OCR PDF

Commande pour OCRiser un fichier pdf image :

```
ocrlauncher img2ocrpdfcs <fichier source> <fichier destination> <ocrmode>
<jobqueueid> <fichier texte> <fichier config>
```

- <fichier source> : le fichier à OCRiser.
- <fichier desination > : le nom du fichier après OCRisation.
- <ocrmode> : le numéro de la tentative.
- <fichier config > : le fichier de configuration opcsdk.conf

Conversion vers Office

Voici comment tester la conversion en ligne de commande :

```
IMG2OCRWORDCS <fichier source> <fichier destination> <ocrmode>  <jobqueueid>
<fichier texte> <fichier config> <format>
```

- <fichier source> : le fichier à convertir.
- <fichier desination > : le nom du fichier une fois qu'il sera converti.
- <ocrmode> : le numéro de la tentative (pour tester via un terminal mettre 0)
- <jobqueueid> : numéro de travail (pour tester via un terminal mettre 0)
- <fichier texte> : fichier vers lequel sera extrait le texte
- <fichier config > : le fichier de configuration opcsdk.conf
- <format> : format vers lequel la conversion sera effectuée

/!\ Si vous souhaitez convertir vers le format **docx** vérifiez que le **.Net Framework 3.5** est bien installé sur votre serveur.

Sinon en testant la conversion en ligne de commande vous aurez un joli message d'erreur :

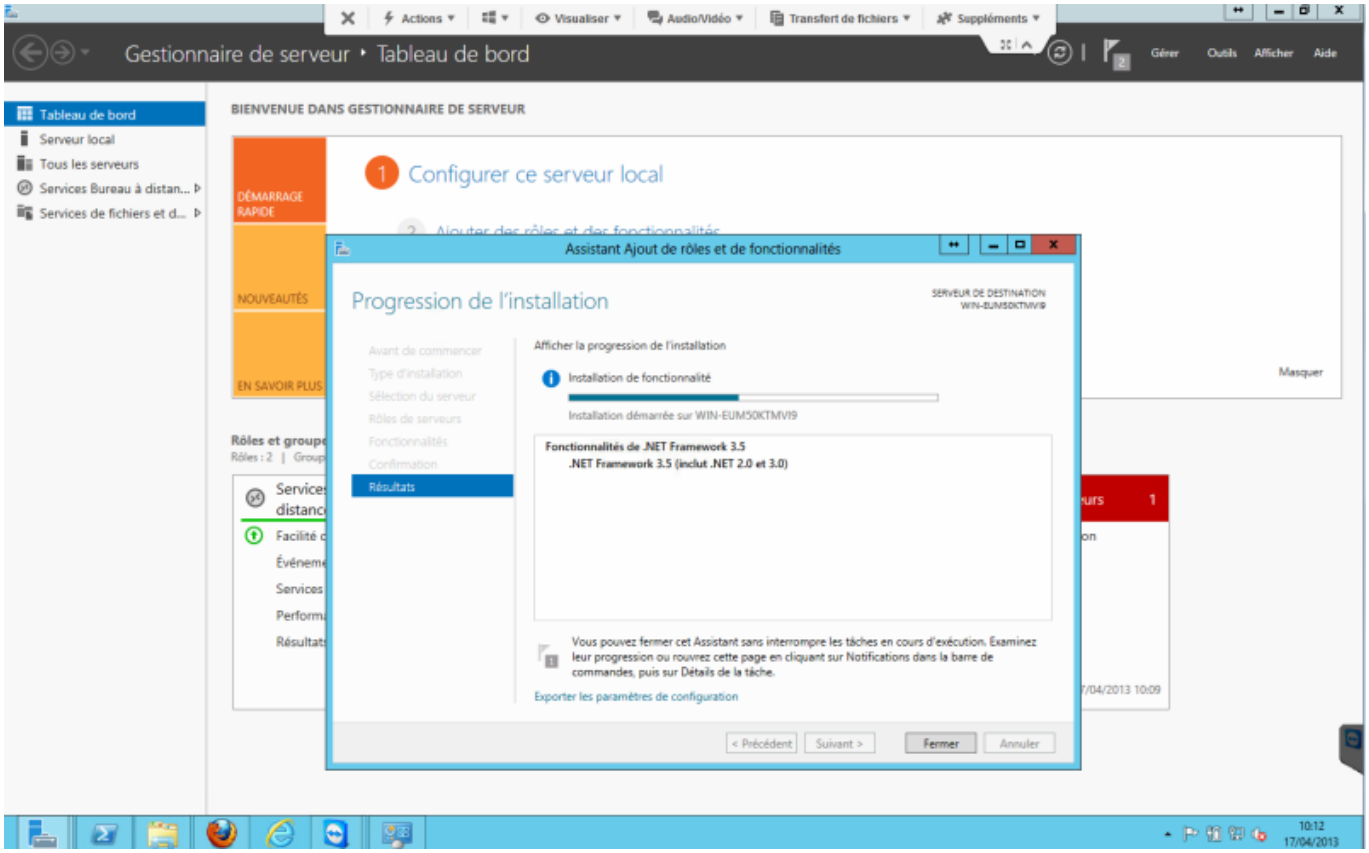
```
Erreur 18 :API_ERROR_ERR
```

Problèmes d'OCR et résolutions

Pas d'OCR sur certains documents

Symptôme : Des documents types .docx, .msg voire certains pdf, ne sont pas OCRisés.

Résolution: Vérifiez que le .NET Framework 3.5 est bien installé.



From:
http://wiki.ezdev.fr/ - **EzGED Wiki**

Permanent link:
http://wiki.ezdev.fr/doku.php?id=ocr&rev=1389882974

Last update: **2023/03/17 09:56**