

Table des matières

L'exécution d'Indri.....	2
Indexation	2
Recherche avec Indri	4
L'exécution OpenEphyra.....	7
Permettre OpenEphyra pour soutenir des questions en français.....	8
L'exécution de MEAD.....	9
Codes ALT pour les caractères spéciaux utilisés en français.....	10

Ouvrir un navigateur Web – Chrome, IE

Taper : <http://www.site.uottawa.ca/~diana/tal>

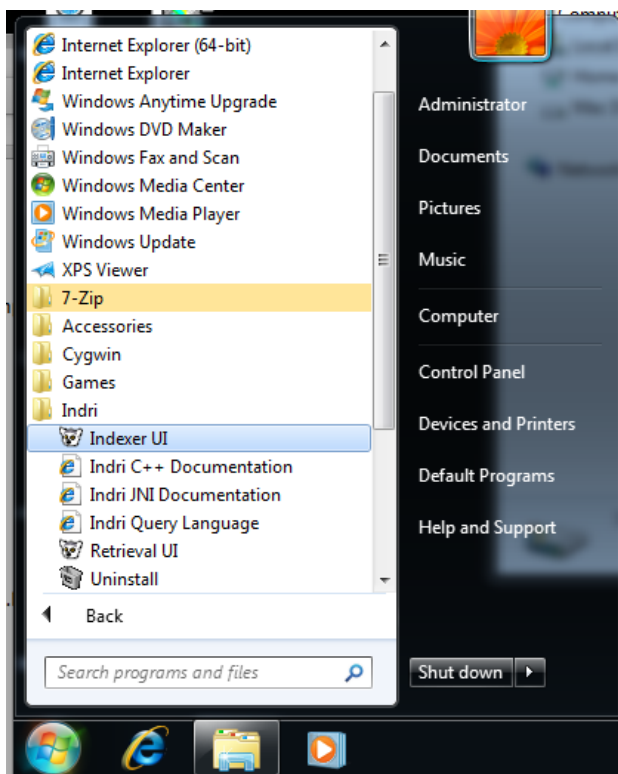
Télécharger les instructions en cliquant sur [InstallationInstructionsFr.txt](#)

L'exécution d'Indri

Utilisation d'Indri est composée de deux parties. Premièrement indexer un corpus qui sera utilisé pour la recherche. Puis en recherchant le corpus.

Indexation

1. Télécharger le corpus français de <http://www.site.uottawa.ca/~diana/tal/AutoSetup.bat>
2. Pour exécuter Indexer UI, dans le menu de démarrage (Start) click Programs > Indri > Indexer IU



3. Dans le premier écran définir les paramètres suivants:

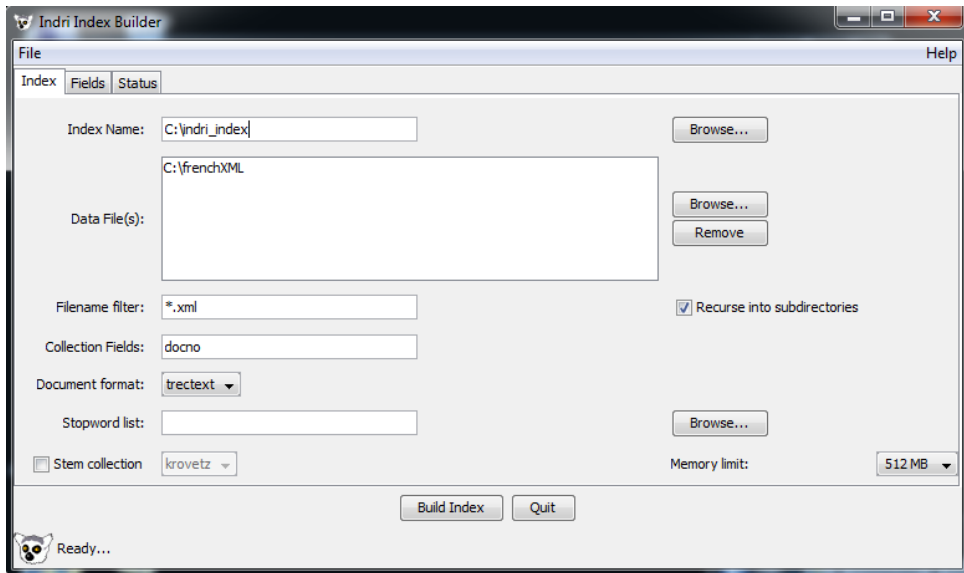
Index Name – c'est là que l'indice sera enregistré. *Retrieval UI* et *OpenEphyra* vont utiliser ce chemin d'accès.

Data File(s) – il s'agit de l'emplacement du corpus à indexer. N'ajouter pas chaque fichier, juste le dossier qui contient tous les fichiers.

Parcourir les sous-répertoires (Recurse into subdirectories) – cocher cette boîte.

Document Format: trectext – le corpus a été élaboré dans ce format.

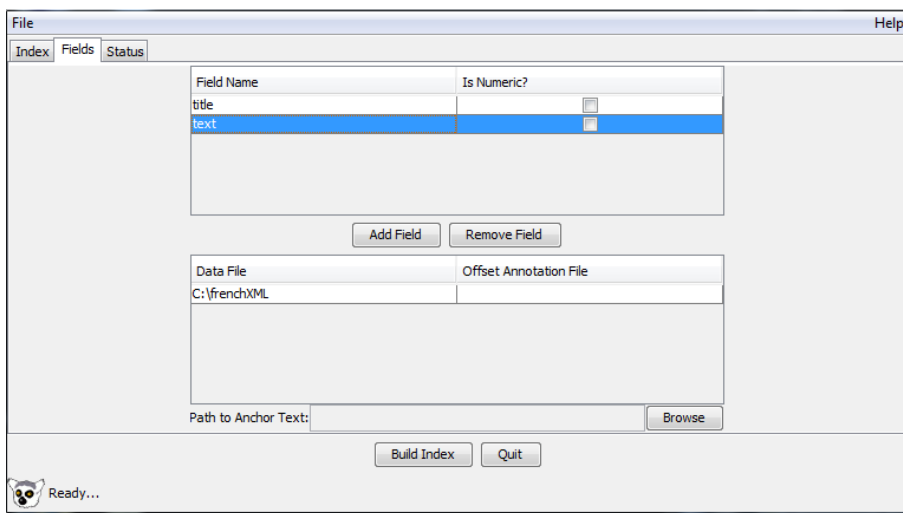
Stemmer: nul ou décocher la boîte – Algorithme de racinisation fourni est seulement pour l'Anglais.



4. Passer à l'onglet de **field** et définir les paramètres suivants

Add Field > title

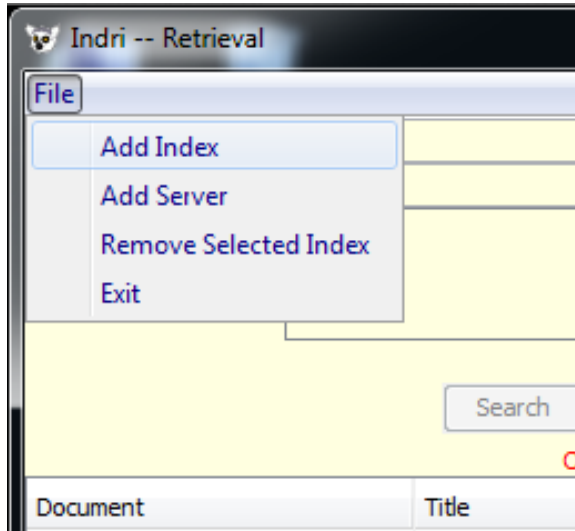
Add Field > text



5. Cliquer Build Index

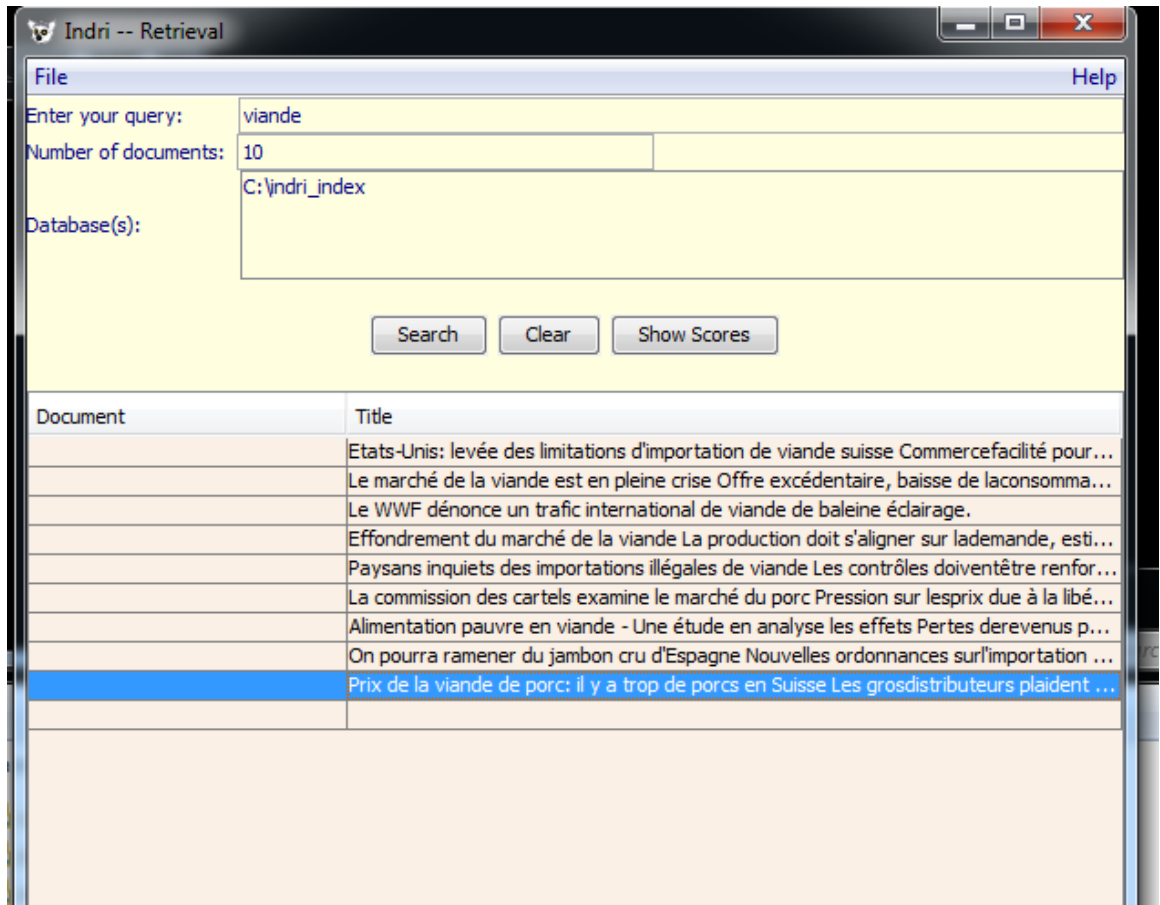
Recherche avec Indri

1. Démarrer le *Retrieval UI*
Dans le menu de démarrage Programs > Indri > Retrieval UI
2. Sélectionner *Add Index* dans le menu de fichier (File)



3. Sélectionner l'emplacement où vous avez créé l'index Indri.
4. Taper termes de recherche dans le champ *Enter your query:*
Soit cliquer *enter* ou cliquer *Search*

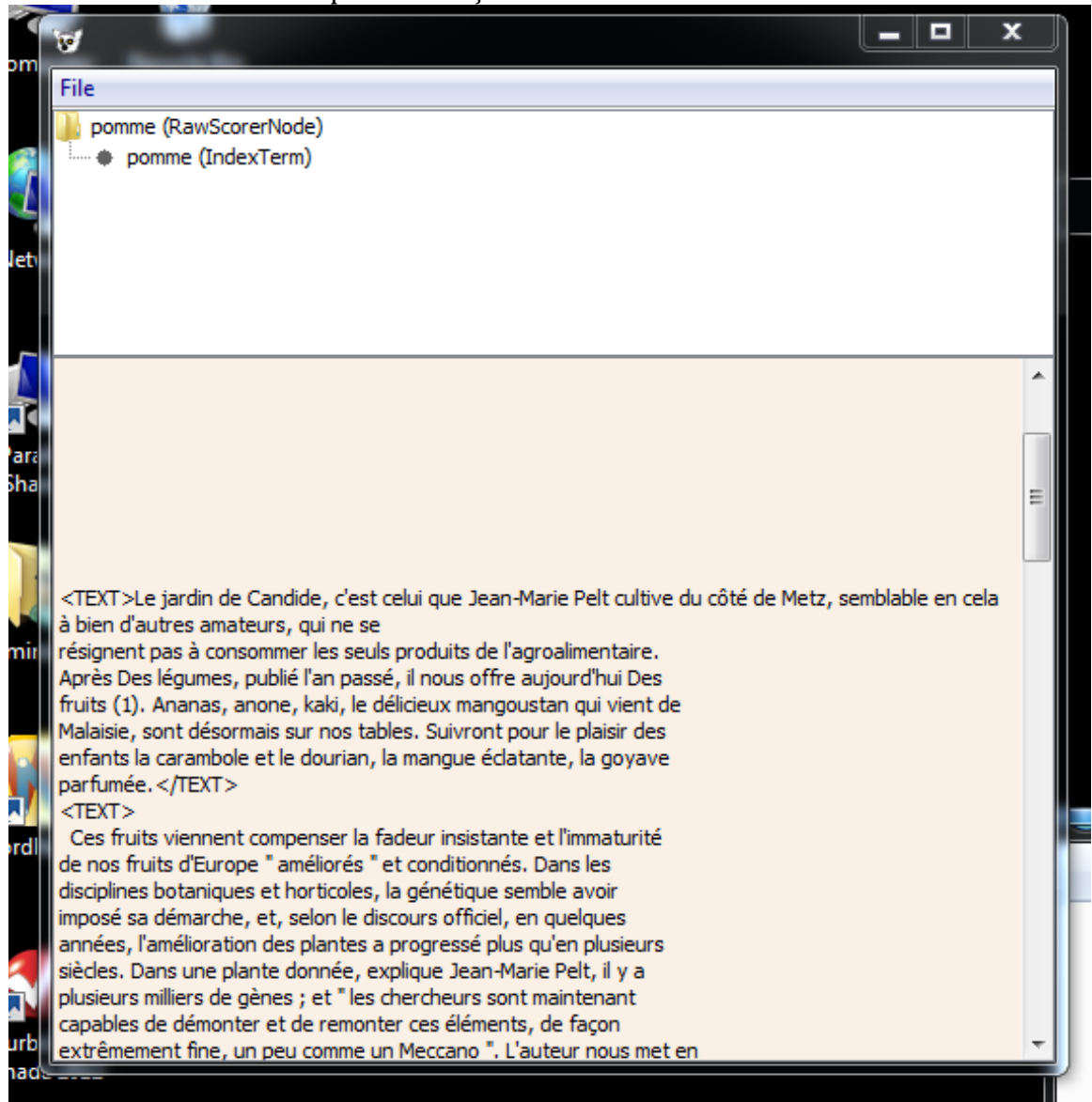
5. Les résultats sont affichés dans le tableau ci-dessous



The screenshot shows the 'Indri -- Retrieval' application window. The interface includes a menu bar with 'File' and 'Help', a search input field containing 'viande', a 'Number of documents' field set to '10', and a 'Database(s)' field containing 'C:\indri_index'. Below these fields are 'Search', 'Clear', and 'Show Scores' buttons. The search results are displayed in a table with two columns: 'Document' and 'Title'. The table contains ten rows of results, with the last row highlighted in blue.

Document	Title
	Etats-Unis: levée des limitations d'importation de viande suisse Commercefacilité pour ...
	Le marché de la viande est en pleine crise Offre excédentaire, baisse de laconsomma...
	Le WWF dénonce un trafic international de viande de baleine éclairage.
	Effondrement du marché de la viande La production doit s'aligner sur lademande, esti...
	Paysans inquiets des importations illégales de viande Les contrôles doiventêtre renfor...
	La commission des cartels examine le marché du porc Pression sur lesprix due à la libé...
	Alimentation pauvre en viande - Une étude en analyse les effets Pertes derevenus p...
	On pourra ramener du jambon cru d'Espagne Nouvelles ordonnances surl'importation ...
	Prix de la viande de porc: il y a trop de porcs en Suisse Les grosdistributeurs plaident ...

6. Cliquer sur un résultat pour voir le document complet
Une boîte de dialogue avec le document original s'affiche.
Le document mettra en évidence où les termes de recherche sont apparus.
Notez cela ne fonctionne pas en français a cause des caractères accentués.



L'exécution OpenEphyra

1. Déclencher un fenêtre de l'explorateur (⌘ + E), choisir Downloads
2. Double-cliquez **launchOE.bat**
3. Taper: **launchOE.bat**
4. Attendre: **Question:**
5. Taper votre question: **Who is Colin Powell?**
Notez que le début de la question doit être en anglais "Qui est" parce que OpenEphyra a été seulement conçu pour une utilisation en anglais.

Permettre OpenEphyra pour soutenir des questions en français

La principale raison pour laquelle OpenEphyra ne peut pas traiter des questions françaises, c'est qu'il ne peut pas identifier le type de réponse attendue. Pour corriger cela, nous devons fournir des modèles pour les types de réponses françaises.

Le type de réponse existe déjà dans OpenEphyra nous avons simplement besoin de faire des motifs avec les expressions régulières (décrites ailleurs)

Dans un éditeur de texte, ouvrir

```
C:\openephyra-0.1.2\res\answertypes\patterns\answertypepatterns
```

La "C:\openephyra-0.1.2" partie du chemin d'accès dépendra de l'endroit où vous installez OpenEphyra.

Le fichier est organisé par type. Si nous faisons défiler à mi-chemin le dossier, nous allons voir la ligne suivante.

```
NEperson                    who(m|se)?
```

La première section indique le type, la seconde est un modèle dans la question qui suggérerait ce type.

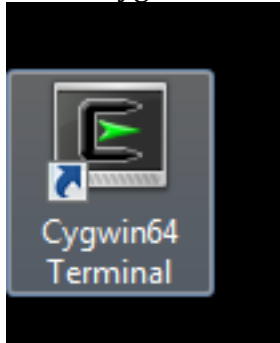
Un équivalent en français pourrait être

```
NEperson                    qui
```


L'exécution de MEAD

Mead doit être exécuté de la ligne de commande Cygwin.

1. Lancer Cygwin



2. Taper: `cp /cygdrive/c/meadrc /cygdrive/q/Mead-3.12/mead/data/mlp2011/meadrc`
3. Changer au répertoire mead. Taper: `cd /cygdrive/q/Mead-3.12/mead`
4. Exécuter Mead. Taper:
`bin/mead.pl -rc data/mlp2011/meadrc -o test.txt`
`data/mlp2011/mlp000`

`bin/mead.pl` – appeler le summarizer Mead

`-rc data/mlp2011/meadrc` – indique d'utiliser le fichier de configuration meadrc. Vous pouvez éditer ce fichier pour modifier la sortie.

`-o test.txt` – enregistré le résumé - vous pouvez voir le résumé en tapant: `more test.txt`

`data/mlp2011/mlp000` – indique d'utiliser l'ensemble des données (ou cluster) mlp2011. Il y a des groupes allant de mlp000 à mlp009

5. Le résumé peut être consulté en tapant: `more test.txt`

Codes ALT pour les caractères spéciaux utilisés en français

Character	ALT code	Character	ALT code
à	0224	À	0192
ç	0231	Ç	0199
è	0232	È	0200
é	0233	É	0201
ê	0234	Ê	0202
ë	0235	Ë	0203
î	0238	Î	0206
ï	0239	Ï	0207
ô	0244	Ö	0212
ù	0249	Ù	0217
û	0251	Û	0219
ü	0252	Ü	0220