**CNIRevelator**

**Documentation du logiciel, version 3.1.1 final release**

Révision 1

# **Licences et mentions légales**

*CNIRevelator 3.1.1 final release*

*Copyright © 2018-2019 Adrien Bourmault (neox95)*

*CNIRevelator* est un logiciel libre : vous avez le droit de le modifier et/ou le distribuer dans les termes de la GNU General Public License telle que publiée par la Free Software Foundation, dans sa version 3 ou ultérieure.

*CNIRevelator* est distribué dans l'espoir d'être utile, sans toutefois impliquer une quelconque garantie de QUALITÉ MARCHANDE ou APTITUDE À UN USAGE PARTICULIER. Référez-vous à la **GNU General Public License** pour plus de détails à ce sujet.

Vous devriez avoir reçu une copie de la GNU General Public License avec CNIRevelator. Si cela n'est pas le cas, jetez un œil à <https://www.gnu.org/licenses>.

*Le module d'OCR Tesseract 4.0 est soumis à l'Apache License 2004.*

*Les bibliothèques python et l'environnement Anaconda 3 sont soumis à la licence BSD 2018-2019.*

*Le présent document est publié sous licence Creative Common CC-BY-SA.*

# **Accès au code source du logiciel**

Le code source de ce programme est disponible sur Github à l'adresse <https://github.com/neox95/CNIRevelator>.

En cas de problèmes ou demande particulière, ouvrez-y une issue ou bien envoyez un mail à [neox@os-k.eu](mailto:neox@os-k.eu) !

# Crédits

Documentation écrite par Adrien Bourmault (neox95) en août 2019.

*CNIRevelator* est un logiciel développé depuis juillet 2018 de façon active par Adrien Bourmault. Ce projet a été initié au service conformité de la DIAC (une filiale de RCI Banque & Services) dans le cadre de la modernisation de la lutte contre la fraude par usurpation d’identité.

Cette documentation est publiée sous licence Creative Commons CC-BY-SA

# Présentation du logiciel et avertissement

*CNIRevelator* est un logiciel d’analyse de document d’identités. A ce titre, il se base sur les normes du Document 9303 de l’OACI, ayant pour but d’harmoniser mondialement la facture des pièces et documents d’identité, et la norme ISO/IEC 18745-1:2014.

Ce document n’est pas exhaustif et il est possible que certains pays n’appliquent pas ces normes, rendant alors impossible l’analyse de ces documents par le présent logiciel. C’est pourquoi *CNIRevelator* ne clame pas son fonctionnement sur toutes les pièces d’identité mais uniquement sur celles supportées par le Document 9303.

Notez bien que *CNIRevelator* ne produit qu’une vérification du document de façon formelle, c’est-à-dire que l’analyse ne porte que sur la conformité de la bande lisible à la machine nommée ZLA ou MRZ, et n’effectue aucune vérification de validité courante comme pourrait le faire un système connecté aux fichiers du Ministère de l’Intérieur. Ainsi, le présent logiciel ne peut détecter la présence d’une carte d’identité volée, ou bien d’une carte d’identité fausse mais conforme (le cas des documents extrêmement bien falsifiés). Cette analyse se base sur la vérification des sommes de contrôle ainsi que la pertinence des données lues.

De plus, notez que la conformité d’une pièce n’implique pas forcément sa validité (comme vu avec les faussaires talentueux, ou bien les remplacements de photographie où seule l’inspection visuelle peut déceler une anomalie). Cependant, la non-conformité implique évidemment l’invalidité du document et permet de conclure à l’usurpation d’identité.

# Au sujet des données personnelles

*CNIRevelator* possède un système de mise-à-jour automatique. Toutes les versions du logiciel sont téléchargées à partir du dépôt Github où est hébergé le code.

Github, aujourd’hui organe du groupe Microsoft, est une plateforme utilisée par des développeurs du monde entier pour héberger leur code et leurs version compilées de logiciels. Cela en fait un dépôt de confiance pour la mise à jour du présent logiciel. De plus, le téléchargement est toujours authentifié à l’aide de sommes de contrôles afin d’éviter tout problème de corruption.

Le dépôt n’enregistre aucune donnée anonyme ou non anonyme concernant le poste de travail effectuant le téléchargement et n’effectue aucune statistique à son sujet, les adresses IP n’étant pas enregistrées par le dépôt.

*CNIRevelator* n’enregistre aucune donnée personnelle de l’utilisateur excepté ses identifiants d’accès à internet et ses paramètres de langue et de mise-à-jour. Les identifiants sont stockés sous forme encrypté sur le poste local, dans le fichier de configuration adapté (voir plus loin). Les autres paramètres sont stockés en clair et sont totalement anonymes.

*CNIRevelator* n’échange aucune donnée personnelle sur le réseau, ni n’en transmet ou en enregistre sur quelconque serveur distant, excepté lors de rapports d’erreurs. Cependant ces données sont anonymes et ne permettent pas d’effectuer un quelconque traçage.

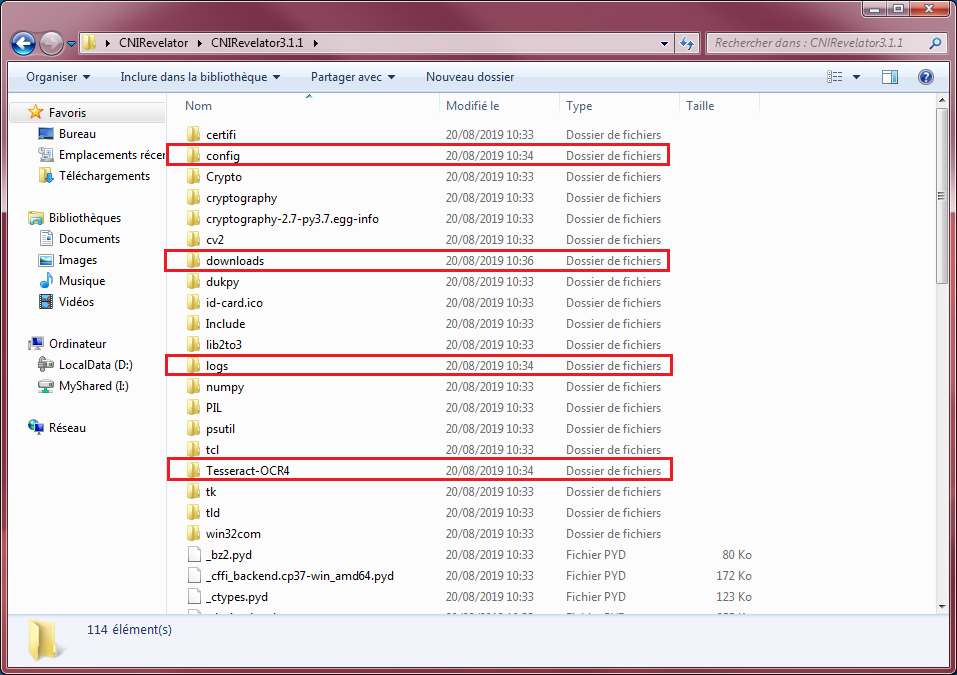
*CNIRevelator* ne nécessite pas de connexion à internet pour fonctionner convenablement. Notez que la connexion à internet est nécessaire pour recevoir les mises-à-jour du logiciel ou envoyer des rapports d’erreurs, bien que cela ne soit pas requis pour utiliser le logiciel.

# Structure du logiciel

*CNIRevelator* est constitué d’un dossier, souvent nommé CNIRevelatorX.X.X où X.X.X est le numéro de version complet. Par exemple, au moment de l’écriture de ce document, le dossier se nomme CNIRevelator3.1.1. Cependant, ce dossier peut être renommé, déplacé, sans perturber le fonctionnement du logiciel.

La création d’un raccourci est toujours proposée par le système de mise-à-jour lors du téléchargement d’une nouvelle version. Ce raccourci se place sur le bureau. Si vous déplacez le logiciel, pensez à modifier le raccourci pour refléter vos changements.

Dans le dossier de *CNIRevelator,* on trouve différentes choses :



* De nombreux fichiers et dossiers, tous nécessaires au bon fonctionnement de l’application (des bibliothèques logicielles surtout, des fichiers de configuration du système, des images utilisées par le programme).
* Les quatre dossiers encadrés en rouge. Ce sont des dossiers créés à la première exécution de *CNIRevelator*.

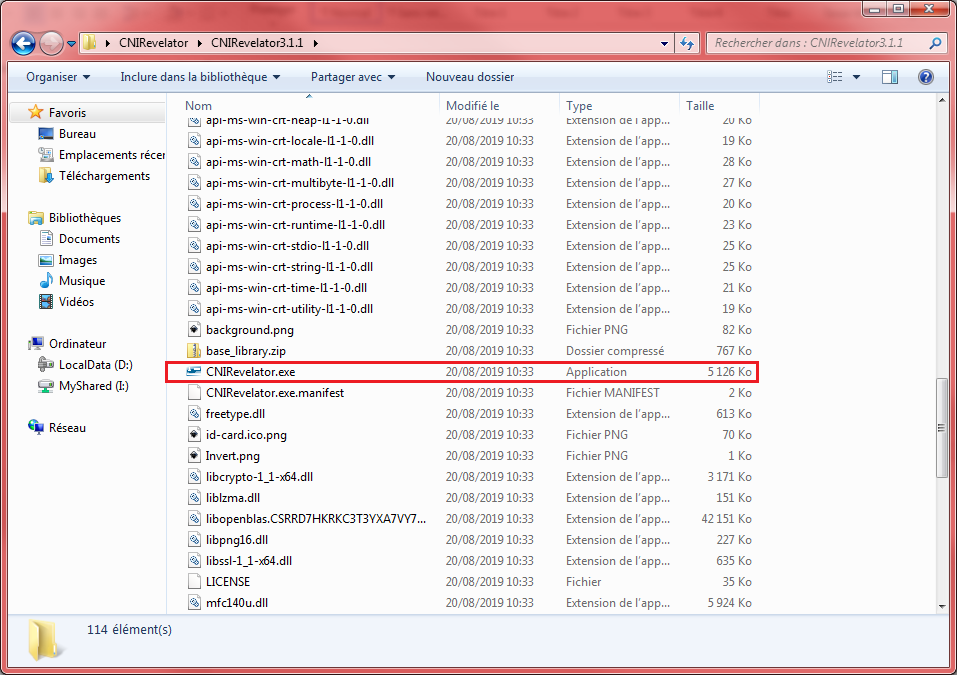
Le dossier *config* contient les fichiers de configuration utilisateur du logiciel.

Le dossier *logs* contient les journaux du logiciel. Ces deux dossiers sont les seuls à contenir des données personnelles.

Le dossier *downloads* contient les téléchargements de nouvelles versions effectués par le système de mise-à-jour.

Le dossier *Tesseract-OCR4* contient le module de reconnaissance optique de caractère utilisé pour lire les MRZ sur des scans.

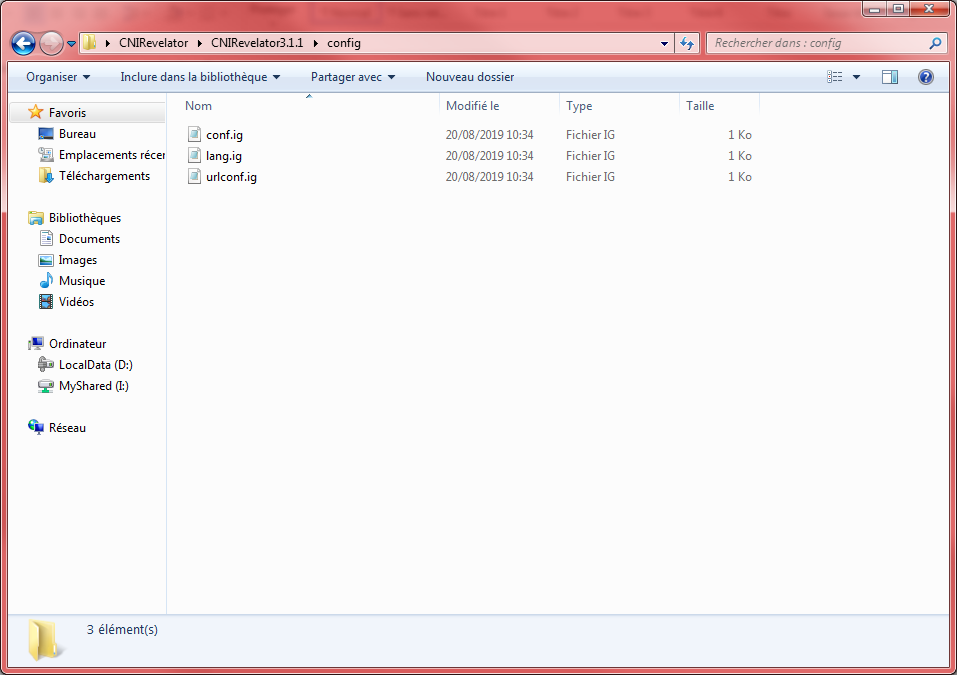
Le dossier de *CNIRevelator* contient également son exécutable, ici encadré :



C’est cet exécutable qui lance le logiciel. Vous pouvez faire un raccourci sur votre bureau vers ce programme si vous le souhaitez, sachant qu’à chaque mise-à-jour il vous sera proposé de le faire automatiquement.

# Configuration de *CNIRevelator*

*CNIRevelator* utilise trois fichiers de configurations tous localisés dans le dossier *config* :



Le fichier *conf.ig* contient les identifiants cryptés d’accès à internet de l’utilisateur. Ceux-ci sont demandés à la première utilisation pour rechercher les mises-à-jour à travers le proxy si présent.

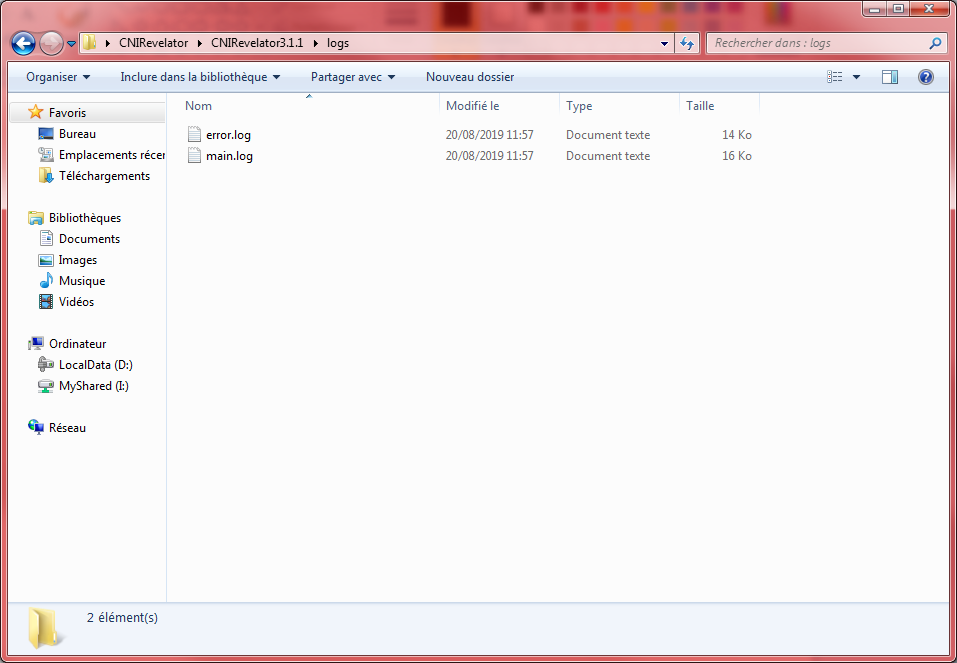
Le fichier *lang.ig* contient les paramètres de langue (français ou anglais au moment de l’écriture du document). Pour le français il contiendra « fr » sinon « en ». Par défaut, l’application est en français.

Le fichier *urlconf.ig* contient les paramètres de mise-à-jour c’est-à-dire l’emplacement du dépôt de mise-à-jour. Cela peut-être utile pour le débogage ou l’utilisation d’un dépôt autre que le dépôt officiel.

Il n’est normalement pas nécessaire de modifier manuellement ces fichiers. En cas de problème, il peut être utile de les supprimer pour forcer leur recréation. Si un problème insolvable se présente, ne pas hésiter à ouvrir une issue sur le Github officiel (voir *Accès au code source du logiciel*).

# Journalisation

*CNIRevelator* possède un système de journalisation (logs) créant des fichiers journaux dans le dossier *logs* :



Le fichier *main.log* est un fichier regroupant tous les évènements de la **dernière utilisation** de *CNIRevelator*. Il est purgé à chaque lancement du logiciel.

Le fichier *error.log* est un fichier qui regroupe uniquement les évènements correspondant à des erreurs du logiciel lors de sa **dernière utilisation.** Il est purgé également à chaque lancement et ne se crée que si une erreur se produit. Il est inclus dans *main.log* mais permet de localiser plus facilement les erreurs car ne contient que celles-ci.

Ces deux fichiers étant toujours purgés, aucune donnée concernant les documents d’identité analysés, excepté lors de la dernière utilisation et jusqu’à la suivante uniquement, ne sont stockés par *CNIRevelator*.

# Démarrage de CNIRevelator

Afin de lancer *CNIRevelator* et commencer à l’utiliser, double-cliquez sur son raccourci ou bien son exécutable (dans le dossier) :

|  |  |  |
| --- | --- | --- |
|  | OU |  |



L’application se lance alors et affiche son écran de démarrage.



Puis, si aucune mise à jour n’est nécessaire, *CNIRevelator* affichera l’écran suivant, puis ouvrira sa fenêtre principale.

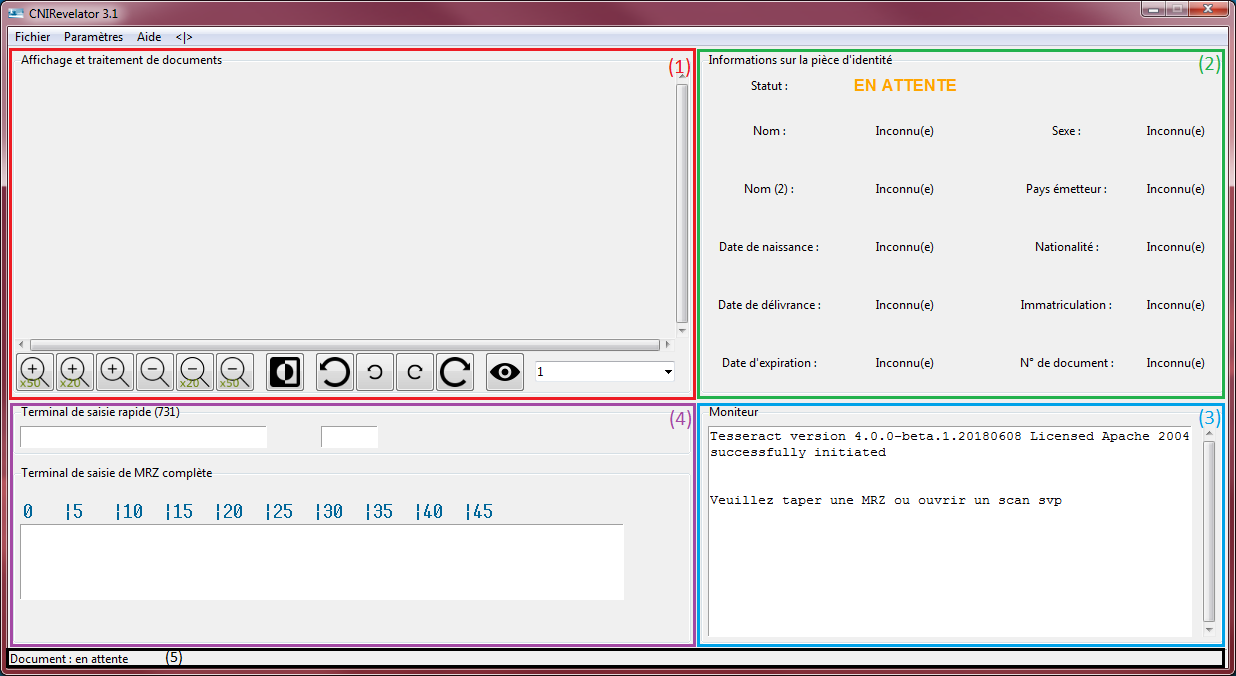


Sinon, si une mise-à-jour est nécessaire, un téléchargement aura lieu et cet écran vous permettra d’en suivre l’avancée.



Enfin, lorsque le téléchargement est terminé, ce dernier écran vous informe de l’installation (qui en général prend moins d’une minute). L’application se lance immédiatement après.

# Fenêtre principale de l’application

La fenêtre principale de *CNIRevelator* se présente ainsi :

On distingue cinq parties distinctes :

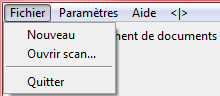
1. Zone d’affichage et traitement de documents. Cette zone, vide au démarrage de l’application est destinée à accueillir des documents d’identité pour les analyser. La barre d’outil qui y est présente comporte diverses fonctions utiles au traitement d’image afin de rendre lisible la pièce d’identité : rotation, zoom, et négatif (pour accentuer le contraste)
2. Zone d’informations sur la pièce d’identité. Cette zone, affichant « Inconnu(e) » au démarrage de l’application est destinée à afficher les informations à propos de la pièce d’identité, à partir de la bande MRZ actuellement dans le terminal de saisie complète. Elle affiche également un avertissement si ces données ont un caractère non conforme.
3. Zone de terminaux. Cette zone contient deux terminaux de saisie :

- un terminal de saisie rapide permettant la saisie d’une portion de MRZ quelconque et le calcul de sa somme immédiatement. Cela peut être utile pour une vérification rapide d’une zone que l’on connait.

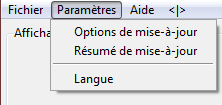
- un terminal de saisie complète permettant la saisie des deux lignes (ou une si une seule) de la MRZ d’un document d’identité et de déterminer sa conformité. Lors de l’utilisation de la fonction de reconnaissance optique de caractère, c’est dans ce terminal que la bande reconnue sera insérée.

1. Zone de moniteur. Cette zone affiche des messages concernant les actions actuelles, les résultats de vérifications (sommes par exemple) et d’éventuelles erreurs.
2. Barre de statut. Cette zone affiche des messages rapides à propos de l’analyse du document.

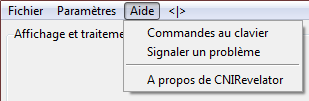
En dehors de ces zones, la fenêtre contient également une barre de menu permettant d’interagir avec l’application.

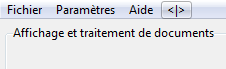


Le menu **Fichier** permet de commencer une nouvelle analyse (*Nouveau*), d’ouvrir un document afin d’analyser son scan (*Ouvrir scan…*) et de quitter l’application.

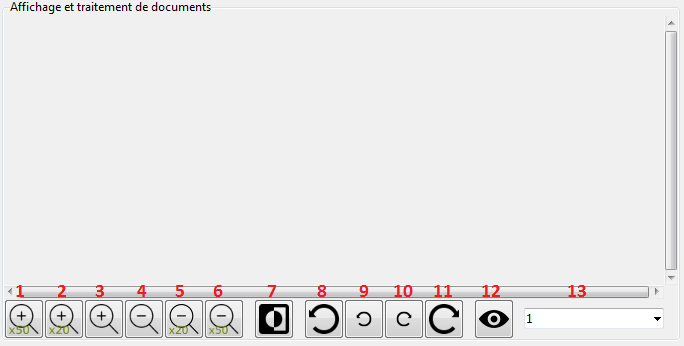


Le menu **Paramètres** permet de régler des options de configurations comme les options de mise-à-jour et la langue, mais également d’afficher la liste des versions précédentes et leurs changements (*Résumé de mise-à-jour*)

Le menu **Aide** permet d’afficher le résumé des commandes clavier de l’application, de signaler un problème en ouvrant la page du Github officiel, et d’afficher des informations sur la version actuelle du logiciel

Le bouton **<|>** permet de basculer entre la vue complète et la vue réduite de la fenêtre principale, occupant moins d’espace sur l’écran.

# Zone d’affichage et traitement de documents

La zone d’affichage et traitement de document contient une barre d’outils avec de nombreuses fonctions.

Ces fonctions permettent de modifier le scan afin de le rendre lisible pour obtenir une lecture efficace.

1. Zoom en avant +50%
2. Zoom en avant +20%
3. Zoom en avant +1%
4. Zoom en arrière -1%
5. Zoom en arrière -20%
6. Zoom en arrière -50%
7. Réduire le bruit de l’image (2 modes)
8. Rotation vers la gauche 90°
9. Rotation vers la gauche 1°
10. Rotation vers la droite 1°
11. Rotation vers la droite 90°
12. Effectuer la reconnaissance optique de caractères (OCR)
13. Changer de page (dans le cas d’une image multipage comme les fichiers TIFF)

# Utilisation de *CNIRevelator* sans reconnaissance optique

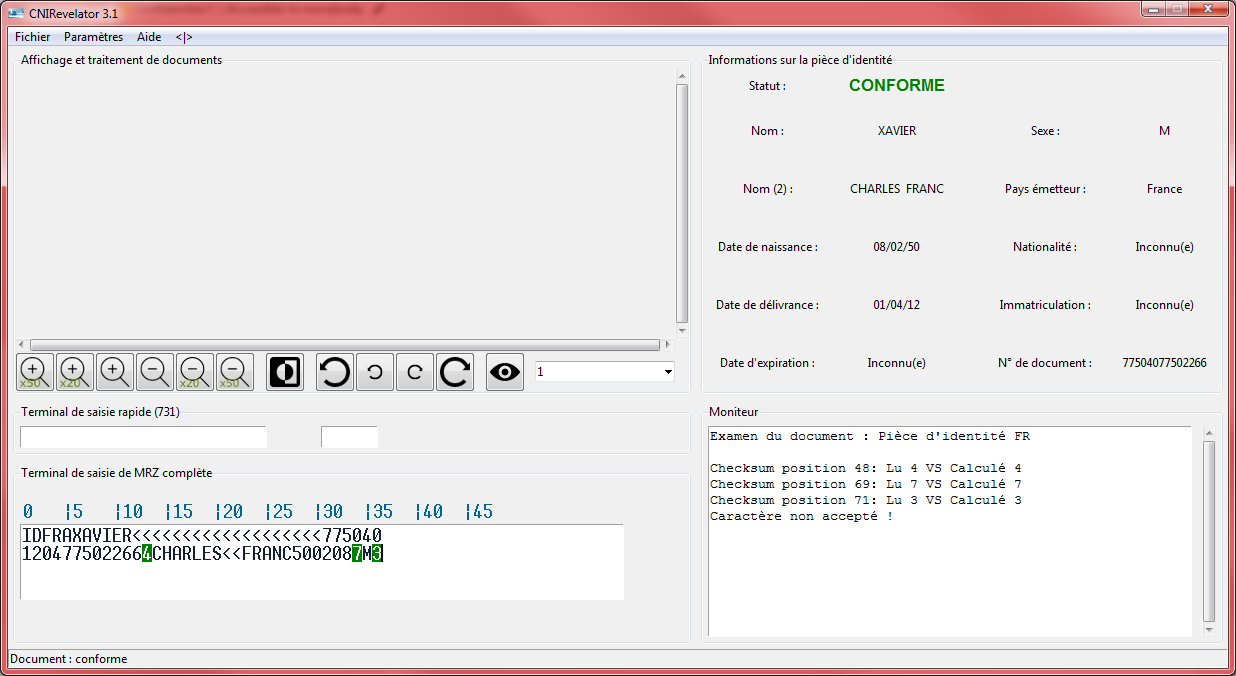
L’utilisation de CNIRevelator sans reconnaissance optique de caractère peut parfois être fastidieuse, mais est le seul moyen si on ne possède pas de scan numérique du document d’identité, ou bien que le scan est de très mauvaise qualité. Il faut pour cela taper la MRZ du document dans le terminal de saisie complète, ou bien en taper des portions dans le terminal de saisie rapide ***EN MAJUSCULES ET UNIQUEMENT CARACTERES ALPHANUMERIQUES AINSI QUE LE CARACTERE <***.

Attention cependant : seul le terminal de saisie complète peut statuer sur la conformité de la pièce. De plus, rappelez-vous que les contrôles claviers sont indiqués dans **Aide** > *Commandes au clavier*.

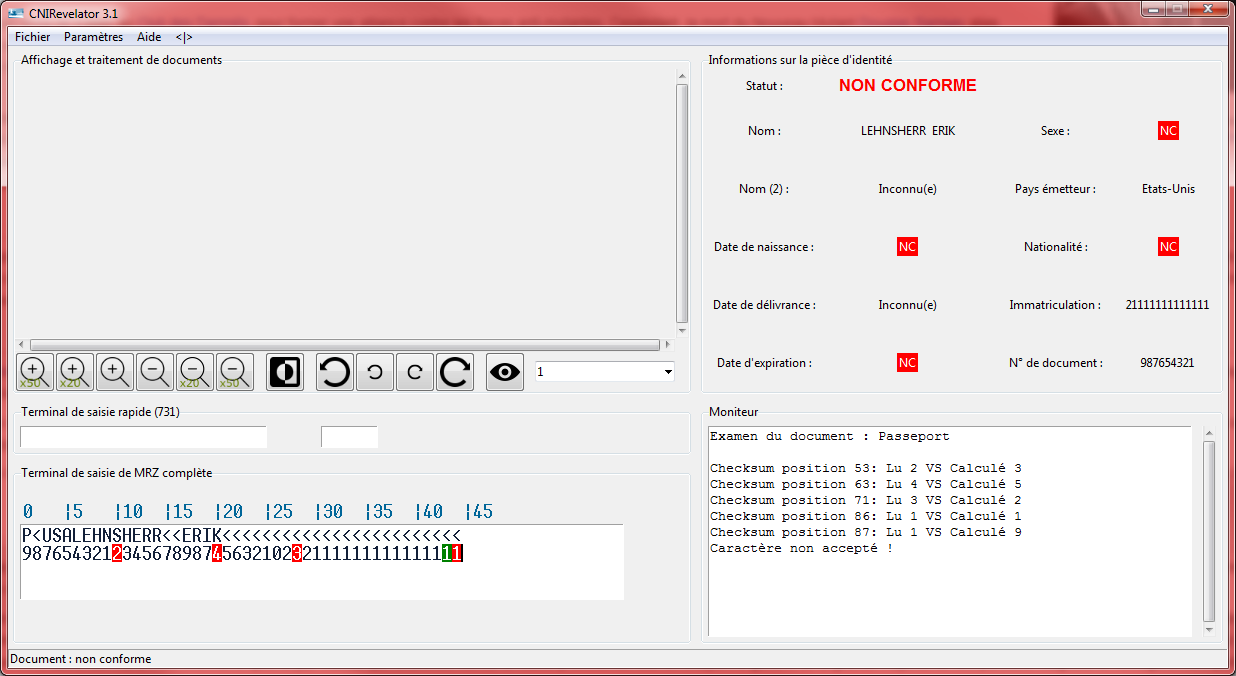
On distingue deux phases dans la saisie du document dans le terminal de saisie complète : une phase de détection du document ainsi qu’une phase d’analyse simultanée.

Lors de la première phase, à chaque caractère entré, *CNIRevelator* tente de détecter le type de document actuellement saisi. La détection d’un document peut parfois différer de la réalité. Dans ce cas, pas de panique ! Il suffit de redemander la détection du document une fois qu’on en a saisi une part plus significative à l’aide de la touche dédiée.

Une fois le document détecté commence la seconde phase : l’analyse simultanée. Durant cette phase, le à chaque caractère entré, *CNIRevelator* analyse ce qui a été précédemment soumis dans le terminal complet. Pour forcer l’analyse, vous pouvez presser la touche dédiée au calcul de résultat. Vous devez forcer l’analyse lorsque vous entrez le dernier caractère, afin de signifier que votre saisie est terminée.

Voici un exemple des résultats montrés par l’application avec une carte d’identité française spécimen :

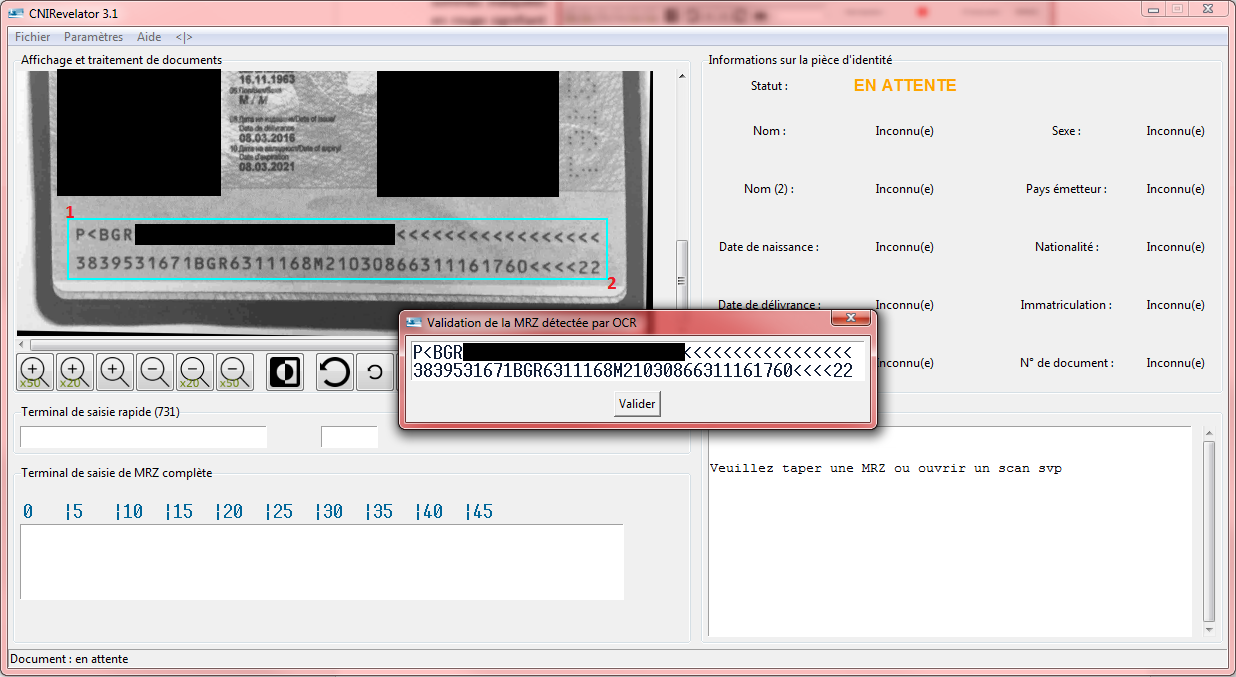
Les sommes validées sont indiquées en vert, et le moniteur indique la concordance. La barre de statut indique le document conforme ainsi que la zone d’informations. On a diverses infos affichées également sur le possesseur de la carte.

Voici un autre exemple, avec un passeport spécimen :

On remarque de nombreuses sommes indiquées en rouge signifiant leur non validité. Des informations sont marquées NC dans la zone d’informations signifiant qu’elles sont non conformes à ce qui devrait s’y trouver (par exemple une lettre dans une date de naissance, ou bien une date non valide, ou bien un sexe autre que « homme, femme ou non précisé ». La barre de statut indique également le caractère non conforme de la pièce.

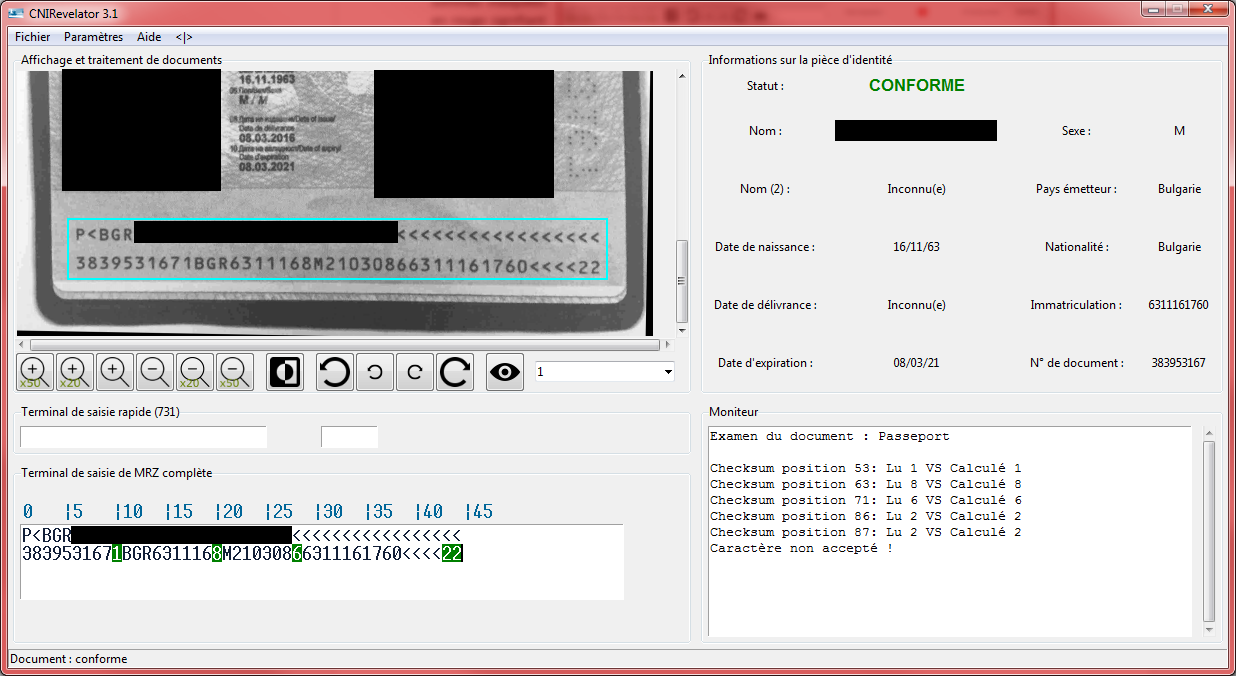
# Utilisation de *CNIRevelator* avec reconnaissance optique

L’utilisation de l’application avec la reconnaissance optique de caractère peut s’avérer pratique lorsqu’on a un scan numérique du document à analyser. Cependant, la qualité du scan peut grandement influencer les résultats. Blâmer *CNIRevelator* ne servira à rien si le scan est de mauvaise qualité, on ne peut pas créer de matière quand il n’y en a pas !

Pour commencer, ouvrez un scan avec le menu **Fichier** > *Ouvrir un scan…* et choisissez le fichier TIFF ou JPEG que vous souhaitez ouvrir. Une fois ceci fait, vous pouvez utiliser les outils de traitement pour obtenir une image bien droite et claire de la MRZ à analyser. Vous devez alors effectuer la sélection de la MRZ en dessinant un rectangle autour de celle-ci. Commencez par cliquer en haut à gauche de la MRZ (1), puis en bas à droite (2). Voilà, vous avez un rectangle. Cliquez à tout autre endroit pour recommencer le dessin. Appuyez alors sur le bouton OCR pour commencer la lecture, et obtenez l’écran ci-contre (les informations personnelles sensibles ont été censurées au noir).

On constate l’ouverture d’une fenêtre de validation. Elle contient le texte qui a été détecté par le module OCR. Selon la qualité du scan, sa qualité peut varier de totalement exacte à complètement erronée.

Vous êtes invité à relire ce qui a été détecté pour vous assurer de l’exactitude et à corriger si nécessaire. Cela est beaucoup plus rapide que de tout écrire car vous n’avez qu’à comparer, mais cela peut vous sembler fastidieux. Encore une fois, le résultat dépend uniquement de la qualité du scan.

Une fois ceci fait, vous pouvez appuyer sur *Valider* et faire s’afficher l’écran suivant. Le texte que vous avez validé est maintenant dans le terminal de saisie. Vous pouvez encore effectuer des modifications dessus si vous le souhaiter. Ensuite, appuyez sur la touche de calcul de résultat. L’écran ci-contre s’affiche, avec les resultats de l’analyse, affichés exactement de la même manière qu’en saisie manuelle (sans reconnaissance optique).

# Types de documents analysables par *CNIRevelator*

*CNIRevelator* supporte six formats de documents d’identités :

* DVOLM TD1 (53,98 mm × 85,6 mm) à 30 caractères/ligne sur trois lignes. Attention : seules les deux premières sont utiles.
  + Carte d’identité européenne
  + Certificat de membre d’équipage
  + Carte-passeport (surtout USA)
  + Titre de séjour européen
* DL TD1 (53,98 mm × 85,6 mm) à 30 caractères/ligne sur une ligne.
  + Permis de conduire européen (norme 2013)
* DVOLM TD2 (74,0 mm × 105,0 mm) à 36 caractères/ligne sur deux lignes.
  + Carte d’identité française
  + Titre d’identité ou de voyage
  + Titre de séjour français
* DVLM TD3 (88,0 mm × 125,0 mm) à 44 caractères/ligne sur deux lignes.
  + Passeport lisible à la machine
* VLM-A (80,0 mm × 120,0 mm) à 44 caractères/ligne sur deux lignes.
  + Visa lisible à la machine de type A
* VLM-B (74,0 mm × 105,0 mm) à 36 caractères/ligne sur deux lignes.
  + Visa lisible à la machine de type B

La pièce d’identité la plus sécurisée est le passeport, qui est la plus rigoureuse en matière de conformité avec le plus haut nombre de sommes de contrôles : de quatre à cinq pour une seule pièce. C’est donc une pièce d’identité à privilégier si un choix peut être fait.

# En cas de problème

Si un problème survient avec CNIRevelator, merci de transmettre un rapport d’erreur lorsque l’application vous le demande.

Sinon, n’hésitez pas à ouvrir un ticket (issue) sur le Github officiel de *CNIRevelator* à l’adresse : <https://github.com/neox95/CNIRevelator/issues>.

Dans la plupart des cas, supprimer le dossier *config* suffira à résoudre le problème. Si celui-ci persistait, vous êtes invité à supprimer les dossiers *config*, *Tesseract-OCR4*, et *downloads*.

Si malgré tout, un problème subsiste avec votre version de *CNIRevelator* et que celle-ci ne peut se mettre à jour automatiquement, vous pouvez télécharger la version la plus récente du logiciel à l’adresse <https://github.com/neox95/CNIRevelator/releases>. Il vous suffira de télécharger la *Latest Release* sous la forme de CNIRevelator.zip et de l’extraire où vous le souhaitez.