Le thésaurus EUROVOC de la communauté européenne présente l'avantage d'être complet et disponible en de nombreuses langues. S'il est bien adapté à la recherche dans un corpus des débats parlementaires, il l'est moins, en revanche, dans d'autres domaines.

La version 4.7 a été mise en ligne le 15 décembre 2017.

Installation pour Oracle

Télécharger la version 4.7 pour le Français en format iso-2788 : eurovoc_fr.4.7

Déverrouiller le schéma CTXSYS

```
Alter user ctxsys account unlock;
```

Changer le password de CTXSYS

```
Alter user ctxsys identified by <PASSWORD> ;
```

Charger le thesaurus avec la commande en ligne ctxload. Ne pas oublier de positionner la variable NLS_LANG auparavant (essentiel dans le cas d'un thésaurus en Français, par exemple)

set NLS_LANG=French_France.UTF8 ctxload -user ctxsys/<PASSWORD> @hostname -thes -file C:\temp\eurovoc_fr.4.7.txt -name default

Pour supprimer un thesaurus existant, il faut utiliser l'une des commandes

fournies dans le package CTX_THES.

```
begin
ctx_thes.drop_thesaurus('default');
end;
```

Annexes

La version ISO de ce thésaurus a été réalisée avec un traitement sql/Xml disponible dans le sgbd Oracle.

Les sources XML sont accessibles à l'adresse suivante: http://data.europa.eu/euodp/fr/data/dataset/eurovoc

Author



Patrick

GPM Factory