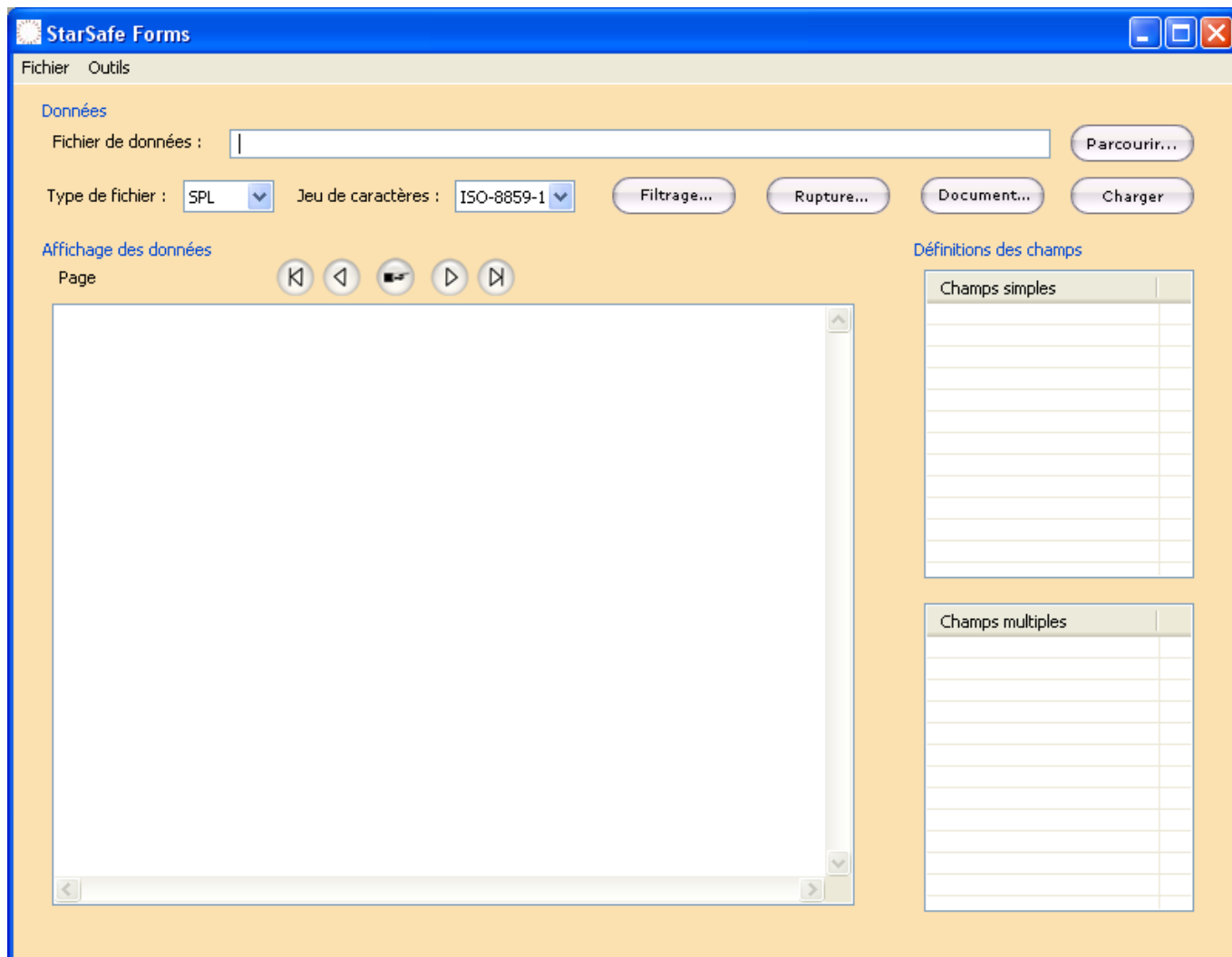
```
\> java -Xmx100M -jar StarSafeForms.jar
```



```
C:\> Invite de commandes - java -Xmx100M -jar StarSafeForms.jar
Microsoft Windows XP [version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\Documents and Settings\Documentation.APPLIC.0000>cd c:\starjet\bin32
C:\Starjet\Bin32> java -Xmx100M -jar StarSafeForms.jar
-
```

L'INTERFACE GRAPHIQUE



L'interface est divisée en trois sections.

1. La section *Données* contient le chemin du fichier de données de format SPL, XML ou PDF, ainsi que la définition des options relatives à ces données.
2. La section *Affichage des données* permet de visualiser sous forme d'un fichier texte les fichiers de données de format SPL ou, sous forme d'un fichier XML, les fichiers de données de format XML, PDF ou OFFICE 2007.
3. La section *Définition des champs* permet de visualiser le nom du (ou des) champ(s) simple(s) et/ou multiple(s) défini(s).

Un menu *Fichier* permet d'ouvrir un fichier de paramétrage .XML déjà existant (option *Ouvrir*), d'enregistrer le fichier de paramétrage .XML nouvellement créé (option *Enregistrer ou Enregistrer sous*), de mettre à blanc l'interface afin de créer un nouveau fichier de paramétrage (option *Nouveau*), ainsi que de quitter le logiciel (option *Quitter*).

Le fichier enregistré dans un répertoire s'accompagne toujours d'un second fichier de sauvegarde comportant un tilde à la fin du nom.



Si, lors du paramétrage, un champ d'importation a été créé, deux fichiers sont créés à l'enregistrement

1. un fichier de paramétrage .XML à utiliser lors de la création de l'archive
2. un fichier de paramétrage IF.XML à utiliser lors du chargement de cette même archive.

Ces deux fichiers s'accompagnent toujours d'un second fichier de sauvegarde comportant un tilde à la fin du nom.

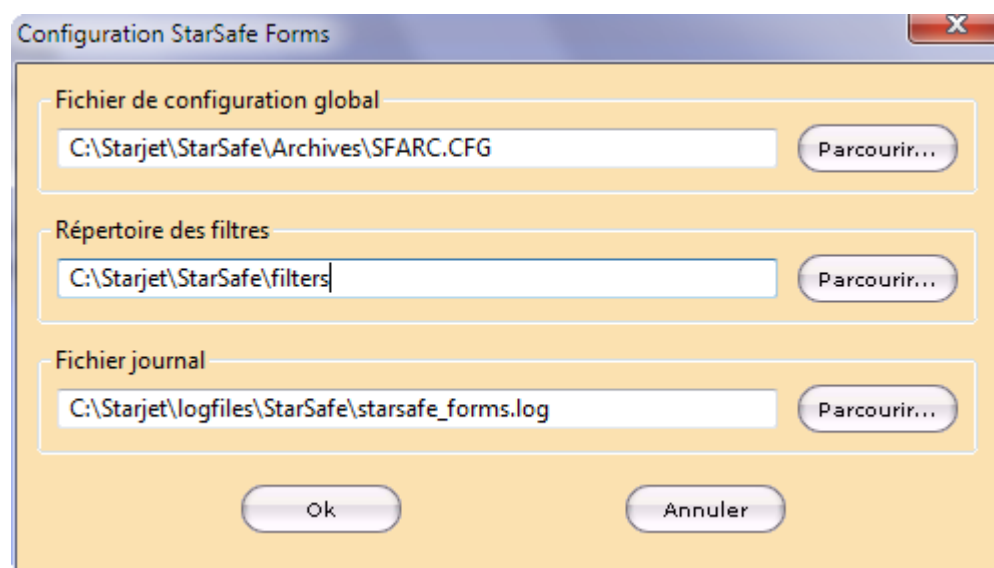


Il est possible, après ouverture d'un fichier de paramétrage .XML déjà existant, de procéder à une mise à jour dynamique de ce fichier en l'enregistrant après modification (modification, ajout ou suppression de champs).

La création d'un nouveau champ provoque l'affectation d'une valeur nulle en ce qui concerne ce champ pour des documents déjà chargés.

Attention donc à ne pas définir comme obligatoire de nouveaux champs dans le cadre d'une mise à jour du fichier de configuration d'une archive existante.

DEFINITION DES PARAMETRES



Cliquer sur l'option *Paramètres* du menu *Outils* affiche la boîte de dialogue *Configuration StarSafe Forms* permettant de sélectionner, à l'aide du bouton *Parcourir*, le fichier de configuration global .cfg, le répertoire des filtres ainsi que le fichier .log, fichier texte comprenant les messages d'erreur générés par l'application, informations indispensables à analyser en cas d'erreur.

Le contenu de ce journal des événements peut en outre être consulté directement depuis l'interface, en cliquant sur l'option *Messages du journal* du menu *Outils*.

Enfin, l'option *Vider le journal* permet de supprimer le contenu de ce fichier .log.



Afin d'éviter des problèmes de fonctionnement du module, les informations de configuration doivent obligatoirement être renseignées. Si ce n'est pas le cas, la boîte de dialogue ne peut être validée.

OPTIONS DES DONNEES

Options

Séparateur des milliers

Séparateur des décimales

☒ Notion de document (valeurs des champs simples identiques)

☒ Grouper uniquement sur les valeurs des champs simples obligatoires

☒ Remise à zéro de la valeur des champs simples d'une page à l'autre

Ok Annuler

Cliquer sur le bouton *Document...* dans la section *Données* affiche la boîte de dialogue *Options* permettant de définir les délimiteurs utilisés pour les conversions de nombres, ainsi que la méthode de regroupement des pages en documents.

Pour faciliter la lecture des nombres, certains logiciels les affichent dans différents formats. Par exemple, le nombre 125235.58 sera affiché 125 235,58 ou encore 125.235,58.

Séparateur des milliers : indiquer quel caractère est utilisé pour délimiter les milliers. Par défaut, le caractère espace est utilisé.

Séparateur des décimales : indiquer quel caractère est utilisé pour délimiter les décimales. Par défaut le caractère *','* (virgule) est utilisé.

Cocher la case *Notion de document (valeur des champs simples identiques)* si *StarSafe* doit rassembler en documents toutes les pages ayant des champs simples identiques. Si cette option est désactivée, les pages ne sont pas rassemblées en documents.

Cocher le bouton *Grouper uniquement sur les valeurs des champs simples obligatoires* si *StarSafe* doit regrouper les pages en documents lorsque **seuls** les champs simples obligatoires sont identiques sur ces pages. Si cette option est désactivée, *StarSafe* regroupera en document les pages dont **tous** les champs simples (obligatoires et facultatifs) sont identiques.

Cocher le bouton *Remise à zéro de la valeur des champs simples d'une page à l'autre* afin que, dans un document de *n* pages, un champ initialement vide le demeure. Sinon, par défaut, un champ vide sur une page d'un document est renseigné avec la valeur attribuée à ce champ dans la page précédente.



Le bouton *Rupture* n'est disponible que dans le cas de fichiers .SPL ou .XML. Si

.PDF est sélectionné dans le menu déroulant *Type de fichier*, le bouton *Rupture*

est grisé et donc, indisponible.

Si le type de fichier .SPL est sélectionné, cliquer sur le bouton *Rupture* dans la section *Données* affiche la boîte de dialogue *Rupture* permettant de définir un des 3 procédés de séparation de page.

Chaîne de rupture	Ligne	Colonne
-------------------	-------	---------

1. Si chaque page possède le même nombre de lignes, cocher le bouton *Nombre de lignes par page* puis, indiquer le nombre de lignes par page.
2. Si chaque page est séparée par un code d'éjection, cocher le bouton *Code d'éjection* puis, indiquer la valeur décimale du code d'éjection.
3. Si chaque page commence par une chaîne de caractères, cocher le bouton *Chaîne de rupture* puis, cliquer sur *Ajouter* pour afficher la boîte de dialogue *Chaîne de rupture*.

Saisir la chaîne de séparation dans le champ *Valeur*. Si plusieurs chaînes sont utilisées (ex : FACTURE ou AVOIR), saisir les chaînes séparées par le caractère "|" (ex: FACTURE|AVOIR).

Cocher *Les chaînes de rupture sont sensibles à la casse* si le programme ne doit pas traiter de la même façon les données, selon qu'elles sont saisies en majuscules ou en

minuscules.

Indiquer l'indice de la ligne dans le champ *Ligne* (0 est l'indice par défaut indiquant que la chaîne n'est pas sur une ligne précise).

Indiquer l'indice de la colonne dans le champ *Colonne* (-1 est l'indice par défaut indiquant que la chaîne n'est pas dans une colonne précise).

Cliquer sur *OK* pour ajouter la chaîne de rupture dans le tableau.

Sélectionner la chaîne de rupture active les boutons *Modifier* et *Supprimer*.

Cliquer sur *Modifier* affiche une boîte de dialogue permettant de modifier la chaîne de rupture. Cliquer sur *Supprimer* efface définitivement la ligne sélectionnée.

Si le type de fichier .XML est sélectionné, cliquer sur le bouton *Rupture* dans la section *Données* affiche la boîte de dialogue *Rupture* permettant de définir le niveau de rupture XML.

Ensuite, saisir le niveau de rupture ou de découpage de ce fichier XML (pour

de plus amples informations, [cf § page XML](#)).



Le bouton *Filtrage* n'est disponible que dans le cas d'un fichier .SPL. Dans les autres cas, le bouton *Filtrage* est indisponible.

Cliquer sur le bouton *Filtrage* dans la section *Données* affiche la boîte de dialogue *Filtrage* permettant de sélectionner un filtre permettant, si nécessaire, de décoder les fichiers.

Le chemin d'accès au fichier à traiter est indiqué dans la section *Données*.

Filtre inutile est coché par défaut, indiquant que le fichier n'a pas besoin d'être filtré.

Cocher *Librairie spécifique* et sélectionner dans le menu déroulant le filtre dll à utiliser si nécessaire

Cocher *Programme filtre* et sélectionner, à l'aide du bouton *Parcourir*, le nom du fichier .prf à utiliser si nécessaire. La boîte de dialogue de sélection de fichier s'ouvre sur le répertoire spécifié dans la boîte de dialogue *Configuration StarSafe Forms* ([cf § Définition des paramètres](#)).

Cocher le bouton *Afficher tout le fichier* pour travailler sur l'intégralité du fichier ou, indiquer, dans le champ *Traiter*, la taille en octets à extraire.

TRAITEMENT DU FICHIER DE DONNEES

Saisir le chemin complet du fichier de données ou, cliquer sur le bouton *Parcourir* afin de naviguer dans l'explorateur de fichiers puis, sélectionner un fichier et cliquer sur *Ouvrir* pour fermer la boîte de dialogue.

Le fichier à traiter est affiché dans le champ *Fichier de données*.

Dans le champ *Type de fichier*, sélectionner à l'aide du menu déroulant, un des types de fichiers pouvant être traités par le logiciel.

Si SPL est sélectionné, il est possible de choisir, à l'aide du menu déroulant du champ *Jeu de caractères*, un des jeux de caractères suivants ISO 8859-1, UTF-8, UTF-16, UTF-16LE ou UTF16-BE ([cf § Unicode](#)), dans lequel est encodé ce fichier SPL.

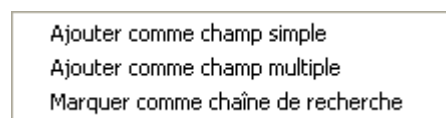


Les fichiers Unicode ne sont actuellement pas supportés par le moteur StarPage couplé avec StarSafe. Il est donc fortement déconseillé de choisir des fichiers de données unicode pour les associer à un scénario StarPage, en vue d'une indexation et d'un archivage.

Enfin, après avoir défini les options relatives aux données ([cf § Options des données](#)), cliquer sur le bouton *Charger* pour afficher le fichier de données dans la fenêtre prévue à cet effet.

Dans le cas d'un fichier .SPL, positionner le curseur de la souris sur le début du champ. Cliquer sur le bouton gauche et faire glisser le curseur jusqu'à la fin du champ. Relâcher le bouton de la souris. Le champ sélectionné est surligné en bleu.

Cliquer avec le bouton droit de la souris sur la sélection donne accès à un menu contextuel permettant d'ajouter cet élément en tant que champ simple ou champ multiple ou, de le marquer comme chaîne de recherche.



Si *Ajouter comme champ simple* est sélectionné, le champ est surligné en jaune et un nouveau champ, affecté d'un nom par défaut, est affiché dans la section *Champs simples*.

Si *Ajouter comme champ multiple* est sélectionné, le champ est surligné en jaune et un nouveau champ, affecté d'un nom par défaut, est affiché dans la section *Champs multiples*.



Si le champ multiple est défini sur plusieurs lignes ([cf § Propriété des champs d'indexation](#)), toutes les lignes sont surlignées en jaune.

Cliquer sur *Marquer comme chaîne de recherche* si la localisation d'un champ simple ou multiple est définie par la présence de la chaîne de caractères sélectionnée. Le champ à marquer est alors surligné en vert.

Dans le cas d'un fichier .XML, PDF ou OFFICE, cliquer avec le bouton droit de la souris sur un nœud donne accès à un menu contextuel permettant d'ajouter cet élément en tant que champ simple ou champ multiple.

Ajouter comme champ simple
Ajouter comme champ multiple
Ajouter comme champ d'importation

Si l'option *Ajouter comme champ simple* est sélectionnée, un nouveau champ, affecté d'un nom par défaut, est affiché dans la section *Champs simples*.

Si l'option *Ajouter comme champ multiple* est sélectionnée, un nouveau champ, affecté d'un nom par défaut, est affiché dans la section *Champs multiples*.

Dans le cas d'un fichier XML, si l'option *Ajouter comme champ d'importation* est sélectionnée – après clic avec le bouton droit sur un nœud *Path_file* – un nouveau champ pré-défini, *IMPORT_FILE*, exclusivement utilisé dans le cas d'un chargement effectué par le loader externe, est ajouté dans la section champs simples.

Celui-ci fait référence, par le biais d'une balise *Path_file* dans le fichier de données XML, au chemin du (ou des) fichier(s) à importer.

PROPRIETES DES CHAMPS D'INDEXATION

Cliquer avec le bouton droit de la souris sur un champ d'indexation donne accès à un menu contextuel permettant de supprimer ce champ, de modifier sa position dans la liste des champs à l'aide des options *Monter* ou *Descendre* ou, d'en afficher les propriétés – comme après un double clic – via une boîte de dialogue *Propriétés des champs* quelque peu différente selon le type de fichier de données SPL, XML ou PDF.

Propriétés des champs

Nom :

Nom Alias :

Nombre de caractères extraits :

Masque :

☒ Obligatoire (Valeur non nulle)

☒ Recherche optimisée dans la base de données

Type

☒ Caractères

☐ Date

☐ Nombre entier

☐ Nombre décimal Chiffres après la virgule :

Localisation

☒ Le champ ne dépend pas d'une chaîne

Le champ est situé entre les lignes et colonne

(Si aucune ligne n'est définie, un masque est requis)

☐ Le champ dépend d'une chaîne

La chaîne

est entre les lignes et colonne

Le champ est situé

☐ sur la ligne où est la chaîne

☐ lignes la chaîne

colonnes

Ok Annuler

Pour définir un champ – **hormis dans le cas de documents office 2007** –, saisir son nom (lettres, chiffres et underscore sont acceptés alors que les caractères espace sont automatiquement remplacés par des underscores). Par défaut, un champ simple est nommé *SFieldn*, n étant le numéro d'incrémentation du champ tandis que le champ multiple est nommé *MFieldn*, n étant le numéro d'incrémentation du champ.

Saisir éventuellement un alias dans le champ *Nom Alias*, la désignation d'un même champ par un nom différent permettant d'étendre les possibilités d'affichage et de recherche dans l'interface de consultation StarSafe.

- Exemple :

Un même champ est appelé Numéro-facture, Number_invoice ou Numero_factura selon la langue du pays d'implantation des agences d'une société. L'utilisateur peut choisir de faire une recherche sur l'alias qui est Number_invoice dans les trois cas.



Si le fichier de données est un fichier office, il est impossible de renommer un champ simple ou multiple. Personnaliser l'intitulé du champ en renseignant le *Nom Alias*.

Dans le cas de fichier de données XML ou PDF, l'adresse XPath ([cf § Fonction XPath](#)) du champ est automatiquement affichée dans le champ *Localisation* mais il est possible de la modifier manuellement.



Les chemins XPath sont sensibles à la casse c'est à dire qu'ils sont pas traités de la même façon selon qu'ils soient saisis en majuscules ou en minuscules.

- Exemple :

Nous avons sélectionné un nœud dont l'adresse XPath est la suivante :

`/message/header[1]/addressee[1]`

Supprimer le dernier indice pour obtenir l'adresse XPath suivante

`/message/header[1]/addressee` afin que tous les destinataires soient indexés.

Il est possible d'affiner les propriétés d'un champ simple ou multiple par

- la saisie d'un nombre maximal de caractères à extraire dans le champ *Nombre de caractères extraits*.
- Exemple :

Dans le cas d'un champ multiple correspondant aux destinataires d'un mail, saisir

8 provoque la troncature des noms de destinataires excédant 8 caractères.

- la saisie d'un masque dans le champ *Masque*.
- Exemple : Dans le cas d'un champ multiple correspondant aux destinataires d'un mail, saisir KER aboutit à ce que seuls les destinataires dont le nom commence par KER soient indexés.

Cocher le bouton *Obligatoire (Valeur non nulle)* aboutit à ce que ce champ soit obligatoirement dans la page pour que le document soit archivé (par exemple, le numéro de client dans le cas d'une archive Facture). Si le champ est facultatif, StarSafe archive le document même si ce champ n'est pas renseigné.



Cette option est indisponible dans le cas d'un champ multiple.

Cocher la case *Recherche optimisée dans la base de données* permet d'accélérer les temps de réponse lors des recherches mais, augmente la taille de la base de de données.



Cette option est indisponible dans le cas d'un champ multiple. Sélectionner le type du champ parmi les quatre possibles :

- *Caractères*

Ou

- Cocher le bouton *Date* puis, sélectionner le format dans la liste.

Il existe 10 formats de date pré-définis.

Si votre format de date ne correspond à aucun des fichiers pré-définis, vous pouvez créer un nouveau format en modifiant le fichier prstock.ini localisé dans le répertoire C:\WINDOWS.

Jour/Mois/Année en chiffres

jj/mm/aa Année réduite (ex : 10/01/94). jj/mm/aaaa Année complète (ex : 10/01/1994).

Mois/Jour/Année en chiffres

mm/jj/aa Année réduite (ex : 01/10/94). mm/jj/aaaa Année complète (ex : 01/10/1994).

Année/Mois/Jour en chiffres

aa/mm/jj Année réduite (ex : 94/01/10). aaaa/mm/jj Année complète (ex : 1994/01/10).

Autres formats de date :

JJ 'MM AA : Jour en chiffres, Mois en lettres, Année en chiffres (ex : 10 Janvier 1994).

'MM' JJ AA : Mois en lettres, Jour en lettres, Année en chiffres (ex : Dix Janvier 1994).

JJ 'MM' AA : Jour en chiffres, Mois en lettres, Année en lettres (ex : 10 Janvier Mille neuf cent quatre vingt quatorze).

'JJ' 'MM' AA : Jour en lettres, Mois en lettres, Année en lettres (Dix Janvier Mille neuf cent quatre vingt quatorze).

Une date occupe 4 octets.

Pour les dates au format court, autrement dit les années réduites, StarSafe les interprète de la manière suivante :

Si l'année est supérieure à 80, il convertit en 1900

- Exemple : 01/12/81 est converti en 1^{er} décembre de l'année 1981.

Si l'année est inférieure à 80, il convertit en 2000

- Exemple : 01/12/79 est converti en 1^{er} décembre de l'année 2079.

Ou

- *Nombre entier*

La valeur d'un champ entier peut varier de -32768 à 32767. Un entier occupe 2 octets.

Pour archiver des valeurs entières supérieur à 32767, utilisez le type Réel autrement dit Nombre décimal.

Ou

- Cocher le bouton *Nombre décimal* puis, saisir le nombre de chiffres après la virgule.

La valeur d'un champ réel peut varier de 1.17×10^{-308} à $1.17 \times 10^{+308}$. Un réel occupe 8 octets.

Dans le cas d'un fichier de données SPL, la section *Localisation* est automatiquement renseignée.

Un champ simple ou multiple peut être localisé de 2 manières différentes :

- soit, par des coordonnées (ligne, colonne),
- soit, par une chaîne de caractères.

Le champ ne dépend pas d'une chaîne

Le champ est situé entre les lignes

0 et 0 colonne 15

(Si aucune ligne n'est définie, un masque est requis)

Si le champ est localisé par des coordonnées, sa position est fixe (coordonnées fixes) ou variable (recherche entre plusieurs lignes). Dans le cas d'un champ simple, unique sur un document, on cherche la première occurrence valide et seule celle-ci sera archivée. Un champ multiple peut être présent plusieurs fois sur un document.

Le champ dépend d'une chaîne

La chaîne

est entre les lignes et colonne

Le champ est situé

☒ sur la ligne où est la chaîne

☐ lignes la chaîne

colonnes

La présence de la chaîne définit la localisation du champ c'est à dire que ce dernier est localisé par rapport à la chaîne dont on connaît la position.

PROPRIETES DES CHAMPS D'IMPORTATION

Cliquer avec le bouton droit de la souris sur un champ d'importation – dans le cas d'un chargement à l'aide du loader externe – donne accès à un menu contextuel permettant de supprimer ce champ, de modifier sa position dans la liste des champs à l'aide des options *Monter* ou *Descendre* ou, d'en afficher les propriétés, comme après un double clic, via une boîte de dialogue *Propriétés des champs*.

Propriétés des champs

Nom :

Fichier spool :

Nombre de caractères extraits :

Masque :

☒ Obligatoire (Valeur non nulle)

☒ Recherche optimisée dans la base de données

Type

☒ Caractères

☐ Date

☐ Nombre entier

☐ Nombre décimal Chiffres après la virgule :

Ok Annuler

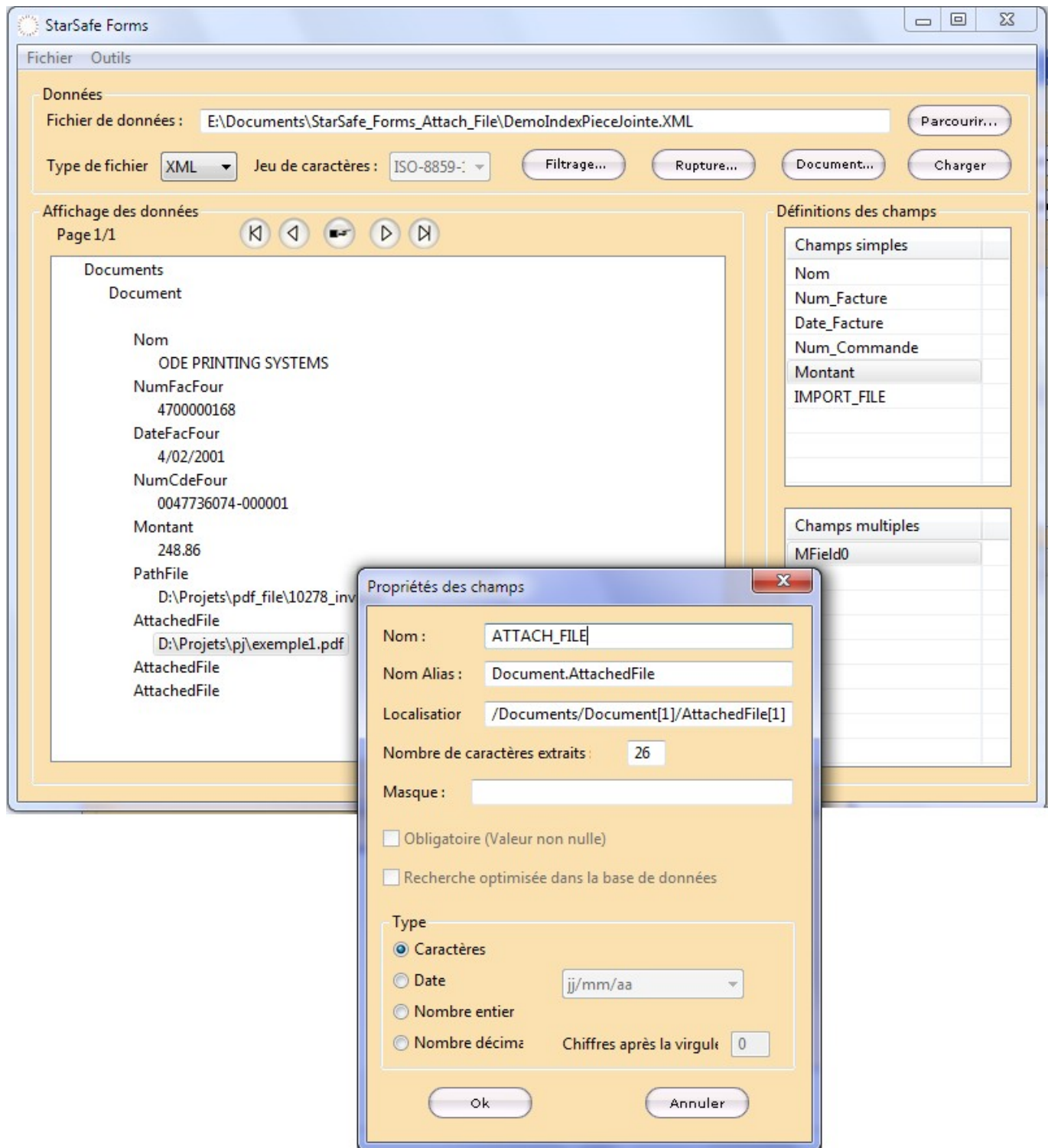
Pour définir un champ d'importation, créer dans le fichier de données XML, une balise *Path_file* faisant référence au chemin complet du fichier contenant le chemin du (ou des) fichier(s) de données à importer.

Cliquer avec le bouton droit de la souris sur le contenu de la balise *Path_file* donne accès à un menu contextuel permettant d'ajouter cet élément en tant que champ d'importation.

Il est possible d'affiner les propriétés d'un champ d'importation par

- la saisie d'un nombre maximal de caractères à extraire dans le champ *Nombre de caractères extraits*.
- la saisie d'un masque dans le champ *Masque*.

Pour définir des pièces jointes dans le cas d'un chargement à l'aide du loader externe, cliquer avec le bouton droit de la souris sur un nœud faisant référence au chemin de la pièce jointe à importer donne accès à un menu contextuel permettant d'ajouter cet élément en tant que champ multiple en lui assignant *ATTACH_FILE* en tant que nom.



Dans le cas de plusieurs pièces jointes et afin de prendre en compte toutes les balises des pièces jointes, supprimer la valeur [1] dans le champ *Localisation*.

ICÔNES DE NAVIGATION

Première page



Icône associée dans la barre d'outils.

Dans le cas où le fichier de données est divisé en plusieurs pages, cliquer sur l'icône *Affichage de la première page* permet d'afficher rapidement la première page du fichier.

Dernière page



Icône associée dans la barre d'outils.

Dans le cas où le fichier de données est divisé en plusieurs pages, cliquer sur l'icône *Affichage de la dernière page* permet d'afficher rapidement la dernière page du fichier.

Page précédente



Icône associée dans la barre d'outils.

Dans le cas où le fichier de données est divisé en plusieurs pages, cliquer sur l'icône *Affichage de la page précédente* permet d'afficher la page précédant celle qui est actuellement affichée dans la section *Affichage des données*.

Page suivante



Icône associée dans la barre d'outils.

Dans le cas où le fichier de données est divisé en plusieurs pages, cliquer sur l'icône *Affichage de la page suivante* permet d'afficher la page suivant celle qui est actuellement affichée dans la section *Affichage des données*.

Atteindre une page



Icône associée dans la barre d'outils.

Dans le cas où le fichier de données est divisé en plusieurs pages, cliquer sur l'icône *Atteindre page* affiche une boîte de dialogue vous permettant la saisie du numéro de page à visualiser

