

## Manuel de référence Module scanner version Sphinx iQ

*Imprimé le 04/11/2011*



### **Le Sphinx Développement**

Parc Altaïs - 74650 Chavanod - Tél. : 04 50 69 82 98 - Fax : 04 50 69 82 78

E-mail : [contact@lesphinx.eu](mailto:contact@lesphinx.eu) - Internet : [www.lesphinx.eu](http://www.lesphinx.eu)

*Copyright © Sphinx Développement 1986 – 2011. Tous droits réservés*



# Sommaire

<b>Sommaire .....</b>	<b>2</b>
<b>Introduction .....</b>	<b>3</b>
1. Description du matériel et des services nécessaires .....	3
2. Quelques questions fréquentes .....	4
<b>Mise en œuvre.....</b>	<b>7</b>
1. Détail du processus.....	7
2. Préparation des documents .....	8
3. Déroulement de la lecture .....	9
4. Réglages du scanner .....	14
5. Fonctions avancées .....	19
<b>Synthèse.....</b>	<b>24</b>
1. Règles de construction.....	24
2. Recommandations pour l'impression et la saisie.....	25
3. Validation et reproduction du formulaire .....	26
4. Spécificités et réglages du scanner .....	27
5. Tester la numérisation (1 <sup>ère</sup> prise en main) .....	28
6. Que faire si... ..	30
<b>Table des matières détaillée .....</b>	<b>34</b>

*Un sommaire détaillé peut être consulté à la fin de ce document.*

# Introduction

## 1. Description du matériel et des services nécessaires

### 1.1. Le scanner

Le scanner doit être un scanner de production, rapide (minimum 10 pages par minute), équipé d'un chargeur fixe de documents et compatible à 100% avec la norme Twain 32 bits.

Les modèles que nous recommandons (au 01/09/2010) sont décrits dans le tableau ci-après.

Modèle	Chargeur	Interface	Recto / verso	Vitesse nominale	Capacité effective	Garantie
Fujitsu FI-6130	A4	USB2.0	Oui	40 ppm *	2 000 ppj *	1 an / 24h
Fujitsu FI-5530C2	A3	USB2.0 et SCSI	Oui	35 ppm *	2 000 ppj *	1 an / 24h

\* ppm = pages par minutes. ppj = pages par jour.

D'autres modèles plus anciens sont compatibles : Fujitsu FI-4110CU, FI-4120C2, FI-5120C.

- Les scanners recommandés ont un chargeur de documents, mais pas de vitre à plat permettant la numérisation de livres.
- La capacité effective est estimée en tenant compte de la capacité du chargeur du scanner et de la capacité d'analyse du logiciel Sphinx. Cette capacité est moindre si des réponses à des questions ouvertes textuelles doivent être saisies par l'opérateur.

Vous pouvez acquérir ces matériels soit par notre intermédiaire soit par votre revendeur habituel.

Ces scanners sont garantis sur site et proposent des extensions de garantie (pour augmenter la durée ou pour réduire le temps d'intervention). Ces extensions ne peuvent s'acquérir qu'au moment de la commande.

Vous pouvez consulter sur Internet les documentations de ces matériels : <http://www.fujitsu-europe.com>

Contactez Sphinx Développement pour vous assurer de la compatibilité d'un autre scanner.

## 1.2. Installation – Formation – Assistance : options conseillées

Le système fait communiquer plusieurs composants : le logiciel Sphinx Plus<sup>2</sup> (conception et lecture), une imprimante, un scanner.

Pour garantir le fonctionnement sans faille de l'ensemble, nous recommandons de nous confier l'installation des matériels et logiciels. Cette installation est complétée par une formation qui permet de réaliser le processus complet d'après une enquête exemple.

En cas de souscription au contrat annuel de maintenance, notre service d'assistance téléphonique, sera ensuite à votre disposition pour répondre à vos questions sur l'utilisation de la solution scanner.

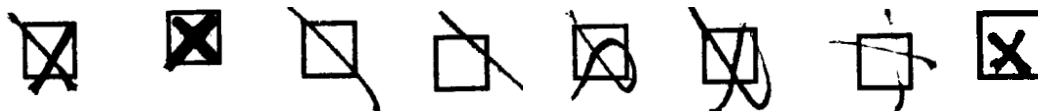
## 2. Quelques questions fréquentes

### Faut-il noircir complètement les cases ?

Non, il suffit de cocher la case. D'ailleurs les cases complètement noircies pourront être considérées comme des ratures.

### Comment coche-t-on une case ? Que se passe-t-il si le trait dépasse de la case ?

La coche n'a pas besoin d'être précise. Voici quelques exemples (agrandis) de coches reconnues :



... et quelques cases non reconnues.



*Par défaut, ceci sera considéré comme une tâche*



*Le trait est trop près de la limite de la case*



*Par défaut, ceci sera considéré comme une rature*



*La machine n'est pas assez intelligente!*

Il est possible de régler le pourcentage de remplissage qui sera considéré comme une coche effective.

### Peut-on utiliser un crayon ?

Oui. Il suffit que le trait soit suffisamment épais pour être repérable par le scanner. La consigne donnée au répondant doit être d'utiliser un stylo à encre noire. Avec les scanners recommandés, le système prend en compte les traits rouges et bleus, ainsi que les crayons à papier (HB).

Les traits de couleur verte seront difficilement reconnus. Il faut aussi éviter les stylos à encre trop claire.

### Peut-on raturer ?

Oui. Si une case est cochée par erreur, il suffit de la noircir intégralement et elle ne sera pas considérée comme cochée (la sensibilité de cette détection de rature peut être réglée).

### Que se passe-t-il si deux cases sont cochées ?

Le module détecte si le nombre de cases cochées est incorrect compte tenu du type de la question. La page correspondante sera repérée en rouge dans la liste des pages lues. Le module prend en compte la coche la plus probable (la plus proche de la coche idéale). La vérification et la modification éventuelle pourront être effectuées manuellement par l'opérateur après la lecture d'un paquet.

### Quelle fiabilité peut-on espérer ?

100% ! Si les questionnaires sont remplis avec un minimum de soin, il n'y aura aucune erreur de reconnaissance.

### Quelle vitesse peut-on espérer ?

La vitesse dépend de la vitesse du scanner et de la puissance de l'ordinateur qui va analyser les images scannées.

Par exemple avec un ordinateur de bureau standard (Pentium IV, 1Go de RAM) et un scanner Fujitsu FI-6130, on aura une vitesse de 300 pages à l'heure (pour une vingtaine de questions sur la page).

### Comment traiter les questions ouvertes ?

Après la lecture d'un paquet de feuilles, vous pourrez consulter les images des zones des questions ouvertes et, par simple clic, vous taperez le texte sans avoir besoin de revenir au questionnaire papier.

### Peut-on "coder" un paquet de feuilles?

Souvent un lot de questionnaires a des caractéristiques communes (n° de service, date, code de formation, ...). Cette caractéristique correspond à une question ouverte dans le questionnaire. Cette question ouverte ne sera pas interprétée automatiquement mais il est possible – en un seul clic - d'attribuer une même valeur à tous les questionnaires du paquet.

### Le questionnaire peut-il être sur plusieurs pages ?

Oui, sans aucun problème, le numéro de la page apparaîtra sous une forme codée dans le bas de la page. Les pages d'un même questionnaire devront être placées à la suite dans le chargeur. L'ordre des pages n'a pas d'importance. Les pages manquantes seront ignorées.

### Peut-on modifier la mise en forme dans un logiciel de PAO ?

Le module reconnaît automatiquement l'emplacement des cases parce qu'il a lui-même lancé l'impression. Si vous remaniez le document, la position de ces cases aura changée et le module ne pourra plus lire la page. Les nombreuses options de mise en page devraient toutefois satisfaire les plus exigeants.

### Peut-on préremplir les questionnaires (avec le nom par exemple) ?

Oui. Les questionnaires peuvent être imprimés avec certains champs préremplis. Lors de la lecture, l'identification du questionnaire par un numéro de client par exemple sera automatique.

### L'installation du scanner est-elle complexe ?

L'installation du scanner se compose de deux phases :

- 1) Relier le scanner à l'ordinateur. La connectique est du type USB.
- 2) Installation logicielle du pilote du scanner. Ce logiciel est fourni par le constructeur du scanner et s'installe simplement.

Le matériel et le logiciel sont à présent convenablement installés et configurés.

Le module configure automatiquement le scanner avec les bonnes options (taille de l'image à scanner, résolution optimale, ...) et aucune intervention n'est nécessaire à ce stade.

Cette installation peut toutefois paraître compliquée et nous proposons à cet effet une solution "clefs en main" comprenant le module de lecture, le scanner et une intervention sur site pour installer le matériel et vous guider dans la découverte du logiciel.

### Quel scanner faut-il choisir ?

Il faut un scanner muni d'un chargeur automatique de documents fonctionnant avec la norme Twain 32 bits. La plupart des grands constructeurs proposent des scanners de ce type adapté à la gestion électronique de documents (GEIDE). Nous avons sélectionné un scanner de milieu de gamme, Fujitsu FI-6130, dont les caractéristiques (40 pages par minute, chargeur de 50 pages) nous paraissent adaptées au traitement des enquêtes Sphinx tout en permettant une utilisation classique pour le scan de documents.

### J'ai un scanner, convient-il ?

Attention ! La plupart des scanners dits "bureautique", que l'on trouve par exemple dans la grande distribution, ne conviennent pas. Si le scanner n'est pas muni d'un chargeur fixe de documents, le bénéfice de la lecture par scanner est complètement contrecarré par le temps nécessaire et la présence constante pour poser le questionnaire sur la vitre puis lancer le scan. Certains scanners permettent l'adjonction d'un chargeur optionnel. Dans de nombreux cas, ce chargeur n'est pas 100% compatible TWAIN et ne fonctionnent qu'avec les applications OCR fournies avec le scanner. Enfin, si ces scanners sont parfaits pour effectuer des images en couleurs avec une excellente résolution, leur vitesse opérationnelle est trop lente pour une application destinée à faire gagner du temps.

# Mise en œuvre

## 1. Détail du processus

Le module de lecture par scanner se présente comme une option du logiciel Sphinx iQ dans lequel il s'intègre totalement. Dès lors que le module est installé, les fonctions "Scanner" sont activées dans les Formulaires personnalisés et au niveau du stade « Diffusion et Collecte » :

- Au stade de la définition des formulaires personnalisés, pour imprimer les éléments nécessaires à la lecture automatique.
- Au stade de la saisie, pour lancer la lecture proprement dite.

Le chargé d'études élaborera sur son poste de travail le questionnaire puis le formulaire avec les options de scanner.

La lecture est éventuellement effectuée avec une autre machine, à laquelle le scanner a été connecté. Sur cette seconde machine, on installera le module opérateur qui permettra de lancer l'opération.

Sur le document imprimé, l'interviewé coche une ou plusieurs cases à chaque question fermée et inscrit une réponse (nombre ou texte) en face de chaque question ouverte.

Les documents sont introduits dans le scanner. Le logiciel fait l'acquisition de l'image de la page. Cette image est ensuite examinée pour retrouver les zones de saisie puis analyser leur contenu.

Le logiciel Sphinx enregistre la position des zones de saisie pour pouvoir les retrouver automatiquement. Malheureusement, la mécanique du scanner ne permet pas d'obtenir une image exactement conforme à l'original : en effet, les feuilles aspirées par l'introducteur se mettent forcément un peu en travers, et la vitesse de défilement n'est pas constante.

Il est donc nécessaire d'imprimer un cadre tout autour du formulaire. Ce cadre sera facilement détecté et le logiciel pourra rectifier l'image. On comprend donc que rien ne doit être inscrit à l'extérieur du cadre.

En outre, le logiciel doit avoir une information sur la page à analyser : s'agit-il d'une feuille correspondant bien à l'enquête ouverte ? Quel est le numéro de la page (si le questionnaire en comporte plusieurs) ? Ces informations, que nous appelons les codes de page, apparaissent sous la forme de cases blanches et noires en bas du document imprimé.

Lorsque la page du questionnaire a été identifiée, commence la détection des réponses. Chaque zone de saisie est analysée. Pour les cases à cocher, on regarde l'intérieur de la case : Est-elle cochée ? Est-elle raturée ? On mesure



pour cela le pourcentage de noir à l'intérieur de la case et on voit bien qu'il faut appliquer des seuils : en-dessous d'une certaine valeur, on considérera qu'il s'agit d'une tâche, au-delà d'une autre valeur, on considérera que la case est raturée, entre ces deux seuils, on considérera que la case est cochée.

Lorsque toutes les questions ont été analysées, le logiciel montre à l'écran les réponses dans la même mise en page que le document imprimé. Pour les questions fermées, les modalités cochées apparaissent avec une croche. Pour les questions ouvertes, c'est l'image scannée de la zone de saisie qui est affichée.

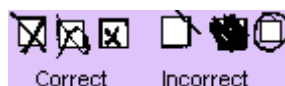
L'utilisateur peut alors interactivement saisir les réponses aux questions ouvertes.

## 2. Préparation des documents

### 2.1. Mise en forme des formulaires

Les fonctions d'élaboration du questionnaire, de mise en page du formulaire et d'impression sont décrites par ailleurs. Dans le cadre de la lecture par scanner, on trouvera page 28, les opérations à faire et à ne pas faire pour assurer une fiabilité optimale.

Si les questionnaires sont remplis par les répondants eux-mêmes, il faut qu'ils soient avertis de la lecture automatique pour les inciter à cocher lisiblement leurs réponses. On pourra à cet effet insérer une image "ConsigneScan.gif" que l'on trouve dans le répertoire Multimédia et qui est reproduite ci-dessous.



### 2.2. Impression et reproduction des formulaires

#### a) Photocopie traditionnelle

Le formulaire est imprimé sur une imprimante laser de bonne qualité (noir et blanc ou couleur) puis reproduite en photocopie.

Il faut faire très attention à la qualité de la photocopie. Par exemple une tâche sur la vitre ou le tambour peut amener à une rupture dans le cadre noir qui entoure la feuille, ce qui peut empêcher la lecture.

Il faut toujours faire un essai avec un formulaire photocopie avant de lancer la reproduction massive.

#### b) Photocopie numérique

Le formulaire est transmis à l'imprimeur sous forme d'un fichier informatique. Le format le plus couramment admis par tous les professionnels de l'imprimerie est le format PDF d'Adobe.

Pour obtenir un fichier au format PDF, il est nécessaire de disposer d'un logiciel permettant la création de tels fichiers. Le logiciel Acrobat d'Adobe en est un. S'il est installé, il suffira d'imprimer le formulaire en choisissant l'imprimante virtuelle Adobe PDF Writer.

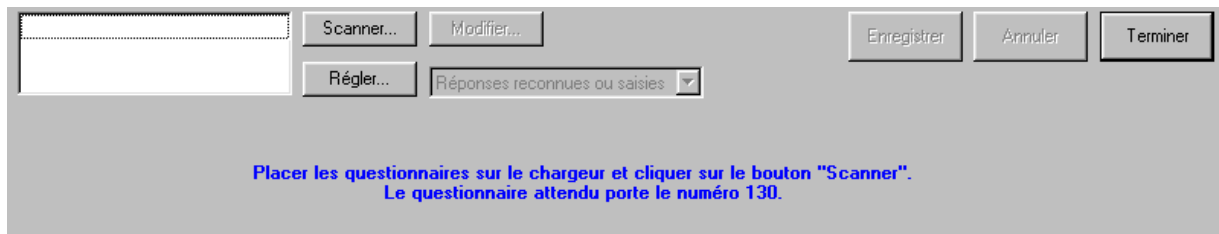
Un fichier reproduisant fidèlement le formulaire sera créé. L'imprimeur pourra l'utiliser pour la reproduction numérique en noir et blanc ou même pour fabriquer les différents films nécessaires à une reproduction en couleurs.

## 3. Déroulement de la lecture

### 3.1. Environnement

Lancer le module de lecture. Dans l'écran principal, choisir l'option **Scanner** qui n'est présente que si le formulaire a été défini comme un formulaire "Scanner".

L'environnement de lecture est présenté dans l'écran suivant :



### 3.2. Lancer la lecture

Cliquer sur le bouton **Scanner** pour lire successivement toutes les feuilles placées dans l'introduit.

Si on ne souhaite lire qu'une seule feuille (par exemple dans une phase de test), enfoncer la touche **CTRL** en cliquant sur le bouton **Scanner**.

Les feuilles sont lues et la liste des pages lues se développe, chaque ligne mentionnant le numéro de l'observation et la page.

### 3.3. Affichage du résultat de la lecture



Après la lecture, la première page lue est affichée à l'écran.

Il ne s'agit pas de l'image mais du questionnaire reconstitué après reconnaissance des cases cochées. Seules les zones de saisie des questions ouvertes sont remplacées par la zone correspondante de l'image scannée.

Si des incohérences ont été détectées, la ligne apparaît en rouge dans la liste et les questions en défaut sont dessinées en rouge. Ces incohérences sont des contrôles de saisie non respectés sur le formulaire. Par exemple, une question déclarée comme obligatoire dans le questionnaire n'a pas été cochée. Ou bien plusieurs coches sont détectées pour une question définie comme fermée unique.

Le module essaye d'opérer une correction intelligente et ce sont les cases les plus probables qui seront prises en compte.

Il est alors possible de faire apparaître l'**Image complète de la page** (raccourci F7) pour lever une ambiguïté ; il arrive que deux cases soient cochées et qu'une "flèche" ou un autre signe indique la bonne réponse, ce que l'ordinateur est bien incapable d'appréhender.

Si plusieurs erreurs sont détectées dans la page, il est possible de faire apparaître les **Diagnostics de lecture** (raccourci F8) pour en avoir la liste détaillée.

Lorsqu'une page avec des erreurs a été visualisée, elle apparaît en magenta dans la liste. On saura que l'on a regardé toutes les ambiguïtés (en effectuant des corrections si nécessaire) dès que la liste ne comporte plus de ligne rouge.

### 3.4. Corriger et compléter

Si des ambiguïtés ont été détectées, une bonne technique est de basculer entre l'**Image complète de la page** (en choisissant l'option dans la liste déroulante ou en utilisant le raccourci F7) et l'affichage des **Réponses reconnues ou saisies** (raccourci F5). Si la lecture est à corriger, il suffit de cliquer sur la bonne case pour modifier la page.

Pour les zones de saisie des questions ouvertes, on voit la portion correspondante de l'image. Il suffit de cliquer pour faire apparaître en dessous une boîte de texte dans laquelle il sera facile de saisir au clavier. Il peut même être intéressant à ce stade de procéder à un début de reformulation pour faciliter l'analyse à venir.

Si l'option de reconnaissance automatique de l'écriture est installée, on basculera entre les **Réponses reconnues ou saisies** (raccourci F5) et l'**Image des zones de saisie** (raccourci F6) pour rectifier le cas échéant les erreurs de lecture.

### 3.5. Erreurs de lecture

L'analyse de l'image du formulaire peut parfois présenter des difficultés et causer une erreur. On s'apercevra alors peut-être que la page est introduite du mauvais côté ou que le formulaire ne correspond pas à l'enquête en cours ou bien que la version du formulaire est incorrecte.

Il y a deux niveaux d'erreurs de lecture :

#### a) Interruption de la lecture

Le message affiché peut être :

- « coin [inférieur/supérieur] [droit/gauche] absent »

Le coin du cadre mentionné est probablement absent de l'image ou trop proche du bord de l'image scannée. Vous pouvez vous en assurer en visualisant l'image. Si ce défaut est systématique, la position et la taille de la zone de scan sont probablement à améliorer. Il est également possible qu'une surcharge placée à proximité d'un coin empêche la reconnaissance de celui-ci.

- « cadre de page déformé »

Les quatre coins du cadre du formulaire ont été détectés correctement, mais l'ensemble de ces quatre coins ne forment pas un rectangle. La feuille est probablement passée en travers dans le chargeur, en butant contre un des montants latéraux, ce qui a provoqué une image très déformée. Repassez la feuille dans le chargeur.

- « lecture de la case x du numéro d'identification impossible »

L'analyse de la zone d'identification du formulaire (zone centrale) dans les codes de pages a rencontré un problème. Le message précise le numéro de la case en défaut dans cette zone (le numéro 1 est la case la plus à droite).

- « lecture de la case x du numéro d'observation impossible », « lecture de la case x du numéro de page impossible ». L'analyse d'une de ces zones dans les codes de page a rencontré un problème. Le message précise le numéro de la case en défaut dans la zone (le numéro 1 est la case la plus à droite).

- « La dernière page lue ne peut être analysée. (numéro de questionnaire attendu : xxx).

Un identificateur valide de formulaire a été reconnu, mais il ne correspond pas à celui qui était attendu. Vérifiez que le formulaire a bien le numéro attendu et qu'il s'agit du formulaire de l'enquête en cours.

- « lecture de la question n°x impossible »

Le cadre du formulaire ainsi que les codes de pages ont été correctement reconnus, mais une question ne peut être analysée. Repassez la feuille. Si la question en défaut est toujours la même, il y a probablement un décalage entre la position des cases attendues et la position des cases détectées pour cette question. Utiliser **Zones de lecture** et visualisez les positions attendues/détectées pour les cases à cocher de cette question.

## b) Avertissements

L'analyse de la page se passe bien, mais des incohérences ou des non-respects de consignes ont été détectées. La couleur de l'observation dans la liste des observations en haut à gauche est alors rouge ou magenta. Par exemple, plusieurs cases ont été cochées pour une question fermée unique, ou trop de cases ont été cochées pour une question fermée multiple.

Vous pourrez visualiser le détail des erreurs détectées par **diagnostics de lecture (F8)**.

Vous trouverez page 29 la liste des obstacles susceptibles d'empêcher le logiciel de lire la page et comment les éviter ou les contourner.

Le mieux est alors de mettre de côté cette page et de relancer la lecture.

Les pages non lues seront repassées à la fin après vérification et éventuellement rectification. Si la lecture s'avère définitivement impossible, l'observation sera saisie manuellement.

Il peut aussi arriver que des erreurs soient détectées dans la lecture d'une question alors que, en observant l'image complète de la page, tout semble normal. En choisissant l'option **Zones de lecture** dans la liste déroulante, on fait apparaître l'image avec pour chaque case à cocher la case attendue (en rouge) et la case détectée (en bleu) assortie du pourcentage de la surface de la case qui a été noircie.

Il faudra peut-être modifier les seuils minimal et maximal admissibles pour ce pourcentage.

### 3.6. Affecter une valeur à toutes les feuilles du paquet

Il est possible de donner une même valeur pour une question donnée à toutes les feuilles du lot.

Prenons un exemple :

- Les questionnaires sont des questionnaires de satisfaction dans un établissement de santé ayant plusieurs dizaines de services.
- L'étude doit permettre de comparer les services entre eux.
- Le service doit donc être une des variables.
- Cette variable n'est pas une question fermée car le nombre de modalités (ici les services) est trop élevé.
- Dans le questionnaire, la variable service est associée à une question de type ouverte codée.
- Cette information est rentrée sous forme manuscrite et le scanner ne pourra pas la lire.

Il n'est pas nécessaire de taper le code du service pour chaque questionnaire. En effet, il est probable que les questionnaires sont classés par service lorsqu'ils sont scannés.

Après chaque lot de feuilles, cliquer sur le bouton **Modifier** puis sur le bouton **Affecter**. Choisir ensuite la variable et indiquer la valeur à attribuer à toutes les feuilles du lot.

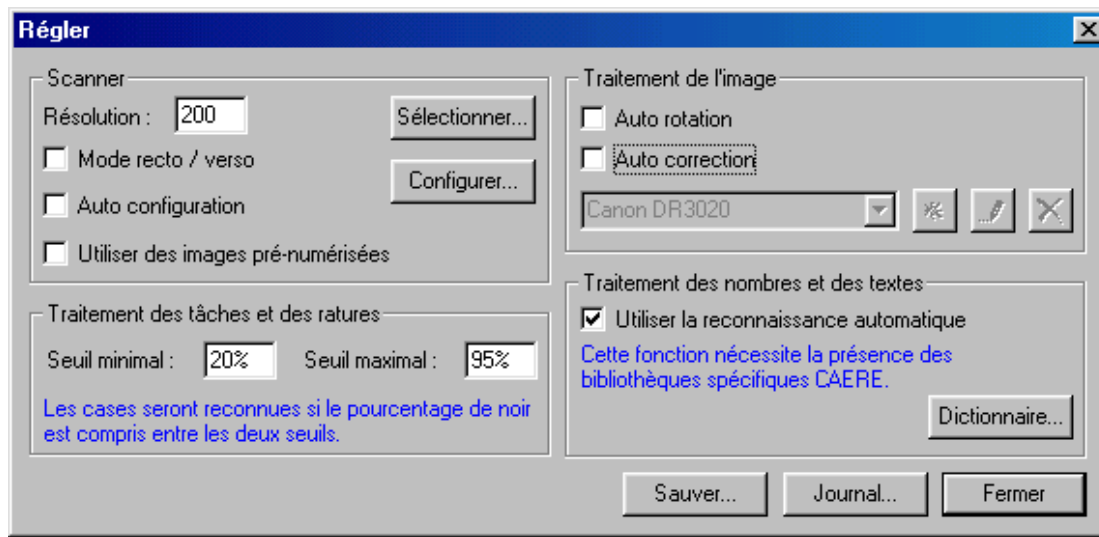
### 3.7. Accepter ou refuser le lot de feuilles

Enfin le lot de feuilles devra être enregistré ou annulé. Dans le premier cas, les observations seront définitivement ajoutées au fichier de l'enquête.

Il est encore possible ici de refuser un lot de feuilles. En cas de bourrage dans le chargeur par exemple, il est difficile de repérer quelles sont les feuilles qui ont été effectivement lues. Plutôt que de procéder à ce repérage, il est souvent préférable de refuser le lot et de le réintroduire dans le chargeur.

## 4. Réglages du scanner

Plusieurs réglages sont nécessaires avant de scanner. Certains sont dépendants de la configuration matérielle utilisée (modèle du scanner), d'autres sont dépendants des particularités du document à analyser. Ces réglages sont enregistrés par le module de lecture.



### 4.1. Réglages de la méthode de scan

#### a) Sélectionner le scanner

Si plusieurs scanners (ou pilotes de scanners) ont été installés sur la machine, le bouton **Sélectionner** permet de choisir le scanner à utiliser.

#### b) Résolution

200 (points par pouce) est un bon compromis entre la rapidité de l'acquisition de l'image et sa résolution. Dans certains cas, il peut être nécessaire de pousser cette résolution à 300 (par exemple si les traits des cases sont très fins et peu marqués).

#### c) Recto-Verso

Les scanners que nous recommandons peuvent scanner en recto-verso.

#### d) Auto-configuration

Lorsque cette case est cochée, le logiciel positionne automatiquement les réglages du scanner : taille de la feuille de papier, sélection du chargeur, etc. On peut désactiver cette auto-configuration pour gérer certains scanners particuliers, comme par exemple pour faire des tests sur un scanner non muni d'un chargeur de documents.

### e) Utiliser des images pré-numérisées

On peut utiliser cette fonction si on dispose d'un ensemble ordonné d'images de feuilles déjà scannées (par exemple scan0.tif, scan1.tif, scan2.tif, scan3.tif ou bien test10.bmp, test11.bmp, ...).

Les images peuvent être au format bitmap non compressé (.bmp) ou compressé (.tif). Dans ce cas, lorsque le bouton scanner est activé, le chemin du premier fichier image de la série est demandé. Un numéro est présent à la fin de chaque fichier. Les images seront analysées tant que le fichier avec le numéro suivant sera trouvé.

### f) Configurer le scanner

Configurer le scanner amène au dialogue de réglages fourni par le constructeur du scanner ou de son pilote. C'est dans cet environnement que l'on fera les réglages nécessaires pour augmenter ou diminuer le contraste et la luminosité afin d'obtenir l'image la plus lisible possible. Si les modifications effectuées ne sont pas prises en compte, faire les réglages en utilisant une application fournie avec le scanner (par exemple ScandAllPro). Si le questionnaire a été imprimé en couleurs, il sera peut-être nécessaire de programmer le traitement des trois couleurs de base pour ne retenir que les éléments utiles (par exemple ignorer le fond de la case si celui-ci a été imprimé en couleurs).

Certains scanners demandent des réglages particuliers pour obtenir l'image adéquate. Vous trouverez dans le dernier chapitre de ce manuel le détail des configurations à appliquer pour la plupart des scanners utilisés.

## 4.2. Traitement de l'image

Après la phase d'acquisition de l'image par le scanner, il peut être nécessaire de lui appliquer certains traitements avant de l'analyser.

### a) Auto-rotation

Le logiciel suppose par défaut que la feuille a été introduite dans le sens naturel, c'est-à-dire le haut de la feuille en premier. Si les éléments nécessaires ne sont pas détectés (voir les explications sur les codes de page plus loin), une erreur sera signalée. En cochant la case auto-rotation, on indique au logiciel d'essayer de réorienter automatiquement l'image si besoin est.

Cette fonction est très utile dans les cas suivants :

- Les formulaires ne sont pas tous rangés dans le même sens et on veut s'épargner la tâche de les remettre tous à l'identique.
- On veut pouvoir délibérément introduire les feuilles à l'envers pour faciliter leur introduction : par exemple si les feuilles ont été agrafées, la déformation des pages lors de l'agrafage peut gêner l'introducteur qui n'arrivera pas à séparer les différentes feuilles d'un même questionnaire.
- Les formulaires ont été réalisés au format paysage et il n'y a pas de sens "évident" pour le placement des questionnaires sur l'introducteur. On laissera donc le système déterminer le bon sens.



## b) Corriger les déformations dues au scanner

Les scanners recommandés récents n'ont pas besoin d'une correction.

Pour les scanners plus anciens (Fujitsu 3091, 3092, 4110), une correction est nécessaire. En effet ils utilisent un moteur pour faire défiler les feuilles. Sur certains scanners, le défilement vertical n'a pas une vitesse constante et il est nécessaire d'appliquer une correction lors de la détection des cases.

Un jeu de corrections prédéfinies est proposé pour les scanners que nous recommandons, choisir la ligne correspondante dans la liste déroulante.

Il est possible d'étalonner son scanner s'il n'apparaît pas dans la liste ou même de ré étalonner un scanner si une utilisation intensive ou un traumatisme a modifié son comportement ou bien encore si le document est imprimé sur un support spécial (papier très épais par exemple) qui modifie le passage de celui-ci à travers la mécanique du chargeur.

Cliquer sur le bouton **Nouvelle correction** et donner un nom à ce nouvel élément.

Cliquer sur le bouton **Modifier** pour **Imprimer, Scanner**, puis **Rectifier** la page d'étalonnage.

Pour vérifier si le scanner est correctement étalonné, il suffit d'examiner les **Zones de lecture** (option de la liste déroulante de l'environnement de lecture), les cases attendues doivent être presque exactement superposées aux images des cases. On peut toutefois tolérer un décalage de l'ordre de  $\frac{1}{4}$  de la taille de la case entre la position de la case attendue et celle de la case détectée.

### 4.3. Traitement des tâches et des ratures

Lorsqu'une case à cocher est analysée, le pourcentage de remplissage de celle-ci est calculé. Il sert à évaluer si la case est non cochée, cochée ou raturée.

Les deux seuils de reconnaissance, seuil minimal et seuil maximal sont exprimés en pourcentage (0 à 100%).

En dessous du seuil minimal, la case sera considérée comme non cochée.

Les valeurs normales pour le seuil minimal sont de 15% à 25 % (par défaut 20%). Celui-ci permet d'absorber les bavures provoquées par l'imprimante ou les petites tâches dues au répondeur ou à la phase de reproduction.

Entre le seuil minimal et le seuil maximal, la case sera considérée comme cochée.

Au-delà du seuil maximal, la case sera considérée comme raturée, c'est à dire initialement cochée puis totalement noircie afin d'invalidier le choix. Les valeurs normales pour le seuil maximal sont de 85% à 100% (par défaut 95%).

Deux méthodes de remplissage des cases peuvent être utilisées :

- La première (à utiliser préférentiellement) consiste à utiliser une croix pour signifier "cochée" et à noircir totalement pour signifier "raturée". Le seuil maximal (environ 95%) sert alors à faire la différence entre une case cochée et une case raturée.
- La seconde est de répondre "cochée" en noircissant complètement la case. Le seuil maximal est réglé à 100%. Il n'est alors pas possible de faire la distinction entre une case normalement cochée et une case raturée.

Quel que soit la méthode utilisée pour le remplissage des cases, il est indispensable que tous les répondants utilisent la même. Il est sage de décrire au début du formulaire quelle est la méthode à utiliser pour cocher les cases.

Si, par exemple, tous les documents ont été cochés avec un trait très fin ou très épais, les seuils définis ont peut-être besoin d'être ajustés. Après le scan d'une page, l'option **Zones de lecture** permet de voir les pourcentages repérés pour chaque case, et donc d'en déduire les valeurs convenables pour les seuils de reconnaissance.

#### 4.4. Configurations spécifiques des scanners

Pour les différents scanners que nous recommandons, voici les choix à effectuer dans les différents dialogues.

##### a) Scanners FUJITSU FI4110, FI4120

###### Dialogue Régler

Cocher la case **Mode recto / verso** pour scanner le document en mode recto/verso.

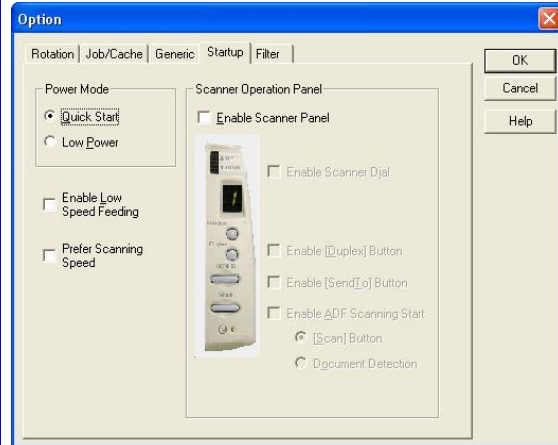
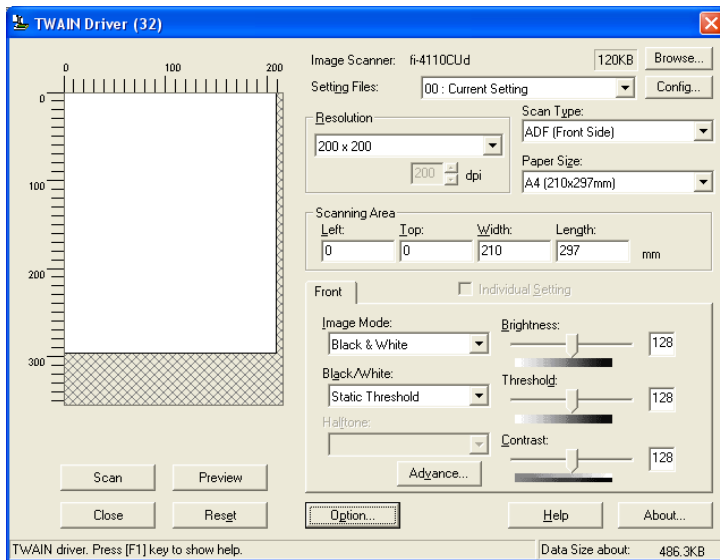
Cocher la case **Auto-configuration**.

Cocher la case **Auto-correction** et choisir la ligne

###### Dialogue Sélectionner

Sélectionner la ligne

###### Dialogue Configurer



## b) Autres scanners

### Dialogue Régler

Décocher la case **Auto-correction**.

Décocher la case **Auto-configuration**.

### Dialogue Sélectionner

Si le scanner n'apparaît pas dans le dialogue **Sélectionner**, c'est qu'il n'est pas compatible Twain et vous ne pourrez pas l'utiliser pour la lecture automatique.

### Dialogue Configurer

Chaque scanner a ses propres caractéristiques. Les éléments à paramétrer sont

- Le pilotage du chargeur de document
- La définition de la zone de scan
- Le réglage du contraste et de la luminosité

## 5. Fonctions avancées

### 5.1. Enregistrement

L'enregistrement des réponses est effectué dès que l'opérateur clique sur le bouton **Enregistrer**. Il est également possible de garder la trace des images produites, des réponses du seul dernier paquet scanné et de tenir un journal.

#### a) Enregistrement des images



Les images sont stockées temporairement dans un répertoire indiqué dans le dialogue **Régler / Sauver**. Si aucun répertoire n'est mentionné, c'est le répertoire **Scanner** du répertoire de l'enquête en cours qui sera utilisé et éventuellement créé.

Ces images sont conservées pour permettre la visualisation de l'image scannée à des fins de vérification et pour permettre la saisie des questions ouvertes en voyant la zone concernée de l'image.

Dans une optique "zéro papier", on peut enregistrer définitivement ces images. Attention toutefois à la place requise. Une page au format A4 avec une résolution de 200dpi occupe environ 500 Ko sur le disque dur.

#### b) Enregistrement du dernier paquet

Une option permet d'enregistrer toujours le dernier paquet scanné non seulement dans l'enquête en cours mais également dans une autre enquête dont le nom est indiqué. L'objet de cette option est de permettre d'obtenir un dépouillement immédiat du dernier paquet de feuilles scannées sans avoir à recourir à des strates.

#### c) Tenir un journal des opérations

Le journal contient, à des fins de vérification, toutes les opérations qui ont été effectuées. Le bouton **Journal** des opérations permet de consulter le journal, de le réinitialiser ou d'arrêter sa tenue.

## 5.2. Codes de page

Propriétés du formulaire

Utilisation | Taille | Fonctionnement

Nom du formulaire :  
Formulaire papier n°1

Utilisation : Papier  Lecture par scanner

Pour permettre la lecture automatique des réponses par scanner, des éléments (un cadre et des identificateurs) doivent être imprimés.

Marges spéciales

Gauche : 0 cm Droite : 0 cm  
Haut : 0 cm Bas : 0 cm

Certains scanners ne permettent pas une reconnaissance sur la totalité de la page. Indiquer le cas échéant des marges supplémentaires.

Utilisation des codes de pages

Identification spéciale des observations :

Lorsque des questionnaires préremplis sont scannés, l'identification de l'observation peut être son numéro ou une des variables numériques.

Numérotation manuelle du questionnaire 128

OK Annuler

Lorsque le formulaire est imprimé, le système lui attribue automatiquement un numéro (qui avance de 1 à chaque nouvelle impression lorsque la mise en page a changé). Ce numéro est enregistré dans le formulaire et est inscrit sur le document imprimé sous la forme de carrés noirs et blancs, les codes de page. Le programme vérifiera ainsi que la feuille qui a été introduite dans le scanner correspond effectivement au formulaire actif de l'enquête ouverte.

Lorsque le formulaire comporte plusieurs pages, un autre jeu de codes de page, sur la droite de la feuille, permet au programme de différencier les différentes pages du formulaire.

Enfin, lorsque les formulaires sont pré remplis avec certaines informations connues pour chaque répondant, l'identification de l'observation est représentée sous la forme d'un troisième jeu de codes de pages.

Observation n°1

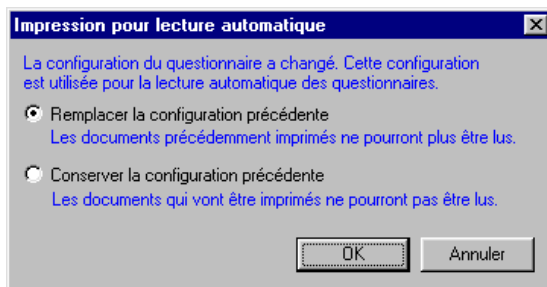
130

Page 2 / 2

On peut toutefois se passer des codes de pages dans le cas où le formulaire ne comporterait qu'une seule page et que l'utilisateur s'assure lui-même de la conformité de la page lue et du formulaire attendu. L'esthétique du formulaire sera améliorée avec toutefois un risque d'erreur.

Par exemple, si la feuille est introduite à l'envers (le bas de la feuille en premier), le cadre de la page sera bien détecté mais les codes de page ne seront pas reconnus. Une erreur sera donc signalée. Si les codes de page sont absents, le système ne pourra pas déceler l'erreur d'introduction et les détections de coches seront erronées.

On a vu que la numérotation est incrémentée automatiquement à chaque impression différente du formulaire. Le système demande alors confirmation de la validité de la nouvelle mise en page parce que les anciens formulaires ne pourront plus être lus.



Malgré cette demande de confirmation, il peut arriver que par mégarde, le questionnaire soit réimprimé avec une mise en page différente et donc avec un nouveau numéro. Les anciens formulaires ne pourront plus être lus... sauf si l'on arrive à remettre exactement les mêmes options que sur l'ancien formulaire. Alors, à ses risques et périls, on cochera la case **Numérotation manuelle du questionnaire** en tapant le numéro qui est imprimé sur les anciens documents.

### 5.3. Cas particuliers

#### a) Marges spéciales

Certains scanners "mangent" le bord de la feuille et le cadre, s'il est placé trop près du bord, ne sera pas scanné et la page ne pourra pas être reconnue. Pour pallier cette déficience mécanique de certains appareils, il est possible de forcer une réduction de la zone d'impression par des marges spéciales.

Le problème n'existe pas avec les scanners recommandés et ces marges spéciales pourront être ignorées.

#### b) Scanner au format "paysage"

Si le formulaire a été défini en format « paysage », Il faut introduire le formulaire dans le chargeur du scanner comme s'il était au format « portrait ». Le logiciel fait alors automatiquement une rotation de 90° de l'image obtenue avant de l'analyser.

Il est également possible d'utiliser la fonction auto-rotation en cochant la case auto-rotation dans le dialogue **Régler**, pour autoriser une recherche automatique de la bonne position de l'image. Il est préférable, si possible, d'éviter utiliser cette fonction, car elle augmente le temps d'analyse de l'image du formulaire.

## 5.4. Rectification de la position des cases

Si des erreurs de lecture se produisent, ce peut être parce que des déformations de l'image ont eu lieu lors de la reproduction. Il est possible de rectifier manuellement la position attendue des cases.

On fait apparaître les **Zones de lecture** (dans la liste déroulante à côté du bouton Régler) puis on choisit dans la liste déroulante l'option **Modifier la position attendue des cases**. Pour chaque case, on utilisera la souris pour déplacer la case attendue (en bleu) jusqu'à une complète superposition avec l'image de la case.

Cette rectification sera conservée pour toutes les lectures à venir tant que le document ne sera pas réimprimé.

## 5.5. Traitement des questionnaires nominatifs

Un protocole de questionnement souvent utilisé consiste à envoyer des questionnaires nominatifs pour pouvoir ensuite confronter les informations reçues à d'autres informations présentes dans une base de données ou un panel.

### a) Identification préalable des observations

Les individus à interroger sont importés dans une nouvelle enquête. Les questions socio-démographiques sont donc présentes pour chacun des individus. On peut donc, sur le formulaire, faire figurer certaines des informations connues, par exemple le nom de la personne ou son numéro de dossier.

La seconde étape consiste à imprimer l'ensemble des formulaires. Dans l'environnement **Consulter / modifier**, on pourra imprimer l'ensemble des formulaires.

Sur ces formulaires, des codes de page permettront par la suite une identification de l'observation. Cette identification peut être faite par le numéro d'enregistrement de l'individu mais il ne sera pas possible de modifier l'ordre des enregistrements ou de supprimer un enregistrement sans quoi toutes les observations seraient décalées.

Une possibilité est d'avoir dans le questionnaire une variable numérique identifiant chaque observation. En cochant la case **Identification spéciale des observations** et en choisissant cette variable dans la liste, les codes de page représenteront la valeur de cette variable (à charge pour le chargé d'études de s'assurer de l'unicité de cette variable).

Lorsque les formulaires seront lus, le système détectera automatiquement l'observation correspondante qu'il ira documenter avec les informations lues sur le formulaire.

### b) Identification a posteriori des observations

Les questionnaires sont reproduits à l'atelier d'imprimerie qui dispose d'un système pour imprimer une information différente sur chaque exemplaire sous forme d'un tampon. Le chargé d'études communique à l'atelier la liste des noms/numéros à faire figurer sur chaque questionnaire.

Le module de lecture permet de traiter de tels questionnaires grâce à la lecture automatique des nombres imprimés (voir plus loin les modalités d'utilisation de cette technologie).

Une variable correspondant au numéro de dossier sera créée dans le questionnaire et positionnée sur le formulaire. C'est dans la zone de saisie correspondant à cette question que l'imprimeur devra faire apparaître les numéros. On veillera à ce que l'imprimeur utilise une police de caractères "facile" à reconnaître (pas de pleins et de déliés, police classique, taille raisonnable, etc.).

Lorsque les questionnaires reviennent, ils sont lus automatiquement (numéro de dossier et variables de l'enquête proprement dite).

Cette enquête est ensuite fusionnée avec la base de données contenant les informations correspondant à chaque numéro de dossier. Le chargé d'études est désormais prêt à confronter les réponses aux questions de l'enquête avec les informations connues sur chaque individu.

## 5.6. Que faire pour...

### a) Lire une date

Plusieurs méthodes peuvent être utilisées :

Avec des cases à cocher : Cette méthode offre la meilleure fiabilité en terme de reconnaissance.

Un exemple complet d'une étude utilisant ce codage d'une date est disponible dans les exemples livrés avec le logiciel, en **\Enquêtes\Exemples\Lecture par scanner\Dates\Date cases a cocher**.

Trois variables de type échelle JOUR, MOIS, ANNEE sont définies. Une variable V\_Date de type date concatène ces trois variables et calcule la date correspondante.

Après la saisie au scanner, il faut simplement mettre à jour cette variable V\_DATE en la recalculant.

### b) Utiliser les services d'un imprimeur

La meilleure solution est de transmettre à l'imprimeur un fichier informatique représentant le formulaire élaboré. Pour cela, vous devez disposer sur votre machine du logiciel **Adobe Acrobat Writer**. Lorsque vous imprimez le formulaire, un item supplémentaire **Acrobat Writer** apparaît dans la liste des imprimantes accessibles depuis votre machine. Un fichier .pdf est alors créé. Il vous suffit de le transmettre à l'imprimeur.

L'autre solution est d'imprimer un original de très bonne qualité et de le faire reproduire chez un imprimeur. Cette solution présente l'inconvénient d'une perte éventuelle du contraste ou de la qualité des couleurs, ainsi que d'une légère modification de la géométrie du document.



# Synthèse

## 1. Règles de construction

Voici un ensemble de règles à observer pour une utilisation satisfaisante du module de lecture par scanner.

### 1.1. Créer son formulaire scanner :

Après avoir créé les questions dans la partie « Rédiger mon questionnaire »,  
Au stade « Rédiger mon questionnaire », cliquez sur « Formulaires personnalisés... »,  
Vous arrivez dans l'assistant de création des formulaires personnalisés,  
Cliquez sur le bouton « Saisir » situé dans l'onglet « Formulaire » puis « Suivant »,  
Terminez l'assistant.

### 1.2. Règles à suivre concernant la mise en page du formulaire

Voici quelques Règles à suivre sur les Zones de lecture (Code de page, case à cocher, zone à reconnaissance optique, écriture manuscrite ou impression par une machine) pour un fonctionnement optimal de votre formulaire scanner :

- Veiller à toujours conserver un espace libre minimum de 1mm autour de ces zones,
- Ne pas mettre de couleur de fond,
- Utiliser du noir pour la couleur de trait,
- Ne pas mettre un effet ombré,

Autres règles à suivre :

- Vérifier que chaque question est présente au maximum une seule fois sur le formulaire,
- Vérifier que les questions ne sont pas superposées entre elles,
- Vérifier qu'il n'y a pas d'objets du formulaire qui dépassent du cadre.

## 2. Recommandations pour l'impression et la saisie

### 2.1. Impression du formulaire

Toujours utiliser le même ordinateur et la même imprimante pendant la phase de mise au point du formulaire. En effet, il peut y avoir des différences minimales entre les deux systèmes d'exploitation ou entre les jeux de polices utilisés par les imprimantes. Ces différences vont amener des changements dans la mise en page et donc dans les coordonnées des cases.

### 2.2. Reproduction du formulaire

La qualité de la reproduction du formulaire est très importante. Il est conseillé d'utiliser un matériel de type professionnel en veillant au respect des points suivants :

- La marge autour du cadre doit être de l'ordre de 5 mm tout autour de celui-ci.
- Le document doit être le plus fidèle possible à l'original (pas de distorsion de l'image).
- Le document doit être exempt d'annotations, de taches ou de bavures, notamment à l'extérieur du cadre, de la zone des codes de pages et au niveau des zones de lecture automatique. Si, toutefois, des annotations manuelles doivent être rajoutées pour les besoins propres de la personne qui anime l'enquête, celles-ci devront être à l'intérieur du cadre et hors des zones de lecture automatique.
- Attention au pliage du document : si le document est destiné à être plié (par exemple pour être mis dans une enveloppe), il est préférable de faire réaliser par l'imprimeur un pré-rainurage sur les lignes de pliage. Il faut vérifier que les lignes de pliage ne passent pas par une case à cocher.

### 2.3. Remplissage par les enquêteurs ou les interviewés

Deux cas se présentent : soit la personne qui remplit le document est un enquêteur à qui on peut avoir enseigné quelques règles de base pour la saisie des réponses, soit c'est l'interviewé. Le deuxième cas laisse la porte ouverte à tous types de fantaisie et peut éventuellement perturber la bonne reconnaissance des zones.

Il faut toutefois essayer d'appliquer les règles suivantes :

- Utiliser un stylo noir, ni trop fin, ni trop épais.
- Pour les cases à cocher :
  - Faire des croix lisibles en les centrant et sans trop dépasser.
  - Si on a coché par erreur une case, il faut la noircir complètement.

## 2.4. Lecture automatique par scanner

Préparer les documents :

- Enlever les agrafes.
- Si le formulaire est sur plusieurs pages, regrouper les pages d'une observation. Celles-ci peuvent être dans le désordre pour une observation donnée.
- Si les pages ont été pliées ou cornées pendant l'enquête, les mettre sous une presse afin qu'elles retrouvent leur géométrie initiale.
- Effeuille le tas de pages pour éviter le collage entre elles.

Mettre les pages dans le chargeur de documents du scanner. La position par défaut des documents dans le chargeur sera la position "normale" d'introduction (exemple pour le scanner Fujitsu FI-6130 : la tête du document en bas dans le chargeur).

## 3. Validation et reproduction du formulaire

### 3.1. Validation du formulaire avant l'impression finale :

- 1) La mise en page terminée,
- 2) Imprimez le formulaire en plusieurs exemplaires, (env 10)
- 3) Après l'impression effectuez un enregistrement,
- 4) Cochez des cases sur les impressions obtenues et réalisez des tests de récupération via le scanner,
- 5) Si récupération correcte, vous pouvez lancer l'impression finale (via un imprimeur par exemple).
- 6) Dans le cas contraire, il faut comprendre l'origine du problème et apporter la correction appropriée. Recommencez ensuite à l'étape 3.

### 3.2. Transmettre le formulaire finalisé à un service de reproduction :

La phase de reproduction du formulaire avant sa diffusion doit être faite dans un souci de qualité.

Le formulaire peut être transmis au service de reproduction sous deux formes :

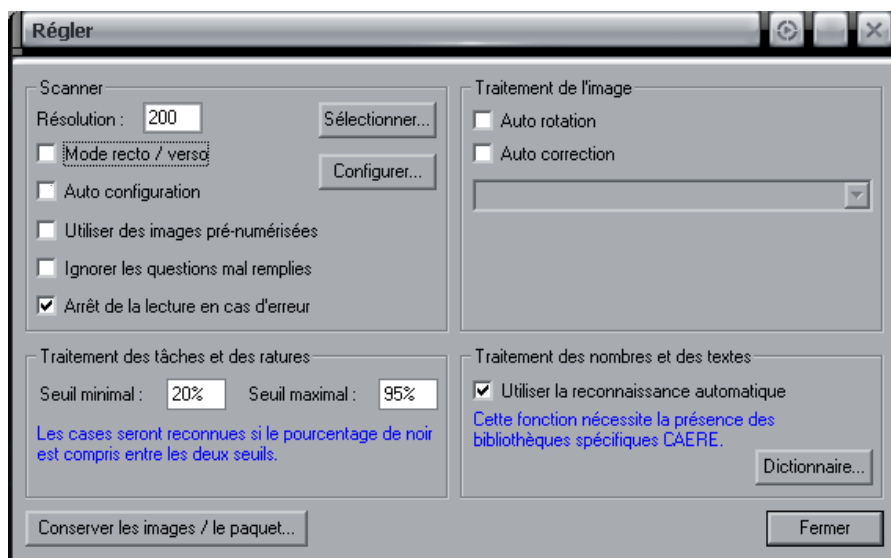
- Un document papier original, qui est le résultat d'une impression sur une imprimante laser de bonne qualité. C'est une solution valable pour les formulaires en noir et blanc.
- Un fichier informatique au format Adobe Acrobat (.pdf) accompagné d'une impression papier. Dans ce cas, il faut posséder le logiciel Adobe Acrobat Writer. C'est la solution la plus intéressante car c'est un fichier qui est transmis à l'imprimeur, donc sans nouvelle étape de numérisation du document imprimé. En outre, elle permet la séparation des couleurs pour une reproduction en polychromie.

## 4. Spécificités et réglages du scanner

### 4.1. Réglage conseillé des options du scanner sous Sphinx :

Procédure pour accéder à cette fenêtre :

- Au niveau des formulaires personnalisé, cliquez sur l'icône « Saisir » situé dans l'onglet « Formulaire », ou à partir du deuxième stade « Diffusion et Collecte » en cliquant sur « Acquisition scanner ».
- Vous arrivez sur une interface grise, cliquez alors sur le bouton « Régler »,
- Vérifiez que vos options correspondent à l'exemple ci-dessous :



- Ceci est une configuration conseillée, mais bien sur pas obligatoire.

### 4.2. Scanner de bureau ou Photocopieur réseau, comment faire ?

**Dans le cas d'une utilisation d'un scanner de bureau** (scanner branché directement sur le poste sur lequel est installé le Sphinx via un câble USB ou SCSI), il vous suffit de placer les formulaires sur le scanner puis à partir de Sphinx, de lancer l'option « Scanner » afin de commencer la numérisation. Le Sphinx ira directement lire les fichiers placés sur le scanner.

**Dans le cas d'un photocopieur réseau**, la procédure n'est pas la même, en effet, vous devez scanner au préalable les formulaires via votre photocopieur afin de générer des images (en BMP ou TIFF). Ces images seront ensuite placées dans un dossier réseau par exemple, puis il ne vous restera plus qu'à scanner ces images obtenues via Sphinx (En utilisant l'option « Utiliser des images pré numérisées »).

### 4.3. Pré-requis sur les images pré-numérisées :

Les images que vous allez utiliser dans le cas de l'utilisation d'un photocopieur, doivent suivre les conditions suivantes :

Format des images reconnues	BMP ou TIFF
Résolution de l'image	200 dpi
Couleur de l'image	Noir et Blanc
Contenu des images	1 image par page <i>Exemple : 1 recto/verso = 2 images</i>

Enfin, il est conseillé de nommer les images créées avec le même nom et une incrémentation à la fin :

Exemple : « image001 », « image002 », etc...

De cette manière, le Sphinx récupèrera toutes les images ayant le même nom lors du scan.

## 5. Tester la numérisation (1<sup>ère</sup> prise en main)

### 5.1. Avec un scanner branché directement sur votre ordinateur :

- Lancez le logiciel Sphinx,
- Vous pouvez soit ouvrir votre propre étude soit utiliser l'enquête « Automobiles » proposée à titre de test,
  - o Pour ouvrir l'enquête « Automobiles », cliquez sur « Exemples » situé au niveau du deuxième stade « Projet Sphinx Ouvrir »,
  - o Ouvrez le dossier « Automobiles » dans la liste proposée, puis cliquez sur « automobiles.sphx » pour ouvrir l'enquête.
- Au premier stade « Conception du questionnaire », cliquez sur « Rédiger mon questionnaire »,
- Une fois dans cet environnement, cliquez sur « Formulaires personnalisés... » puis « Lancer le mode personnalisé »
- Vous devez arriver sur votre formulaire scanner (dans le cas contraire, si vous avez plusieurs formulaires, assurez-vous de bien avoir à l'écran le formulaire scanner),
- Cliquez sur le menu « Fichier » puis « Imprimer »,
- Une fenêtre peut alors apparaître, « Impression pour lecture automatique ». Si c'est le cas, cochez l'option « Remplacer la configuration précédente » puis validez avec « Ok »,
- Une autre fenêtre peut ensuite se présenter, « Analyse du formulaire », ne vous en occupez pas, cliquez simplement sur « Continuer »,
- Imprimez plusieurs formulaires.

- Votre impression réalisée, faites au stylo quelques coches afin de simuler une saisie puis placez les documents renseignés sur votre scanner.
- Attention : au sens des feuilles :
- Les documents placés dans le bac doivent être la tête en bas et le verso en face de vous (le recto se retrouve donc contre le bac du scanner).
- Toujours au niveau des formulaires personnalisés, cliquez sur l'icône « Saisir » située dans l'onglet « Formulaire »,
- Vous arrivez sur une page grise,
- Cliquez sur le bouton « Scanner » pour lancer la récupération.

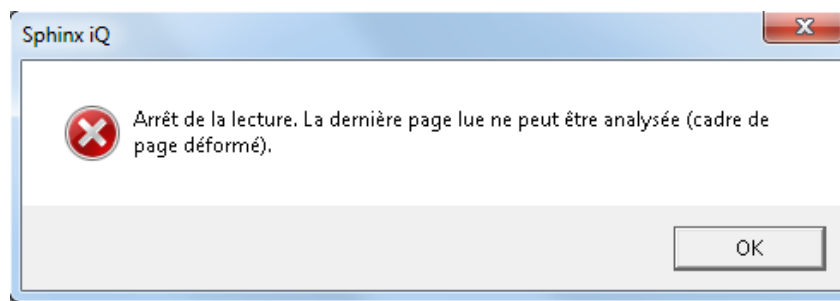
## 5.2. Avec un scanner réseau (photocopieur) :

- Lancez le logiciel Sphinx,
  - o Vous pouvez soit ouvrir votre propre étude soit utiliser l'enquête « Automobiles » proposée à titre de test,
  - o Pour ouvrir l'enquête « Automobiles », cliquez sur « Exemples » situé au niveau du deuxième stade « Projet Sphinx Ouvrir »,
- Ouvrez le dossier « Automobiles » dans la liste proposée, puis cliquez sur « automobiles.sphx » pour ouvrir l'enquête.
- Au premier stade « Conception du questionnaire », cliquez sur « Rédiger mon questionnaire »,
- Une fois dans cet environnement, cliquez sur « Formulaires personnalisés... » puis « Lancer le mode personnalisé »
- Vous devez arriver sur votre formulaire scanner (dans le cas contraire, si vous avez plusieurs formulaires, assurez-vous de bien avoir à l'écran le formulaire scanner),
- Cliquez sur le menu « Fichier » puis « Imprimer »,
- Une fenêtre peut alors apparaître, « Impression pour lecture automatique ». Si c'est le cas, cochez l'option « Remplacer la configuration précédente » puis validez avec « Ok »,
- Une autre fenêtre peut ensuite se présenter, « Analyse du formulaire », ne vous en occupez pas, cliquez simplement sur « Continuer »,
- Imprimez plusieurs formulaires.
- Votre impression réalisée, faites au stylo quelques coches afin de simuler une saisie puis placez les documents renseignés sur votre scanner.
- Sous ce type de configuration, vous ne pouvez pas scanner directement vos formulaires dans le Sphinx (puisque votre scanner est en réseau). Vous devez tout d'abord réaliser des images pré numérisées.
- Pour cela, vous devez scanner vos documents via votre scanner de manière à produire des images au format BMP ou TIF (Concernant la procédure, veuillez vous référer à la documentation de votre scanner ou à votre service informatique),
- Ces images doivent être enregistrées dans un dossier accessible depuis le poste où est installé le Sphinx,
- Vos images doivent suivre certaines conditions (voir chapitre 4.3)
- Vos images enregistrées, revenez sur le poste avec le Sphinx,

- Ouvrez votre enquête, puis au stade « Diffusion et collecte », cliquez sur « Acquisition scanner », pour revenir sur le formulaire scanner que vous aviez imprimé,
- Si vous avez plusieurs formulaires, assurez-vous de bien sélectionner le formulaire scanner voulu,
- Vous arrivez sur une page grise, Cliquez sur le bouton « Régler »,
- Dans la nouvelle fenêtre, cochez l'option « scanner des images pré numérisées » et quittez en cliquant sur le bouton « Fermer »,
- Cliquez sur le bouton « Scanner » pour lancer la récupération.
- A ce moment, une fenêtre va s'ouvrir afin de vous permettre d'indiquer l'emplacement des images que vous avez préparées.

## 6. Que faire si...

Voici quelques exemples de messages d'erreurs que vous pourrez rencontrer durant le scan :

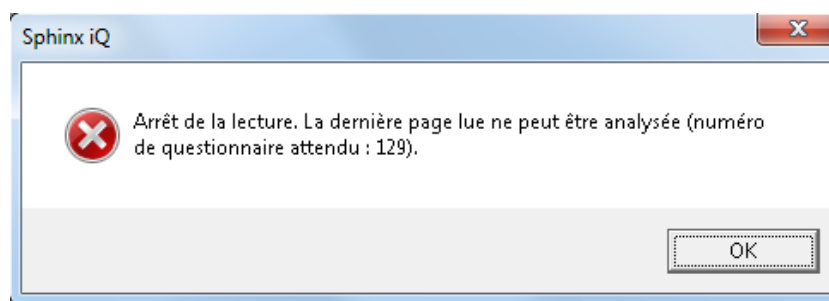


Ce message survient lorsque Sphinx ne réussit pas à détecter (complètement ou en partie) le cadre du formulaire. Suite à ce message, on vous propose de voir l'aperçu, validez avec « Oui » afin de cibler l'origine du problème. Vous devez alors vérifier les points suivants :

Y a-t-il bien un cadre autour de votre formulaire,

- Si non, vous devez refaire vos formulaires scanner de manière à avoir ce cadre,
- Si oui, vous devez avoir un (ou plusieurs) coin(s) qui n'a (ont) pas été détecté(s) correctement. Les coins posant problèmes n'auront pas le coin rouge du Sphinx sur le coin du cadre. Voici les éléments pouvant empêcher la reconnaissance des coins :
  - o Trait du cadre pas assez net ou prononcé (de nombreux points blancs apparaissent dans le trait du cadre via l'aperçu) :

- Vérifiez la qualité d'impression de votre document, si celle ci est correcte, il faut alors vérifier les options de votre scanner (luminosité / contraste) pour obtenir un formulaire au niveau de l'aperçu avec un trait plus net.
- Le coin est complètement ou partiellement effacé,
  - Retracez proprement le trait avec un stylo noir et une règle, puis passez de nouveau le formulaire dans le scanner.
- Des éléments empiètent sur le cadre (commentaire du répondant, coup de tampon, etc...)
  - Supprimez les éléments perturbateurs (à l'aide d'un tipp-ex blanc par exemple) afin de ne laisser que le cadre.

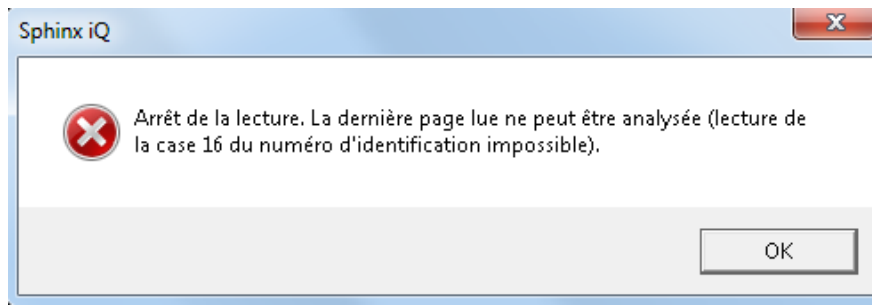


Ce message survient lorsqu'il y a un problème avec le code de page. Suite à ce message, on vous propose de voir l'aperçu, validez avec « Oui » afin de cibler l'origine du problème. Vous devez alors vérifier les points suivants :

- Le code de page que vous avez sur le document papier correspond-il bien au code de page présent sur le formulaire écran sous Sphinx ?
  - Si non, êtes vous sur que le formulaire papier correspond bien au formulaire de l'enquête en cours ? Si oui, il faut alors indiquer au Sphinx que le code n'est plus le même. Pour cela :
    - Au niveau des formulaires personnalisés,
    - Cliquez sur le menu « Formulaires » puis « Propriétés formulaire »,
    - Dans l'onglet « Utilisation », cochez l'option « Numérotation manuelle du questionnaire » et indiquez le numéro que vous avez sur votre formulaire papier,
    - Validez avec « Ok » et scannez de nouveau votre formulaire.
  - Si oui, passez à l'étape suivante,
- Le code de page correspond bien mais il n'est peut être pas bien reconnu. Suite au message d'erreur, on vous propose de voir l'aperçu, validez avec « Oui » afin de cibler l'origine du problème. Vous devez alors vérifier les points suivants :



- Vérifiez tout d'abord si les 4 coins du cadre sont bien reconnus. (voir page précédente),
- Si les 4 coins sont bien reconnus par Sphinx, vous devez alors vérifier si le code de page en bas est lui aussi bien identifié. Des cases rouges doivent être placées sur les cases du code de page. Vérifiez le placement,
  - Si celui ci n'est pas correct, il est conseillé alors de nous contacter directement par téléphone.
  - S'il est correct, vérifiez la qualité des traits des cases. Si les traits ne sont pas précis et net, il faut alors vérifier l'impression ou les options du scanner (luminosité / contraste) afin d'obtenir un trait plus prononcé.



Ce message survient lorsqu'il y a un problème avec le code de page. Suite à ce message, on vous propose de voir l'image scannée, validez avec « Oui » afin de cibler l'origine du problème. Vous devez alors vérifier les points suivants :

- Vérifiez tout d'abord si les 4 coins du cadre sont bien reconnus. (détail de la procédure en page 40),
- Si les 4 coins sont bien reconnus par Sphinx, vous devez alors vérifier si le code de page en bas est lui aussi bien identifié. Des cases rouges doivent être placées sur les cases du code de page.
  - Vérifiez le placement,
  - Si celui ci n'est pas correct, il est conseillé alors de nous contacter directement par téléphone.

S'il est correct, vérifiez la qualité des traits des cases. Si les traits ne sont pas précis et net, il faut alors vérifier l'impression ou les options du scanner (luminosité / contraste) afin d'obtenir un trait plus prononcé.

## 6.1. Le scanner ne démarre pas

Vérifier :

- Les branchements (câble USB, alimentation secteur).

- Dans Sphinx, dialogue **Régler**, bouton **Sélectionner**, vérifier que le scanner est visible.
- Tester le fonctionnement du scanner avec un autre logiciel que Sphinx. **Microsoft Word** dans le menu **Insertion / Image / A partir d'un scanner**, ou tout autre logiciel capable d'utiliser le scanner.

## 6.2. Le chargeur de documents n'est pas activé

Si le scanner n'est pas un scanner que nous recommandons, il est probable que le chargeur n'est pas 100 % compatible TWAIN.

## 6.3. Plusieurs pages sont avalées en même temps

L'introducteur n'arrive pas à séparer les pages les unes des autres. Elles ont été "collées" par un agrafage top fort, par un pliage durable, ...Si le modèle du scanner le permet, régler la force de l'introducteur de documents.

Comme avec une photocopieuse il est prudent d'"effeuiller" le paquet de feuilles pour les aligner et les séparer. N'oubliez pas que la qualité de l'introducteur est une forte composante du prix d'un scanner.

## 6.4. Message "Impossible de lire la question n °"

Vérifier la position des cases attendues et celle des cases détectées (option **Zones de lecture**).

Si les cases ont été effacées, il faudra les reconstituer. La position des cases pourra être manuellement rectifiée si nécessaire.

## 6.5. Affichage à l'écran incohérent avec l'image

Le formulaire a peut-être été modifié (sans être imprimé) depuis la reproduction du document papier. Rectifier les options pour retrouver la même configuration que le document imprimé.

Le formulaire a peut-être été réalisé sur une machine configuré différemment. Remettez la machine en configuration standard (petites polices).

En tout état de cause, les formulaires imprimés pourront être lus même si l'affichage à l'écran est différent.

## 6.6. Erreurs de détection des coches

En vérifiant les cases détectées, on s'aperçoit d'erreurs notoires entre les cases cochées et ce que l'on voit sur le papier. Vérifier ce que le scanner a "vu" en examinant l'image scannée. Il se peut que le réglage du contraste et de la luminosité ne soit pas correct. Certaines couleurs peuvent avoir été ignorées, vérifier que le stylo utilisé n'est pas de cette couleur.

# Table des matières détaillée

<b>Sommaire .....</b>	<b>2</b>
<b>Introduction .....</b>	<b>3</b>
1. Description du matériel et des services nécessaires .....	3
1.1. Le scanner .....	3
1.2. Installation – Formation – Assistance : options conseillées .....	4
2. Quelques questions fréquentes .....	4
<i>Faut-il noircir complètement les cases ? .....</i>	<i>4</i>
<i>Comment coche-t-on une case ? Que se passe-t-il si le trait dépasse de la case ? .....</i>	<i>4</i>
<i>Peut-on utiliser un crayon ? .....</i>	<i>4</i>
<i>Peut-on raturer ? .....</i>	<i>5</i>
<i>Que se passe-t-il si deux cases sont cochées ? .....</i>	<i>5</i>
<i>Quelle fiabilité peut-on espérer ? .....</i>	<i>5</i>
<i>Quelle vitesse peut-on espérer ? .....</i>	<i>5</i>
<i>Comment traiter les questions ouvertes ? .....</i>	<i>5</i>
<i>Peut-on "coder" un paquet de feuilles ? .....</i>	<i>5</i>
<i>Le questionnaire peut-il être sur plusieurs pages ? .....</i>	<i>5</i>
<i>Peut-on modifier la mise en forme dans un logiciel de PAO ? .....</i>	<i>5</i>
<i>Peut-on préremplir les questionnaires (avec le nom par exemple) ? .....</i>	<i>6</i>
<i>L'installation du scanner est-elle complexe ? .....</i>	<i>6</i>
<i>Quel scanner faut-il choisir ? .....</i>	<i>6</i>
<i>J'ai un scanner, convient-il ? .....</i>	<i>6</i>
<b>Mise en œuvre .....</b>	<b>7</b>
1. Détail du processus .....	7
2. Préparation des documents .....	8
2.1. Mise en forme des formulaires .....	8
2.2. Impression et reproduction des formulaires .....	8
a) Photocopie traditionnelle .....	8
b) Photocopie numérique .....	8
3. Déroulement de la lecture .....	9
3.1. Environnement .....	9
3.2. Lancer la lecture .....	9
3.3. Affichage du résultat de la lecture .....	10

3.4. Corriger et compléter	11
3.5. Erreurs de lecture	11
a) Interruption de la lecture .....	11
b) Avertissements .....	12
3.6. Affecter une valeur à toutes les feuilles du paquet	13
3.7. Accepter ou refuser le lot de feuilles	13
4. Réglages du scanner .....	14
4.1. Réglages de la méthode de scan	14
a) Sélectionner le scanner .....	14
b) Résolution .....	14
c) Recto-Verso .....	14
d) Auto-configuration .....	14
e) Utiliser des images pré-numérisées .....	15
f) Configurer le scanner .....	15
4.2. Traitement de l'image	15
a) Auto-rotation .....	15
b) Corriger les déformations dues au scanner .....	16
4.3. Traitement des tâches et des ratures	16
4.4. Configurations spécifiques des scanners	17
a) Scanners FUJITSU FI4110, FI4120 .....	17
<i>Dialogue Régler</i> .....	17
<i>Dialogue Sélectionner</i> .....	17
<i>Dialogue Configurer</i> .....	17
b) Autres scanners .....	18
<i>Dialogue Régler</i> .....	18
<i>Dialogue Sélectionner</i> .....	18
<i>Dialogue Configurer</i> .....	18
5. Fonctions avancées .....	19
5.1. Enregistrement	19
a) Enregistrement des images .....	19
b) Enregistrement du dernier paquet .....	19
c) Tenir un journal des opérations .....	19
5.2. Codes de page	20
5.3. Cas particuliers	21
a) Marges spéciales .....	21
b) Scanner au format "paysage" .....	21
5.4. Rectification de la position des cases	22
5.5. Traitement des questionnaires nominatifs	22
a) Identification préalable des observations .....	22
b) Identification a posteriori des observations .....	22
5.6. Que faire pour...	23
a) Lire une date .....	23
b) Utiliser les services d'un imprimeur .....	23

## Synthèse ..... 24

1. Règles de construction .....	24
1.1. Créer son formulaire scanner :	24
1.2. Règles à suivre concernant la mise en page du formulaire	24
2. Recommandations pour l'impression et la saisie .....	25
2.1. Impression du formulaire	25
2.2. Reproduction du formulaire	25

2.3. Remplissage par les enquêteurs ou les interviewés	25
2.4. Lecture automatique par scanner	26
3. Validation et reproduction du formulaire .....	26
3.1. Validation du formulaire avant l'impression finale :	26
3.2. Transmettre le formulaire finalisé à un service de reproduction :	26
4. Spécificités et réglages du scanner .....	27
4.1. Réglage conseillé des options du scanner sous Sphinx :	27
4.2. Scanner de bureau ou Photocopieur réseau, comment faire ?	27
4.3. Pré-requis sur les images pré-numérisées :	28
5. Tester la numérisation (1 <sup>ère</sup> prise en main) .....	28
5.1. Avec un scanner branché directement sur votre ordinateur :	28
5.2. Avec un scanner réseau (photocopieur) :	29
6. Que faire si... .....	30
6.1. Le scanner ne démarre pas	32
6.2. Le chargeur de documents n'est pas activé	33
6.3. Plusieurs pages sont avalées en même temps	33
6.4. Message "Impossible de lire la question n °"	33
6.5. Affichage à l'écran incohérent avec l'image	33
6.6. Erreurs de détection des coches	33

## **Table des matières détaillée .....34**