# Indri, OpenEphyra, and Mead

Préparé par: Martin Scaiano and Diana Inkpen  
Dans le cadre d'une introduction aux outils de Natural Language Processing (NLP)

Chaque section de ce document décrit la procédure d'installation de l'un de ces trois outils de traitement de text OpenSource.

Mead est un système de résumé des textes.

Indris est un système de fouille des textes.

OpenEphyra est un système questions-réponses.

# Table of Contents

[Indri, OpenEphyra, and Mead 1](#_Toc373320346)

[Table of Contents 2](#_Toc373320347)

[Indri 3](#_Toc373320348)

[Résumé de l'installation: 3](#_Toc373320349)

[Instructions détaillées de l’installation: 3](#_Toc373320350)

[Vérifier les travaux d'installation et trouble de dépannage: 6](#_Toc373320351)

[OpenEphyra 10](#_Toc373320353)

[Résumé de l'installation: 10](#_Toc373320354)

[Instructions détaillées de l’installation: 10](#_Toc373320355)

[Mead 12](#_Toc373320356)

[Résumé de l'installation: 12](#_Toc373320357)

[Instructions détaillées de l’installation : 12](#_Toc373320358)

[Étapes pour installer Cygwin 12](#_Toc373320359)

[Étapes pour télécharger Mead 15](#_Toc373320360)

[Tester Mead 15](#_Toc373320361)

[Utilisation de Mead 15](#_Toc373320362)

[Ajout du support français 15](#_Toc373320363)

[Ajout du document Français Pour Summarization 15](#_Toc373320364)

[Installer Java 17](#_Toc373320365)

[Rouge 21](#_Toc373320366)

# Indri

Indri est un système de recherche d'information (RI). Il y a des installateurs de Windows, mais pour installer sur d'autres systèmes d'exploitation, vous devez installer à partir des sources.

Pour exécuter l'interface utilisateur graphique Indri (GUI) Java doit être installé.

Indri se trouve à

Page d'accueil: <http://lemurproject.org/indri.php#download>

Source: <http://sourceforge.net/projects/lemur/>

## Résumé de l'installation:

1. Install Java
2. Download and Install Indri

## Instructions détaillées de l’installation:

1. Aller à <http://sourceforge.net/projects/lemur/>

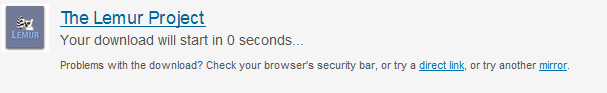
Macintosh HD:Users:martinscaiano:Desktop:Screen Shot 2013-09-21 at 3.59.36 PM.png

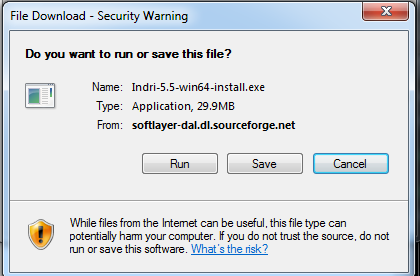
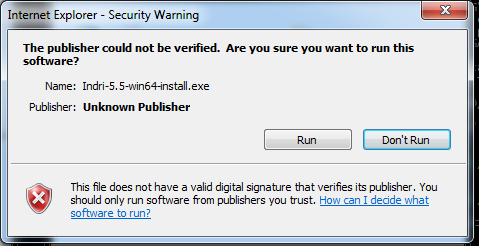
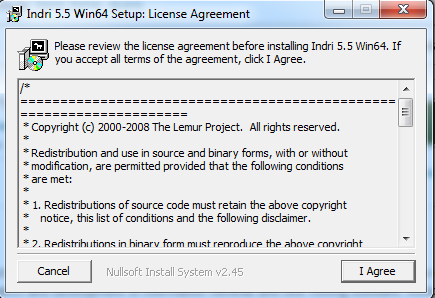
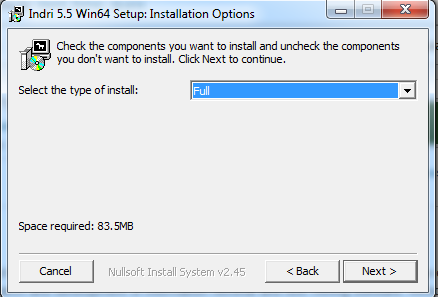
1. Cliquer **Download**

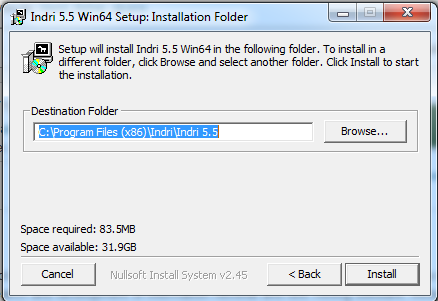
Cela doit sélectionner une version qui devrait fonctionner sur votre système d'exploitation.

Pour sélectionner manuellement la version à télécharger (source, Window 32 bit, or Windows 64 bit) aller à <http://sourceforge.net/projects/lemur/files/lemur/indri-5.5/>

1. Attendre quelques secondes pour que le téléchargement commence, s’il ne démarre pas cliquer sur le lien direct

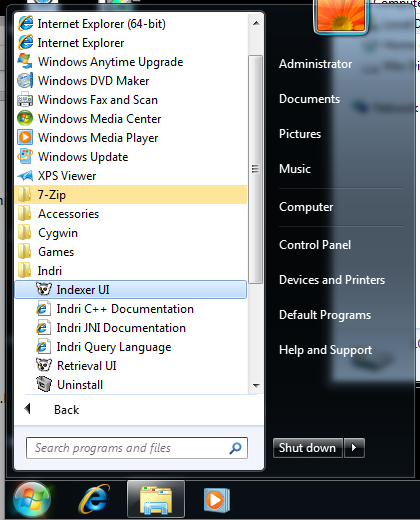


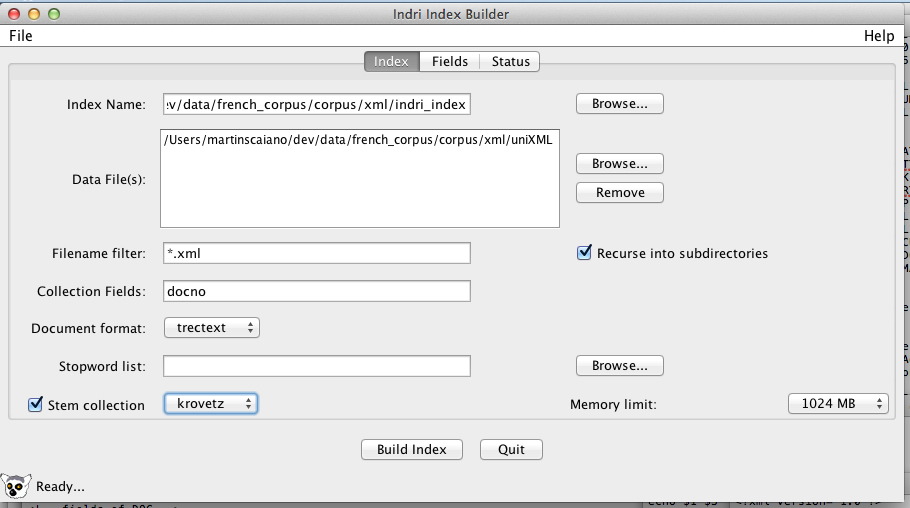
1. Cliquer Run
2. Cliquer Run de nouveau si nécessaire 
3. Accepter la licence 
4. Sélectionner l'installation complète (Full) 
5. Installer à l'emplacement par défaut

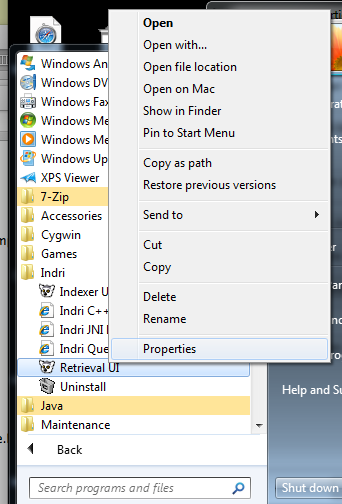
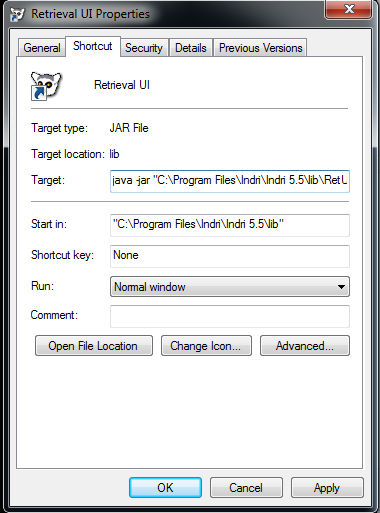
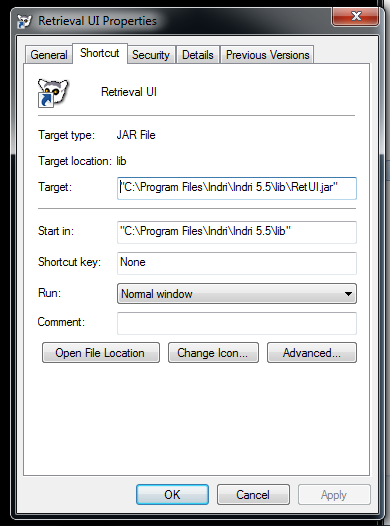
Rappelez-vous où Indri est installé parce que OpenEphyra nécessitera certains fichiers à copier.

## Vérifier les travaux d'installation et trouble de dépannage:

Pour vérifier l'installation, nous allons exécuter l'interface utilisateur d'indexation Indri. Si cela ne fonctionne pas, mais vous avez installé Java correctement il y a une solution simple.

1. À partir du menu Démarrer (Start) cliquer: Programs > Indri > Indexer UI
2. Si vous voyez l'écran ci-dessous, Indri fonctionne bien; ne suivez pas d'autres instructions.  
   Si vous ne voyez pas l'écran ci-dessous, suivez les étapes 3 à 5 pour configurer Windows pour exécuter l'indexeur et l'interface de récupération

utilisant Java. 

1. Faire un clic avec la souris de droite sur *Indexer UI* dans le menu de départ (Start): Programs > Indri > Indexer UI   
   Puis cliquer sur *properties*  
   **
2.  Au début de cible (Target) ajouter: java –jar   
   Le premier écran est avant le changement, la seconde après.

# Répéter les instructions 3-4 pour *Retrieval UI*.

# OpenEphyra

OpenEphyra est un système questions-réponses Open Source écrit en Java. OpenEphyra n'a pas eu de mises à jour récemment et n'est plus compatible avec Google, Bing, ou les API de recherche de Yahoo. Ainsi OpenEphyra ne peut pas fonctionner comme un système questions-réponses en ligne, ni sa fonction première de projection. OpenEphyra fonctionne encore bien sur l'indice corpus déconnecté.

La version de OpenEphyra que nous fournissons a été modifiée pour supprimer les fonctions qui ne fonctionnent pas. Ce document décrit comment installer notre version modifiée de OpenEphyra.

Code source OpenEphyra et documents peuvent être trouvés à:

Page d'accueil: <http://www.ephyra.info>

Documentation: <https://mu.lti.cs.cmu.edu/trac/Ephyra/wiki/WikiStart>

Source: <http://sourceforge.net/projects/openephyra/?source=navbar>

Installation: <https://mu.lti.cs.cmu.edu/trac/Ephyra/wiki/Docs/Tutorials/UsingIndri>

## Résumé de l'installation:

1. Installer Indri – voir la section Indri
2. Installer Java
3. Installer OpenEphyra

## Instructions détaillées de l’installation:

1. Télécharger notre fichier modifié de OpenEphyra à :  
   <http://scaiano.com/download/openephyra-0.1.2.zip>
2. Décompresser le fichier à c:\
3. Copier les fichiers Indri API à OpenEphyra
   1. Si vous avez installés Windows 32 bits Indri
      1. Copier bin/indri\_jni.dll à OpenEphyra/lib/search/
      2. Copier Indri 5.5/src/swig/src/java/indri.jar  
         à OpenEphyra/lib/search/
   2. Si vous avez installés Windows 64 bits Indri
      1. Copier bin/indri\_jni.dll à OpenEphyra/lib/search/
      2. Copier Indri 5.5/swig/obj/java/indri.jar  
         à OpenEphyra/lib/search/

# Mead

Mead est destiné à être installé sur les systèmes Linux ou Unix. Mead est facile à installer, mais difficile à configurer et à utiliser sur de nouveaux ensembles de données. Dans cette démonstration, nous allons fournir des fichiers prétraités à utiliser avec Mead. Les instructions sur la façon de préparer un document se trouvent dans la documentation de l'utilisateur Mead.

<http://www.summarization.com/mead/documentation/meaddoc.pdf>

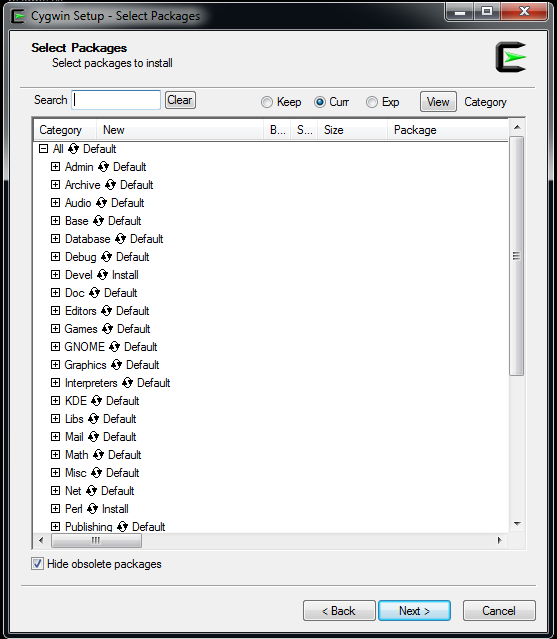
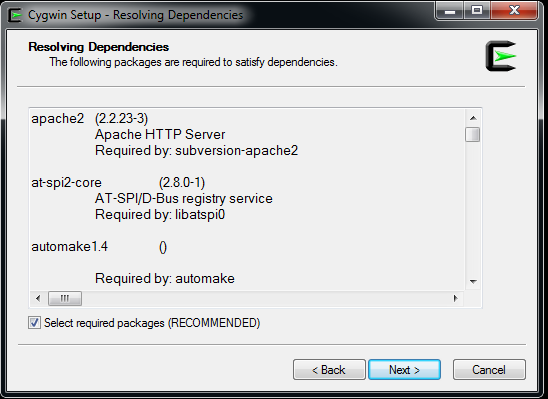
Ce document décrit comment installer Mead sur Windows.

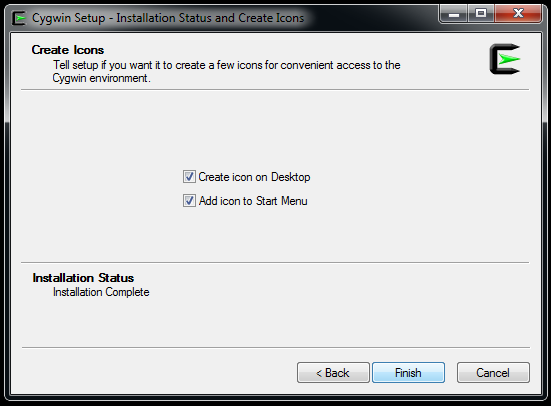
## Résumé de l'installation:

1. Installer Cygwin avec les fichers *devel* et *perl*
2. Copier et décompresser MEAD-3.12.tar.gz sur c:
3. Exécuter Install.pl

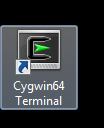
## Instructions détaillées de l’installation :

### Étapes pour installer Cygwin

1. Télécharger le logiciel de Cygwin à partir de <http://cygwin.com>
2. Exécuter le programme d'installation
3. Sélectionner et installer le**s fichers devel** and **perl**
4. Accepter toutes les dépendances 
5. Cocher **Create Icon on Desktop** et **Add icon to Start Menu**



### Étapes pour télécharger Mead

1. Télécharger Mead à partir de <http://www.summarization.com/mead/>
2. Décompresser le fichier à C:\mead
3. Ouvrir Cygwin invite de commande (l'icône est illustré ci-dessous)
4. 
5. Changer de répertoire. Taper: cd /cygdrive/c/mead
6. Exécuter le script d'installation. Taper: perl Install.PL

### Tester Mead

1. Exécuter le résumé GA. Taper: bin/mead.pl GA3
2. 20 phrases doivent être retournées comme un résumé du texte GA3.

## Utilisation de Mead

1. Changer de répertoire. Taper: cd /cygdrive/c/mead
2. Pour exécuter Mead sur les données GA3 taper: bin/mead.pl GA3
   1. GA3 et tout autre document se trouvant dans le répertoire de données dans le répertoire mead: /cygdrive/c/mead/data/GA3

## Ajout du support français

1. Télécharger le document idf (*Inverse Document Frequency*) français. <http://Scaiano.com/edu/fridf>
2. Copier fridf dans mead/bin/

## Ajout du document Français Pour Summarization

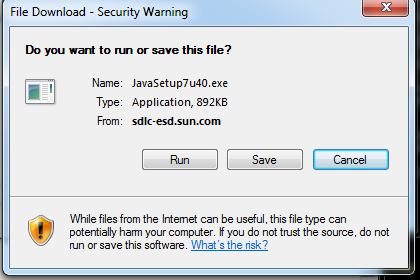
1. Télécharger des documents en français pour résumé: <http://Scaiano.com/edu/mlp2011.zip>
2. Décompresser dans c:\mean\data\mlp2011
   1. Notez que nous avons déjà prétraité les fichiers pour vous et créé un fichier de configuration de base.
   2. Pour les instructions de préparation de votre propre document voir la section 18 de la documentation utilisateur Mead.
   3. Pour obtenir des instructions de configuration, voir les sections 3 et 4
   4. Les fichiers de configuration ont déjà été installés pour utiliser le fichier idf français (fridf) que vous avez téléchargé précédemment
3. Pour exécuter mlp2011 (MultiLing Pilot 2011)
   1. Changer au répertoire mead. Taper: cd /cygdrive/c/mead
   2. Taper: bin/mead.pl -rc data/mlp2011/meadrc data/mlp2011/mlp000

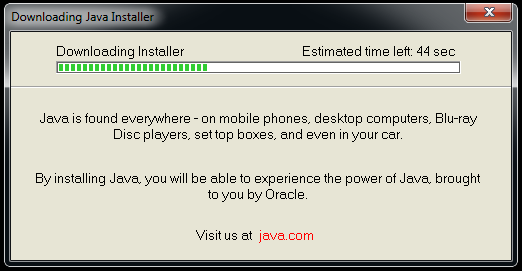
# Installer Java

1. Aller à <http://www.java.com/>

Macintosh HD:Users:martinscaiano:Desktop:Screen Shot 2013-09-21 at 2.54.02 PM.png

1. Cliquer *Free Java Download*
2. Macintosh HD:Users:martinscaiano:Desktop:Screen Shot 2013-09-21 at 2.55.07 PM.pngCliquer *Agree and Start Free Download*
3. Choisir *Run* si l'option est présentée



1. Cliquer Install
2. Attendre la fin du téléchargement 
3. Décocher *Include Free McAfee…*  
   N'installez pas plus que vous souhaitez ou avez besoin. Les offres supplémentaires ne vous aideront pas davantage.



1. Attendre que l'installation soit terminée
2. Installation terminée 

# Rouge

Rouge est une évaluation automatique de résumé automatique. Il permet des comparaisons faciles mais pas intelligentes des divers résumés à un ensemble de résumés humains (ou).

Rouge se trouve à:

<http://www.berouge.com/Pages/default.aspx>