

CONSIGNES DE TRAITEMENT

Réf : CO_PRD_5554

Date: 13/09/2016

Arald - OCR Etape : OcrView

Page 1 sur 6

Ver : 00

Date	Version N°	Nature des modifications	Rédigé par	Vérifié par	Validé par
13/09/2016	00	Création	Annie	Mariolla/Bebi	Mariolla/Bebi

Outil:

- OcrView for IsakoStudio

Tâche:

- Vérification/Correction des textes douteux en rouge et les textes non coloriés. Seuls les titre, sous-titre, intertitre, auteur sont à lire

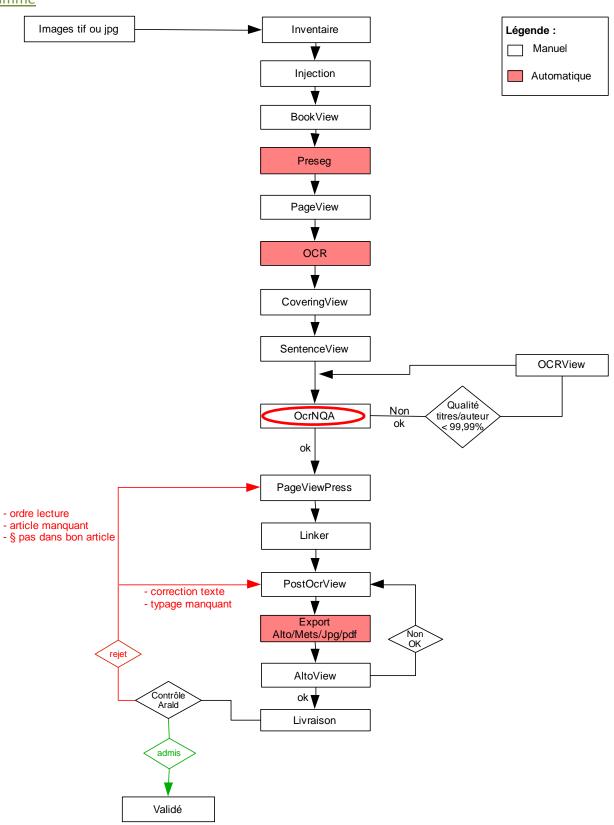


Réf : CO_PRD_5554		
Date: 13/09/2016		
Ver : 00		

Etape : OcrView

Page 2 sur 6

Logigramme

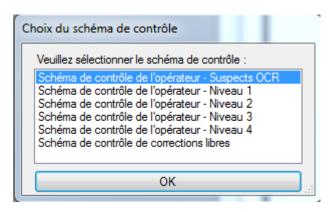




CONSIGNES DE TRAITEMENT	Réf : CO_PRD_5554	
	Date: 13/09/2016	
Arald - OCR	Ver : 00	
Ftane : OcrView	Page 3 sur 6	

OcrView (380 p/h)

L'application OcrView permettra plusieurs niveaux d'échantillonnage (profils d'ouverture)



Chaque niveau permet d'échantillonner "un peu plus" jusqu'au dernier niveau qui échantillonne tous les mots qui n'ont pas déjà été contrôlés aux passes précédents.

L'application OcrView affiche, en pourcentage, la progression du travail de correction par rapport au niveau en cours mais ne dit pas quel est le taux réel.

Quand l'opérateur arrive à 100% des mots échantillonnés, il passe à la page suivante.

Si l'opérateur veut forcer le passage à la page suivante sans finir de corriger les mots échantillonnés dans la page en cours, OcrView lui affiche un message d'avertissement avant passer à la page suivante.

A chaque niveau (i.e., à chaque échantillonnage) les mots déjà contrôlés par l'opérateur ne sont pas échantillonnés à nouveau

L'échantillonnage est effectué chaque fois que la page est ouverte (ou ré-ouverte) dans OcrView. Cela veut dire que si l'opérateur revient en arrière à une page déjà vue, des mots supplémentaires peuvent être échantillonnés à nouveau.



CONSIGNES	DE TR	AITEMENT
-----------	-------	-----------------

Arald - OCR

Etape: OcrView

Réf : CO_PRD_5554

Date: 13/09/2016 Ver: 00

Page 4 sur 6

Les niveaux existants

Mots pour lequels le moteur OCR a détecté ou marqué :	plusieurs caractères suspects	Hors dictionnaire ou Noms propres (selon le moteur OCR)	Mots commençant par un Majuscule	Mots avec hyphen/césure en fin de ligne	Mots qui correspondent à certaines RegExp	Pourcentage des mots restants, i.e., mots qui n'ont pas été échantillonés	Commentaires
Niveau 1	100%	100%	100%	100%			
Niveau 2	х	х	х	x	100% de [RegEx-A]		
Niveau 3	х	х	х	х	100% de [RegEx-A] et de [RegEx-B]	30%	
Niveau 4						100%	

Les RegEx référencées sont :

RegEx-A

Mots ayant les caractères ou les combinaisons des caractères suivants :

 $[IL] \,,\, [II],\, [ii] \,,\, [ii] \,,\, [11] \,,\, [11] \,,\, [11] \,,\, [Y] \,,\, [11] \,,\, [m] \,,\, [nn] \,,\, [ri] \,,\, [cl] \,,\, [<\!char-alpha>']$

RegEx-B

[e], [c]



CONSIGNES DE TRAITEMENT

Arald - OCR

Etape: OcrView

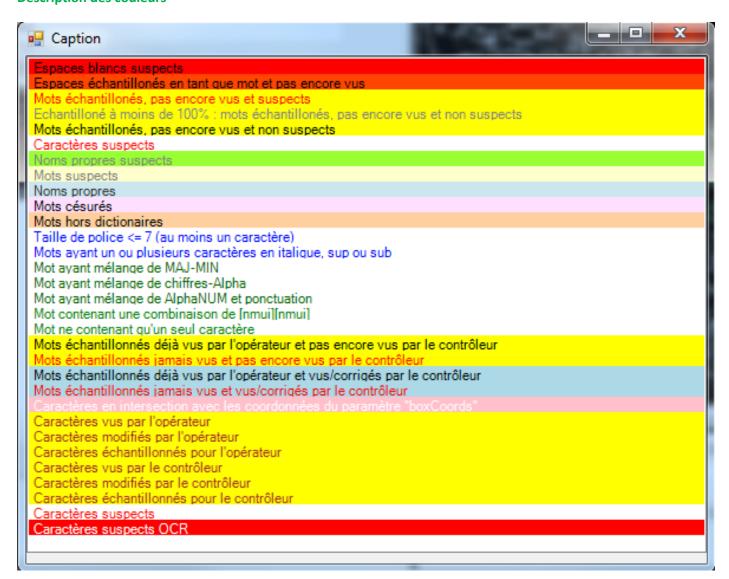
Réf: CO PRD 5554

Date: 13/09/2016

Page 5 sur 6

Ver: 00

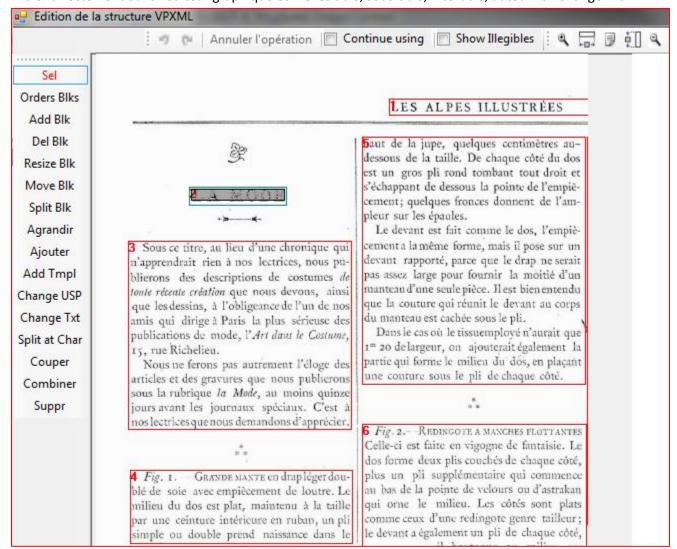
Description des couleurs





CONSIGNES DE TRAITEMENT	Réf : CO_PRD_5554	
	Date: 13/09/2016	
Arald - OCR	Ver : 00	
Etape : OcrView	Page 6 sur 6	

✓ Entrer directement dans l'éditeur graphique et lire les titre, sous-titre, intertitre, auteurvia "Change Txt"



✓ Fermer l'application à la fin de la lecture