

## CAPTURE D'INFORMATIONS ( Source Internet )

Ce premier document répond à deux objectifs :

- ① Présenter des compétences comme "inventeur de solutions".

Afin d'aborder cette compétence, l'idée retenue met au premier plan la logique de lecture (l'analyse) d'un problème posé. Internet permettant d'accéder à de nombreuses sources d'informations, il s'agit avant tout de démontrer que tout assemblage d'informations peut être interprété pour en reconstituer la source de données.

- ② Présenter des compétences de "formation".

Afin d'aborder cette compétence, j'ai conçu cette note à l'image d'un support de cours technique. Les différentes manipulations ont donc été détaillées en raccourcis clavier afin d'être pratiquées comme exercice de formation à Microsoft EXCEL.

Nous prendrons l'exemple d'un annuaire institutionnel produit par le "Syndicat des Energies Renouvelables".  
[http://www.enr.fr/docs/2010140532\\_CPannuairePVannuaire.pdf](http://www.enr.fr/docs/2010140532_CPannuairePVannuaire.pdf)

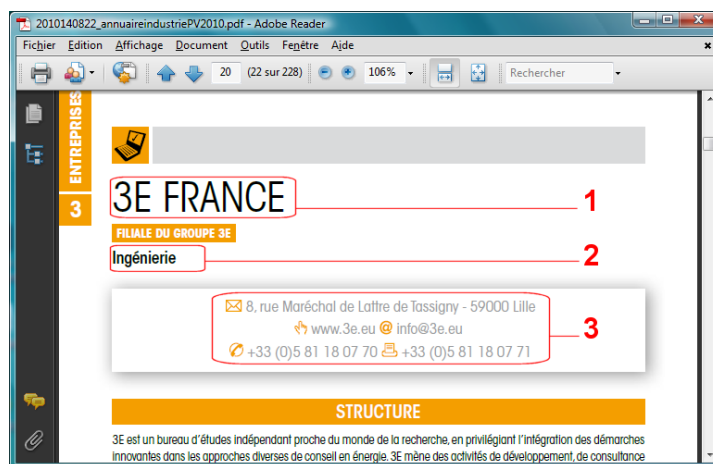
A la lecture du document PDF de 228 pages, l'on peut remarquer que la position des adresses est récurrente :

- 1 - Le nom de l'entreprise
- 2 - L'activité de l'entreprise
- 3 - L'adresse de l'entreprise organisée en trois lignes (Adresse postale, Internet, Téléphones)
- 4 - Le mot "STRUCTURE", qui ponctue systématiquement ces trois blocs d'informations.

Notons que cette récurrence est bien souvent nécessaire à la lisibilité du document.

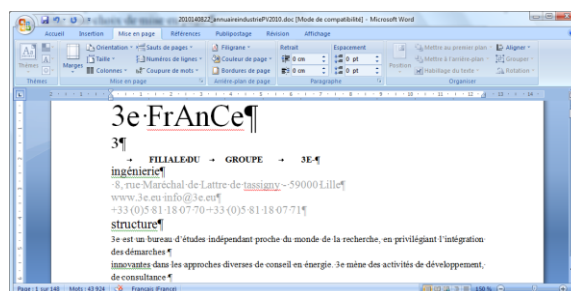
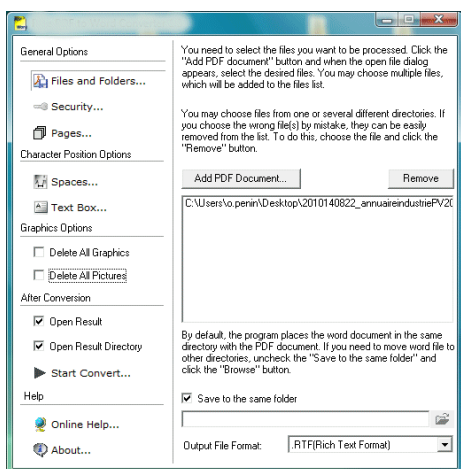
Elle correspond donc à un choix de mise en page que toute publication de type annuaire impose.

*(Cet exemple est donc transposable à d'autres supports)*

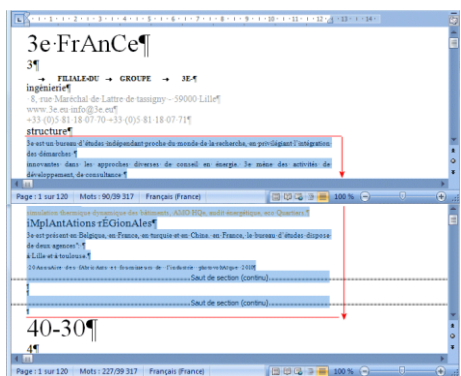


## RECUPERATION DE L'INFORMATION EN TRAITEMENT DE TEXTE

A l'aide d'un simple logiciel de conversion de fichiers PDF en Texte, et un logiciel de traitement de texte,



supprimer les textes issus  
des 19 premières pages  
(introduction au document)  
et des 19 dernières pages (Index et Notes).



Se positionner après la fin de paragraphe du mot "Structure".

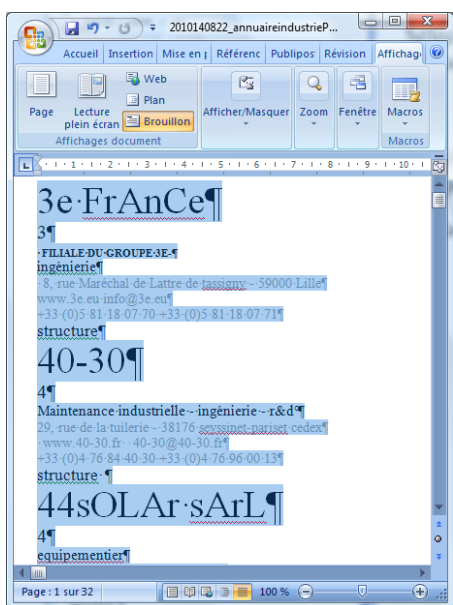
Sélectionner à l'aide du clavier ([SHIFT]+[FLECHE BAS])  
jusqu'au dernier paragraphe avant l'entreprise suivante.

Supprimer cette sélection ([SUPPR])

Réitérer autant de fois que nécessaire en se positionnant  
après la fin de paragraphe du mot "Structure" suivant.

*17 minutes suffisent pour réduire ce document de 120 à 32 pages.*

Placer ensuite ces données dans votre tableur...

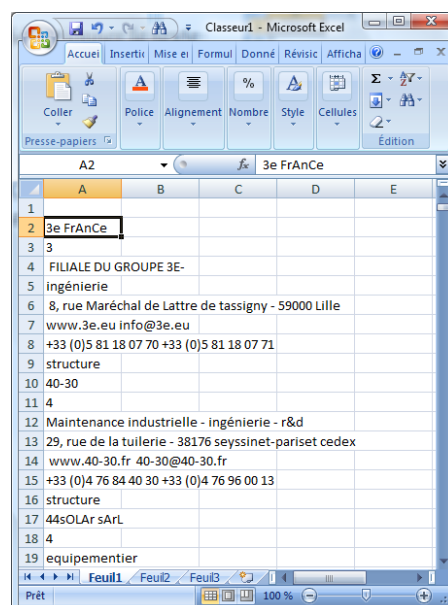


Sélectionner le texte  
([CTRL]+[A])

Copier le texte  
([CTRL]+[C])



Ouvrir votre tableur  
et coller ce contenu  
à partir de la cellule [A2]  
([CTRL]+[V])



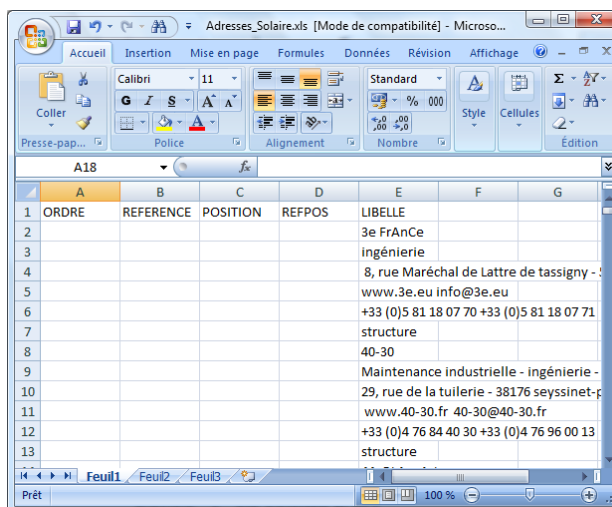
## CONSTITUTION DE LA LISTE D'ADRESSES SOUS MICROSOFT EXCEL

Insérer quatre colonnes...

*L'une est destinée à fixer l'ordre initial de ces lignes, les trois autres pour repérer la position attendue de la ligne dans un tableau croisé.*

Ecrire les libellés

"ORDRE" en cellule [A1],  
 "REFERENCE" en [B1],  
 "POSITION" en [C1],  
 "REFPOS" en [D1],  
 et "LIBELLE" en [E1],



**Dans la colonne [A] (Ordre)**

Ecrire "1" en cellule [A2]  
 et la formule '=A2+1' en cellule [A3]

Se positionner en cellule [A3],

Copier la formule ..... ([CTRL]+[C])

Etendre la zone de collage

Pour atteindre la dernière cellule, ..... [CTRL]+[FLECHE BAS]

[FLECHE DROITE]

[FLECHE DROITE]

[FLECHE DROITE]

[FLECHE DROITE]

[CTRL]+[FLECHE HAUT]

[FLECHE GAUCHE]

[FLECHE GAUCHE]

[FLECHE GAUCHE]

[FLECHE GAUCHE]

Pour sélectionner la zone de collage, ..... [CTRL]+[SHIFT]+[FLECHE HAUT]

Coller la formule ..... [ENTREE]

Copier la sélection ..... [CTRL]+[C]

Coller les valeurs ..... [ALT]+[E] , [G] , [V] , [ENTREE]

### Première série de nettoyage...

Trier les données par ordre croissant sur 'LIBELLE'.

Supprimer les suites de lignes inutiles (chiffres seuls, lettres seules, les références aux filiales)  
*(Autres que les noms d'entreprises, les adresses, les références Internet, les références aux numéros de téléphone et le mot "structure")*

Trier les données par ordre croissant sur 'ORDRE'.

**Dans la colonne [C] (Position)**

Ecrire "1" en cellule [C2]

et la formule '=SI(E3="structure";0;C2+1)' en cellule [C3]

Se positionner en cellule [C3],

Copier la formule .....([CTRL]+[C])

Etendre la zone de collage

Pour atteindre la dernière cellule, ..... [CTRL]+[FLECHE BAS]

[FLECHE DROITE]

[FLECHE DROITE]

[CTRL]+[FLECHE HAUT]

[FLECHE GAUCHE]

[FLECHE GAUCHE]

Pour sélectionner la zone de collage, ..... [CTRL]+[SHIFT]+[FLECHE HAUT]

Coller la formule ..... [ENTREE]

Copier la sélection.....[CTRL]+[C]

Coller les valeurs..... [ALT]+[E] , [G] , [V] , [ENTREE]

**Dans la colonne [B] (Référence)**

Ecrire la formule '=SI(C2=1;E2;B1)' en cellule [B2]

Se positionner en cellule [B2],

Copier la formule .....([CTRL]+[C])

Etendre la zone de collage

Pour atteindre la dernière cellule, ..... [CTRL]+[FLECHE BAS]

[FLECHE DROITE]

[CTRL]+[FLECHE HAUT]

[FLECHE GAUCHE]

Pour sélectionner la zone de collage, ..... [CTRL]+[SHIFT]+[FLECHE HAUT]

Coller la formule ..... [ENTREE]

Copier la sélection.....[CTRL]+[C]

Coller les valeurs..... [ALT]+[E] , [G] , [V] , [ENTREE]

**Dans la colonne [D] (RefPos)**

Ecrire la formule '=CONCATENER(B2;C2)' en cellule [D2]

Se positionner en cellule [D2],

Copier la formule .....([CTRL]+[C])

Etendre la zone de collage

Pour atteindre la dernière cellule, ..... [CTRL]+[FLECHE BAS]

[FLECHE DROITE]

[CTRL]+[FLECHE HAUT]

[FLECHE GAUCHE]

Pour sélectionner la zone de collage, ..... [CTRL]+[SHIFT]+[FLECHE HAUT]

Coller la formule ..... [ENTREE]

Copier la sélection.....[CTRL]+[C]

Coller les valeurs..... [ALT]+[E] , [G] , [V] , [ENTREE]

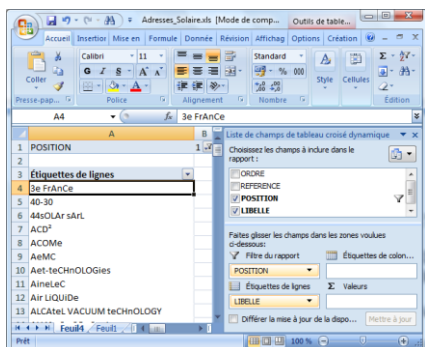
**Ensuite,**

Se placer dans le haut du tableau .....[Début]

Trier les données par ordre croissant sur 'REFPOS'.

### Générer un tableau croisé

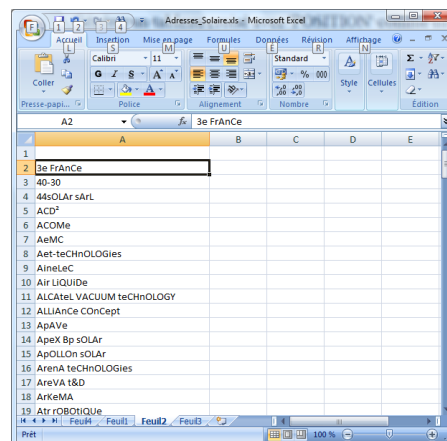
avec 'POSITION' comme filtre de rapport avec la seule valeur "1"  
 et 'LIBELLE' en étiquettes de lignes.



[CTRL]+[SHIFT]  
 +[FLECHE BAS]

[FLECHE HAUT]

Copier les valeurs par [CTRL] + [C]

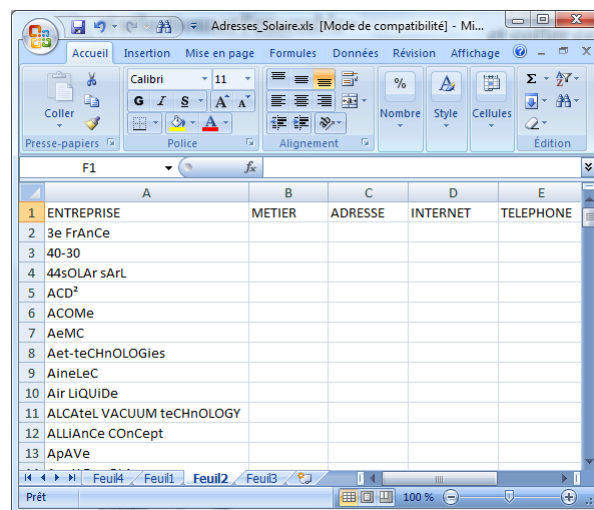


Se positionner sur la cellule [A4],  
 et sélectionner l'ensemble  
 des libellés affichés.

Se placer sur une feuille  
 vide du classeur  
 et coller ce contenu  
 ([CTRL]+[V])

Ecrire les libellés

- "ENTREPRISE" en cellule [A1],
- "METIER" en cellule [B1]
- "ADRESSE" en cellule [C1]
- "INTERNET" en cellule [D1]
- "TELEPHONES" en cellule [E1]



**En cellule [B3]**

Ecrire la formule

'=RECHERCHEV(CONCATENER(\$A2;COLONNE());Feuil1!\$D\$2:\$E\$973;2;0)'

Se positionner en cellule [B3],

Copier la formule ..... ([CTRL]+[C])

Etendre la zone de collage

Pour atteindre la dernière cellule, ..... [CTRL]+[FIN]

Pour sélectionner la zone de collage, ..... [CTRL]+[SHIFT]+[FLECHE HAUT]

[FLECHE BAS]

[CTRL]+[SHIFT]+[FLECHE GAUCHE]

[FLECHE DROITE]

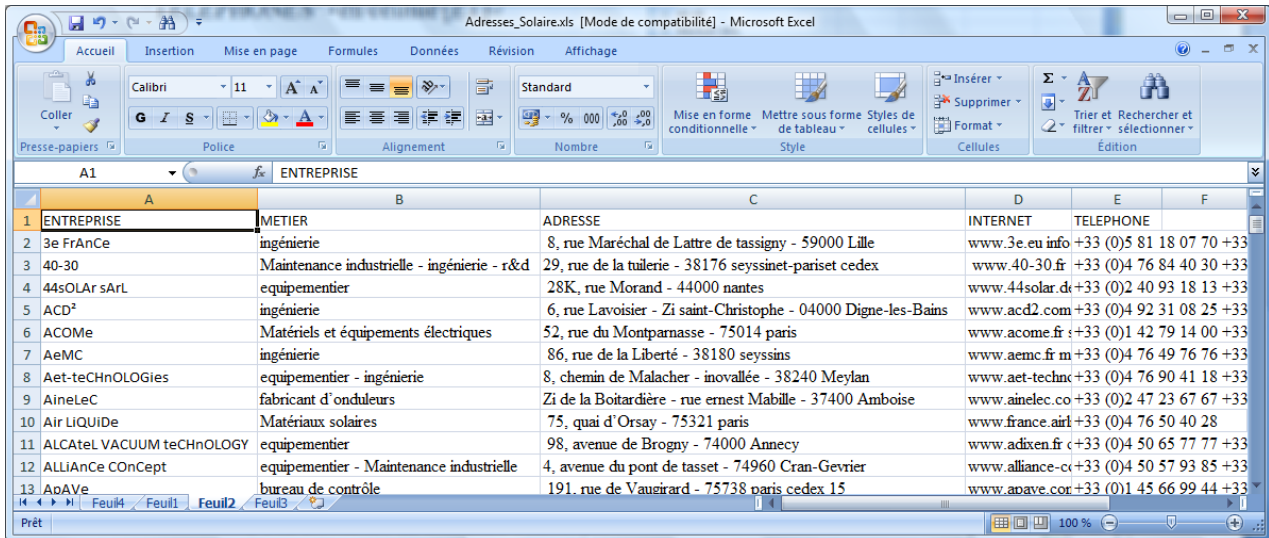
Coller la formule ..... [ENTREE]

Copier la sélection ..... [CTRL]+[C]

Coller les valeurs ..... [ALT]+[E] , [G] , [V] , [ENTREE]

Moins de dix minutes ont été nécessaires pour constituer cette liste de 158 adresses sous Excel... (Ajouter les dix-sept minutes de manipulations antérieures sur traitement de texte... Cf. page 2)

A compter de la récupération de la source des données, moins d'une heure devrait donc suffire pour constituer une liste d'adresse exploitable, et cela quelque soit le nombre de lignes... (Sous réserve des capacités mémoires du poste informatique)



Une seconde série de nettoyage restera nécessaire...

Afin de rendre cette liste réellement exploitable, une autre étape de manipulations devra être appliquée...

...  
*Mise en majuscule, Suppression des espaces inutiles,*  
*Reformatage des adresses, des références Internet et des numéros de téléphones,*  
 ...

