

ABBYY FlexiCapture 8.0
Professional
Guide de l'utilisateur

Cher utilisateur,

Ce guide explique les principes de base d'ABBYY FlexiCapture 8.0 Professional. Veuillez le lire attentivement avant toute utilisation du programme.

Pour plus d'informations, consultez la documentation suivante :

- Accédez à l'aide en ligne à partir du menu ou en appuyant sur F1. Le fichier d'aide est également disponible dans Démarrer > Programmes > ABBYY FlexiCapture 8.0 Professional > Aides.
- Le guide de l'administrateur système ABBYY FlexiCapture 8.0 Professional est disponible dans Démarrer > Programmes > ABBYY FlexiCapture 8.0 Professional > Guides > Guide de l'administrateur système.
- Le guide de l'opérateur est disponible dans Démarrer > Programmes > ABBYY FlexiCapture 8.0 Professional > Guides > Guide de l'opérateur.
- Le guide de création de formulaires lisibles par machine est disponible dans Démarrer > Programmes > ABBYY FlexiCapture 8.0 Professional > Guides > Création d'un formulaire lisible par machine.
- Accédez à l'aide en ligne de FlexiLayout Studio 8.0 et de FormDesigner 8.0 à partir du menu de l'application correspondante ou en appuyant sur F1. Le fichier d'aide est également disponible dans Démarrer > Programmes > ABBYY FlexiCapture 8.0 Professional > Aides.
- Les didacticiels ABBYY FlexiLayout Studio sont disponibles dans le menu Démarrer > Programmes > ABBYY FlexiCapture 8.0 Professional > Guides > Enseignement par les modèles.

Pour en savoir plus sur la façon d'utiliser efficacement le programme, nous vous recommandons non seulement de lire la documentation correspondante mais également de créer un projet pour traiter des documents d'un certain type. Vous pouvez utiliser à titre d'exemple un formulaire de *Questionnaire* dont une copie se trouve sur <votre disque>:\Documents and Settings\All Users\Application Data\ABBYY\FlexiCapture\8.0\Samples\FormDesigner\French\Questionnaire_fr (Sous Microsoft Windows Vista : <votre disque>:\Users\Public\ABBYY\FlexiCapture8.0\Samples\FormDesigner\French\Questionnaire_fr). Tout au long de ce guide, vous accéderez à des instructions situées dans des zones de texte. Ces instructions vous aideront à configurer le programme et à traiter les documents. Elles vous permettront également de visualiser les résultats une fois le traitement du questionnaire terminé. Si vous rencontrez des difficultés dans la configuration du projet, ouvrez un projet prêt à l'emploi appelé Questionnaire.fcproj et qui se trouve sur <votre disque>:\Documents and Settings\All Users\Application Data\ABBYY\FlexiCapture\8.0\Samples\FlexiCapture\French\Questionnaire_fr (Sous Microsoft Windows Vista : <votre disque>:\Users\Public\ABBYY\FlexiCapture8.0\Samples\FormDesigner\French\Questionnaire_fr).

Cet exemple ainsi que d'autres exemples de projets sont disponibles dans Démarrer > Programmes > ABBYY FlexiCapture 8.0 Professional > Aides > FlexiCapture projets des modèles.

Exportez les résultats au format XML et envoyez-les nous. Vos commentaires sont les bienvenus car ils nous permettent d'en apprendre davantage sur nos utilisateurs et ainsi d'améliorer notre produit pour mieux satisfaire vos besoins.

Nous espérons que vous apprécierez notre produit et que son utilisation vous sera agréable.

Sommaire

1. Introduction.....	5
1.1. Finalité de l'acquisition de données et de documents	5
1.2. Acquisition automatisée de documents et de données.....	5
2. Fonctions dédiées à l'administrateur et à l'opérateur	6
3. Quels documents l'application ABBYY FlexiCapture 8.0 Professional peut-elle traiter ?	7
4. Configuration du système pour les formulaires fixes d'acquisition.....	8
4.1. Création d'un modèle de formulaire fixe	8
4.1.1. Éléments de formulaire	10
4.1.2. Marquage des champs de données.....	12
4.2. Création d'un projet.....	13
4.3. Création d'un modèle de document.....	14
4.3.1. Éditeur de modèles de document	15
4.3.2. Utilisation des éléments pour marquer des objets dans le formulaire	16
4.3.2.1. Groupes de champs	18
4.3.2.2. Champs sans marquage.....	18
4.3.2.3. Marquage des tableaux	19
4.3.2.4. Champs avec régions multiples	19
4.3.2.5. Champs avec instances multiples.....	20
4.3.2.6. Exclusion d'une région de la reconnaissance	20
4.3.2.7. Suppression des champs	21
4.3.3. Éléments statiques.....	21
4.3.3.1. Particularités des codes barres	22
4.3.4. Propriétés des champs.....	22
4.3.4.1. Propriétés des champs communs	22
4.3.4.2. Types de données.....	24
4.3.4.2.1. Types de données d'un champ de saisie de texte	24
4.3.4.2.2. Types de données des coches.....	27
4.3.4.2.3. Types de données pour un groupe de coches.....	28
4.3.4.3. Propriétés de reconnaissance des champs.....	29
4.3.4.3.1. Propriétés de reconnaissance d'un champ de saisie de texte.....	29
4.3.4.3.2. Propriétés de reconnaissance des coches et des groupes de coches.....	30
4.3.4.3.3. Propriétés de reconnaissance de code barre.....	32
4.3.4.3.4. Propriétés de reconnaissance d'image.....	32
4.3.4.4. Paramètres de vérification.....	32
4.3.4.5. Options d'exportation des images	33
4.3.4.6. Vérifications basées sur les règles	34
4.3.5. Création d'un modèle pour un document multipage	36
4.3.6. Création d'un modèle pour un document possédant des pages annexes	39
4.3.7. Configuration de l'exportation des données	40
4.3.7.1. Exportation vers des fichiers.....	40
4.3.7.2. Exportation vers une base de données	41
4.3.7.3. Exportation de données vers une bibliothèque de documents SharePoint.....	43
4.3.7.4. Exportation personnalisée.....	45
4.3.8. Configuration de la vue des données reconnues	46
4.3.9. Modification et publication d'un modèle	47
4.1.10. Test du modèle.....	47
4.4. Configuration de l'importation des images	48

5.	Configuration du système pour le traitement des documents flexibles	50
6.	Particularités de l'acquisition de documents non structurés.....	50
7.	Utilisation d'un projet de configuration	51
7.1.	Ajout d'images	52
7.2.	Reconnaissance	53
7.3.	Vérification	53
7.4.	Exportation.....	59
8.	Conclusion	59

1. Introduction

1.1. Finalité de l'acquisition de données et de documents

De nos jours, des types de documents très variés sont utilisés, que ce soit par les entreprises, les industries ou les services. Une part importante du travail réalisé par une entreprise ou une institution provient des applications, des enquêtes et des factures. Les technologies de l'information et les normes utilisées empêchent tout simplement d'utiliser uniquement des documents papier : la plupart des données doivent être converties au format électronique à des fins de stockage, d'analyse et de traitement ultérieur.

La principale difficulté dans le traitement d'un document papier concerne la saisie des données qu'il contient dans un système informatique. Une saisie des données manuelle et traditionnelle ne peut se justifier que si le volume d'informations est relativement réduit. Si le volume d'informations est conséquent, la saisie manuelle devient inefficace car elle est trop lente, monopolise trop de personnel et se révèle trop onéreuse. La saisie manuelle n'en devient pas plus efficace pendant la nuit. Le temps et l'argent consacrés à une saisie plus importante équivalent plus ou moins en termes de coût à recommencer l'intégralité du cycle de traitement.

Autant dire que la saisie de données manuelle est loin d'être idéale. Il existe cependant une alternative bien moins problématique et bien plus efficace : un système d'acquisition automatisée de données et de documents. ABBYY FlexiCapture 8.0 Professional est exactement ce type de système.

1.2. Acquisition automatisée de documents et de données

ABBYY FlexiCapture 8.0 Professional est un système logiciel dédié à l'acquisition de données provenant de documents de types très variés : formulaires fixes, documents semi-structurés et documents non structurés.

Plusieurs étapes interviennent dans l'acquisition automatisée de données :

- la **numérisation** d'un lot de page au moyen d'un scanner ;
- l'assemblage automatique des pages numérisées en documents ;
- la **reconnaissance** automatique de caractères ;
- si le programme a généré des possibilités différentes au sujet de certains caractères (c'est-à-dire que la reconnaissance de ces caractères est incertaine), ils sont envoyés à l'opérateur pour **vérification** ;
- l'**exportation** des données vérifiées dans un fichier ou une base de données et l'enregistrement des images du document dans le dossier indiqué. Les images peuvent être enregistrées au format graphique ou en tant que PDF indexé.

ABBYY FlexiCapture 8.0 Professional est une solution efficace pour l'acquisition automatisée de données. Elle vous permet de surveiller et de gérer le traitement des données, ainsi que de contrôler la qualité des données saisies.

Avantages de l'utilisation d'ABBYY FlexiCapture 8.0 Professional

- **Un faible coût.** Il est possible de configurer le système pour que la plupart des opérations soient effectuées automatiquement, quasiment sans aucune intervention humaine. L'opérateur se contente alors uniquement d'insérer les lots de documents dans le chargeur du scanner et de vérifier les caractères pour lesquels le système n'est pas sûr. Au lieu de disposer de plusieurs

postes de travail dédiés à la saisie manuelle des données, un seul poste de travail équipé d'ABBYY FlexiCapture 8.0 Professional suffit.

- **Une saisie plus rapide.** En automatisant l'intégralité du processus, vous accélérez considérablement le traitement des documents.
- **Une qualité de reconnaissance élevée.** Les technologies d'ABBYY maintes fois primées autorisent une reconnaissance des caractères rapide et de très grande qualité. Il est extrêmement difficile de parvenir à une qualité comparable avec une saisie manuelle des données sans une diminution considérable de la vitesse de traitement.
- **Une utilisation conviviale et une grande facilité d'apprentissage.** ABBYY FlexiCapture 8.0 Professional offre une interface conviviale, tant à l'administrateur du système qu'à la personne en charge du traitement des documents. Aucun cours de formation prolongé n'est nécessaire pour utiliser le programme. Les utilisateurs peuvent apprendre à maîtriser le programme en toute simplicité en s'appuyant sur notre documentation technique et nos documents de référence.
- **Une évolutivité remarquable.** À la différence d'un système de saisie manuelle, ABBYY FlexiCapture 8.0 Professional offre une modularité exemplaire. Pour accroître votre rendement, il vous suffit d'installer le système sur un poste de travail supplémentaire.

2. Fonctions dédiées à l'administrateur et à l'opérateur

Pour acquérir des données à l'aide d'ABBYY FlexiCapture 8.0 Professional, vous devez configurer le système et lui permettre de traiter un certain type de document. Les administrateurs indiquent les paramètres nécessaires et les opérateurs procèdent à l'acquisition des documents. Le nombre de postes de travail dépend du volume de documents à traiter.

Les **administrateurs** indiquent les paramètres et surveillent le système. Leurs tâches sont les suivantes :

- création de nouveaux formulaires (cette tâche peut être réalisée par un concepteur) ;
- configuration des paramètres relatifs à l'importation et à la numérisation des images ;
- préparation des modèles de document, notamment :
 - paramétrage des paramètres de reconnaissance ;
 - définition des règles de vérification et d'assemblage des documents ;
 - configuration de l'exportation des données.
- surveillance du processus de saisie des données dans le système préalablement configuré et conseils apportés aux opérateurs.

Les **opérateurs** sont responsables de la récupération des documents. En utilisant le système tel que configuré par l'administrateur, les opérateurs traitent les documents finalisés d'un certain type. Leurs tâches sont les suivantes :

- numérisation et importation des documents ;
- vérification des données ;
- exportation des données.

La reconnaissance des documents est automatique.

Deux modes sont proposés : le mode administrateur et le mode opérateur. L'administrateur dispose d'un accès sans restriction à l'ensemble des fonctions du programme. Le mode administrateur peut être

protégé par mot de passe. Un opérateur ne peut pas accéder à la création de modèles ni à l'édition ou à la création de profils d'importation. Un opérateur peut travailler sur un seul projet ou un seul lot à la fois.

Lorsque vous installez le programme, vous pouvez choisir entre installer uniquement la Station Opérateur ou le programme dans son intégralité. Si vous sélectionnez une installation complète, vous pourrez passer du mode administrateur au mode opérateur et vice-versa. Pour sélectionner l'un de ces deux modes, choisissez **Démarrer > Tous les programmes > ABBYY FlexiCapture 8.0 Professional > Station Administrateur ABBYY FlexiCapture 8.0** ou **Démarrer > Tous les programmes > ABBYY FlexiCapture 8.0 Professional > Station Opérateur ABBYY FlexiCapture 8.0**.

3. Quels documents l'application ABBYY FlexiCapture 8.0 Professional peut-elle traiter ?

ABBYY FlexiCapture 8.0 Professional est un système logiciel dédié à l'acquisition de données provenant de documents de types très variés :

Vous pouvez traiter des documents de types très variés dans un seul et même lot, de même que vous pouvez configurer le système pour traiter des documents de type mixte. Un type de document n'affecte que la méthode utilisée pour la création du modèle de document. Le type des documents traités n'affecte en rien le travail de l'opérateur.

Passons en revue les différents types de documents qu'il est possible de traiter à l'aide d'ABBYY FlexiCapture 8.0 Professional.

- **Documents structurés.** Les documents structurés sont des documents contenant un certain nombre de champs de données spécifiques dont l'emplacement et le marquage sont identiques sur toutes les copies du document. De tels documents sont appelés « formulaires fixes ». Les questionnaires, les enquêtes et les formulaires de demande sont habituellement considérés comme des formulaires fixes et sont le plus souvent des formulaires papier qui doivent être renseignés manuellement. Pour identifier un formulaire spécifique dans un flux de documents de différents types ainsi que pour extraire des données d'un tel formulaire, vous devez créer un modèle fixe uniforme qui renseigne le programme sur l'endroit où trouver les champs de données nécessaires. Certains formulaires fixes sont traités plus efficacement que d'autres car ils ont été créés pour répondre à des besoins spécifiques d'acquisition des données. De tels formulaires sont appelés « formulaires lisibles par machine ». ABBYY FormDesigner 8.0 est un outil efficace de création de formulaires lisibles par machine. Il est fourni avec ABBYY FlexiCapture 8.0 Professional. Pour plus d'informations sur la création de formulaires à l'aide d'ABBYY FormDesigner 8.0, consultez l'aide d'ABBYY FormDesigner 8.0 ainsi que toute autre documentation disponible. Une description des principales étapes de la création d'un modèle vous est proposée, tout particulièrement concernant les documents structurés.
- **Documents semi-structurés.** Il s'agit de documents contenant un certain nombre de champs de données dont la quantité, le marquage et l'emplacement peuvent varier d'une copie à l'autre du document. Par exemple, les factures sont des documents semi-structurés. En effet, les factures envoyées par différentes sociétés diffèrent souvent concernant le nombre de champs de données et leur format. Toutes les factures possèdent un numéro et mentionnent le montant total du

paiement mais ces champs de données ne se situent pas tous au moment endroit sur le document. Pour identifier des documents semi-structurés et pour en extraire des données, ABBYY FlexiCapture 8.0 Professional utilise des modèles flexibles (FlexiLayouts). Pour créer un FlexiLayout, vous devez utiliser le module spécial ABBYY FlexiLayout Studio. Pour tout complément d'informations sur ce modèle, consultez l'aide d'ABBYY FlexiLayout Studio et le guide de l'utilisateur. Le traitement des documents flexibles diffère du traitement des documents fixes uniquement concernant la création d'un modèle et son rattachement.

- **Documents non structurés.** L'application ABBYY FlexiCapture 8.0 Professional peut être utilisée pour traiter des documents non structurés tels que des contrats, des lettres ou des commandes, c'est-à-dire des documents dans lesquels les informations sont présentées dans un style relativement libre. Le programme peut automatiquement identifier les documents non structurés en tant que pièces jointes à des documents fixes ou semi-structurés, ou au moyen d'un FlexiLayout. Il est alors possible d'exporter ces documents sous forme de fichiers PDF indexés ou de fichiers graphiques. Les données provenant des champs d'index des documents non structurés peuvent être extraites manuellement ou bien automatiquement, à l'aide d'un FlexiLayout. Parmi les scénarios types de traitement de documents non structurés, citons par exemple la conversion d'une archive papier au format électronique et l'extraction de plusieurs champs d'index pour accélérer les recherches d'attributs.

Les chapitres suivants décrivent comment configurer ABBYY FlexiCapture 8.0 Professional pour traiter des documents de types variés, notamment la procédure d'acquisition de données automatisée, l'amélioration de la qualité de la reconnaissance et l'organisation de l'exportation des données.

L'acquisition de documents structurés y est largement développée. Toutes les étapes du traitement y sont décrites et des formulaires fixes sont utilisés à titre d'exemple. Les particularités des autres types de documents et les différences constatées dans la création de modèle pour ces documents y sont également expliquées.

4. Configuration du système pour les formulaires fixes d'acquisition

ABBYY FlexiCapture 8.0 Professional vous permet d'acquérir et de traiter des formulaires fixes, le tout rapidement et efficacement. La procédure pour travailler avec des formulaires fixes est la suivante :

- Concevez des formulaires ou rendez les formulaires existants lisibles par machine.
- Créez un modèle : décrivez la mise en page géométrique des objets, précisez les propriétés des objets, créez des règles de vérification et d'assemblage des documents, configurez l'exportation des données.
- Configurez la méthode d'ajout d'images et de création de profils d'importation d'image.
- Procédez à l'acquisition des documents, une fois tous les paramètres nécessaires définis.

4.1. Création d'un modèle de formulaire fixe

La conception d'un formulaire papier est très importante car son apparence détermine si un utilisateur doit renseigner ce formulaire ou si ce formulaire convient pour une acquisition automatisée des données. L'utilisateur doit pouvoir repérer immédiatement les zones de saisie des données. Seul un formulaire clair peut réduire considérablement les risques d'erreur. Au moment de concevoir un formulaire, vous devez

faire en sorte qu'il soit intuitif à renseigner. Les données saisies doivent également être clairement et aisément reconnaissables.

Si un formulaire remplit l'ensemble des critères d'acquisition de données, il est considéré comme étant **lisible par machine**. De tels formulaires possèdent les propriétés suivantes :

- Ils possèdent des **ancres** (ou **marques de référence**). Il s'agit d'éléments secondaires particuliers qui aident le programme à déterminer l'orientation du formulaire, à associer le modèle et à réaligner l'image numérisée si nécessaire. Les ancres sont représentées par des carrés noirs, des croix ou des coins.
- Tous les champs et tous les éléments graphiques (séparateurs, ancres, etc.) doivent se trouver exactement aux mêmes emplacements sur toutes les copies du formulaire. Tous les formulaires d'un lot d'impression donné doivent être créés au moyen d'une copie principale unique.
- L'ensemble des informations fournies à titre explicatif ne doit en aucun cas gêner l'extraction des données depuis les champs.

Pour plus d'informations sur les exigences propres aux formulaires lisibles par machine, consultez l'aide d'ABBYY FlexiCapture 8.0 Professional ainsi que le guide à la création de formulaires lisibles par machine.


ABBYY FormDesigner 8.0 est un outil spécial de création de formulaires lisibles par machine. Il est fourni avec ABBYY FlexiCapture 8.0 Professional. Pour plus d'informations sur la création de formulaires lisibles par machine dans ABBYY FormDesigner, consultez l'aide d'ABBYY FormDesigner et le guide de l'utilisateur. ABBYY FormDesigner vous permet de créer simplement et aisément des formulaires lisibles par machine. Un modèle de formulaire créé dans ABBYY FormDesigner peut être importé dans ABBYY FlexiCapture 8.0 Professional. Vous disposez alors d'une mise en page complète de tous les champs de données et des éléments graphiques. La seule chose qu'il vous reste à faire est de définir les propriétés des champs et de configurer l'exportation.

Pour rendre un formulaire existant lisible par machine, il vous suffit de procéder à quelques légères modifications de conception et d'ajouter un certain élément (le plus souvent des ancres). Si vous ne parvenez pas à modifier les formulaires vierges pour une raison quelle qu'elle soit, vous pouvez configurer le programme pour un traitement des formulaires sans ancre. Vous pouvez utiliser des éléments de formulaire autres que les ancres, par exemple les lignes verticales et horizontales, du texte explicatif ou des codes barres. L'association du modèle sera plus efficace si vous utilisez des ancres standard.

4.1.1. Éléments de formulaire

Analysons à présent les principaux *éléments de formulaire*. (Figure 1)

Titre de formulaire



QUESTIONNAIRE CLIENT

Date de saisie

Prénom et nom :

Profession :

Adresse électronique :

Nom de la société :

Pays :

Activité de la société :

Informatique <input type="checkbox"/>	Santé <input type="checkbox"/>	Enseignement <input type="checkbox"/>	Finance/Banque <input type="checkbox"/>
Externalisation <input type="checkbox"/>	Secteur public <input type="checkbox"/>	Assurance <input type="checkbox"/>	Production <input type="checkbox"/>

Autre :

Types de document à traiter :
(plusieurs choix possibles)

Factures <input type="checkbox"/>	Tests <input type="checkbox"/>	Rapports <input type="checkbox"/>	Formulaires de demande <input type="checkbox"/>
Bordereaux <input type="checkbox"/>	Commandes <input type="checkbox"/>	Enquêtes <input type="checkbox"/>	Questionnaires <input type="checkbox"/>

Autre :

Volume de traitement (pages par mois) :

De quelle façon les documents sont-ils actuellement traités dans votre entreprise :

Aucun traitement <input type="checkbox"/>	Saisie manuelle des données <input type="checkbox"/>	Numérisation <input type="checkbox"/>	FormReader 6.5 <input type="checkbox"/>
---	--	---------------------------------------	---

Autre :

Nom du fabricant du dispositif que vous prévoyez d'utiliser pour la numérisation :
(plusieurs choix possibles)

Konica Minolta® <input type="checkbox"/>	Fujitsu® <input type="checkbox"/>	Kodak® <input type="checkbox"/>	Xerox® <input type="checkbox"/>	HP® <input type="checkbox"/>	Kyocera® Mita <input type="checkbox"/>
--	-----------------------------------	---------------------------------	---------------------------------	------------------------------	--

Autre :

Quel type d'application de gestion ou de stockage comptez-vous utiliser en complément ?
(plusieurs choix possibles)


MS Access® <input type="checkbox"/>	MS SQL Server® <input type="checkbox"/>	Documentum® <input type="checkbox"/>	IBM Content Manager® <input type="checkbox"/>	MS SharePoint® <input type="checkbox"/>
-------------------------------------	---	--------------------------------------	---	---

Autre :

Vos commentaires et vos attentes :

Je voudrais faire partager mon expérience d'ABBY FlexiCapture. Veuillez me contacter par courrier électronique

Merci d'avoir renseigné ce formulaire. Nous vous invitons à procéder à son traitement au moyen d'ABBY FlexiCapture et à envoyer les données extraites à FlexiCapture_feedback@abby.com.



© 2009 ABBYY. Que n'aurait et SQL Server sont des marques déposées de Microsoft Corporation aux États-Unis et dans d'autres pays. Xerox est une marque commerciale de Xerox Corporation aux États-Unis et dans d'autres pays. Kodak est une marque commerciale d'Eastman Kodak Company. IBM est une marque déposée d'IBM aux États-Unis et dans d'autres pays. Toutes les marques commerciales mentionnées dans ce questionnaire sont la propriété de leurs détenteurs respectifs.




Figure 1. Exemple de formulaire lisible par machine contenant les principaux éléments

- **Champs de données.** Tous les formulaires destinés à recueillir des informations contiennent des champs de données. Ces champs sont habituellement accompagnés d'un texte explicatif. Les types de champ de données sont les suivants :

Champs de texte : utilisés pour saisir des données de texte. De tels champs correspondent à des groupes de cellules de caractères destinés à la saisie. La conception d'un champ de texte invite la personne qui remplit le formulaire à utiliser des caractères séparés.

Coches : également conçues pour recueillir des informations. Toutefois, aucun besoin cette fois-ci de saisir du texte, il suffit de cocher les éléments nécessaires. Une coche possède habituellement un contour fermé (un carré, un cercle ou un polygone) et les informations sont saisies au moyen d'un signe (par exemple, une marque de pointage ou une croix) inséré dans ce contour. Parfois, les coches ne disposent d'aucun contour. Auquel cas, l'utilisateur doit insérer le signe au niveau de l'arrière-plan de couleur blanche, à l'emplacement prévu sur la page. Si vous souhaitez autoriser les corrections apportées aux coches, sélectionnez l'option correspondante au moment de créer le modèle. Dans ce cas, une coche raturée sera considérée comme décochée.

Groupes de coches. Un groupe de coches consiste en plusieurs coches situées les unes près des autres et reliées de manière logique. Dans le cadre de la règle établie, les réponses associées aux coches d'un groupe sont mutuellement exclusives. En d'autres termes, il ne peut être coché qu'un seul élément du groupe.

Les champs de données peuvent également être représentés par des tableaux.

- **Ancres.** Les ancres sont utilisées pour déterminer l'orientation de la page et associer le modèle. Le programme utilise également des éléments de référence pour contrôler et corriger (réaligner) les distorsions d'image qui peuvent apparaître durant la numérisation. Les ancres sont représentées par des carrés noirs (le plus souvent), des croix ou des coins. Nous vous recommandons d'utiliser 5 ancres par formulaire : quatre à chaque coin et un sur l'un des côtés, de façon à pouvoir détecter de manière fiable l'orientation de la page. L'association du modèle est alors rapide et précise. En outre, vous pourrez acquérir plusieurs formulaires à la fois, des formulaires imprimés sur des imprimantes différentes ou reçus par fax.
- **Identificateurs.** Les identificateurs sont utilisés pour détecter le formulaire auquel une page appartient et pour sélectionner le modèle nécessaire si plusieurs modèles existent dans le lot, avec des jeux d'éléments de référence similaires. Si vous traitez plusieurs formulaires à la fois, vous devez insérer un élément unique sur chacune des pages d'un formulaire. Cet élément informera le programme du formulaire auquel la page appartient. Les identificateurs peuvent être des codes barres, des ancres, des séparateurs ou du texte statique (par exemple, un titre de formulaire ou une partie d'un texte explicatif).
- **Images graphiques.** Il vous faudra peut-être enregistrer certains objets sous forme d'images, par exemple des photos, des signatures, un cachet ou un timbre. ABBYY FlexiCapture 8.0 Professional peut enregistrer les objets images et les exporter dans des fichiers ou des bases de données.
- **Éléments décoratifs.** Un formulaire peut contenir certains éléments décoratifs, par exemple le logo d'une société.

les cellules les unes des autres, le programme détecte les lignes verticales et horizontales de la mise en page et les sépare du contenu du champ.

DÉTAILS PERSONNELS

Prénom

Nom

Ville de naissance

Le questionnaire que vous avez créé dans ABBYY FlexiCapture 8.0 Professional constitue un bon exemple de formulaire fixe lisible par machine. Il est clairement identifiable comme étant un formulaire simili détourné couleur. Veuillez analyser la juxtaposition des principaux éléments sur ce formulaire.

4.2. Création d'un projet

Un projet contient tous les paramètres d'acquisition de document nécessaires : modèles de document, profils d'importation d'image, paramètres du programme et documents traités.

Les documents sont regroupés en *lots*. Le nombre de lots dépend de votre approche de traitement : vous pouvez traiter tous les documents dans un lot ou trier les documents en lots en fonction de leur date d'importation ou de leur date de numérisation.

Les documents sont traités par *lots de travail*. Seuls les lots de travail sont accessibles en mode opérateur. Les *lots de test* sont utilisés par l'administrateur à des fins de test et de réglage des modèles. La différence principale entre les deux types de lots est que les modèles locaux sont utilisés pour le traitement des documents de lots de test, tandis que les modèles publiés sont utilisés pour le traitement des documents de lots de travail.

Un document contient des images d'une ou de plusieurs pages (documents à une page ou documents multipages), ainsi que les données extraites de ces pages.

Un projet peut contenir plusieurs modèles. Dans ce cas, des documents de différents types sont traités au sein d'un seul et même projet. Vous n'avez pas besoin de trier le document à traiter car des documents de types différents font l'objet d'une acquisition dans un flux unique. Toutefois, si les flux de document ne se croisent pas, vous pouvez créer un projet séparé pour ces documents.

Pour commencer, l'administrateur doit créer un projet et au moins un modèle de document.

Pour créer un projet, cliquez sur **Créer nouveau...** dans la boîte de dialogue **Ouvrir un projet** qui s'affiche lorsque vous lancez le programme. Vous pouvez également sélectionner **Fichier > Nouveau projet...** dans la fenêtre principale du programme. Indiquez le dossier dans lequel le projet doit être enregistré et donnez-lui un nom.

Pour ajouter des lots au projet, cliquez avec le bouton droit de la souris dans la fenêtre principale du programme qui contient la liste des lots (Figure 2) et sélectionnez **Nouveau lot** dans le menu contextuel. Vous pouvez charger des images sans créer de lot, auquel cas le programme crée un lot automatiquement.

Pour afficher les documents ajoutés au lot, double-cliquez sur le nom du lot. Pour retourner à la liste des lots, sélectionnez **Projet > Liste des lots de travail** ou appuyez sur CTRL+B.

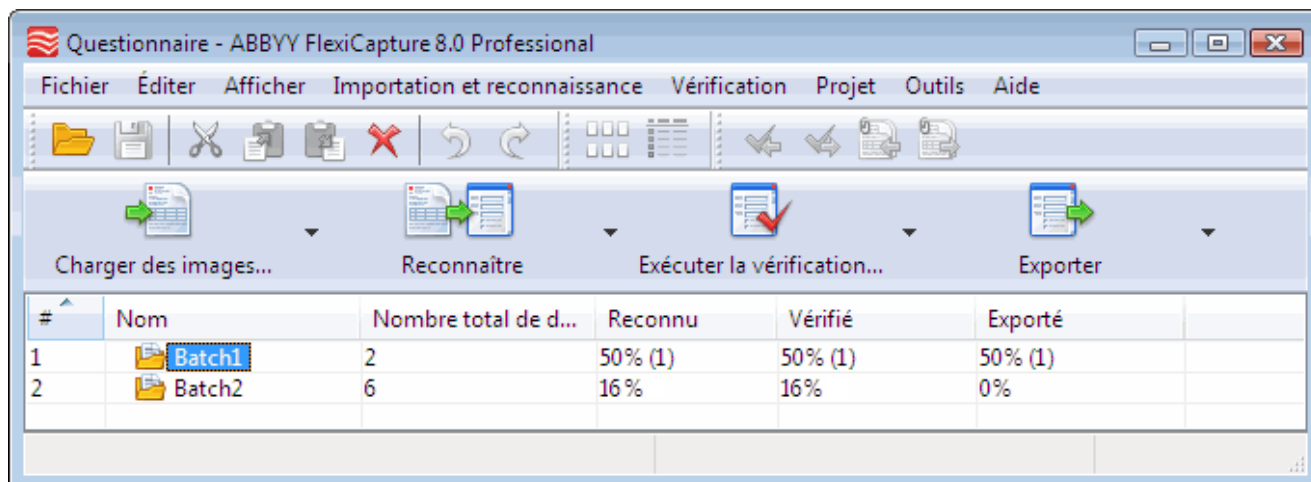


Figure 2. Fenêtre principale d'ABBYY FlexiCapture 8.0 Professional

Lancez ABBYY FlexiCapture 8.0 Professional. Pour créer un projet, cliquez sur Créer nouveau... dans la boîte de dialogue Ouvrir un projet qui s'affiche lorsque vous lancez le programme. Vous pouvez également sélectionner Fichier > Nouveau projet... dans la fenêtre principale du programme. Indiquez le dossier dans lequel le projet doit être enregistré et donnez-lui un nom. Le nouveau projet s'ouvre lorsque vous cliquez sur Créer.

4.3. Création d'un modèle de document

L'étape la plus importante dans la configuration d'un projet est la création d'un modèle. La qualité des données reçues après le traitement des formulaires dépend de l'exactitude du modèle. Pour créer un modèle, plusieurs éléments doivent être indiqués :

- Les *éléments statiques* de l'image : ancrs, séparateurs, texte statique et codes barres. Sélectionnez parmi ces éléments ceux devant être utilisés pour l'association du modèle et l'identification du document. Les ancrs sont automatiquement détectées et marquées.
- L'emplacement de tous les *champs*. Les champs doivent correspondre aux zones de l'image à partir desquelles les données sont extraites.
- Propriétés de chaque champ : sélectionnez les types de données à rechercher dans chacun des champs (cela améliore considérablement la qualité de la reconnaissance) et indiquez les champs devant être envoyés à l'opérateur pour vérification, etc.
- Les règles utilisées pour la vérification des valeurs de champ. De telles règles aident le programme à détecter les documents dont les valeurs ne correspondent pas à certaines conditions définies, par exemple une valeur de champ qui ne correspond pas aux valeurs de la base de données.
- Méthode d'exportation des données. Les données peuvent être exportées dans un fichier ou dans une base de données, ou bien selon la procédure décrite dans un script.

Une fois créé, le modèle doit être publié pour pouvoir être utilisé dans des reconnaissances de document ultérieures.

Pour créer un nouveau modèle, sélectionnez **Projet > Modèles de document...** Cliquez sur **Nouveau...** dans la boîte de dialogue qui apparaît. Vous lancez alors l'Assistant de création des modèles de document. Utilisez la boîte de dialogue **Créer un nouveau modèle de document** pour indiquer les


propriétés principales du modèle : son nom, sa description, ses paramètres régionaux et son style d'écriture. Sélectionnez le type de texte : **ICR (manuscrit)** si la plupart des champs du document sont renseignés manuellement ou **OCR (imprimé)** si les valeurs de la plupart des champs du document sont imprimées. Dans le deuxième cas, sélectionnez le type d'impression dans la liste déroulante. Le type de texte défini à ce niveau sera le type de texte par défaut mais vous pourrez le modifier pour des champs en particulier.

Ensuite, chargez ou numérisez l'image sur la base de laquelle le modèle est créé. Si votre document comprend plusieurs pages, chargez la première page. Au moment d'ajouter les pages restantes, consultez les recommandations définies dans le chapitre Création d'un modèle pour un document multipage. Vous pouvez numériser l'image d'une page vierge ou la charger depuis un fichier. Si vous prévoyez de traiter des documents semi-structurés, vous devez créer un modèle en utilisant une description flexible. Si tel est le cas, sélectionnez l'option **Charger FlexiLayout**, puis indiquez le chemin du fichier AFL contenant la description flexible créée dans ABBYY FlexiLayout Studio.

Vous pouvez à présent sélectionner les types de champ automatiquement détectés. Indiquez les coches et les champs de saisie de texte. Les recherches automatiques de champs de texte possédant un type de marquage et des coches rectangulaires se révèlent très efficaces. Toutefois, si les champs de texte de votre formulaire ne contiennent aucun marquage et que les coches doivent être apposées sur un arrière-plan de couleur blanche, nous vous recommandons de marquer ces champs manuellement.

Les ancres seront automatiquement détectées et marquées si l'image en contient.

Dans la fenêtre principale du programme, sélectionnez **Projet > Modèles de document...** Cliquez sur **Nouveau...** dans la boîte de dialogue **Modèles de document**.

1. Dans la boîte de dialogue qui apparaît, indiquez les paramètres du modèle, son nom et sa description. Dans le champ Langue (lieu), indiquez la langue dans laquelle le formulaire doit être renseigné. Dans le champ Style d'écriture, sélectionnez le pays correspondant. Cette procédure s'explique par le fait que certains caractères, par exemple les chiffres, s'écrivent différemment d'un pays à un autre. Sélectionnez le type de texte : ICR (manuscrit). Cliquez sur Suivant.
2. Sélectionnez l'option Numériser, puis numérisez le formulaire vierge sans les filtres de sorte que la couleur de l'arrière-plan soit conservée. Cette procédure vous aide à plus facilement marquer les champs de données de l'image sur le formulaire (par la suite, lorsque vous numériserez les images de la page pour reconnaître les données, vous sélectionnerez un mode de numérisation qui supprime l'arrière-plan). Le modèle de page se crée alors d'après cette image. Vous pouvez également sélectionner l'option Charger à partir du/des fichier(s) et charger l'image depuis un fichier. (Vous trouverez une image d'un questionnaire sur <nom du disque>:\Documents and Settings\All Users\Application Data\ABBYY\FlexiCapture\8.0\Samples\FormDesigner\French\Questionnaire_fr, sous Microsoft Windows Vista : <nom du disque>:\Users\Public\ABBYY\FlexiCapture8.0\Samples\FormDesigner\French\Questionnaire_fr). Cliquez sur Suivant.
3. Sélectionnez les options Texte et Coches pour marquer automatiquement les champs de texte et les coches. Cliquez sur Terminer.
4. Au terme de la procédure de l'Assistant, la fenêtre Éditeur de modèles de document s'ouvre. Le questionnaire apparaît dans la fenêtre et tous les champs de texte ainsi que les coches sont marqués. Les noms des champs du modèle qui sont automatiquement détectés s'affichent quant à eux dans la fenêtre Structure de document. Si vous passez en mode Éléments statiques (bouton ) , vous accédez à la présentation des éléments statiques. Les ancres seront également automatiquement marquées.

4.3.1. Éditeur de modèles de document

Toutes les principales actions de création et d'édition d'un modèle s'effectuent dans la fenêtre **Éditeur de modèles de document** (Figure 3), qui s'ouvre au terme de la procédure de l'Assistant de création de modèle. Pour ouvrir l'Éditeur de modèles de document depuis la fenêtre principale du projet,

sélectionnez **Projet > Modèles de document...**, puis sélectionnez le nom du modèle et cliquez sur **Éditer...**

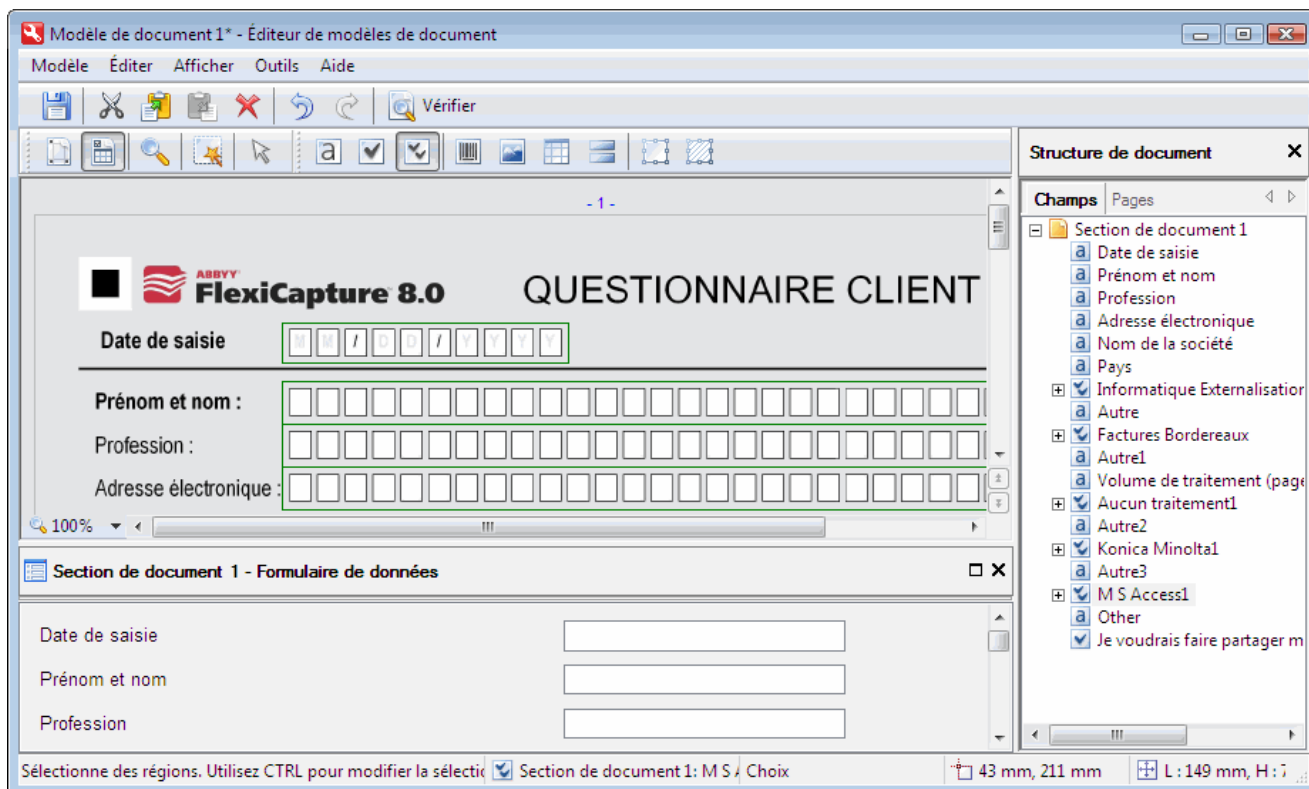





Figure 3. Éditeur de modèles de document

4.3.2. Utilisation des éléments pour marquer des objets dans le formulaire

Au terme de la procédure de l'Assistant de création de modèles, l'image chargée s'affiche dans la fenêtre Éditeur de modèles de document. Les ancrés et les champs de données des types sélectionnés durant les précédentes étapes de la création du modèle sont déjà marqués.

Par la suite, vous pouvez automatiquement marquer les objets en utilisant l'outil  et en cliquant sur la zone de l'élément à marquer. Le programme détecte alors automatiquement le type et l'emplacement de l'élément.

L'Éditeur de modèles de document propose des outils intuitifs et conviviaux destinés à marquer les champs et les éléments statiques. Deux modes sont proposés dans l'Éditeur :








- Mode marquage de champ (outil ) et
- Mode Éléments statiques (outil )

Pour créer manuellement un élément statique ou un champ de données, passez du mode marquage de champ ou mode Éléments statiques et vice-versa, puis cliquez sur le bouton correspondant de la barre d'outils. Pour créer un élément, faites glisser le curseur pour dessiner un rectangle autour de l'objet sur le formulaire. Vous pouvez également sélectionner l'outil nécessaire, puis cliquez à proximité de l'objet tout



en maintenant la touche MAJ enfoncée. La zone du champ de données ou de l'élément statique sera alors automatiquement détectée.

Pour la suite, passez en revue la liste des outils disponibles pour créer des éléments de types différents :

Champs :

-  - champ de saisie de texte
-  - coche
-  - groupe de coches
-  - code barre
-  - élément graphique (image)
-  - tableau
-  - champ de groupe

Éléments statiques :

-  - ancre
-  - séparateur
-  - texte statique
-  - code barre

Un code barre peut être à la fois un champ de reconnaissance et un élément statique. Vous devez faire attention au moment de sélectionner le mode de marquage : si les données doivent être extraites à partir du code barre, utilisez le mode marquage de champ. Si le code barre est utilisé à des fins d'identification du document et d'association du modèle, procédez à un marquage en mode Éléments statiques.


Les champs créés s'affichent dans la liste sous l'onglet **Champs** de la fenêtre **Structure de document**. Par défaut, ces champs sont nommés en fonction du texte explicatif. Vous pouvez renommer un champ en le sélectionnant dans la structure de document et en appuyant sur F2. Pour nommer le champ en fonction du texte explicatif, sélectionnez-le, effectuez un clic droit sur celui-ci, puis sélectionnez **Obtenir le nom de l'image...** dans le menu contextuel. Après cela, faites glisser le curseur pour dessiner un rectangle sur le texte explicatif de l'image.




Le type de champ est marqué d'un pictogramme dans la liste des champs et de la couleur du cadre de l'image. Les éléments statiques ne font pas partie de la liste.

Vous pouvez copier des éléments (y compris dans d'autres sections du document), en supprimer, en déplacer ou en modifier la taille. Si vous copiez des champs, des numéros sont automatiquement ajoutés aux noms des champs.

Pour sélectionner plusieurs éléments en même temps, cliquez tout en maintenant la touche CTRL enfoncée. L'action est alors exécutée sur tous les éléments sélectionnés. Pour sélectionner des éléments,


utilisez l'outil .

1. Les champs de saisie de texte sont déjà automatiquement marqués sur le formulaire. Si vous n'avez pas sélectionné l'option Texte lors de la dernière étape de création du modèle ou si vous souhaitez marquer manuellement les champs de texte, utilisez l'outil  suivi des champs dans lesquels saisir votre texte. Un champ doit intégrer toutes les cellules de caractères, ainsi que des espaces supplémentaires sur les côtés. Intégrez des champs pour la date de saisie du formulaire, le nom, la profession, le pays, l'adresse électronique, le nom de la société, le nombre de pages traité tous les jours et les 4 champs destinés aux informations supplémentaires.

2. Veuillez marquer le champ Vos commentaires et vos attentes. Ce champ n'a pas été détecté automatiquement, il n'est donc pas marqué.
3. Si les coches ont été marquées automatiquement, sélectionnez l'outil  et marquez toutes les coches n'appartenant pas aux groupes. N'oubliez pas d'insérer un espace de chaque côté des coches. Les coches qui appartiennent à un groupe n'ont pas besoin d'être marquées individuellement. Sélectionnez l'outil  pour dessiner un rectangle autour de toutes les coches appartenant au groupe. Chaque coche individuelle sera automatiquement marquée et un nom lui sera affecté. Par la suite, toutes ces coches seront regroupées.
4. Passez en mode Éléments statiques. Les ancres sont déjà marquées sur le formulaire. Sélectionnez l'outil  et marquez le code barre sur l'image.

4.3.2.1. Groupes de champs

Les champs peuvent être regroupés afin d'optimiser la structure de document et de créer des groupes de champs répétitifs. Par exemple, il est possible de regrouper la ville, la rue et le numéro du bâtiment pour créer l'« adresse personnelle ». Si vous souhaitez créer des champs pour l'adresse professionnelle, il vous suffit tout simplement de copier le groupe de champs « Adresse personnelle ».

Pour associer des champs et créer un groupe, utilisez l'outil .

Si un document possède des groupes de champs répétitifs, il vous suffit de créer un seul groupe puis plusieurs *instances* de ce groupe. L'ensemble des propriétés des champs et toutes les règles définies pour le groupe d'origine s'appliquent également à toutes les instances du groupe. Pour plus d'informations, veuillez consulter le chapitre [Champs avec instances multiples](#).

Vous pouvez également copier un groupe de champs. Toutefois, dans ce cas, le nouveau groupe est indépendant.


4.3.2.2. Champs sans marquage

Certains champs ne sont parfois associés à aucune région sur l'image. Les noms des champs non marqués sont indiqués d'un astérisque rouge. De tels champs servent à enregistrer les résultats provisoires de calculs qui utilisent les valeurs des champs de reconnaissance.

Les champs non marqués possèdent toutes les propriétés afférentes à leur type : ils peuvent être envoyés à l'opérateur pour vérification, leur format est vérifiable et les valeurs qu'ils contiennent peuvent être exportées.

Il existe deux méthodes de création d'un champ sans marquage :

1. Dans la fenêtre Éditeur de modèles de document, sélectionnez **Modèle > Créer un champ** dans le menu, puis créez un champ du type souhaité. Le nom du champ apparaît dans la liste et est indiqué d'un astérisque. Un champ est créé dans la structure de document, ce qui n'est pas le cas de sa région sur l'image.
2. La deuxième méthode consiste à supprimer le marquage d'un champ standard. Sélectionnez les champs nécessaires sur l'image ou dans la liste, puis choisissez **Supprimer la région** dans le menu contextuel. Le marquage est supprimé et le nom du champ est alors marqué d'un astérisque rouge.


Pour créer une région sur l'image destinée à un champ sans marquage, sélectionnez l'outil  dans la barre d'outils, puis créez un cadre autour de la région concernée. S'il existe des champs de la liste


marqués d'un astérisque, le programme vous invitera à sélectionner un nom pour l'un de ces champs (par exemple, si vous créez en premier la liste des champs de document et que vous définissez leur emplacement par la suite).



4.3.2.3. Marquage des tableaux

ABBYY FlexiCapture 8.0 Professional vous permet de travailler avec des tableaux. En créant des champs de type « Tableau », vous facilitez grandement la configuration, l'extraction et l'exportation des données de ces tableaux. Un champ Tableau se compose d'un jeu de colonnes de type unique contenant des lignes répétitives.

Le programme possède des outils spéciaux destinés au marquage des tableaux dans un modèle fixe. Ces outils vous aideront à dessiner un tableau et à en marquer les lignes et les colonnes.

Sélectionnez l'outil  pour dessiner la région du tableau. Veuillez noter que cette région n'intègre pas l'en-tête du tableau. Une fois le tableau créé dans l'image, les outils de marquage du tableau s'affichent

dans la barre d'outils. Pour ajouter des séparateurs, marquez les cellules du tableau à l'aide de l'outil . Pour créer des séparateurs verticaux, placez le pointeur dans le tableau, faites glisser le séparateur en pointillé à l'emplacement souhaité et cliquez une fois. Pour créer des séparateurs horizontaux, exécutez la même opération en maintenant la touche ALT enfoncée. Vous pouvez également détecter les séparateurs automatiquement. Pour détecter des séparateurs, sélectionnez le tableau créé, puis sélectionnez l'option **Détecter automatiquement les séparateurs verticaux** ou **Détecter automatiquement les séparateurs horizontaux** dans le menu contextuel. Supprimez les séparateurs à

l'aide de l'outil . Une fois que vous avez ajouté le nombre de séparateurs requis, marquez les colonnes du tableau. Pour cela, sélectionnez l'outil  et utilisez votre curseur pour définir la région d'une colonne. Chaque colonne contient des cellules de type identique : texte, coches, éléments graphiques ou codes barres. Le programme vous invitera alors à sélectionner le type de colonne durant le marquage.

Une fois le tableau marqué, vous devez définir pour chaque colonne, les propriétés de reconnaissance, les propriétés de vérification, ainsi que le type de données. La définition des propriétés des colonnes est identique à celle des propriétés des champs de document standard.

4.3.2.4. Champs avec régions multiples

Si votre formulaire contient des champs de données dont la région se compose de plusieurs parties (par exemple, des tableaux qui commencent sur une page et se terminent sur la page suivante), vous pouvez créer des champs avec régions multiples afin de traiter ces objets sur le formulaire.

Les valeurs de toutes les régions d'un champ sont regroupées et exportées en tant que champ unique. Le retour à la ligne est utilisé comme séparateur.

Pour créer un champ avec régions multiples, commencez par créer l'une des régions des champs, sélectionnez cette région, puis choisissez **Continuer la région...** dans le menu contextuel. Sélectionnez la zone de la page où la région créée doit se poursuivre et répéter la procédure autant de fois que nécessaire.

4.3.2.5. Champs avec instances multiples

Vos documents contiennent peut-être des objets répétitifs, c'est-à-dire des champs ou des groupes de champs qui sont utilisés à plusieurs reprises dans un document et qui décrivent des informations similaires, par exemple des données détaillées concernant les salariés, les enfants ou les factures. Pour traiter de tels objets, vous pouvez créer des champs avec instances multiples.

Tous les champs peuvent posséder des instances multiples et ces instances être espacées d'une distance quelle qu'elle soit, même sur des pages différentes. Les instances des champs possèdent des propriétés identiques. Les champs avec instances multiples sont exportés dans des tables de base de données ou des fichiers séparés.

Les champs avec instances multiples se révèlent très utiles pour créer des groupes de champs répétitifs. Vous pouvez créer un groupe de champs, puis simplement créer le nombre d'instances nécessaire.

Pour créer un champ avec instances multiples, commencez par créer un champ, sélectionnez la région du champ à l'aide du bouton droit de la souris, puis choisissez **Nouvelle instance...** dans le menu contextuel. Déplacez la région créée à l'emplacement souhaité sur la page. Créez autant d'instances du champ que nécessaire et déplacez-les aux emplacements souhaités sur la page.

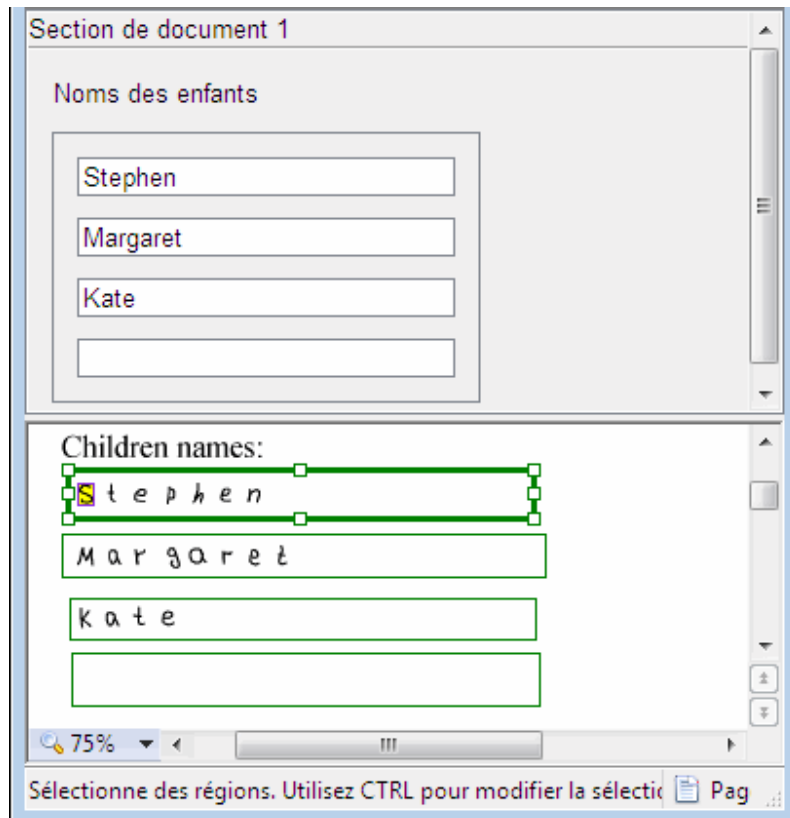



Figure 4. Exemple de l'utilisation d'un champ avec instances multiples

4.3.2.6. Exclusion d'une région de la reconnaissance

Parfois, il vous faudra exclure de la reconnaissance une région en particulier d'un formulaire. Cela arrive par exemple lorsque du texte explicatif masque les données à extraire du formulaire (voir Figure 5). Pour

exclure une région de la reconnaissance, sélectionnez l'outil  et indiquez la région à exclure.

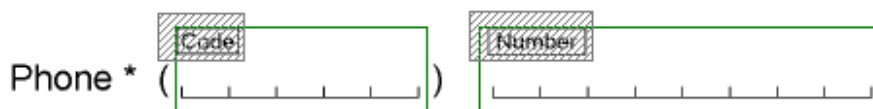



Figure 5. Exemple d'exclusion d'une région de la reconnaissance

4.3.2.7. Suppression des champs

Pour supprimer un champ, sélectionnez-le, puis appuyez sur la touche **SUPPR** ou sélectionnez **Supprimer** dans le menu contextuel. Pour ne supprimer que la région du champ tout en conservant le champ dans la structure de document, appuyez sur les touches **MAJ+SUPPR** ou sélectionnez **Supprimer une région** dans le menu contextuel de ce champ.

4.3.3. Éléments statiques

Les éléments statiques marquent les objets desquels des données ne sont pas extraites. De tels éléments sont utilisés pour l'association du modèle et l'identification de document uniquement.

Pour pouvoir travailler avec des éléments statiques, vous devez passer en mode Éléments statiques en cliquant sur le bouton . Le marquage des éléments statiques apparaît uniquement dans ce mode.

Tous les types d'élément statique peuvent être utilisés pour l'association du modèle. Toutefois, vous obtenez de meilleurs résultats lorsqu'un document possède des éléments spéciaux (ancres standard) : carrés, croix et coins. Ils doivent être marqués automatiquement ou manuellement comme des éléments statiques de type *Ancre*. La forme des ancres peut être sélectionnée dans l'onglet **Général** des propriétés de l'élément.

Vous pouvez utiliser tout type d'éléments statiques comme identificateurs. Le programme utilise l'emplacement des identificateurs ou leurs valeurs pour déterminer à quel document la page actuelle appartient. Si un code barre sert d'identificateur, il est possible de définir les valeurs de ce code barre. Cela se révèle utile pour identifier la page rapidement et précisément.

Pour utiliser un élément statique en vue de l'association du modèle et/ou de l'identification, vous devez sélectionner l'option correspondante dans l'onglet **Général** de la boîte de dialogue Propriétés. Pour ouvrir la boîte de dialogue Propriétés d'un élément, sélectionnez **Propriétés...** dans le menu contextuel de l'élément.

Ancre : Pour utiliser l'ancre en vue de l'association du modèle (recommandé), sélectionnez l'option **Utiliser pour l'association des modèles**. Pour utiliser l'ancre en vue de l'identification, sélectionnez l'option **Utiliser pour l'identification des modèles**. Sélectionnez le type d'ancre (carré noir, croix ou coin).

Texte statique : Pour utiliser le texte en vue de l'association du modèle, sélectionnez l'option **Utiliser pour l'association des modèles**. Pour utiliser le texte en vue de l'identification, sélectionnez l'option **Utiliser pour l'identification des modèles**. Si un texte statique est utilisé pour l'identification, définissez-en la valeur. Notez que vous devez définir la valeur du texte statique uniquement s'il est impossible d'identifier la page par l'emplacement du texte. La valeur du texte statique est alors requise (par exemple, des pages qui diffèrent uniquement par leurs titres alors que l'emplacement et la taille de ces titres sont identiques).

Séparateur : Pour utiliser le séparateur en vue de l'association du modèle, sélectionnez l'option **Utiliser pour l'association des modèles**. Pour utiliser le séparateur en vue de l'identification, sélectionnez l'option **Utiliser pour l'identification des modèles**.

Code barre : Pour utiliser le code barre en vue de l'association du modèle, sélectionnez l'option **Utiliser pour l'association des modèles**. Pour utiliser le code barre en vue de l'identification, sélectionnez l'option **Utiliser pour l'identification des modèles**. Si un code barre est utilisé pour l'identification, définissez-en la valeur. Pour cela, dans l'onglet **Reconnaissance**, indiquez le type de code barre, son orientation, ainsi que les options de traitement de l'image.

Cinq ancrés ont été marqués sur votre formulaire. Veillez à sélectionner le type Carré noir pour toutes les ancrés, ainsi que les options Utiliser pour l'association des modèles et Utiliser pour l'identification des modèles.

Concernant le code barre, sélectionnez les options Utiliser pour l'association des modèles et Utiliser pour l'identification des modèles. Si la valeur du code barre ne s'affiche pas automatiquement dans les champs Valeur du champ, cliquez sur Conseil. La valeur du code barre sera alors insérée.

Le code barre sur cette page correspond au type EAN 13 qui doit être défini dans le champ Type de code de l'onglet Reconnaissance. L'orientation est De gauche à droite.

Cinq ancrés et un code barre suffisent à identifier clairement le modèle et à l'associer, à moins que vous ne prévoyiez de traiter d'autres documents possédant une organisation identique d'ancres dans le même flux.

4.3.3.1. Particularités des codes barres

Si un code barre est utilisé comme identificateur, il s'agit alors d'un code barre ancre et donc d'un élément statique. De tels codes barres doivent être créés en mode Éléments statiques. La boîte de dialogue Propriétés pour un tel code barre contient deux onglets : **Général** et **Reconnaissance**

Si vous prévoyez d'extraire les données d'un code barre, sachez que ce code barre correspond alors à un champ de données et qu'il doit être créé en mode marquage de champ. La boîte de dialogue Propriétés pour un tel code barre contient tous les onglets des champs de données, à savoir **Général**, **Type de données**, **Reconnaissance**, **Vérification** et **Règles**. La valeur de ce code barre sera reconnue et envoyée pour vérification, en fonction des paramètres définis.

4.3.4. Propriétés des champs

Il est très important que les propriétés des champs soient correctement définies. Les propriétés des champs affectent la qualité de la reconnaissance et déterminent si les valeurs des champs doivent être exportées et envoyées à l'opérateur pour vérification. Certaines propriétés sont tout particulièrement importantes pour la reconnaissance des données. Le type de marquage d'un champ de texte constitue un bon exemple. Vous devez définir cette propriété correctement pour que le marquage qui n'est pas supprimé durant la numérisation soit séparé des caractères. Si vous procédez de la sorte, le résultat de la reconnaissance contiendra uniquement le texte saisi dans le champ. Aucune autre acquisition d'objet ne sera effectuée.

Si les propriétés des champs sont correctement définies, les résultats de la reconnaissance de données seront excellents et les opérateurs ne devront pas s'acquitter de tâches supplémentaires. Au moment de définir les propriétés des champs, vous pouvez faire en sorte que la vérification du document soit minimale car ces valeurs seront automatiquement corrigées.

Une fois un champ d'un certain type créé, les propriétés par défaut lui sont affectées. Pour modifier les propriétés d'un élément dans la boîte de dialogue, sélectionnez **Propriétés** dans le menu contextuel de l'élément. Chaque type de champ possède sa liste de propriétés exclusive.

4.3.4.1. Propriétés des champs communs

La boîte de dialogue **Propriétés** d'un champ, quel que soit son type, contient l'onglet **Général** (Figure 6). Le nom des champs et leurs descriptions sont proposés dans cet onglet. Lors de la création d'un champ, le programme génère automatiquement un nom, en fonction du texte d'explication le plus proche du

champ. Vous avez la possibilité de modifier ce nom par défaut. La **légende** est le nom du champ qui apparaît dans le formulaire de données. La définition du type de champ intervient dans cet onglet au moyen d'un pictogramme situé à droite du nom du champ.

En outre, il est possible de sélectionner les options suivantes dans l'onglet **Général** :

- **Exporter la valeur du champ.** Désactivez cette option si vous n'avez pas besoin d'exporter la valeur du champ. La désactivation de cette option peut se révéler utile si la valeur du champ est par exemple utilisée pour obtenir la valeur d'un champ calculé (consultez le chapitre Vérification basée sur les règles) et seul le résultat final doit être exporté.
- **Lecture seule.** Sélectionnez cette option pour passer le champ en lecture seule. L'opérateur ne pourra pas en modifier la valeur. Vous pouvez sélectionner cette option pour les champs dont la valeur doit être calculée automatiquement, à l'aide d'une règle. Par exemple, procédez ainsi pour un champ dans lequel la somme ou les valeurs fusionnées des autres champs sont stockées (consultez le chapitre Vérification basée sur les règles).
- **Afficher dans le formulaire de données.** Désactivez cette option si vous ne souhaitez pas que ce champ apparaisse dans le formulaire de données.
- **Ne peut pas être vide.** Activez cette option si le champ doit être renseigné.
- **Champ d'index.** Sélectionnez cette option si vous prévoyez d'utiliser ce champ pour l'indexation du document. Si vous procédez ainsi, la valeur du champ dans chaque document de la liste sera indexée et l'opérateur pourra utiliser cette valeur pour trier et rechercher des documents.

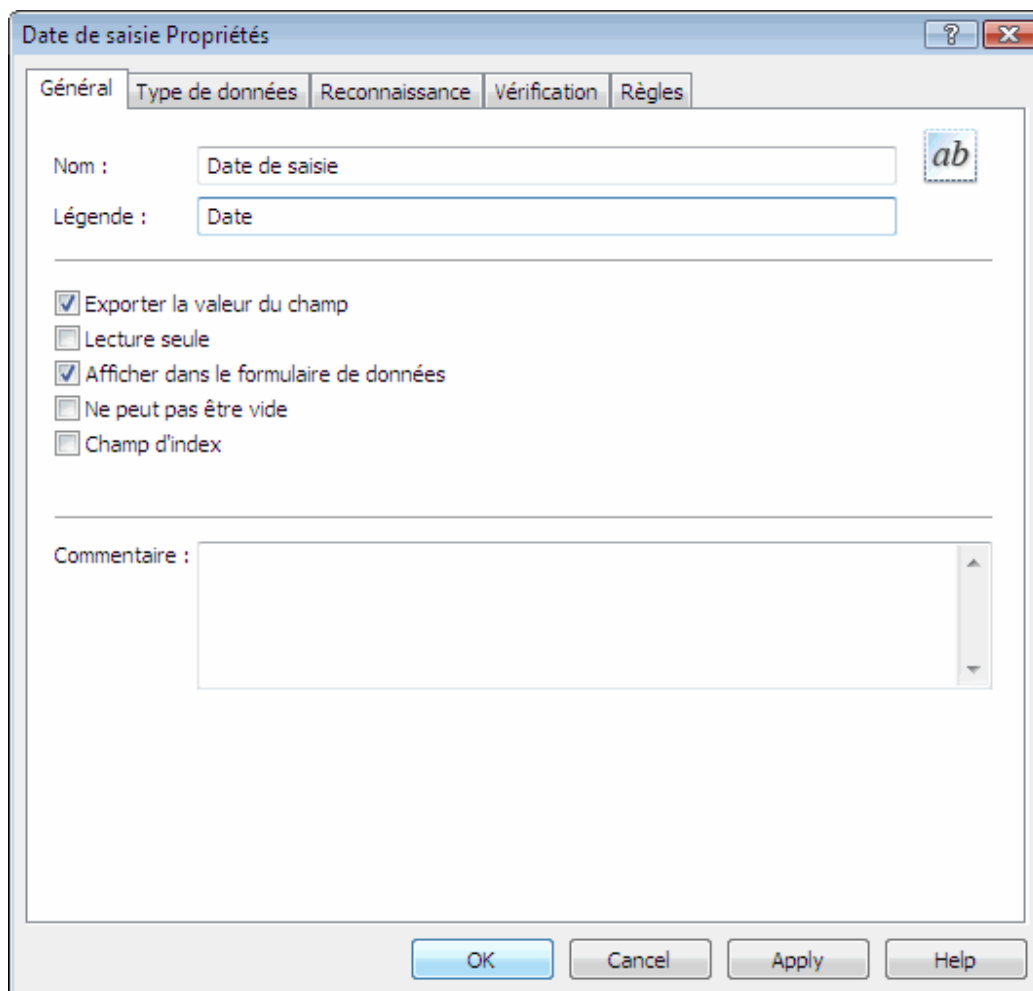


Figure 6. Onglet Général de la boîte de dialogue Propriétés

(pour un champ de saisie de texte)

Au moment de leur création, les champs sont automatiquement nommés en fonction du texte explicatif. Veillez à nommer correctement les champs créés. Si nécessaire, modifiez le nom d'un champ dans l'onglet Général de la boîte de dialogue Propriétés.

Si vous le souhaitez, décrivez l'élément.

Les options Exporter la valeur du champ et Afficher dans le formulaire de données sont sélectionnées par défaut pour tous les champs. Veillez à ce que l'option Exporter la valeur du champ soit sélectionnée pour tous les champs texte, les coches n'appartenant à aucun groupe et les groupes de coches. Notez que les propriétés des coches appartenant à un groupe sont identiques pour toutes les coches du groupe. Ainsi, ces propriétés sont-elles définies une seule fois pour le groupe.

4.3.4.2. Types de données

Un **type de données** détermine les valeurs possibles d'un champ ainsi que son format autorisé. Si une valeur saisie dans un champ ne correspond pas au type de données indiqué, l'opérateur verra s'afficher un message d'erreur au moment de vérifier le champ. Le type de données d'un texte décrit habituellement le regroupement des valeurs autorisées, par exemple la date, l'heure, l'adresse, le montant, etc. Le type de données d'une coche représente les valeurs acquises par le champ si la coche est activée ou désactivée.

4.3.4.2.1. Types de données d'un champ de saisie de texte

Il est très important que le **type de données** des champs de texte soit correctement défini. La définition du type de données informe le programme des données attendues dans ce champ : nombres ou lettres d'un alphabet particulier, caractères d'un jeu donné, date, etc. Le programme dispose d'un mécanisme flexible pour définir les types de données. L'utilisateur dispose d'un jeu de type de données prêt à l'emploi qui intègre les types les plus courants. En outre, les utilisateurs créent leurs propres types de données pour mener à bien des tâches spécifiques.

Au moment de définir des types de données, vous pouvez configurer la vérification du format des valeurs et les contraintes des valeurs, comme par exemple les valeurs maximale et minimale d'un nombre ou d'un intervalle de temps pour une date.

Vous définissez le type de données dans l'onglet **Type de données** de la boîte de dialogue **Propriétés** d'un champ (Figure 7).

Sélectionnez la catégorie souhaitée dans la liste **Contenu**. Dans le champ **Détails**, une description de l'un des types de données de cette catégorie apparaît (type de données par défaut ou type configuré précédemment).

Si l'option **Traiter la valeur comme du texte** est sélectionnée, les valeurs des champs de tous types de contenu seront traitées et exportées comme du texte.

Si vous souhaitez modifier la langue de reconnaissance ou définir un type de données plus précis, cliquez sur le bouton **Éditer...** situé à droite de la description.

Les options suivantes sont disponibles pour les types de données standard (option **Général** sélectionnée dans la liste **Paramètres de contenu**) :

- Pour le type *Texte*, vous pouvez sélectionner plusieurs langues de reconnaissance (le bouton « ... ») et utiliser le dictionnaire intégré et/ou un dictionnaire personnalisé (le programme vérifiera les valeurs de champs dans le dictionnaire que vous indiquez).
- Pour le type *Numéro*, vous pouvez sélectionner l'option **Entier**, si vous pensez que la valeur du champ sera un nombre entier.

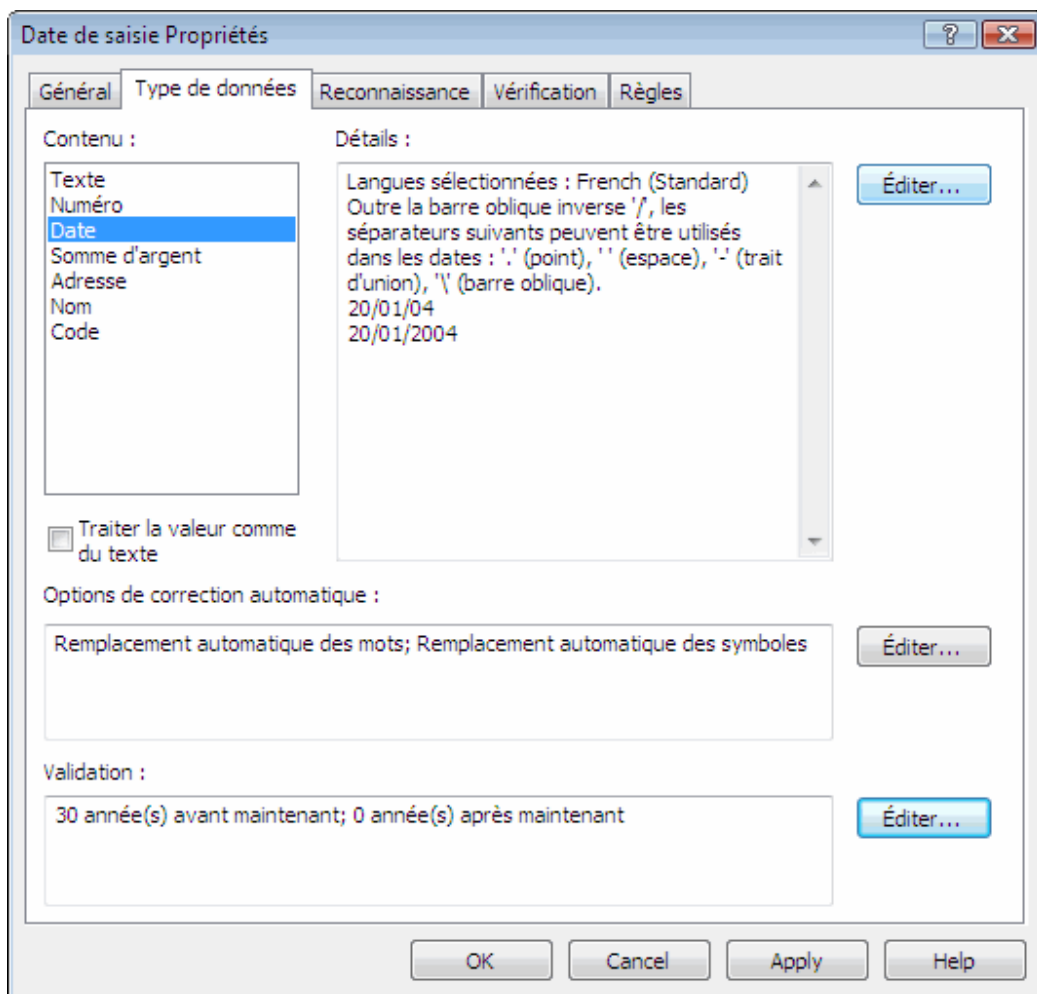
- Pour le type *Date*, vous pouvez sélectionner l'ordre des composants de la date, indiquer si le mois doit être écrit en toutes lettres et déterminer si la date peut inclure l'heure ou le jour de la semaine.
- Pour les types *Adresse*, *Nom* et *Code*, vous pouvez définir un dictionnaire personnalisé.

Les types de données spéciaux (option **Spécial** sélectionnée dans la liste **Paramètres de contenu**) contiennent des types de données prédéfinis à partir desquels vous pouvez sélectionner le plus adapté à un champ spécifique. Consultez les descriptions affichées dans la partie inférieure de la boîte de dialogue lors de la sélection d'un type de données.

Si aucun des types de données de la liste ne vous convient, vous pouvez créer le vôtre.

1. Sur l'onglet **Type de données** de la boîte de dialogue des propriétés du champ, sélectionnez l'un des éléments de la liste **Contenu**. Vous pouvez sélectionner n'importe quel élément correspondant le mieux à votre type de données. Ce choix n'aura aucune répercussion sur la création du type de données : le type de données nouvellement créé sera uniquement stocké dans cette catégorie.
2. Cliquez sur le bouton **Éditer...** situé à droite du champ **Détails**. Dans la boîte de dialogue qui s'ouvre, sélectionnez **Spécial** dans le champ **Paramètres de contenu**. Cliquez sur **Nouveau...** situé à droite du champ **Sélectionner un type de données**.

Suivez les instructions de l'assistant de création de type de données.



**Figure 7. Onglet Type de données de la boîte de dialogue Propriétés
(pour un champ de saisie de texte)**

Le programme peut traiter automatiquement les valeurs entrées quel que soit le type de données : supprimer les espaces en trop, convertir les lettres en majuscules ou en minuscules ou remplacer automatiquement les caractères ou parties de texte spécifiés. Pour activer le traitement automatique des valeurs, cliquez sur le bouton **Éditer...** en regard du champ **Options de correction automatique**. Dans la boîte de dialogue qui s'affiche, sélectionnez les options de traitement automatique requises (Figure 8).

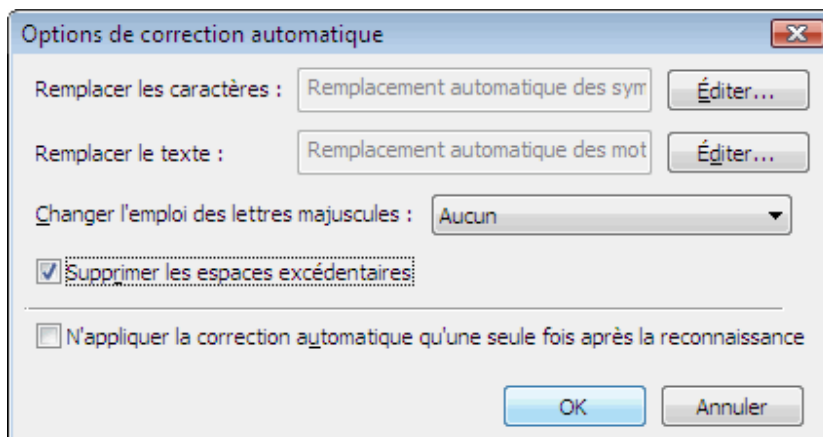


Figure 8. Boîte de dialogue Options de correction automatique

Vous pouvez également configurer une procédure de vérification des valeurs reconnues afin de savoir si elles sont comprises dans un intervalle donné. Pour indiquer un intervalle, cliquez sur **Éditer...** en regard du champ **Validation** (Figure 9).



Figure 9. Boîte de dialogue Paramètres de validation

Indiquez un type de données pour chaque champ de texte. Dans le champ Prénom et nom, sélectionnez le type Nom et indiquez la langue correcte. Concernant le champ Volume de traitement dans lequel le nombre de pages est indiqué, sélectionnez le type Numéro (ce format est un entier).

Dans le champ Date de saisie, sélectionnez Date dans la liste Contenu. Procédez ensuite à une sélection dans la liste ou décrivez le format de la date. Si vous souhaitez vérifier que la date se trouve dans l'intervalle indiqué, cliquez sur Éditer dans la zone Validation et activez l'option La date doit être comprise dans l'intervalle suivant dans la boîte de dialogue qui s'affiche. Vous pouvez définir un intervalle fixe ou relatif. Par exemple, indiquez l'intervalle suivant : 90 jours avant la date actuelle et 0 jour après. Dans ce cas, aucun message d'erreur ne s'affichera dans les questionnaires qui seront renseignés dans les 90 jours précédant la date du jour et qui ne seront pas renseignés après cette date.

4.3.4.2.2. Types de données des coches

Au moment de sélectionner un type de données pour les coches dans l'onglet **Type de données**, vous indiquez les valeurs à affecter aux champs en fonction de l'activation ou de la désactivation des coches (Figure 10).

Les valeurs suivantes peuvent affecter aux champs de coche :

- *Oui* si la coche est activée, *Non* si la coche est désactivée ;
- *1* si la coche est activée, *0* si la coche est désactivée ;
- *Nom de coche* si la coche est activée, *Vide* si la coche est désactivée ;
- vous pouvez également sélectionner l'option *Personnalisé* et proposer vos propres valeurs.

Remarque : Si les coches font partie d'un groupe, elles possèdent des propriétés communes. Les propriétés sont définies pour l'intégralité du groupe uniquement mais les valeurs des coches sont également définies dans les propriétés du groupe.

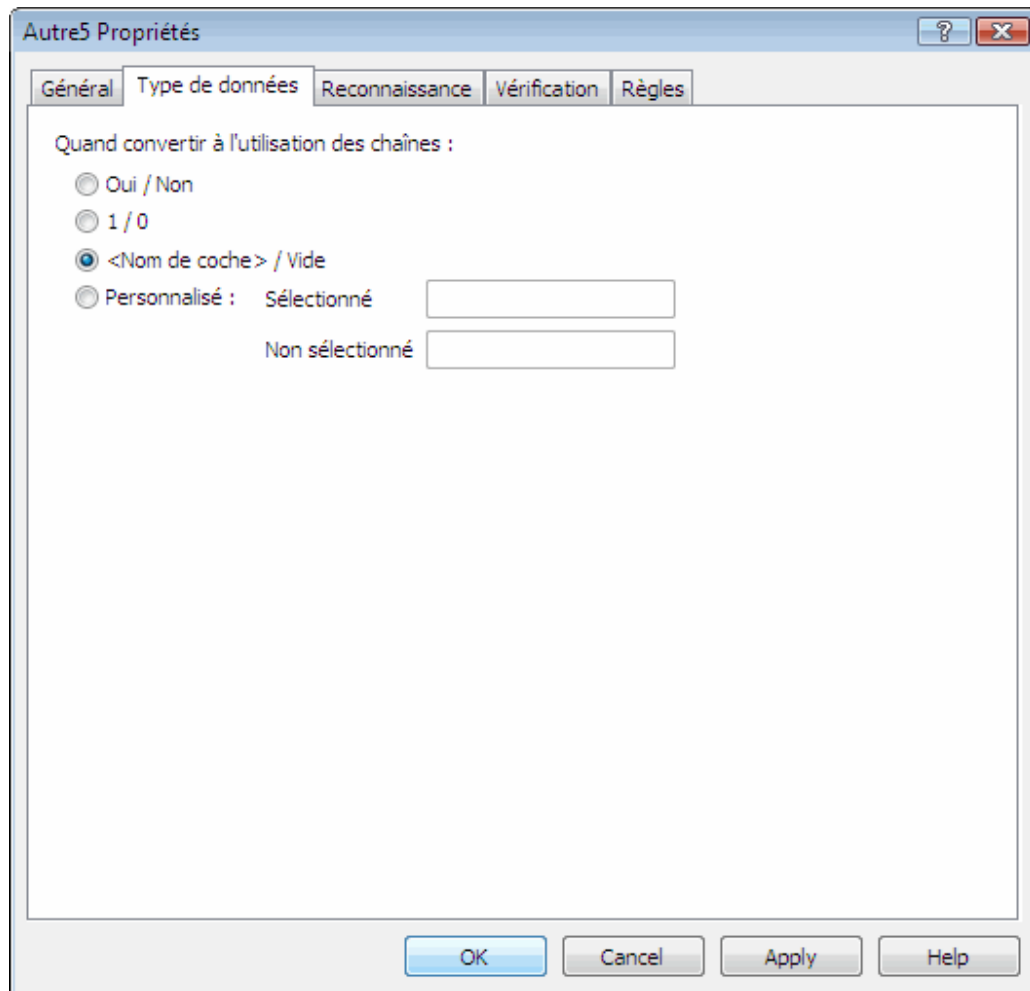


Figure 10. Onglet Type de données de la boîte de dialogue Propriétés (pour une coche qui n'appartient à aucun groupe)

4.3.4.2.3. Types de données pour un groupe de coches

Dans l'onglet **Type de données** de la boîte de dialogue **Propriétés** pour un groupe de coches, vous pouvez afficher une liste de toutes les coches qui font partie de ce groupe (Figure 11).

Sélectionnez **Autoriser une sélection vide** si au moins un champ de coche du groupe doit être sélectionné.

Si plusieurs coches du groupe peuvent être sélectionnées, vous devez activer l'option **Autoriser les sélections multiples**. Si vous n'activez pas cette option et que plusieurs coches sont sélectionnées dans le groupe, un message d'avertissement apparaîtra.

Vous pouvez indiquer les valeurs à exporter si aucun champ de coche n'a été sélectionné ou si plusieurs champs de coche ont été sélectionnés. Sélectionnez **<Sélection vide>** ou **<Sélection multiple>** dans la liste et cliquez sur **Éditer...** dans la boîte de dialogue qui s'ouvre, saisissez la valeur souhaitée dans le champ **Valeur exportée**. Si aucune valeur n'est indiquée, le programme exportera une chaîne vide à chaque fois qu'aucun champ de coche n'est sélectionné, ainsi que les valeurs des champs de coche sélectionnés séparées par des virgules si plusieurs champs sont sélectionnés.

Si l'option **Considérer une erreur de validation comme un avertissement** est sélectionnée, un message d'avertissement apparaîtra au lieu d'un message d'erreur de validation.

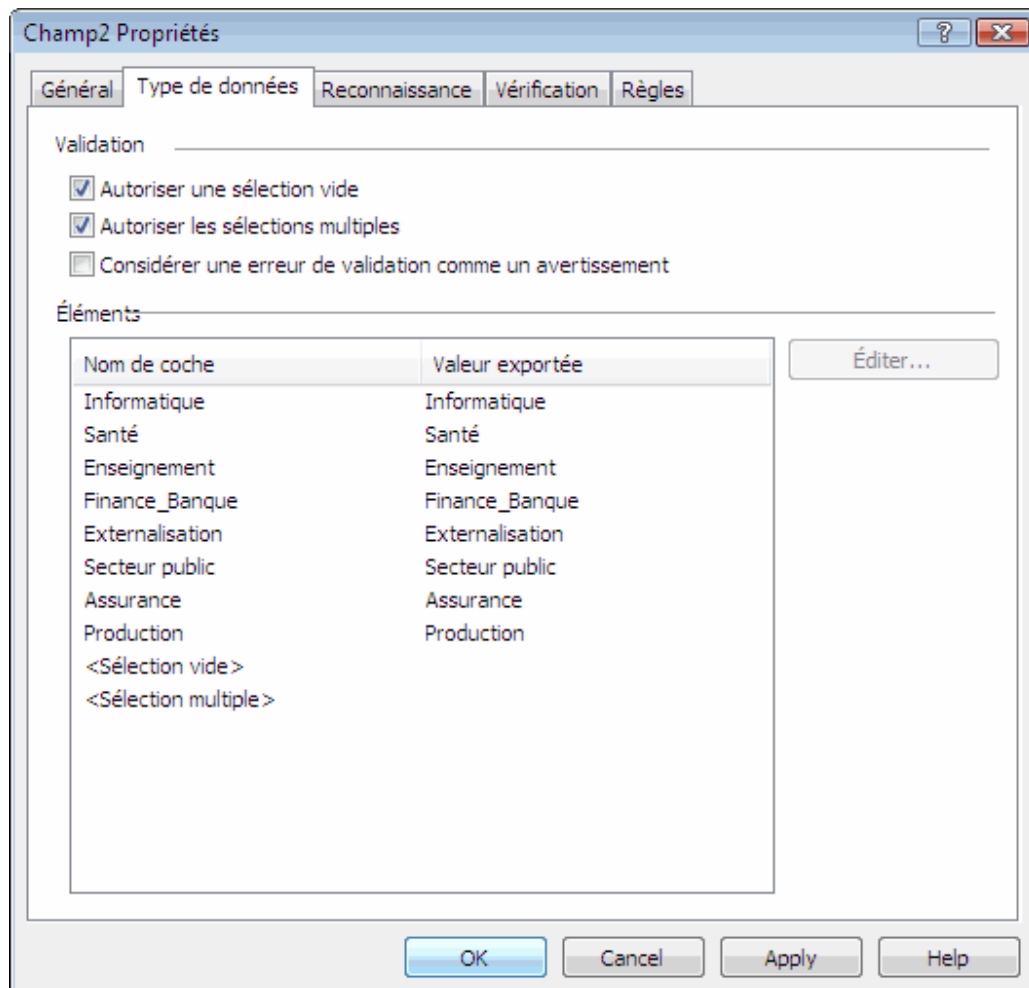


Figure 11. Onglet Type de données de la boîte de dialogue Propriétés (pour un groupe de coches)

Définissez les propriétés des coches et des groupes de coches. Choisissez la méthode <Nom de coche>/Vide pour la conversion des valeurs de coche.

Pour le groupe de coches « Types de documents à traiter », sélectionnez les options « Autoriser une sélection vide » et « Autoriser les sélections multiples ».
Indiquez la valeur à exporter si aucun champ de coche n'est sélectionné, « aucun sélectionné » par exemple.
Indiquez la valeur à exporter si plusieurs champs de coche sont sélectionnés, « plusieurs sélectionnés » par exemple.

Répétez la procédure pour les autres groupes de coches.

4.3.4.3. Propriétés de reconnaissance des champs

ABBYY FlexiCapture 8.0 Professional vous permet de définir des propriétés de reconnaissance pour chacun des champs. Si vous définissez ces propriétés correctement dans l'onglet **Reconnaissance** de la boîte de dialogue **Propriétés**, la qualité de la reconnaissance sera nettement améliorée et les erreurs moins nombreuses. Les propriétés de reconnaissance des différents types de champs diffèrent. Vous trouverez ci-dessous une description détaillée de ces propriétés pour tous les types de champ.

4.3.4.3.1. Propriétés de reconnaissance d'un champ de saisie de texte

Sélectionnez l'option **Ne pas reconnaître (clé du champ d'image - sera entré manuellement)**, uniquement si vous ne souhaitez pas que ce champ et sa valeur soient saisis manuellement par l'opérateur (par exemple, s'il est trop difficile de reconnaître ce champ du fait de l'utilisation de lettres en italique). Si tel est le cas, vous n'avez pas besoin d'indiquer d'autres propriétés de reconnaissance car ce champ ne sera pas reconnu et l'opérateur en charge de la vérification sera invité à saisir les valeurs de ce champ manuellement.

Sélectionnez le type de texte : **ICR (manuscrit)** ou **OCR (imprimé)**. Pour du texte imprimé, sélectionnez le type d'impression (typographique, imprimante matricielle, machine à écrire, etc.). Pour définir plusieurs types de texte ou utiliser un modèle, sélectionnez l'option **Avancé**, puis cliquez sur **Modifier...**

Sélectionnez le type de marquage à l'aide des exemples de marquage fournis dans la liste déroulante. Si le marquage est supprimé durant la numérisation, sélectionnez le type à espacement fixe (Zones grises). Si le marquage n'est pas supprimé et que les cellules de caractères sont présentes, indiquez le nombre total de ces cellules (le nombre de cellules peut être détecté automatiquement).

Vous pouvez sélectionner la **casse**, auquel cas toutes les lettres reconnues seront en majuscules ou en minuscules. Si un champ contient à la fois des lettres majuscules et minuscules, conservez l'option **Automatique**.

Sélectionnez l'orientation du texte (verticale ou horizontale).

Sélectionnez **Une ligne** si le champ ne possède qu'une seule ligne. Si la valeur d'un champ est toujours d'un seul mot (c'est-à-dire s'il ne contient aucun espace vide), sélectionnez **Un mot**.

Vous pouvez sélectionner la casse, auquel cas toutes les lettres reconnues passeront en majuscules ou en minuscules.

Définissez les options de prétraitement de l'image. Le programme peut nettoyer l'image (indiquez pour cela la taille de la corbeille), inverser l'image ou supprimer la texture de l'image.

Remarque : Si l'option **Nettoyer** est activée, la corbeille de l'image sera automatiquement vidée par défaut. Pour définir la taille de la corbeille, activez l'option **Vider la corbeille de taille spécifiée uniquement**, puis saisissez la taille voulue.

Sélectionnez l'orientation du texte : vertical ou horizontal.

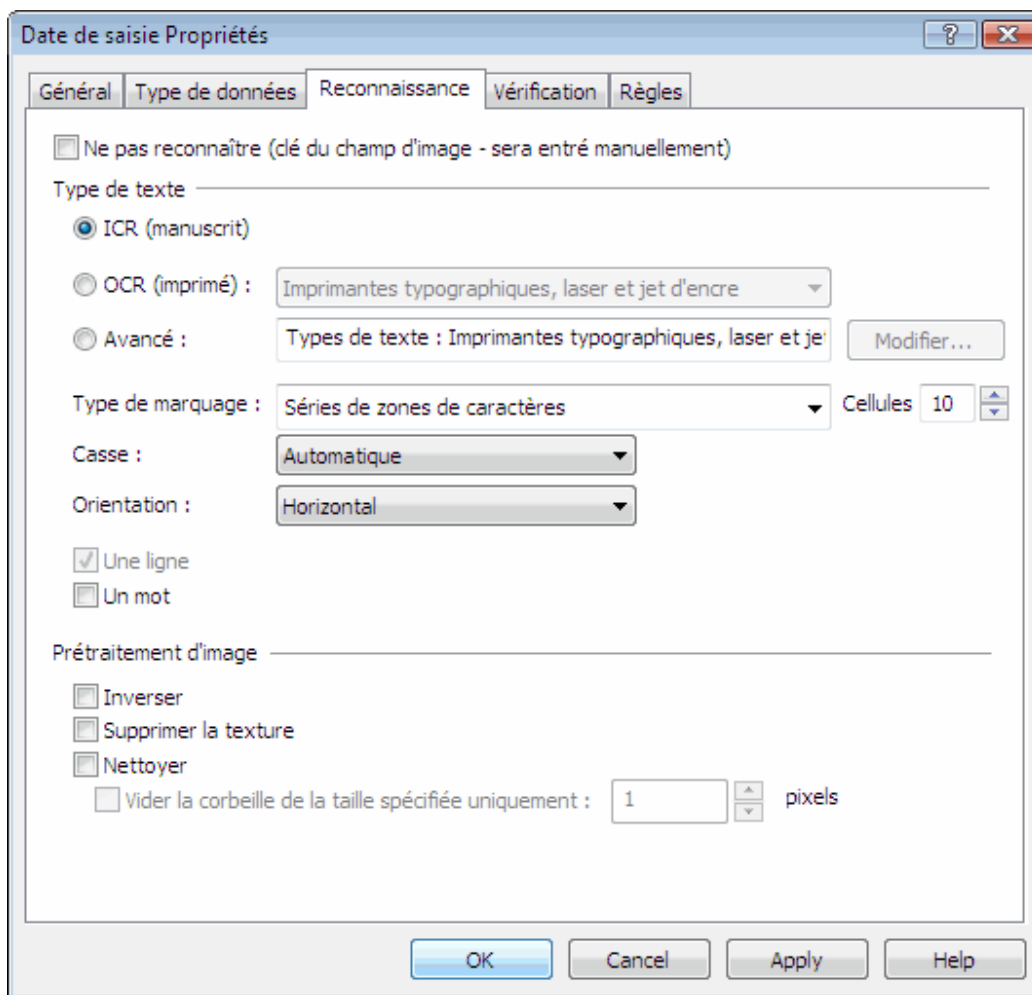


Figure 12. Onglet Reconnaissance de la boîte de dialogue Propriétés (pour un champ de saisie de texte)

1. Définissez les propriétés de reconnaissance pour tous les champs de texte de votre formulaire. Le type de texte pour tous les champs doit être ICR (manuscrit), le type de marquage doit être Séries de zones de caractères, le nombre de cellules doit pouvoir être détecté automatiquement et l'orientation du texte définie sur Horizontal. Ce formulaire ne contient aucun champ de plusieurs lignes. L'option Une ligne peut alors être sélectionnée pour tous les champs. Vous pouvez sélectionner l'option Un mot pour les champs pour lesquels aucun espace vide n'est autorisé, par exemple le champ relatif à l'adresse électronique.
2. Indiquez la casse. Vous pouvez conserver la valeur Automatique de la propriété Casse pour tous les champs de texte.
3. Le champ Vos commentaires et vos attentes est un champ de saisie manuelle. Le texte de ce champ peut apparaître en italique, ce qui l'empêche d'être reconnu automatiquement. Aussi, nous vous recommandons d'activer l'option Ne pas reconnaître (clé du champ d'image - sera entré manuellement) pour ce champ. Par conséquent, l'opérateur devra saisir la valeur de ce champ.
4. Vous pouvez conserver les valeurs par défaut de tous les autres champs.

4.3.4.3.2. Propriétés de reconnaissance des coches et des groupes de coches

Pour assurer une reconnaissance correcte des coches, vous devez définir le type de reconnaissance des coches ou **Type de coche**. Si une coche est entourée d'un carré, sélectionnez le type **Carré**. Si une coche doit être placée sur un fond blanc, sans délimitation (ou avec une délimitation qui sera supprimée au cours de la numérisation), sélectionnez le type **Sans cadre**. Si le type est **Automatique**, le programme détectera automatiquement la forme de la coche. Notez que dans ce cas, la coche sur le formulaire ne doit

pas être sélectionnée car le programme détecte si la coche est ou non sélectionnée en comparant l'image de la coche sur un document traité à l'image du formulaire vide utilisé pour créer le modèle.

Vous pouvez autoriser les corrections apportées à certaines coches. Si la personne renseignant le questionnaire a coché une case par erreur, elle peut simplement l'effacer. Les coches entièrement effacées seront considérées comme étant non sélectionnées. Toutefois, si vous avez sélectionné le type **Automatique** pour une coche, aucune correction n'est autorisée.

Vous pouvez définir les options de prétraitement de l'image pour la reconnaissance des coches. Ces options ressemblent à celles dédiés aux champs de texte.

Si les coches font partie d'un groupe, elles possèdent des propriétés communes. Les propriétés de reconnaissance sont définies de la même manière, mais pour l'intégralité du groupe de coches.

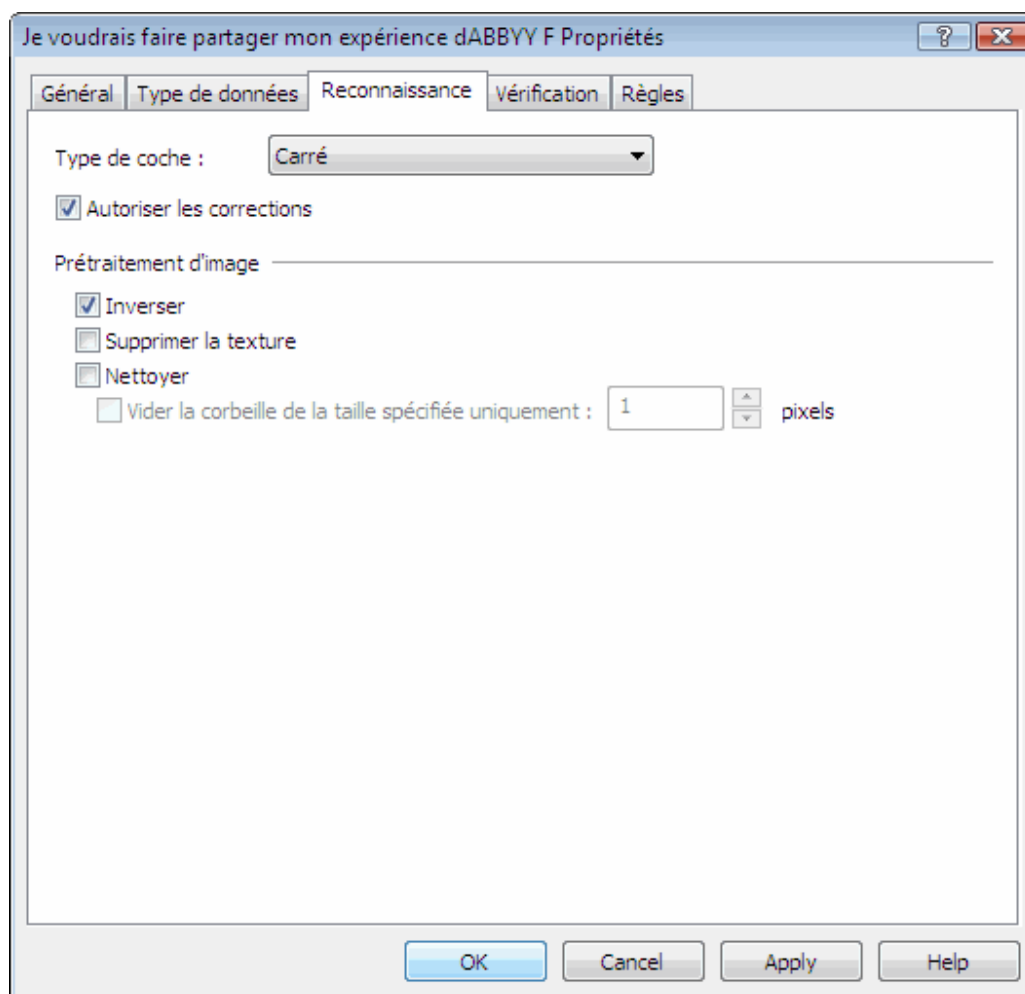


Figure 13. Onglet Reconnaissance de la boîte de dialogue Propriétés (pour une coche)

Définissez les propriétés des coches et des groupes de coches.

Sélectionnez le type de coche Carré :

- pour les coches qui n'appartiennent à aucun groupe ;
- pour les groupes de coches.

Pour autoriser les corrections apportées aux coches, activez l'option Autoriser les corrections.

4.3.4.3.3. *Propriétés de reconnaissance de code barre*

Les propriétés utilisées pour reconnaître un code barre de champ sont identiques à celles des codes barres statiques. Pour un code barre de champ, vous devez définir le type de code barre, son orientation, ainsi que les options de nettoyage de l'image. La seule différence réside dans le fait que l'opérateur peut saisir la valeur du champ manuellement. Pour cela, activez l'option **Ne pas reconnaître (clé du champ d'image - sera entré manuellement)**.

4.3.4.3.4. *Propriétés de reconnaissance d'image*

Dans l'onglet **Reconnaissance**, sélectionnez l'option **Exclure de la reconnaissance** pour le champ Image si la région de l'image doit être exclue des régions de la reconnaissance des autres champs. (Cette option est uniquement disponible pour les champs Image et est rétro compatible avec les modèles FormReader 6.5.)

4.3.4.4. Paramètres de vérification

La **vérification** est le processus réalisé par l'opérateur consistant à vérifier les données reconnues. Au moment de créer un modèle, vous pouvez configurer les options de vérification dans l'onglet **Vérification** de la boîte de dialogue des propriétés des champs (Figure 14). Les caractères non reconnus de manière certaine sont marqués par le programme et envoyés à l'opérateur pour vérification. Toutefois, vous pouvez configurer les propriétés du champ de sorte à ce que le champ soit envoyé pour vérification même s'il ne contient aucun caractère incertain. Vous pouvez également faire en sorte que les champs ne soient pas envoyés pour vérification s'ils contiennent certains caractères incertains. Sélectionnez **Tous** (vérification obligatoire) si aucune erreur dans la valeur de ce champ n'est acceptée.

L'opérateur consultera alors un caractère reconnu et son image, puis validera ou corrigera le résultat de la reconnaissance.

Si vous souhaitez intégrer tous les caractères d'un champ dans la vérification groupée, sélectionnez l'option **Inclure dans la vérification groupée**. Pour inclure un champ dans la vérification du champ, sélectionnez l'option **Inclure dans la vérification du champ**. Tous les champs peuvent être intégrés dans la vérification groupée. Toutefois, nous vous recommandons d'envoyer uniquement les coches et les champs de texte contenant des chiffres et des séparateurs (points, virgules, barres obliques) pour une vérification groupée. Les lettres doivent être vérifiées en tenant compte du contexte du champ (pour plus d'informations au sujet de la vérification des groupes et du contexte, consultez le chapitre [Vérification](#) ci-dessous).

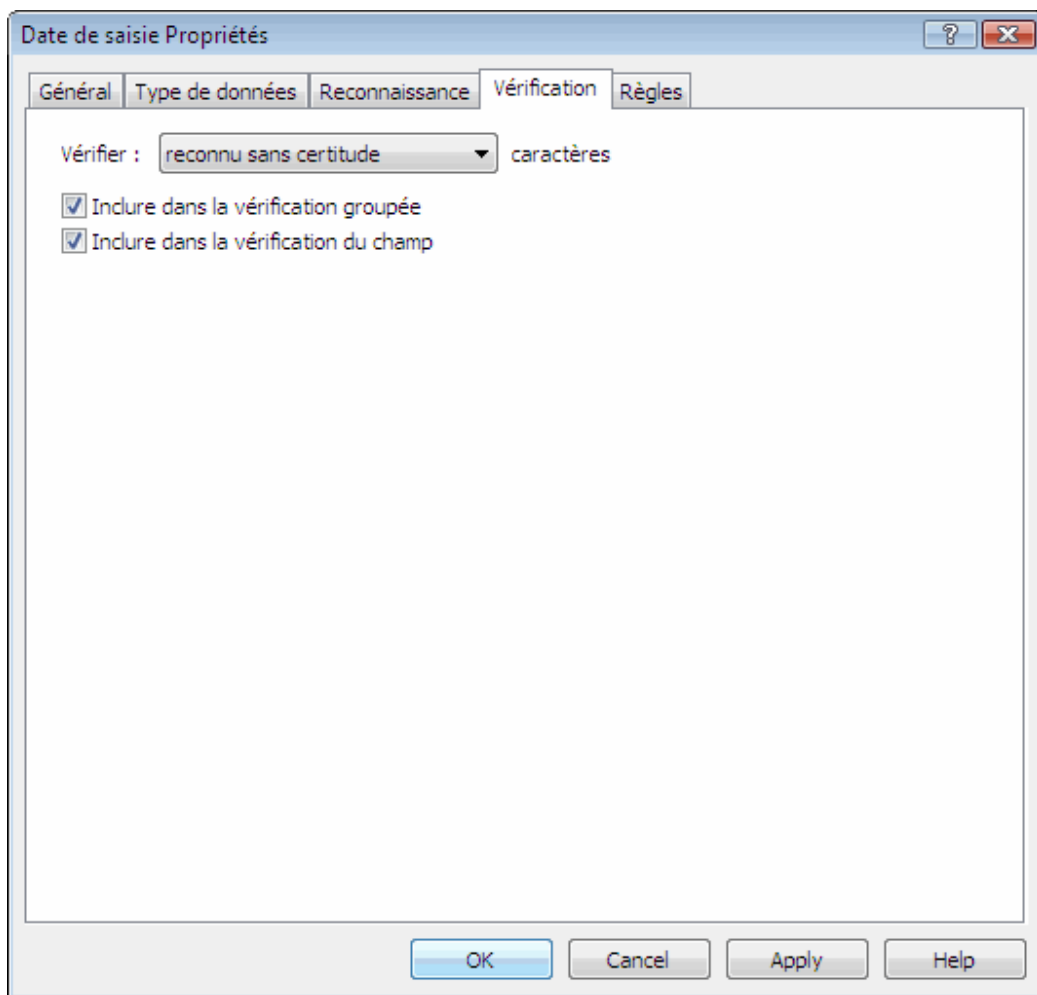


Figure 14. Onglet Vérification de la boîte de dialogue Propriétés

5. Configurez les options de vérification des champs de données. La valeur « reconnu sans certitude » est sélectionnée par défaut, ce qui signifie que les caractères incertains d'un champ seront envoyés à l'opérateur pour vérification. Conservez cette valeur pour tous les champs à l'exception par exemple du champ Prénom et nom, pour lequel vous sélectionnez la valeur Tous. Si vous procédez de la sorte, tous les caractères du champ Prénom et nom seront envoyés pour vérification, sans tenir compte du côté certain ou incertain de leur reconnaissance.
5. Sélectionnez l'option Inclure dans la vérification groupée pour toutes les coches et tous les champs de texte dont la saisie contient des chiffres (champs Date de saisie et Volume de traitement).
6. Sélectionnez l'option Inclure dans la vérification du champ pour les champs dont vous souhaitez vérifier le contexte (par exemple le nom, la profession et le nom de la société).

4.3.4.5. Options d'exportation des images

Dans un modèle de document, vous pouvez sélectionner les options d'exportation de chaque image. Pour sélectionner les options d'exportation des images, cliquez avec le bouton droit de la souris sur le champ souhaité, sélectionnez **Propriétés...**, puis cliquez sur l'onglet **Exporter**. Vous pouvez sélectionner l'une quelconque des options suivantes :

- **Type de fichier** : (TIFF, JPEG, BMP, JPEG2000, PCX packbits, PNG).
- **Qualité** : pour les fichiers TIFF, JPEG et JPEG2000, vous pouvez sélectionner la qualité du fichier obtenu (qualité optimale à minimale).
- **Type de couleur** : couleur, niveaux de gris, demi-ton (noir et blanc).

Vous pouvez également définir la résolution de l'image obtenue : sélectionnez **Changer la résolution en** et saisissez la résolution souhaitée ou sélectionnez-la dans la liste.

4.3.4.6. Vérifications basées sur les règles

Les **règles** s'utilisent pour vérifier automatiquement les résultats de la reconnaissance. À l'instar des types de données, les règles vous permettent de définir les contraintes de données, comme par exemple les conditions que les valeurs de certains champs doivent satisfaire. Si les valeurs des documents renseignés ne répondent pas à ces conditions, les pages sont marquées d'un drapeau et un message s'affiche. L'objectif principal de ces règles est de vérifier l'intégrité des données dans un document. Les règles peuvent également être utilisées pour traiter les données saisies, par exemple pour fusionner les valeurs des différents champs ou pour remplacer les valeurs reconnues par les valeurs correspondantes d'une base de données.

ABBYY FlexiCapture 8.0 Professional vous permet de créer des règles de différents types :

- **Somme de contrôle** : vérification de la somme des valeurs de plusieurs champs. Cette somme est comparée à un nombre ou à une valeur précise d'un autre champ. Par exemple, si le document contient le salaire et les primes d'un salarié et qu'un champ affiche le revenu total de ce salarié, vous pouvez créer une règle qui vérifie si la somme du salaire et des primes correspond au revenu total. Si les montants ne correspondent pas, le message affiche un message d'erreur.
- **Comparer les champs** : comparaison des valeurs de plusieurs champs. Cette règle peut être utilisée si un document contient plusieurs champs dont les valeurs doivent être identiques. Si ces valeurs ne le sont pas, le programme affiche un message d'erreur de la règle pour le document.
- **Règle de vérification de base de données** : vérification des valeurs saisies par rapport à certaines valeurs d'une base de données.
- **Fusionner les champs** : fusion des valeurs de plusieurs champs. Les valeurs des champs fusionnés peuvent être séparées par des points, des espaces vides ou tout autre séparateur. Par exemple, il peut se révéler utile de ne pas reconnaître les données comme un seul élément mais de fusionner les valeurs des champs reconnus séparément, « Jour », « Mois » et « Année ». Les points peuvent être utilisés comme séparateurs de la valeur fusionnée. Le résultat d'une telle fusion peut être stocké dans n'importe quel champ du modèle. Il est tout particulièrement utile d'utiliser des champs sans marquage pour stocker de tels résultats (consultez le chapitre [Champs sans marquage](#)).
- **Somme en chiffres - somme en lettres** (pour le russe uniquement). Cette règle est uniquement disponible si le russe est sélectionné comme langue de reconnaissance pour le modèle.
- **Règle de script** : l'utilisateur peut décrire les contraintes à l'aide d'un script.

Les règles sont définies dans l'onglet **Règles** de la boîte de dialogue **Propriétés** (Figure 15). Une règle peut être affectée à un ou plusieurs champs simultanément.

Il est possible d'affecter un état d'erreur ou d'avertissement à une règle. Si la condition requise n'est pas satisfaite, un champ sera marqué d'un drapeau rouge en cas d'erreur ou d'un drapeau jaune en cas d'avertissement.

Si le programme tente d'exporter un document contenant des erreurs de règles, l'avertissement correspondant s'affiche.

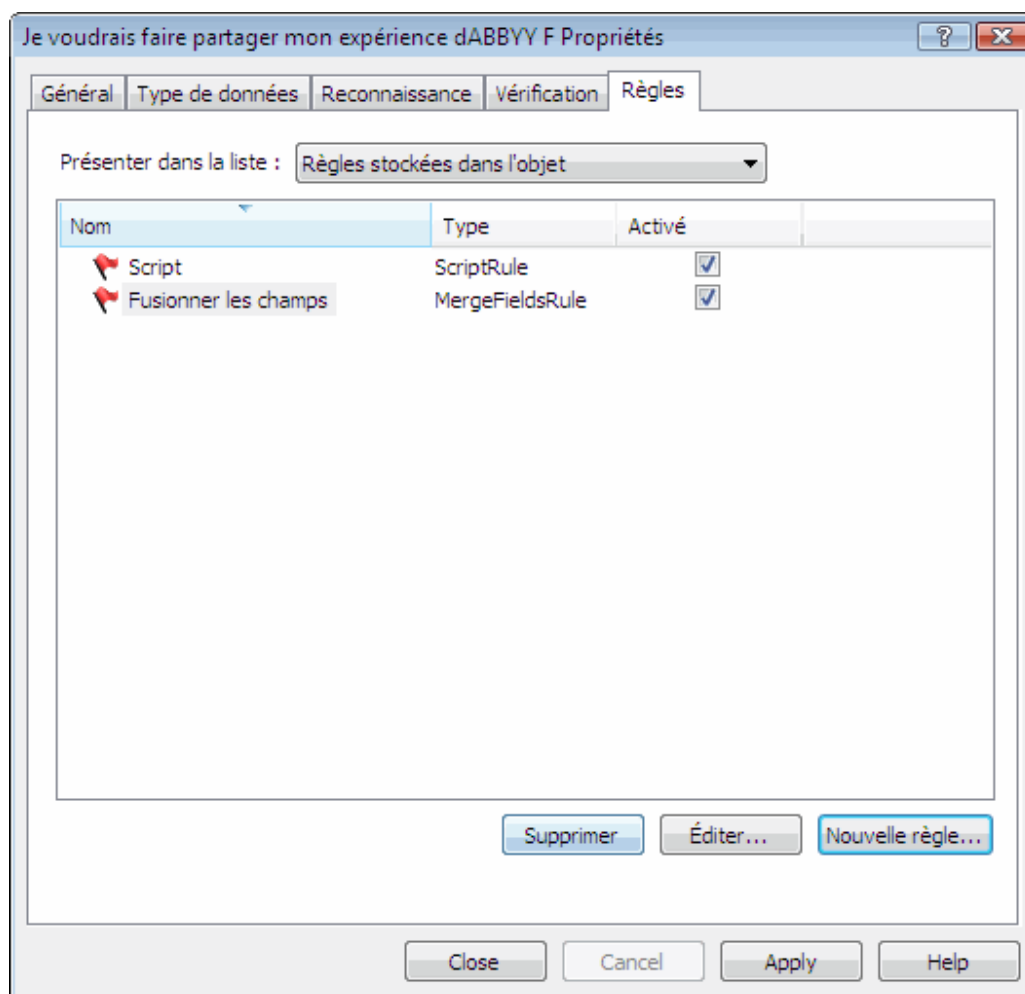


Figure 15. Onglet Règles de la boîte de dialogue Propriétés

Analysons un exemple de règle. Nous allons décrire la date de saisie du questionnaire comme étant la fusion des champs Jour, Mois et An. Pour cela, nous allons procéder de la manière suivante :

1. Supprimez la région du champ date sur l'image du formulaire mais ne supprimez pas le champ lui-même. Pour ce faire, sélectionnez Supprimer une région dans le menu contextuel du champ Date de saisie. Le marquage sera retiré de l'image mais le champ lui-même restera sur le document et sera marqué d'un astérisque rouge.

2. Créez les champs Jour, Mois et An.

3. Définissez les propriétés de ces champs. Dans l'onglet Général, désactivez l'option Exporter la valeur du champ car nous exporterons uniquement la date entière (assurez-vous que l'option Exporter la valeur du champ EST activée pour le champ Date de saisie).

4. Dans l'onglet Type de données, indiquez le type de données correspondant à chacun des champs. Vous devez sélectionner le type de données Numéro, ainsi que le format nombre entier. En outre, indiquez les contraintes relatives à chacun des champs en cliquant sur Éditer dans la zone Validation :

- Jour : de 1 à 31
- Mois : de 1 à 12
- An : en fonction de la période à laquelle le questionnaire est renseigné, par exemple de 2000 à 2020

5. Dans les onglets Reconnaissance et Vérification, définissez des propriétés similaires à celles définies dans le champ Date de saisie.

6. Ouvrez la boîte de dialogue Propriétés et le champ Date de saisie, puis cliquez sur l'onglet Règles pour créer une règle.

7. Cliquez sur Nouvelle règle... Sélectionnez le type de règle Fusionner les champs.

8. Nommez le groupe, par exemple Fusionner les champs - Date de saisie. Sélectionnez Erreur pour affecter l'état Erreur à la règle. L'erreur survient si un champ participant à la règle n'est pas détecté. Cliquez sur Suivant.

9. Ajoutez les champs Jour, Mois et An à la liste des champs en cliquant sur Ajouter. Dans le champ Champ de résultat, sélectionnez le champ Date de saisie. Utilisez les points à titre de séparateurs.

Nous avons configuré la vérification pour les cas où la date du champ Date de saisie figure dans l'intervalle de temps indiqué. Cette vérification sera à présent effectuée pour la valeur reçue en fusionnant les valeurs des trois champs.

Avec cette procédure, nous avons terminé la création des champs et des éléments statiques sur le formulaire, ainsi que la définition de leurs propriétés. Assurez-vous d'avoir créé tous les éléments nécessaires et d'avoir défini correctement leurs propriétés. Une fois l'association du modèle effectuée, si vous n'êtes pas satisfait des résultats de la reconnaissance ou si certains éléments ne sont pas correctement positionnés, vous pouvez toujours revoir l'édition du modèle et modifier le marquage ou les propriétés des champs.

4.3.5. Création d'un modèle pour un document multipage

ABBYY FlexiCapture 8.0 Professional vous permet de créer des modèles pour les documents multipages. Un modèle peut être composé d'un certain nombre de *sections*, chacune contenant une ou plusieurs *pages*. Au moment de créer un modèle de document multipage, vous devez définir la séquence des sections, leur nombre, mais aussi les règles d'assemblage des pages dans les documents.

Au terme de la procédure de l'Assistant de création de modèle, un modèle simple est créé qui ne contient aucune section. Vous pouvez alors ajouter des pages au modèle. Au moment d'ajouter une page, son image est également ajoutée à la zone d'image de la fenêtre Éditeur de modèles de document et tous les champs créés sur la page sont représentés dans la liste globale des champs. Dans ce cas, un document se compose d'une section, section qui peut contenir plusieurs pages.

Si le document contient des champs ou des tableaux qui continuent sur la page suivante, ou bien des blocs répétitifs, vous devez ajouter des pages à une section. Cette méthode est également utilisée pour le chargement d'une description flexible multipage (consultez le chapitre [Configuration du système pour le traitement des documents flexibles](#)). Tous les champs d'une description flexible doivent appartenir à une seule et même section d'un modèle de document.

Il existe également une autre méthode. Elle consiste à ajouter des sections à un modèle. L'exemple le plus simple est celui d'une section qui contient une seule page. Cette méthode peut être utilisée si les jeux de champs des pages sont indépendants. Elle facilite la compréhension de la structure du document car vous pouvez voir à quelle section des champs appartiennent de même que vous pouvez définir une structure dédiée à l'assemblage du document. Par exemple, votre document contient 3 pages, la 1^{ère} et la 3^e pages ne sont utilisées qu'une seule fois mais la 2^e page se répète entre 2 et 5 fois. Vous devez donc créer une section séparée pour chaque page, puis définir la structure du document, c'est-à-dire la séquence et le nombre de répétitions des sections.

Prenons un exemple plus compliqué, comme celui d'un modèle composé de plusieurs sections, certaines d'entre elles contenant plusieurs pages. Il peut s'agir d'un modèle composé de sections fixes et d'une section multipage flexible ou d'un modèle décrivant des documents contenant une page recto-verso se répétant plusieurs fois. Partons du principe que nous devons créer un modèle pour un document contenant une page de titre (page A) et un certain nombre de pages recto-verso (séquence de pages BCBC...). Nous créons la section A contenant la page A et la section BC contenant deux pages, B et C. Le document adopte alors la structure suivante : section A ne se répétant qu'une seule fois, est suivie de la section BC se répétant entre 3 et 7 fois par exemple.

Si un modèle se compose de plusieurs sections, vous ne pouvez pas uniquement définir la séquence et le nombre de répétitions des sections. Vous devez également configurer la vérification des *champs clés*. Pour cela, vous devez définir les champs clés dans chacune des sections du document. Les valeurs de ces

champs clés doivent être identiques. Si le nom de la personne renseignant le formulaire apparaît sur chacune des pages du document, ce nom peut servir de champ clé. De même, vous pouvez utiliser un numéro d'identification.

Durant le traitement des lots, le programme tente d'assembler les pages successives en documents, puis vérifie les règles d'assemblage. Les valeurs des champs clés sont vérifiées à ce stade. Si les champs clés d'un document ne sont pas identiques, un message d'erreur apparaît. Cela se produit parfois si les pages ont été mélangées durant la numérisation. Si tel est le cas, il vous suffit de modifier l'ordre des pages et la règle d'assemblage sera alors satisfaite.

Pour ajouter une page à une section, accédez au menu Éditeur de modèles de document, sélectionnez **Modèle > Ajouter une page...** ou sélectionnez **Ajouter une page...** depuis le menu contextuel de l'image. Ensuite, chargez l'image de la nouvelle page vierge et sélectionnez les types d'objet à détecter automatiquement sur la page.

Pour ajouter une nouvelle section, sélectionnez **Modèle > Ajouter une section de document...** L'Assistant de création de section s'ouvre et vous aide à configurer tous les paramètres nécessaires. Suivez les instructions de l'Assistant, puis procédez aux actions suivantes :

- Nommez la section.
- Chargez l'image de la page vierge (numérisée ou chargée depuis un fichier) ou chargez le fichier de description flexible.
- Sélectionnez les types d'objets à détecter automatiquement.

La nouvelle section et ses champs s'affichent alors dans la fenêtre de la structure du document.

Pour afficher la structure du document conçue durant la création du modèle, allez dans l'onglet **Pages** de la fenêtre **Structure du document** située dans la partie droite de la fenêtre Éditeur de modèles de document. La structure du document est représentée dans l'onglet **Pages** sous forme de vignettes. Dans cet onglet, vous pouvez modifier le nombre des répétitions de section du document (pour cela, saisissez une nouvelle valeur, un nouveau nombre ou un nouvel intervalle dans la zone située à droite du nom de la section), joindre des pages en annexe, modifier la séquence des pages d'une section et déplacer les pages d'une section à une autre par glisser-déposer.

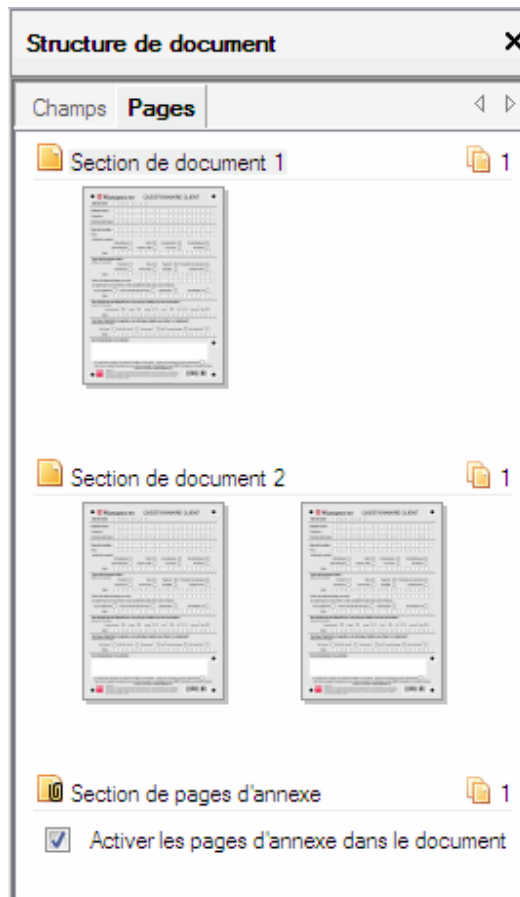


Figure 16. Structure du document, mode vignettes de page

Vous pouvez également définir la séquence et le nombre des répétitions de section en sélectionnant **Modèle > Propriétés du modèle de document...** dans la fenêtre Éditeur de modèles de document. Dans l'onglet **Assemblage** (Figure 17), définissez le nombre minimum et maximum de répétitions de section dans le document (1 par défaut). Si vous souhaitez vérifier les valeurs des champs clés, cochez la case **Vérifier l'égalité des champs clés** et définissez le champ clé sur chacune des pages.

Il se peut que vous souhaitiez désactiver la vérification de l'ordre des sections du document (par exemple, si l'ordre des sections n'affecte pas l'assemblage de document). Le cas échéant, sélectionnez **Désactiver la vérification de l'ordre des sections** pour désactiver la vérification. Le programme continuera de vérifier que toutes les sections sont présentes dans le document mais leur ordre sera ignoré.

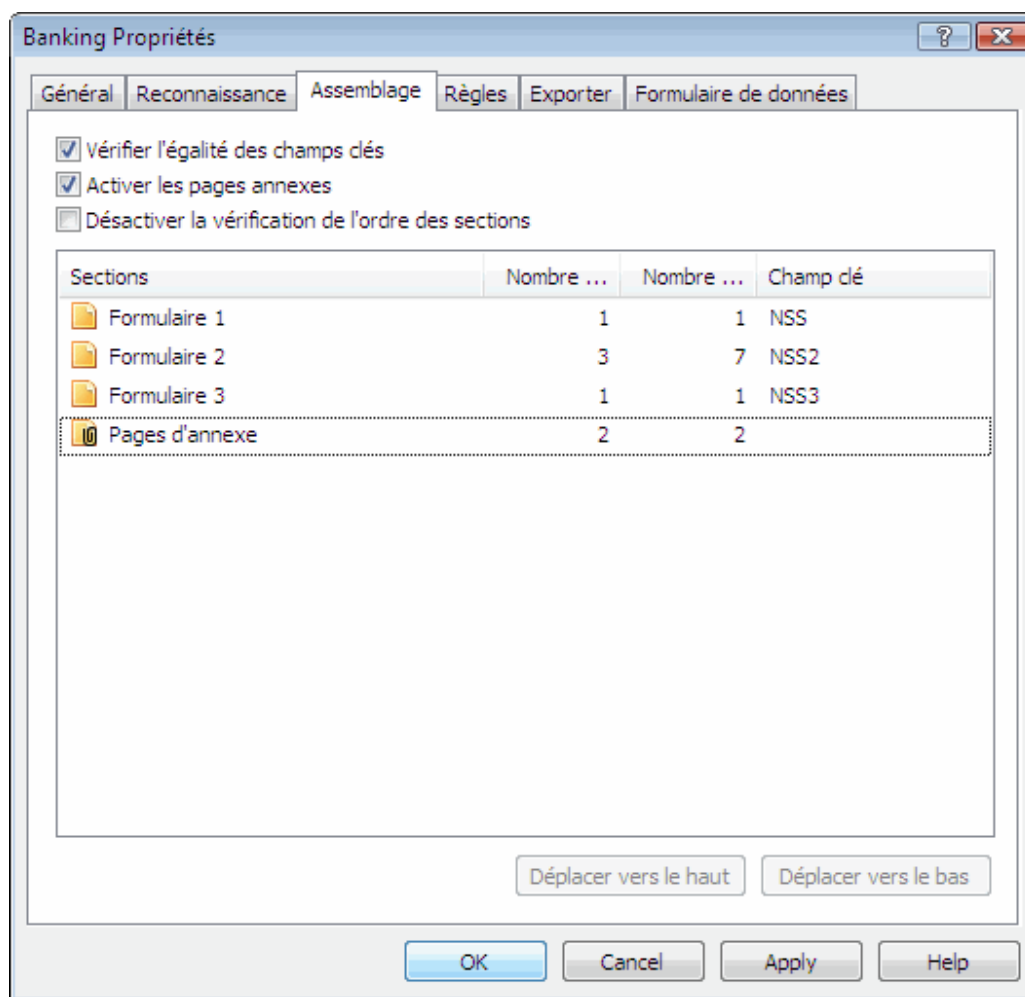


Figure 17. Onglet Assemblage de la boîte de dialogue Propriétés

Notre formulaire n'est pas un document multipage. Son modèle se compose d'une seule page et aucune règle d'assemblage n'est nécessaire. Dans l'onglet Assemblage de la boîte de dialogue Propriétés du modèle de document, vous pouvez définir qu'une seule section se répétant une seule fois s'affichera.

4.3.6. Création d'un modèle pour un document possédant des pages annexes

ABBYY FlexiCapture 8.0 Professional permet de créer des modèles de documents comportant des pages annexes.

Les pages annexes sont des pages supplémentaires qu'il est possible d'intégrer à un document quel qu'il soit. Elles ne contiennent aucun champ de reconnaissance et il n'est pas nécessaire d'y associer un modèle. Toutefois, elles sont prises en compte pour l'assemblage des documents. Par exemple, une demande de crédit est un formulaire fixe. Un certificat attestant du poste occupé écrit en style libre est joint à la demande. Ce certificat peut être traité comme une page annexe.

Pour créer un modèle de document avec annexes :

- Cliquez sur l'onglet **Pages** de la fenêtre **Structure du document** (dans la partie droite de l'Éditeur de modèles de document), puis cochez la case **Activer les pages d'annexe dans le document**. Saisissez le nombre de pages annexes ou leur intervalle dans la zone qui s'affiche à droite du nom de la section. (Figure 16)

- Vous pouvez également ouvrir la boîte de dialogue Propriétés du modèle de document (**Modèle > Propriétés du modèle de document...** dans l'Éditeur de modèles de document), puis cliquez sur l'onglet **Assemblage**. Cochez la case **Activer les pages annexes**. Saisissez le nombre minimum et maximum de pages annexes (Figure 17).

Pour enregistrer l'image d'une page annexe, vous devez indiquer les options d'enregistrement de l'image correspondantes au moment de configurer l'exportation (dans l'onglet **Images** de la boîte de dialogue **Paramètres d'exportation**). L'enregistrement d'une image peut se faire au format graphique ou au format PDF indexé.

4.3.7. Configuration de l'exportation des données

Pour configurer la méthode d'enregistrement des données obtenues après traitement de documents papier, l'administrateur doit configurer l'exportation des données pour chaque modèle de document. Quatre types d'exportation sont possibles. Vous pouvez exporter les données vers un fichier, une base de données ou une bibliothèque Microsoft SharePoint, ou décrire l'exportation à l'aide d'un script. Pour configurer l'exportation des données, sélectionnez **Modèle > Paramètres d'exportation...** dans la fenêtre Éditeur de modèles de document, puis définissez les options nécessaires dans la boîte de dialogue **Paramètres d'exportation**.

Le type d'exportation (fichier, base de données, SharePoint ou personnalisé) apparaît dans le champ **Type d'exportation**. La valeur saisie dans ce champ détermine les paramètres suivants.

4.3.7.1. Exportation vers des fichiers

Pour configurer l'exportation vers un fichier, sélectionnez **Exporter vers des fichiers** dans le champ **Type d'exportation** (Figure 18).

Dans l'onglet **Destination**, sélectionnez le dossier dans lequel les fichiers d'exportation seront stockés. Vous pouvez exporter des documents d'un lot vers un fichier unique (sélectionnez **Créer un dossier distinct pour chaque lot**) ou chaque document vers un fichier séparé (sélectionnez **Créer un dossier distinct pour chaque document**). Si vous ne sélectionnez aucune des options proposées, tous les documents seront exportés dans un seul et même fichier. Cochez la case **Écraser les fichiers existants** si vous souhaitez autoriser l'écrasement du fichier d'exportation.

Indiquez le mode de génération du nom des fichiers d'exportation : cliquez sur le bouton **Options du nom de fichier...** et sélectionnez les options souhaitées dans la boîte de dialogue qui apparaît.

Dans l'onglet **Format**, sélectionnez le type de fichier (les formats suivants sont pris en charge : DBF, TXT, XLS, XML), puis définissez les propriétés d'exportation supplémentaires pour le type de fichier sélectionné. Vous pouvez également préciser le codage du texte.

L'onglet **Images** contient les paramètres d'exportation des images. Pour enregistrer les images, cochez la case **Enregistrer des images de document**. Sélectionnez le dossier et le nom du fichier dans lequel les images traitées seront enregistrées. Vous pouvez également activer le bouton radio **dans le dossier de données** pour enregistrer les images dans le même dossier que celui où les données sont enregistrées.

Choisissez le format dans lequel les images seront stockées. Si vous sélectionnez PDF ainsi que l'option **Créer un PDF indexé**, le programme reconnaîtra l'ensemble du texte du document et enregistrera le texte reconnu au format sélectionné. Dans ce cas, vous pouvez indiquer une langue de reconnaissance : conservez la langue indiquée dans le modèle ou sélectionnez une ou plusieurs langues de la liste (bouton **Sélectionner**).

Si vous souhaitez modifier la résolution des images initiales, par exemple pour réduire la taille des données stockées, cochez la case **Changer la résolution en** et saisissez une nouvelle résolution.

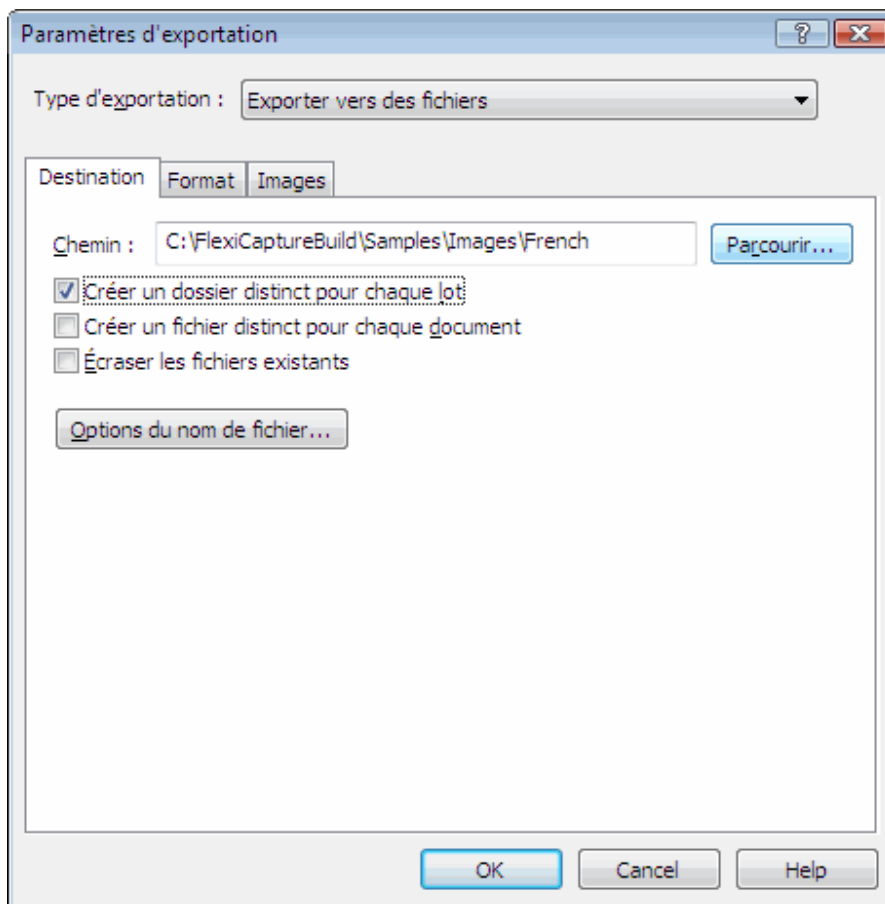


Figure 18. Configuration de l'exportation vers des fichiers dans la boîte de dialogue Paramètres d'exportation

4.3.7.2. Exportation vers une base de données

Pour exporter vers une base de données, sélectionnez l'option **Exporter vers une base de données compatible ODBC** dans le champ **Type d'exportation** (Figure 19).

Dans l'onglet **Connexion**, saisissez les paramètres de connexion dans la fenêtre **Chaîne de connexion** ou cliquez sur **Configurer la connexion...** et définissez les paramètres de connexion dans la boîte de dialogue **Propriétés des liaisons de données**. Sélectionnez le schéma dans la liste déroulante.

Pour tester la connexion avec la base de données, cliquez sur **Tester la connexion**.

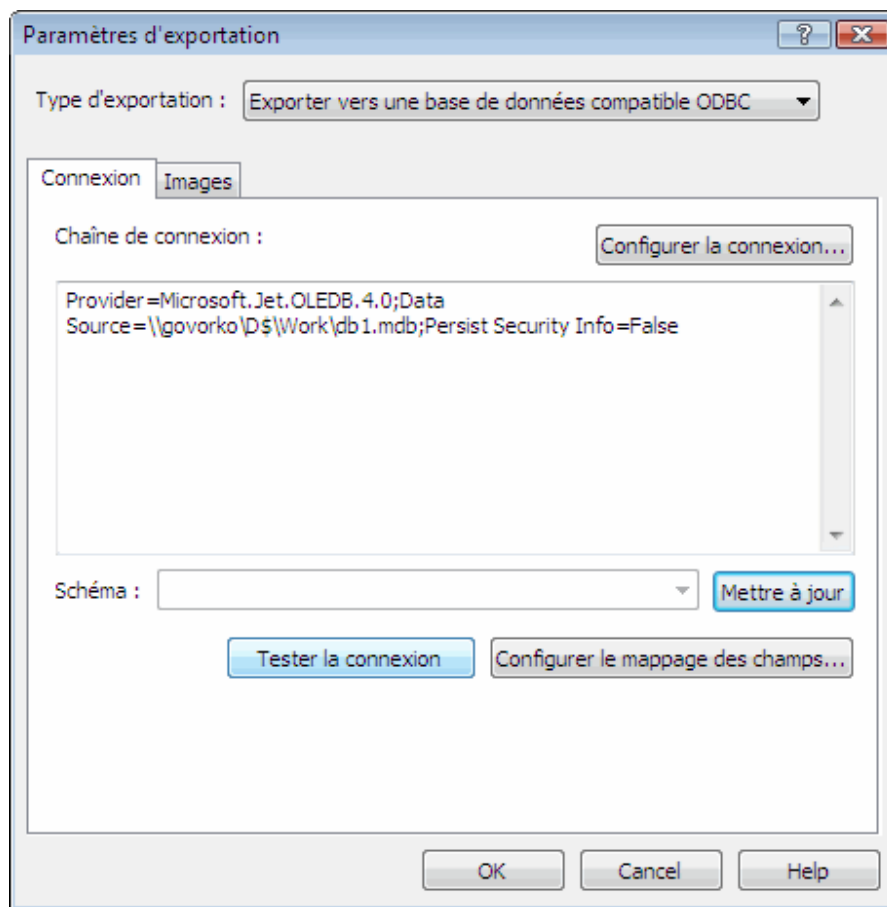


Figure 19. Configuration de l'exportation dans une base de données au moyen de la boîte de dialogue Paramètres d'exportation

À présent, définissez vers quels tableaux et colonnes de tableau de la base de données les valeurs de champ du document seront exportées. Pour cela, cliquez sur **Configurer le mappage des champs...**

La partie gauche de la boîte de dialogue **Mappage de champ** (Figure 20) contient une liste des sections et des champs du document. Dans la partie droite, indiquez les tableaux et les champs correspondants de la base de données. Si la base de données contient déjà des tableaux pour l'exportation des données, utilisez la boîte de dialogue **Mappage de champ** pour sélectionner un tableau de base de données pour chaque section et une colonne de tableau de base de données pour chacun des champs du document.

Si la base de données ne contient aucun tableau dédié à l'exportation du document, vous pouvez en créer automatiquement en cliquant sur **Créer automatiquement des tableaux**. Si vous cliquez sur ce bouton, le programme crée des tableaux dont la structure correspond idéalement à l'exportation. Les lignes du tableau sont alors affectées aux champs correspondants du document.

Les groupes de champs, les champs possédant plusieurs instances, les tableaux de document et les sections sont exportés dans des tableaux de base de données séparés. Deux clés sont utilisées pour lier le tableau parent aux tableaux enfants : la clé primaire et la clé étrangère. Dans le tableau parent, une clé primaire est affectée à chacune des entrées alors que dans le tableau enfant, une clé étrangère contenant la valeur de la clé primaire correspondante est affectée à chacune des entrées. Les clés sont automatiquement ajoutées si nécessaire. Vous ne devez indiquer que le champ dans lequel la clé doit être exportée.

L'option **Afficher les colonnes liées** est sélectionnée par défaut. Si cette option n'est pas sélectionnée, les listes des colonnes du tableau disponible à droite de l'arborescence ne contiendront aucune colonne dont l'exportation a déjà été configurée.

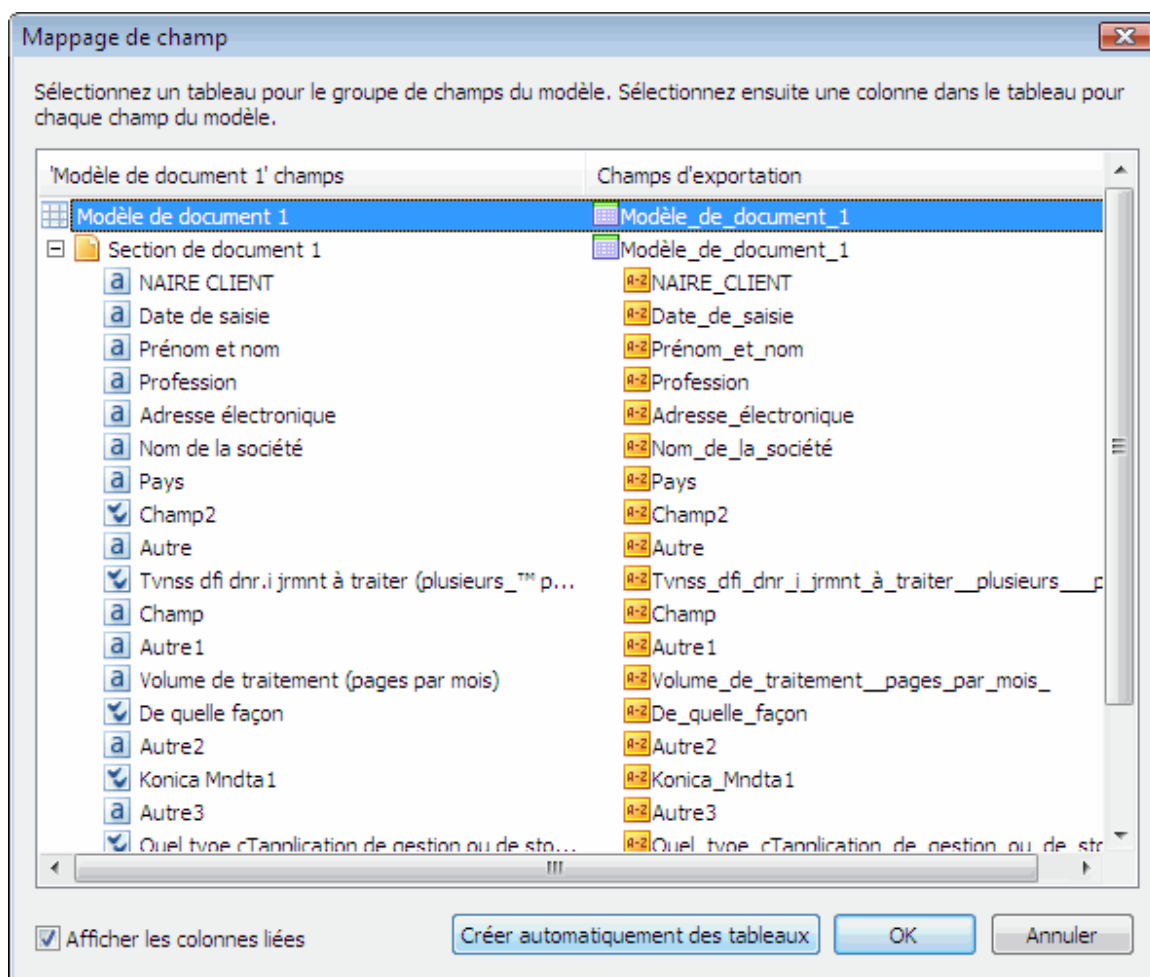


Figure 20. Configuration des liens entre champs de document et champs de base de données durant une exportation

Les paramètres relatifs à l'enregistrement des images peuvent être définis dans l'onglet **Image** de la boîte de dialogue **Paramètres d'exportation**.

Vous pouvez enregistrer des images dans une base de données ou dans un fichier (auquel cas, vous devrez désigner le dossier correspondant).

Choisissez le format dans lequel les images seront stockées. Si vous sélectionnez PDF indexé, le programme ne se contentera pas d'enregistrer l'image du document, il procédera également à la reconnaissance de l'intégralité du texte du document et enregistrera le texte reconnu au format choisi.

Si vous souhaitez modifier la résolution des images initiales, par exemple pour réduire la taille des données stockées, activez la case à cocher **Changer la résolution en** et saisissez une nouvelle résolution.

4.3.7.3. Exportation de données vers une bibliothèque de documents SharePoint

ABBYY FlexiCapture 8.0 vous permet d'exporter des documents vers une bibliothèque de documents Microsoft™ SharePoint. Ce processus vous permet d'obtenir une bibliothèque contenant vos images de document et des colonnes supplémentaires comprenant la valeur de certains champs pour chaque document.

Remarque :

- Vous devez disposer de droits d'administrateur pour configurer l'exportation de données vers SharePoint. Des droits de collaborateur sont suffisants pour exporter des données vers SharePoint.
- Les colonnes SharePoint dans lesquelles les données sont enregistrées doivent être de type *Ligne de texte unique* ou *Lignes de texte multiple*.

Pour configurer l'exportation vers SharePoint, sélectionnez **Exporter vers SharePoint** dans le champ **Type d'exportation** (figure 21).

Sur l'onglet **Connexion SharePoint**, saisissez l'adresse du serveur (**URL serveur**) où sont situées les bibliothèques SharePoint. Utilisez le bouton **Paramètres de connexion...** pour configurer les paramètres d'authentification (les paramètres d'authentification de Windows sont utilisés par défaut) et les paramètres proxy au besoin.

Cliquez sur le bouton **Connecter** pour rétablir la connexion au serveur en cas de coupure.

Sélectionnez une bibliothèque SharePoint dans la liste.

Sélectionnez le type de contenu de document requis dans la liste **Type de contenu** (à partir de SharePoint 2007). Le cas échéant, vous pourrez exporter les valeurs vers les champs correspondant au type de contenu sélectionné. Cliquez sur le bouton **Mappage des champs...** et sélectionnez les options requises dans la boîte de dialogue **Mappage de champ**.

Indiquez les règles de dénomination des fichiers : cliquez sur le bouton **Options du nom de fichier...** et sélectionnez les options souhaitées dans la boîte de dialogue qui apparaît.

Sur l'onglet **Images**, indiquez les paramètres de stockage d'images (format de fichier, qualité d'image, type de couleur, résolution). Si vous sélectionnez le format de stockage d'images PDF ou PDF/A, vous pouvez créer un fichier PDF indexé en sélectionnant l'option **Créer un PDF indexé**. Vous pouvez sélectionner une langue (ou plusieurs langues pour les documents multilingues) différente des langues indiquées dans le modèle : cliquez sur le bouton **Sélectionner...** situé à droite du champ **Langue** et sélectionnez les langues requises. Si vous souhaitez utiliser la langue spécifiée dans le modèle, sélectionnez **Comme dans modèle** dans le champ **Langue**.

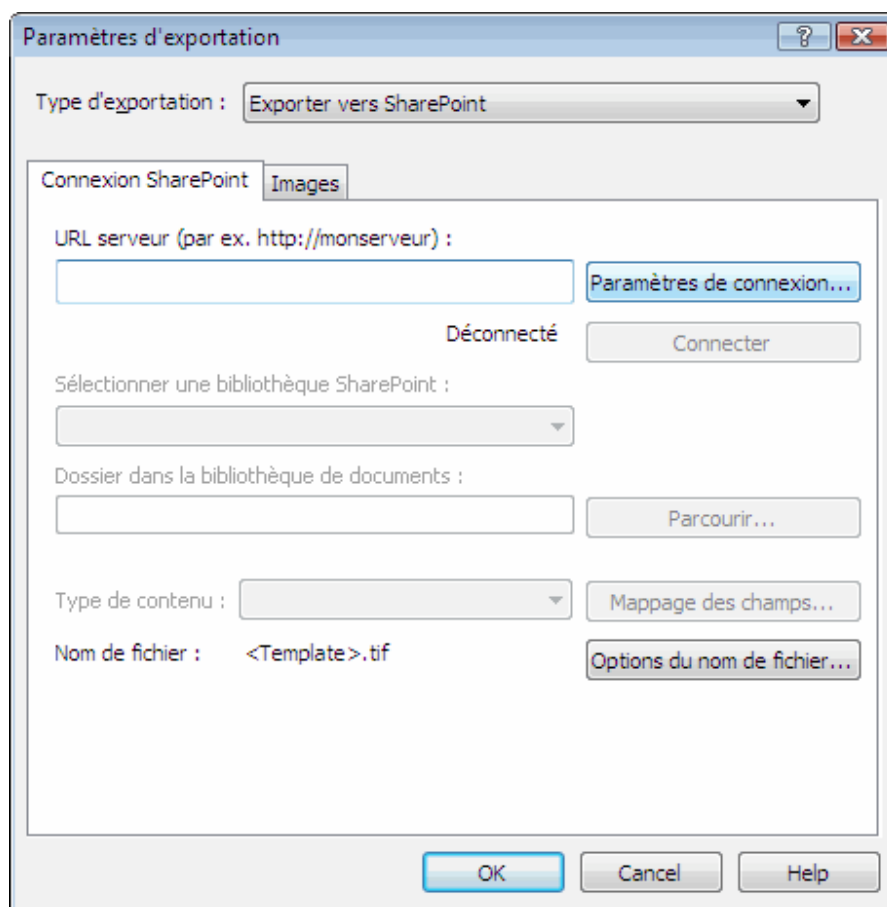


Figure 21. Configuration de l'exportation vers une bibliothèque Microsoft SharePoint

4.3.7.4. Exportation personnalisée

Ce type d'exportation vous permet de définir des procédures d'exportation avancées. Vous utilisez pour cela des outils qui ne sont pas disponibles dans l'interface du programme au moment d'exporter vers des fichiers ou vers une base de données.

Si vous souhaitez configurer une exportation d'après un script, sélectionnez **Exportation personnalisée (script)** dans le champ **Type d'exportation** (Figure 22).

Ensuite, sélectionnez le langage de script (**JScript®** ou **VBScript**), puis saisissez le texte du script dans la fenêtre de l'éditeur qui s'ouvre lorsque vous cliquez sur **Éditer le script...** (Pour une description plus détaillée et des exemples d'utilisation des scripts, nous vous invitons à consulter le fichier d'aide du programme.)

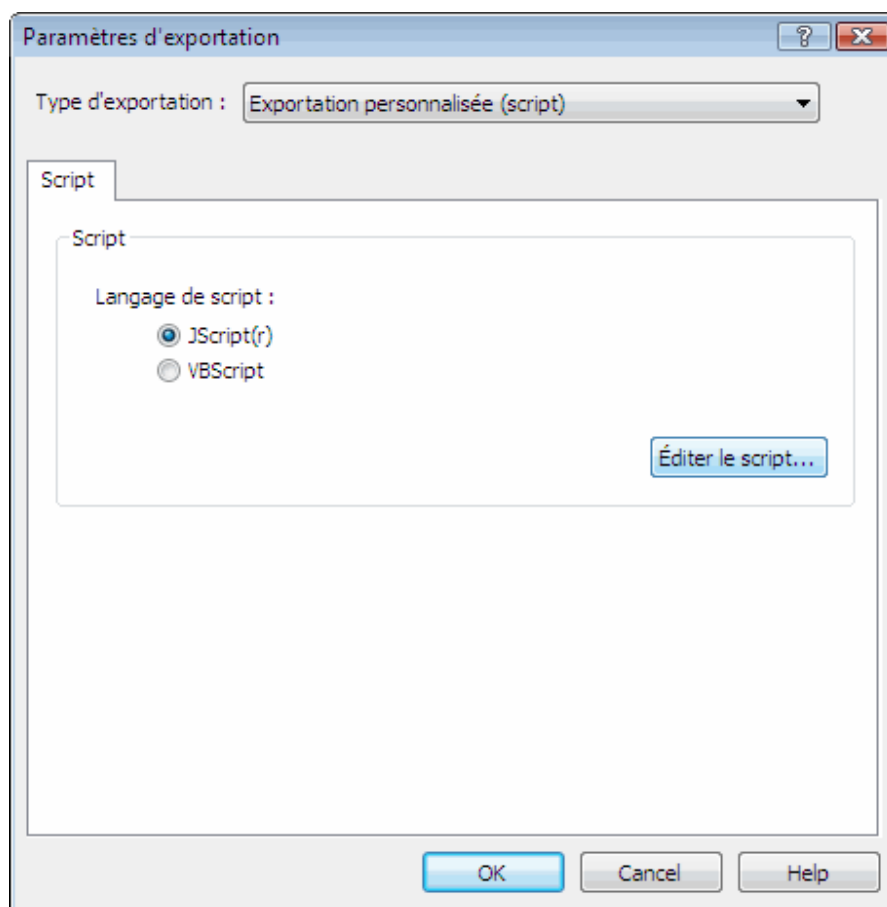


Figure 22. Configuration d'une exportation d'après un script dans la boîte de dialogue Paramètres d'exportation

Configuration des options d'exportation des données :

1. Dans la fenêtre Éditeur de modèles de document, sélectionnez **Modèle > Paramètres d'exportation...** Dans le champ **Type d'exportation** de la boîte de dialogue qui apparaît, sélectionnez **Exporter vers les fichiers**.
2. Dans l'onglet **Destination**, saisissez le chemin du dossier dans lequel le fichier d'exportation doit être enregistré, puis donnez un nom à ce fichier.
3. Dans l'onglet **Format**, sélectionnez le type **Document XML (*.xml)** et cochez l'option **Erreurs** dans la section **Exporter des données supplémentaires**.
4. Si vous souhaitez enregistrer des images de page, allez dans l'onglet **Images**, cochez la case **Enregistrer des images de document**, puis définissez les paramètres nécessaires.
5. Une fois que vous avez défini les paramètres d'exportation, les propriétés principales apparaissent dans l'onglet **Exportation** de la boîte de dialogue **Propriétés du modèle de document (Modèle > Propriétés du modèle de document...)**. Vous pouvez modifier les paramètres d'exportation dans cette boîte de dialogue en cliquant sur **Modifier les paramètres d'exportation**.

4.3.8. Configuration de la vue des données reconnues

Une fois les données reconnues, l'utilisateur les visualisent dans la fenêtre de document. Par défaut, les données sont triées par ordre et les libellés correspondent aux noms des champs du modèle. Toutefois, vous pouvez modifier la vue des données reconnues et la rendre plus lisible, par exemple en modifiant l'ordre de la présentation des données ou en ajoutant un texte explicatif. La partie inférieure droite de la fenêtre Éditeur de modèles de document sert à la modification de la vue des données reconnues.

Vous pouvez déplacer les champs, modifier leurs noms ou toute autre propriété. Pour ajouter un libellé, sélectionnez **Insérer une zone de libellé** dans le menu contextuel.

Vous pouvez modifier la police et la taille du texte des noms des champs et des valeurs des données reconnues dans l'onglet **Formulaire de données** de la boîte de dialogue **Propriétés du modèle de document (Modèle > Propriétés du modèle de document...)**.

Visualisez la façon dont les données reconnues s'affichent dans la fenêtre de document. Vous pouvez conserver la présentation des données par défaut ou bien la modifier.

Cette étape termine la création du modèle.

4.3.9. *Modification et publication d'un modèle*

Une fois que vous avez créé un modèle, enregistrez-le et testez-le sur plusieurs images. Si vous êtes satisfait du résultat, publiez le modèle de sorte à rendre le disponible pour la reconnaissance des documents.

Pour publier un modèle, cliquez sur **Publier** dans la boîte de dialogue **Modèles de document (Projet > Modèles de document...)** dans la fenêtre principale du programme).

Pour revenir à la modification d'un modèle, sélectionnez le modèle concerné dans la boîte de dialogue **Modèles de document**, puis cliquez sur **Éditer**. Le modèle est alors rendu inaccessible et ne peut pas être modifié par d'autres utilisateurs. La fonction de reconnaissance utilisera la dernière version publiée du modèle. Les tâches relatives à la modification du modèle ne généreront ainsi pas le travail des opérateurs. Une fois la modification du modèle terminée, publiez-le. La nouvelle version deviendra immédiatement accessible à tous les utilisateurs. Si vous ne souhaitez pas publier la nouvelle version du modèle, ignorez les modifications effectuées et retournez à la dernière version publiée de celui-ci. Pour cela, sélectionnez le modèle concerné dans la boîte de dialogue **Modèles de document**, puis cliquez sur **Refuser les modifications**.

4.1.10. *Test du modèle*

Avant de lancer le traitement de vos documents, commencez par tester le modèle.

Vous pouvez tester le modèle directement dans l'éditeur de modèles à partir de l'image utilisée pour le créer.

Pour lancer le test, sélectionnez **Document>Exécuter un test**.

Des tests supplémentaires sont exécutés sur les documents ajoutés aux lots de test, car ces lots utilisent la version locale du modèle. Les lots de test sont disponibles dans **Projet>Liste de lots de test**, dans la fenêtre principale. Vous pouvez également accéder aux lots de test à partir de l'éditeur de modèles en sélectionnant **Outils>Passer au test de lot...**

Si le programme détecte des erreurs de règles pendant le test, des propriétés de champs incorrectes, etc., vous devrez modifier le modèle pour les corriger. Une fois toutes les erreurs corrigées, publiez le modèle. Vous pouvez maintenant commencer à traiter vos documents.

Dans un premier temps, testez votre modèle sur l'image source (Document>Exécuter un test), puis sur l'image renseignée de votre questionnaire (voir [Utilisation d'un projet de configuration](#)). Effectuez toutes les modifications nécessaires et publiez le modèle.

4.4. Configuration de l'importation des images

La première tâche de l'opérateur consiste à ajouter de nouvelles images au projet. Ces images peuvent être des documents papier (alors numérisés) ou des images électroniques. Si les images proviennent régulièrement d'une seule et même source, vous pouvez automatiser la procédure d'ajout de sorte à ce que toutes les actions nécessaires soient effectuées automatiquement, d'un simple clic. Une fois les paramètres requis configurés, les images sont numérisées ou importées à partir d'une source prédéfinie et les images reçues sont traitées (elles sont nettoyées, par exemple). Si nécessaire, l'opérateur peut modifier la source de l'importation des images ou ses paramètres.

Si les sources importées changent souvent et si vous ne souhaitez pas en modifier les paramètres à chaque fois, vous pouvez créer plusieurs profils d'importation et passer de l'un à l'autre.

Vous pouvez également configurer l'ajout régulier d'une image en mode arrière-plan.

Pour ajouter un profil d'importation, sélectionnez **Projet > Profils d'importation d'image...** dans la fenêtre principale du programme. Dans la boîte de dialogue qui s'ouvre, vous pouvez créer de nouveaux profils d'importation, de même que vous pouvez copier des profils existants, en modifier et en supprimer. (Figure 23).

Pour créer un nouveau profil d'importation, cliquez sur **Nouveau**. L'Assistant de création de profil s'ouvre.

Au cours de la première étape de l'Assistant, sélectionnez la source d'importation. Il peut s'agir d'un scanner ou d'un dossier actif (par exemple un dossier dans lequel le programme recherchera de nouvelles images). Si vous souhaitez ajouter des images depuis un dossier actif, vérifiez que vous possédez des droits en lecture et en écriture sur ce dossier. Si la source d'importation est un scanner, veillez à ce qu'il soit connecté à votre PC.

Création d'un profil d'importation au moyen d'un scanner

1. Au cours de la première étape de l'Assistant, sélectionnez **Scanner** comme source d'importation.
2. Précisez ensuite les options nécessaires au traitement des images reçues. Par exemple, vous pouvez définir le nettoyage des images (nécessaire si la qualité de l'image est faible et si elle contient un bruit important). Vous pouvez également modifier les options de numérisation. Dans la liste déroulante **Style de la boîte de dialogue des paramètres**, sélectionnez la méthode d'interaction entre le programme et le scanner : la méthode **Paramètres du scanner FlexiCapture** qui utilise la boîte de dialogue ABBYY FlexiCapture pour la configuration des options de numérisation ou la méthode **Options de numérisation** qui utilise la boîte de dialogue du pilote du scanner pour configurer les options de numérisation.
3. Indiquez ensuite les options d'assemblage du document et de prétraitement des images. Vous pouvez par exemple configurer le programme pour nettoyer les images (cela peut s'avérer nécessaire pour les images contenant du bruit), pour convertir les images en noir et blanc, pour faire tourner les images dans un sens donné ou pour désactiver la correction de mise en travers.
4. Pour terminer, modifiez le nom affecté par défaut au profil d'importation et saisissez une description.

Création d'un profil d'importation au moyen d'un dossier actif

1. Dans l'Assistant, sélectionnez **Dossier actif** comme source d'importation, puis indiquez le chemin menant à ce dossier.

2. Définissez ensuite les options de chargement des images. Vous pouvez cocher la case **Vérifier le dossier actif toutes les** et indiquer l'intervalle après lequel le programme devra vérifier le dossier actif à la recherche de nouvelles images. Dans la section **Paramètres du lot**, sélectionnez l'une des options pour indiquer le lot dans lequel de nouvelles images seront ajoutées. Dans la liste déroulante **Nombre de fichiers à ajouter**, sélectionnez l'une des valeurs pour indiquer le nombre de fichiers à ajouter au lot sélectionné ou au lot nouvellement créé.
3. Vous avez ensuite la possibilité de définir des options d'assemblage du document et de prétraitement des images, comme lors de la création d'un profil d'importation pour le scanner.
4. L'Assistant de création des profils d'importation vous invite à présent à configurer les options visant à purger le dossier actif après l'importation. Les images correctement importées et celles dont le traitement a induit une erreur peuvent être supprimées ou déplacées dans un autre dossier.
5. Pour terminer, modifiez le nom affecté par défaut au profil d'importation et saisissez une description.

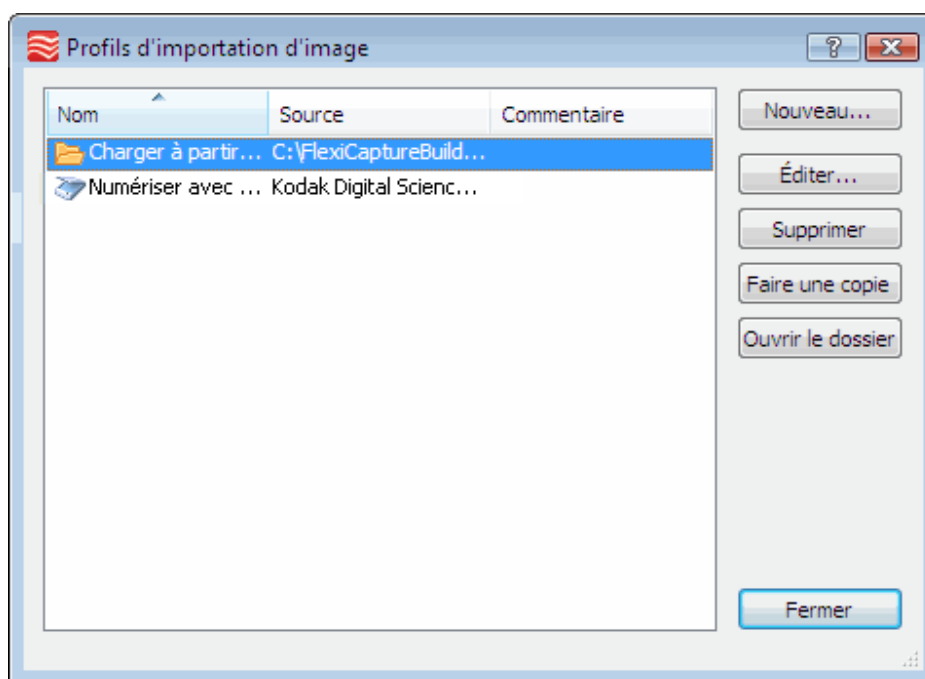


Figure 23. Boîte de dialogue Profils d'importation d'images

1. Configurez un profil d'importation pour vos images. Pour cela, sélectionnez **Projet > Profils d'importation d'images** dans la fenêtre principale d'ABBYY FlexiCapture 8.0.
2. Dans la boîte de dialogue **Profils d'importation d'images**, cliquez sur **Nouveau...** pour créer un nouveau profil. Nommez ce profil et sélectionnez l'option **Scanner** (les images du questionnaire seront reçues du scanner).
3. Configurez les options de traitement de la numérisation et des images.
4. Au cours de la dernière étape de l'Assistant, vous pouvez modifier le nom par défaut affecté au profil d'importation et fournir une description.

5. Configuration du système pour le traitement des documents flexibles

La configuration du traitement des documents flexibles est une tâche que se révèle plus difficile. Étant donné que l'emplacement, la taille et le nombre de champs de données varient d'un document à un autre, vous ne pouvez pas créer de modèle possédant une organisation fixe de champs pour les factures, les ordres de paiement et autres documents similaires. Ainsi, un modèle fixe ne peut pas être associé à toutes les copies de ces documents. Dans de tels documents, une recherche est effectuée sur les champs de données après la reconnaissance de la page dans son ensemble et le modèle est alors créé d'après les mots-clés et l'organisation mutuelle des champs de données. Un composant du système a été tout particulièrement conçu pour créer de tels modèles. Il s'agit d'ABBYY FlexiLayout Studio 8.0. Pour en savoir plus sur la création d'un modèle flexible, consultez l'aide d'ABBYY FlexiLayout Studio 8.0.

Un FlexiLayout est créé et exporté dans un fichier AFL, puis joint à ABBYY FlexiCapture 8.0 Professional. Nous pouvons utiliser une description flexible pour le modèle dans son ensemble ou comme l'une de ses sections. Une description flexible peut comporter plusieurs pages.

Vous pouvez joindre une description flexible au moment de la création d'un modèle de document. Pour cela, ajoutez l'image du document durant la seconde étape de la création du modèle. Ensuite, sélectionnez l'option **Charger FlexiLayout**, puis saisissez le chemin du fichier AFL contenant la description flexible.

Vous pouvez également joindre une description flexible à la fenêtre Éditeur de modèles de document. Pour cela, utilisez la boîte de dialogue Propriétés d'une section. Pour ouvrir la boîte de dialogue **Propriétés** d'une section, sélectionnez **Propriétés...** dans le menu contextuel de la section. Cliquez sur l'onglet **FlexiLayout**, puis sur **Charger...** Sélectionnez le fichier AFL contenant la description flexible.

Une fois la description flexible jointe, tous les champs et leur marquage figureront dans l'image.

Le jeu de champs du modèle doit être identique au jeu de champs de la description flexible. Vous n'êtes pas autorisé à ajouter ni à supprimer des champs (à l'exception des champs sans région de l'image). Vous pouvez modifier le jeu de champs pour qu'il soit reconnu uniquement si vous modifiez la description flexible.

Définissez ensuite les propriétés de champs de données. Ce processus est similaire à la configuration des propriétés de champ d'un modèle fixe. Il s'agit de configurer la reconnaissance, la vérification et les propriétés d'exportation, puis de définir les règles nécessaires.

Pour plus d'informations, consultez le chapitre [Création d'un modèle de document](#).

6. Particularités de l'acquisition de documents non structurés

ABBYY FlexiCapture 8.0 Professional peut vous aider à traiter des documents non structurés contenant des informations écrites dans un style libre, par exemple des contrats, des lettres, des commandes, des annexes, etc. Grâce à ABBYY FlexiCapture 8.0 Professional, les documents non structurés utilisant du texte ou des images séparés par des pages vides ou des pages avec codes barres peuvent être traités, puis exportés dans des fichiers PDF indexés ou des fichiers au format graphique.

Le traitement de tels documents passe le plus souvent par une conversion au format électronique puis par une organisation de la recherche d'après des champs clés.

La recherche est habituellement lancée sur les champs clés au moyen d'une description flexible créée dans ABBYY FlexiLayout Studio.

Si la recherche automatisée des champs clés est indisponible, l'opérateur peut saisir les valeurs des champs clés manuellement. Pour cela, créez un modèle de document ne possédant qu'un seul champ (ou plusieurs champs si nécessaire) et sélectionnez l'option **Ne pas reconnaître (clé du champ d'image - sera entré manuellement)** dans les propriétés de reconnaissance du ou des champs. Au lancement du processus de vérification, l'opérateur sera invité à saisir la valeur du ou des champs clés manuellement.

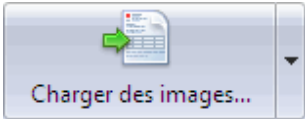
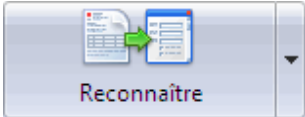
Pour enregistrer le document, vous devez configurer le processus d'exportation. Les valeurs des champs clés peuvent être exportées dans un fichier ou une base de données et l'image d'un document, enregistrée dans le format graphique de votre choix. Pour cela, cochez la case **Enregistrer des images de document** et définissez les paramètres d'enregistrement de l'image dans l'onglet **Images** de la boîte de dialogue **Paramètres d'exportation**. Vous pouvez enregistrer les images de document sous forme de fichiers PDF indexés ou de fichiers dans un format graphique précis.

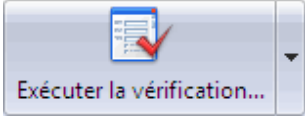
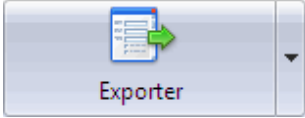
Faites néanmoins attention au moment d'assembler les pages dans les documents : avec des documents non structurés, vous pouvez déterminer à quel document une page en particulier appartient. Pour automatiser l'assemblage des documents non structurés, nous vous recommandons de séparer les documents au moyen de feuilles vierges ou de pages contenant des codes barres. Au moment d'ajouter des images à un lot (numérisation, ajout depuis un fichier ou création d'un profil d'importation), vous devrez alors activer l'option **Images séparées par**, puis sélectionnez la valeur **pages vierges** ou **pages avec code barre** dans la liste déroulante, en fonction des pages devant servir de séparateurs. Les pages sont assemblées en documents automatiquement : les pages seront ajoutées au document actuel jusqu'à la page de séparation suivante.

7. Utilisation d'un projet de configuration

Une fois que l'administrateur a configuré un modèle et défini tous les paramètres nécessaires, vous pouvez lancer le traitement des documents. L'intégralité de la procédure peut être configurée de sorte que les opérations d'ajout d'images, de reconnaissance et d'exportation de données puissent être effectuées quasiment sans intervention. Le travail de l'opérateur ne dépend pas du type de document.

Le traitement d'un document se déroule en quatre étapes. À chaque étape correspond un bouton de la barre d'outils.

<p>1. Importation</p> <p>Cliquez sur la flèche pour sélectionner l'une des options d'importation de l'image : Charger les images..., Numériser les images ou importation d'après l'un des profils définis. Le texte du bouton correspond à l'action accomplie.</p>	
<p>2. Reconnaissance</p> <p>Lance le processus de reconnaissance du document. En cliquant sur la flèche, vous pouvez également lancer l'analyse du document ou</p>	

associer l'un des modèles.	
<p>3. Vérification</p> <p>Lance le processus de vérification. En cliquant sur la flèche, vous pouvez également lancer le processus de vérification des règles.</p>	
<p>4. Exportation</p> <p>Lance l'exportation des données selon les propriétés définies dans le modèle. En cliquant sur la flèche, vous pouvez également exporter les données dans un fichier ou une base de données.</p>	

7.1. Ajout d'images

La première étape du traitement des documents consiste à ajouter des images de pages au projet. Pour cela, sélectionnez ou créez un lot destiné à recevoir les images (pour créer un lot, cliquez dans la fenêtre principale du programme avec le bouton droit de la souris et sélectionnez **Nouveau lot**). Si vous ajoutez des images à un projet qui ne contient aucun lot, un lot sera créé automatiquement.

Il existe plusieurs méthodes pour ajouter des images à un lot :

1. Chargez des images précédemment enregistrées. Pour cela, sélectionnez **Charger les images...** dans le menu déroulant du bouton Importation ou appuyez sur les touches CTRL+O. Dans la boîte de dialogue qui apparaît, sélectionnez les fichiers d'images nécessaires et configurez les options d'importation des images.
 Au moment d'importer des images provenant de fichiers multipages, plusieurs pages seront ajoutées au lot.
2. Numérisez les images. Pour ajouter des images provenant d'un scanner, sélectionnez **Numériser les images....** Vous serez invité à sélectionner un scanner et à numériser les images.
3. Importez des images au moyen d'un profil d'importation des images déjà créé par l'administrateur (consultez le chapitre [Configuration de l'importation des images](#))

Si des profils d'importation ont été configurés, leur nom apparaît sur les boutons du menu d'importation. Pour lancer l'importation des images, sélectionnez le nom de l'un des profils d'importation.

Sélectionnez **Importer les images...** La boîte de dialogue **Sélectionner le profil d'importation** s'ouvre. Dans cette boîte de dialogue, sélectionnez un profil d'importation dans la liste déroulante, puis cliquez sur **Importer** pour lancer l'importation des images. Cette boîte de dialogue vous sert également à créer un nouveau profil d'image ou à éditer des profils existants.

Si des profils d'importation ont été configurés, leur nom apparaît sur les boutons du menu d'importation. Pour lancer l'importation des images, sélectionnez le nom de l'un des profils d'importation.

Sélectionnez **Importer les images...** La boîte de dialogue **Sélectionner le profil d'importation** s'ouvre. Dans cette boîte de dialogue, sélectionnez un profil d'importation dans la liste déroulante, puis cliquez sur **Importer** pour lancer l'importation des images. Cette boîte de dialogue vous sert également à créer un nouveau profil d'image ou à éditer des profils existants.

Si vous avez déjà sélectionné un profil d'importation, vous n'avez pas besoin de le rechercher dans la liste pour l'utiliser à nouveau. Cliquez simplement sur le bouton d'importation portant le nom du profil récemment utilisé.

Les images peuvent être importées en mode arrière-plan si un profil d'importation correspondant est configuré. Les images du dossier actif sont importées automatiquement dans le système. Passé le délai indiqué dans les paramètres, le programme vérifie le dossier actif à la recherche de nouvelles images.

Une fois les nouvelles images ajoutées à un lot, les pages non traitées s'affichent dans la liste.

1. Créez un lot dans lequel l'image sera ajoutée. Pour cela, cliquez dans la fenêtre principale du programme avec le bouton droit de la souris et sélectionnez **Nouveau lot**. Saisissez le nom et la description du nouveau lot. Double-cliquez sur le lot pour l'ouvrir.
2. Renseignez le questionnaire.
3. Numérisez l'image.
4. Importez l'image en sélectionnant **Importer les images...** ou **Numériser les images...** Une page de nom <Inconnu> apparaît alors dans le lot.

7.2. Reconnaissance

Pour reconnaître des données, cliquez sur **Reconnaître**. Le programme associe le modèle en le sélectionnant d'après les ancres et les identificateurs, puis reconnaît les données des régions du document indiquées par les champs du modèle sélectionné.

Il est possible de reconnaître automatiquement et immédiatement les images une fois ajoutées au lot. Pour cela, cochez la case **Reconnaître les images ajoutées automatiquement (Outils > Options... > onglet Traitement de document)**.

Le **Niveau de confiance** affiche le pourcentage de caractères fiables reconnus d'après le nombre total de caractères. Une fois la reconnaissance terminée, l'opérateur peut passer à la vérification du document.

Préalablement à la reconnaissance, vous pouvez analyser une page, c'est-à-dire associer le modèle à cette page. Pour cela, sélectionnez **Analyser** dans le menu déroulant du bouton de reconnaissance. Si le modèle peut être associé, les noms des champs apparaissent pour la page ou le document auquel la page appartient et la reconnaissance des champs s'exécute alors. Si aucun des modèles du projet ne peut être associé à la page, la page demeure non traitée.

Dans la plupart des cas, l'association d'un modèle correctement créé et des pages est automatique. Toutefois, il arrive de devoir sélectionner un modèle manuellement. Pour associer un modèle, sélectionnez le document ou la page concernée et choisissez **Associer un modèle...** dans le menu déroulant du bouton de reconnaissance.

Sélectionnez **Analyser** dans le menu déroulant du bouton de reconnaissance. Si le modèle est correctement associé, son nom s'affichera en remplacement de <Inconnu>.

Cliquez sur **Reconnaître**. La page sera alors reconnue.

7.3. Vérification

La vérification, c'est-à-dire le contrôle des données reconnues, est la phase qui nécessite le plus de travail de la part de l'opérateur. Le processus de vérification d'ABBYY FlexiCapture 8.0 Professional veille à être extrêmement pratique pour l'opérateur et à minimiser le nombre d'erreurs. Concernant les documents multipages, le programme vérifie en premier que les pages ont été correctement assemblées dans les documents. Le processus de vérification peut alors démarrer. Il s'agit d'une vérification du groupe et du

contexte. Vous pouvez également lancer une vérification dans la fenêtre de document. Les règles sont contrôlées durant le processus de vérification.

Vérification de l'assemblage du document. Concernant les documents multipages, le programme vérifie que les pages ont été correctement assemblées dans les documents. Si l'ordre des pages ne correspond pas à l'ordre indiqué ou si les valeurs du champ clé ne sont pas identiques sur toutes les pages, un drapeau rouge apparaît en regard du document et un message d'erreur s'affiche dans la fenêtre de document. L'opérateur doit alors s'assurer que les pages n'ont pas été mélangées durant la numérisation. Il est parfois simplement possible de corriger des erreurs d'assemblage en modifiant l'ordre des pages.

N'hésitez pas à utiliser le mode vignettes de page pour vérifier l'assemblage des pages car il est nettement plus confortable (Figure 24) Il vous suffit alors de modifier l'ordre des pages ou même de les déplacer d'un document à un autre par simple glisser-déposer.

Si le modèle indique qu'une vérification d'un champ clé doit être effectuée pour un assemblage correct du document, les valeurs des champs clés s'afficheront sous l'image, dans chacune des pages. Si les champs clés des pages d'un document ne sont pas identiques, ils apparaissent en rouge. Les champs clés peuvent être différents du fait d'une reconnaissance ou d'un renseignement incorrect. Veuillez vérifier les valeurs de champs clés. Si ces valeurs continuent à être différentes, il est fort probable que les pages concernées appartiennent à d'autres documents. Si l'ordre des pages a été mélangé, localisez les pages possédant des champs clés identiques, puis assemblez-les correctement.

Remarque : Pour effectuer un zoom avant sur les détails des vignettes de page, maintenez la touche CTRL enfoncée et utilisez la molette de la souris.

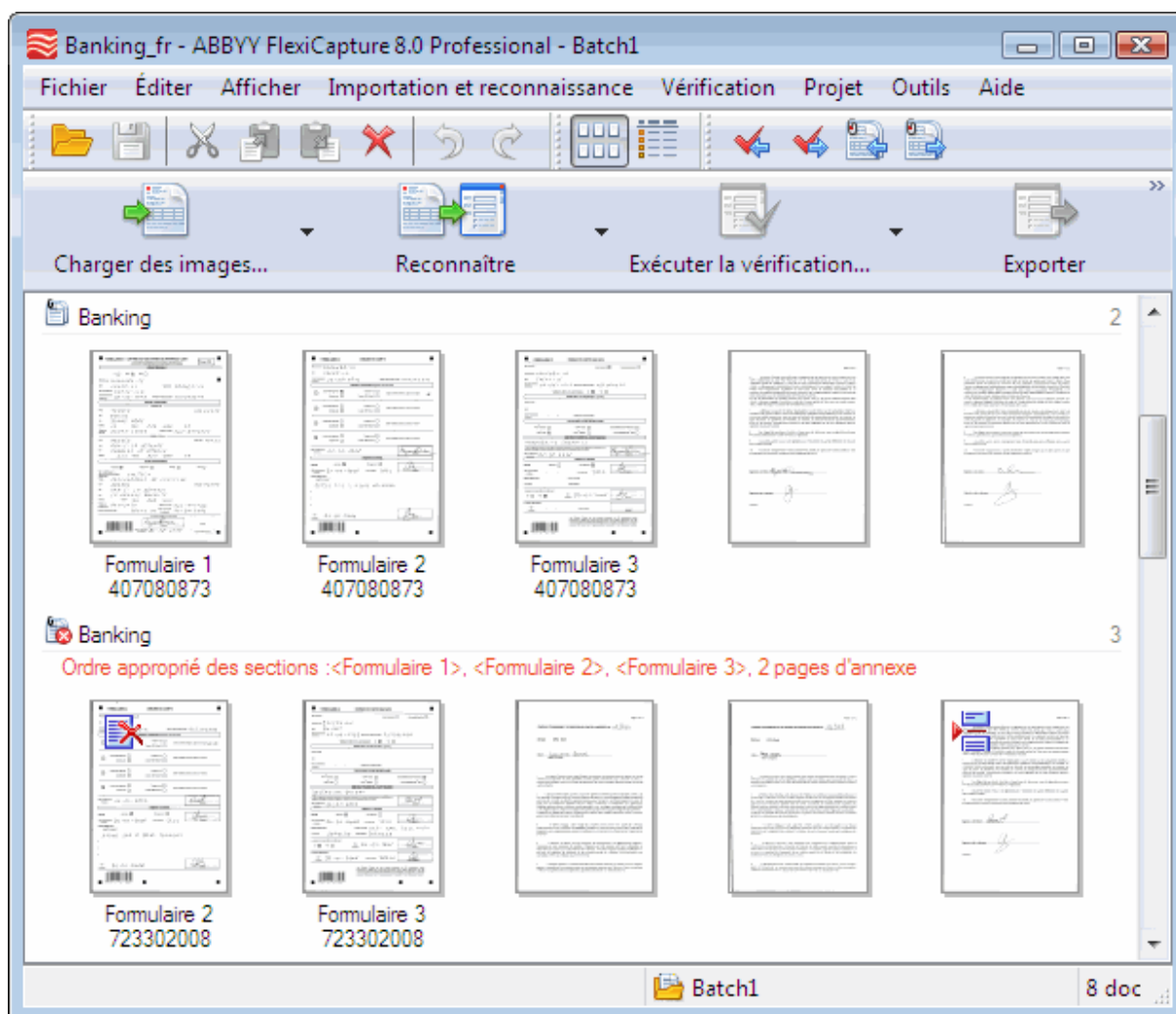


Figure 24. Fenêtre principale du programme, mode vignettes de page

Pour lancer la vérification des données reconnues, cliquez sur **Exécuter la vérification...**

La **vérification groupée** sous-entend le regroupement des images de caractères reconnues comme ayant une valeur identique et leur affichage sur l'écran de vérification. Il s'agit de confirmer les caractères correctement reconnus et de laisser pour la prochaine étape uniquement, les caractères jugés incorrects ou incertains (Figure 25).

Au cours de la vérification groupée, l'opérateur peut afficher l'image du champ dans lequel se trouve le caractère contrôlé. Pour cela, sélectionnez **Afficher l'image du caractère** dans le menu de raccourcis du caractère sélectionné ou appuyez sur F2. En positionnant le curseur sur le caractère vérifié, vous pouvez également activer le mode qui affichera le champ où se trouve le caractère. Pour cela, dans la fenêtre de vérification, sélectionnez **Afficher > Image du champ > Afficher l'image du champ** ou appuyez sur CTRL+I.

Pour corriger les caractères incorrectement reconnus, procédez comme suit : sélectionnez un caractère qui ne correspond pas au caractère du groupe, puis saisissez une valeur que vous savez convenir. La valeur saisie apparaît en vert dans le coin supérieur gauche de l'image du caractère. Si vous n'êtes pas sûr de la valeur du caractère même après l'avoir visualisé dans son contexte, cliquez sur celui-ci pour le marquer d'un point d'interrogation rouge. Vous pouvez également modifier l'état du caractère en utilisant le bouton correspondant de la barre d'outils **Basculer**.

Pour confirmer un caractère correctement reconnu, sélectionnez **Confirmer** dans le menu contextuel. De même, dans la barre d'outils, cliquez sur **Tout confirmer** pour confirmer en une seule fois tous les caractères affichés.

Au cours de la création d'un modèle, vous configurez les options de vérification au moment où vous indiquez les propriétés des champs. La vérification groupée s'exécute pour les caractères des champs pour lesquels vous avez sélectionné l'option **Inclure dans la vérification groupée**, dans l'onglet **Vérification** de la boîte de dialogue **Propriétés**.

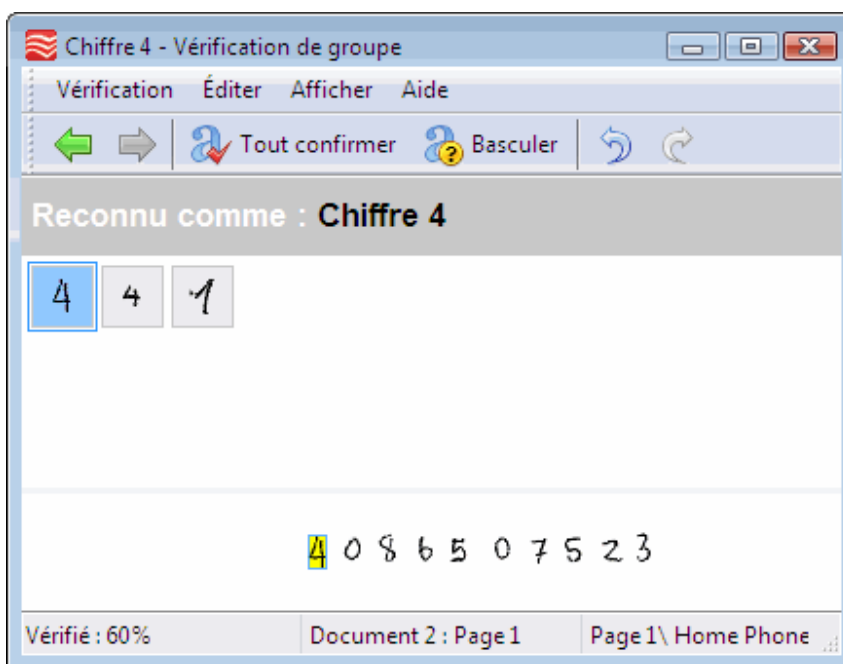


Figure 25. Vérification groupée des chiffres

La **vérification contextuelle** est un mode de vérification utilisé pour corriger le format des champs dont la plage de valeurs est connue ou aisément identifiée. Le nom du pays est l'un de ces champs car nous savons les valeurs que ce champ peut contenir.

Pour décider des changements nécessaires dans ce champ, l'utilisateur doit uniquement visualiser ce champ et connaître le jeu des valeurs disponibles. Les erreurs dues à un mauvais renseignement ou à une erreur de format des champs peuvent alors être corrigées.

La vérification contextuelle s'exécute pour les champs pour lesquels vous avez sélectionné l'option **Inclure dans la vérification du champ**, dans l'onglet **Vérification** de la boîte de dialogue **Propriétés**.

Pour corriger les caractères incorrectement reconnus, utilisez les modes standard de l'éditeur de texte, comme par exemple le mode d'insertion ou le mode de remplacement. Pour passer d'un mode à l'autre, appuyez sur **INSER**.

Les utilisateurs visualisent les résultats de la reconnaissance de chacun des champs, les corrigent et les confirment en appuyant sur la touche **ENTRÉE** ou en cliquant sur **Confirmer le champ**.

Les champs qui ne correspondent pas au type de données indiqué sont marqués d'un drapeau rouge et un message d'erreur s'affiche les concernant. Les champs pour lesquels vous avez indiqué des règles mais dont les valeurs ne satisfont pas aux exigences sont marqués de la même manière. Vous devez corriger les valeurs de ces champs. Si cela n'est pas possible, retardez la reconnaissance de la valeur du champ en cliquant sur **Pause**.

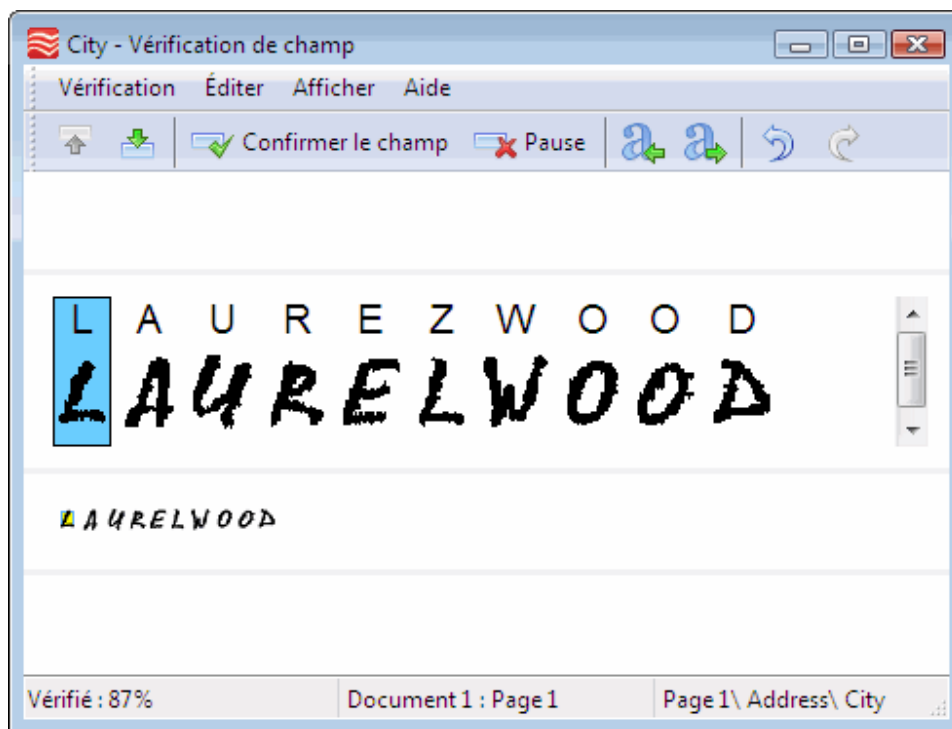




Figure 26. Fenêtre de vérification du champ

La *fenêtre de document* vous permet également de vérifier que la reconnaissance est correcte et de corriger les caractères erronés (Figure 27). Vous ouvrez la fenêtre de document en double-cliquant sur le nom de la page. Cette fenêtre se compose d'une zone de données, de l'image de la page et de la zone des erreurs de la règle (le cas échéant). Vous pouvez configurer l'organisation des fenêtres en vous aidant du bouton **Présentation**. L'opérateur peut visualiser l'intégralité du document dans la fenêtre de document, et pas seulement les groupes de caractères ou les champs séparés.

Dans la zone des données, les caractères incertains sont indiqués en rouge et les champs possédant des types de données erronés ou des erreurs de règle sont mis en surbrillance. Pour passer à l'erreur suivante ou revenir à l'erreur précédente, utilisez les boutons  et . Ces boutons vous permettent de parcourir les erreurs d'assemblage, les caractères incertains, les erreurs de règle, etc.

L'ordre des champs destinés à l'affichage et à la modification peut être séquentiel ou tout autre, selon les besoins de l'utilisateur. Il est possible de modifier cette organisation dans l'Éditeur de modèles de document. Consultez pour cela [Configuration de la vue des données reconnues](#).

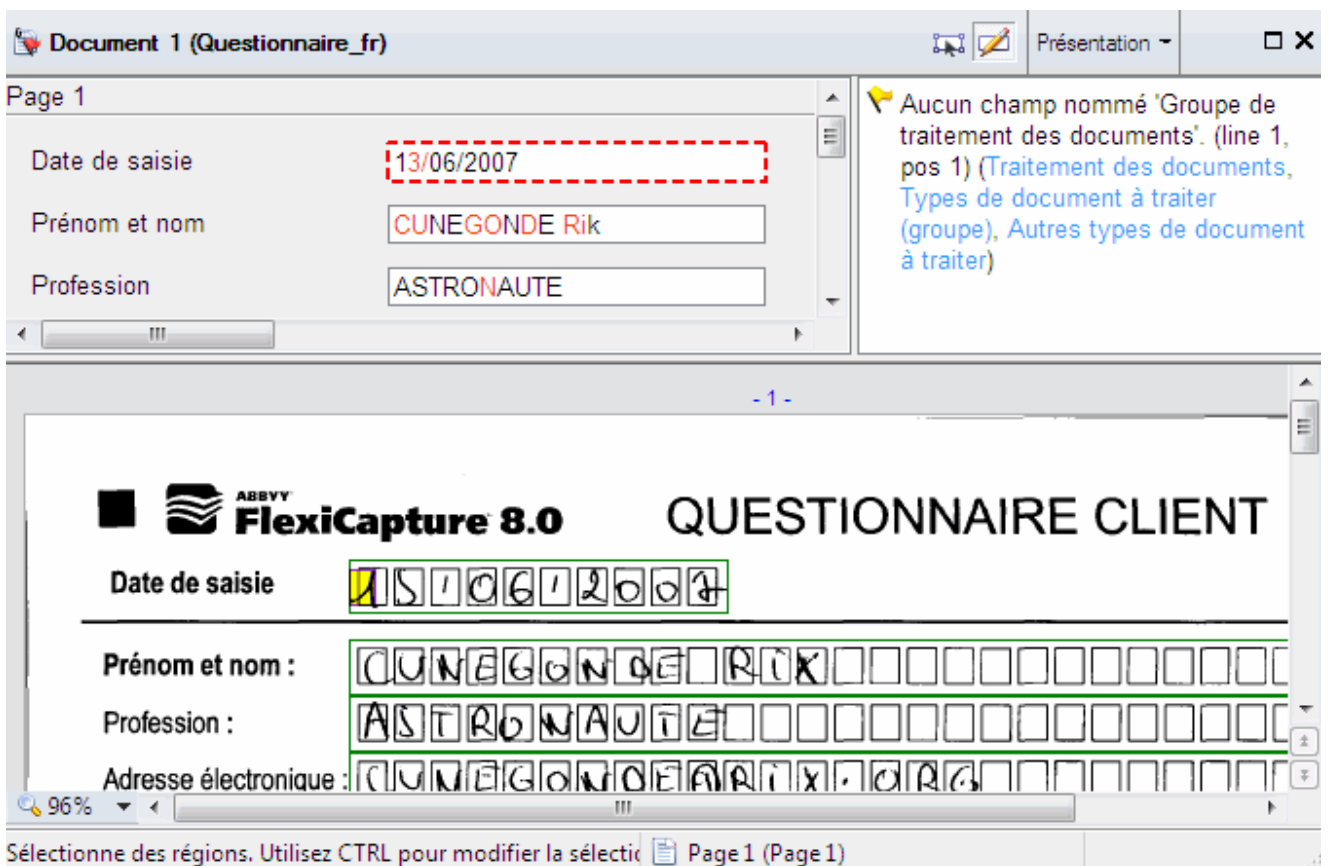


Figure 27. Fenêtre de document

Vérification des règles. Les règles dont les exigences ne sont pas satisfaites sont indiquées d'un drapeau jaune (avertissement) ou d'un drapeau rouge (erreur). Si une règle concerne l'un des champs, ce champ doit être envoyé à l'opérateur pour vérification durant l'étape de vérification contextuelle. Les erreurs de règle s'affichent dans une fenêtre séparée de l'éditeur de document et les documents qui ne satisfont pas aux exigences de la règle sont indiqués par des drapeaux rouges.

Si les exigences d'une règle donnée ne sont pas satisfaites, l'opérateur doit vérifier que les données ont été correctement reconnues et que les erreurs de reconnaissance ont été corrigées. Si l'erreur provient des renseignements fournis et ne peut pas être corrigée, l'opérateur ne doit pas exporter le document.

Pour lancer une nouvelle vérification des erreurs de règle, cliquez sur la flèche située à droite du bouton et sélectionnez **Vérifier à nouveau les règles**. Si les conditions imposées par les règles sont satisfaites après avoir corrigé les valeurs des champs, les drapeaux disparaissent.

1. Cliquez sur Exécuter la vérification... et vérifiez les données reconnues.
2. Observez les caractères et les champs envoyés pour une vérification groupée et contextuelle. Vérifiez-les.
3. Si des champs sont marqués d'un drapeau rouge et qu'un message d'erreur apparaît, veillez à ce que les valeurs de ces champs correspondent au type de données défini.
4. Par exemple, si la date de saisie ne correspond pas au format défini ou se situe en dehors de la période désignée, veillez à ce que les données du champ Date de saisie soient correctement reconnues. Une fois les modifications des données effectuées, le message d'erreur disparaît de la zone correspondante.

7.4. Exportation

Une fois les données reconnues vérifiées, l'opérateur exporte le lot en cliquant sur **Exporter**. L'exportation est effectuée conformément aux paramètres du modèle de document.

Si vous ne souhaitez pas utiliser les paramètres du modèle de document, exportez les données dans un fichier ou une base de données en sélectionnant les éléments correspondants dans le menu déroulant du bouton **Exporter**. En procédant de la sorte, vous pouvez définir les paramètres d'exportation que vous souhaitez.

Exportez les données dans un fichier d'après les paramètres précédemment définis. Pour cela, cliquez sur **Exporter**. Ouvrez le fichier obtenu et analysez les résultats de l'exportation. Envoyez ce fichier à ABBYY par courrier électronique à l'adresse suivante : FlexiCapture_Feedback@abbyy.com.

Nous vous remercions de votre aide et nous vous souhaitons une bonne continuation avec notre logiciel.

8. Conclusion

Cet exemple nous a permis d'aborder toutes les étapes de la configuration du programme et du traitement de documents structurés. Toutefois, les possibilités du programme sont bien plus étendues. Il peut vous aider à traiter des documents multipages simples et complexes, de différents types : documents semi-structurés, non structurés et mixtes. Pour toute question, nous vous invitons à consulter l'aide du programme, ainsi que le guide d'installation.