# Indri, OpenEphyra, and Mead

Prepared by: Martin Scaiano and Diana Inkpen  
As a part of an introduction to Natural Language Processing (NLP) tools

Each section of this document will outline installing one of these three Open Source NLP tools: Indri, OpenEphyra, and Mead

Indri is an information retrieval system.

OpenEphyra is a question answering system.

Mead is a document summarization system.

Table of Contents

Installing and Using Mead, Indri, and OpenEphyra 1

Indri 3

Installation Overview 3

Detailed Installation Instructions: 3

Verify the installation works and trouble shooting: 6

OpenEphyra 10

Installation Overview: 10

Detailed Installation Instructions: 10

Mead 12

Installation Overview: 12

Detailed Installation Instructions: 12

Steps to Install Cygwin 12

Steps to Mead 15

Test Mead 15

Using Mead 15

Adding French Support 15

Adding French Document For Summarization 15

Install Java 17

Rouge 22

# Indri

Indri is an Information Retrieval (IR) system. It has windows installers, but to install on other operating systems you will need to install from source.

To run Indri graphical user interface (GUI) java must be installed.

Indri can be found at

Homepage: <http://lemurproject.org/indri.php#download>

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

## Installation Overview

1. Install Java (see page 16)
2. Download and Install Indri

## Detailed Installation Instructions:

1. Go to <http://sourceforge.net/projects/lemur/>

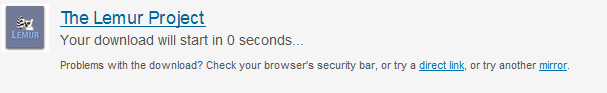
Macintosh HD:Users:martinscaiano:Desktop:Screen Shot 2013-09-21 at 3.59.36 PM.png

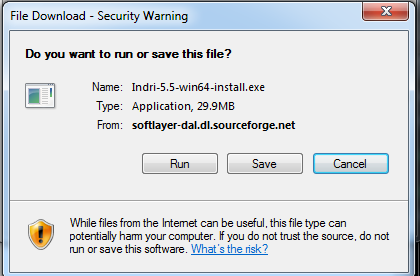
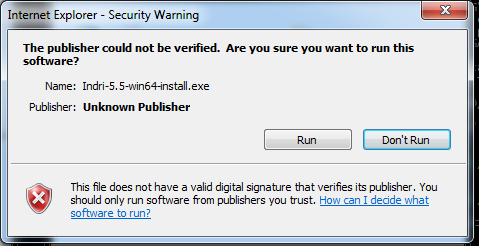
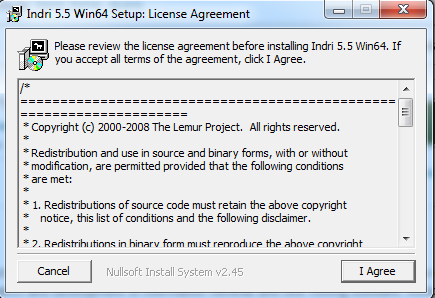
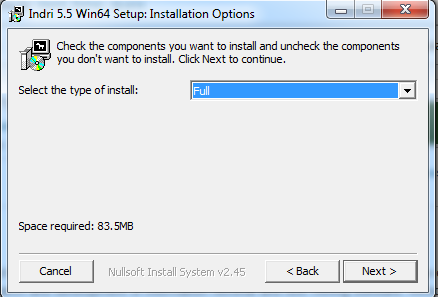
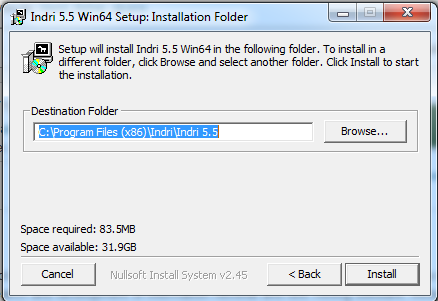
1. Click Download

This will select a version that should work on your operating system.

To manually select the version to download (source, Window 32 bit, or Windows 64 bit) go to <http://sourceforge.net/projects/lemur/files/lemur/indri-5.5/>

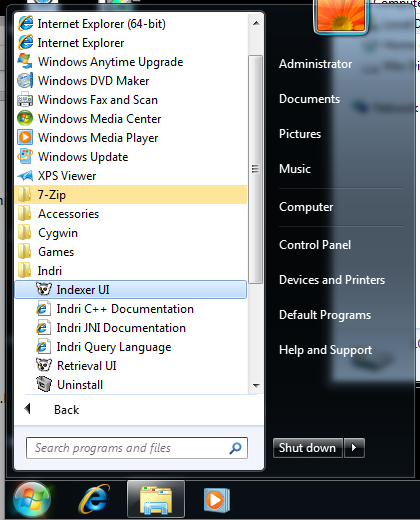
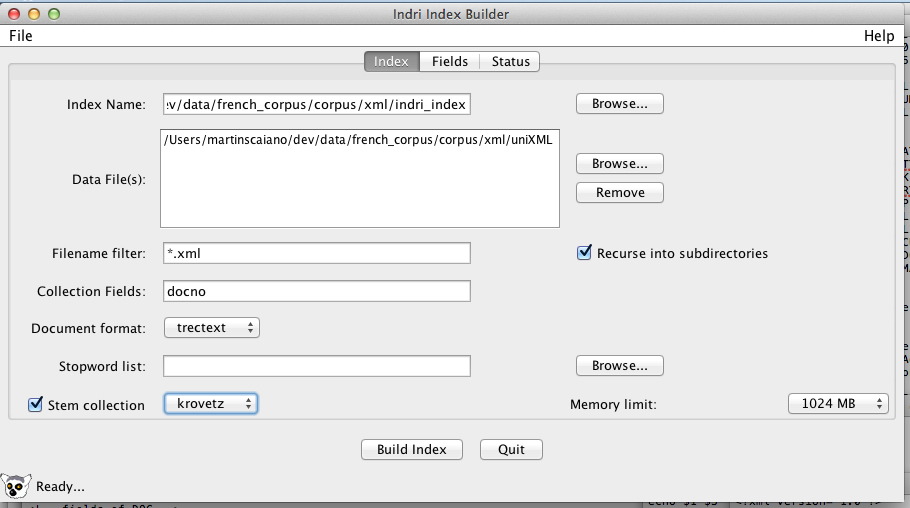
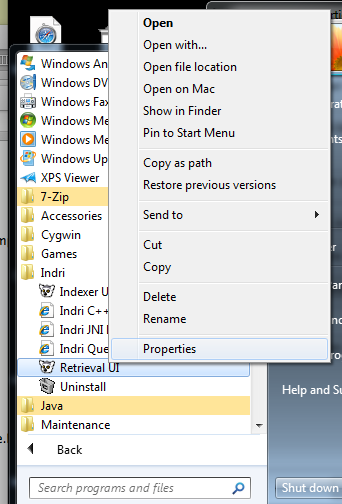
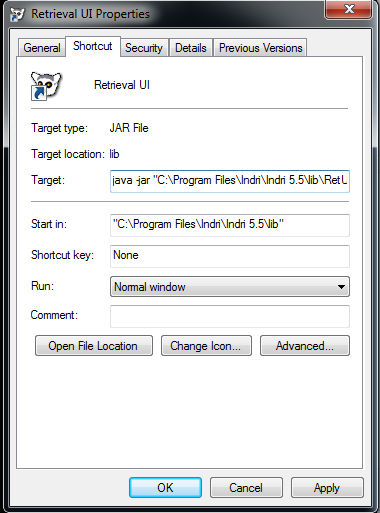
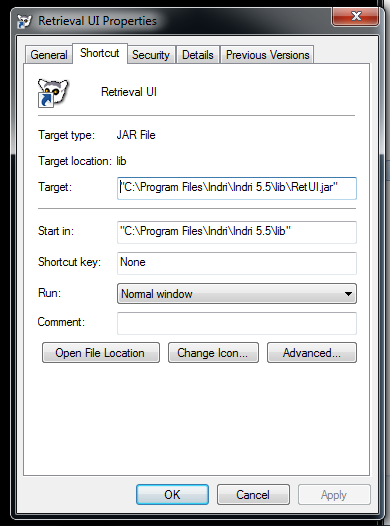
1. Wait a few seconds for the download to start, if it does not start downloading then click the *direct link*



1. Click Run
2. Click Run again if needed
3. Agree to the license
4. Select full installation 
5. Install to the default location  
   Remember where Indri is installed because OpenEphyra will require some of the files to be copied.

## Verify the installation works and trouble shooting:

To verify the Install, we will run the Indri Indexing user interface. If it does not work but you install java correctly there is a simple solution.

1. From the start menu, click: Programs > Indri > Indexer UI
2. If you see the screen below, Indri is working fine; do not follow further instructions.  
   If you do not see the screen below, steps 3 to 5 to set Windows to run the *indexer* and *retrieval* UI using Java. 
3. Right click on *Indexer UI* in the start menu: Programs > Indri > Indexer UI   
   Then click *properties*  
   **
4. At the beginning of target add: java –jar   
   The first screen is before the change and the second after

Repeat Instructions 3-4 for *Retrieval UI*.

# OpenEphyra

OpenEphyra is Open Source Question Answering written in java. OpenEphyra has not been maintained recently and is no longer compatible with Google’s, Bing’s, or yahoo’s search APIs. Thus OpenEphyra cannot function as an online QA system, nor can its key projection feature function. OpenEphyra still functions well on index offline corpora.

The version of OpenEphyra that we provide has been modified to remove the now non-function online search engine related features. This document outlines how to install our modified version of OpenEphyra.

OpenEphyra source code and document can be found at:

Homepage: <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>

## Installation Overview:

1. Install Indri – see Indri section
2. Install Java
3. Install OpenEphyra

## Detailed Installation Instructions:

1. Download Our modified OpenEphyra package at  
   <http://scaiano.com/download/openephyra-0.1.2.zip>
2. Decompress the file to c:\
3. Copy the Indri API files to OpenEphyra
   1. If you Installed the Windows 32 bit Indri
      1. Copy bin/indri\_jni.dll to OpenEphyra/lib/search/
      2. Copy Indri 5.5/src/swig/src/java/indri.jar  
         to OpenEphyra/lib/search/
   2. If you installed the Windows 64 Indri
      1. Copy bin/indri\_jni.dll to OpenEphyra/lib/search/
      2. Copy Indri 5.5/swig/obj/java/indri.jar  
         to OpenEphyra/lib/search/

# Mead

Mead is intended to be installed on Linux or Unix compatible operating systems. Mead is easy to install but difficult to configure and use on new data sets. In this demonstration we will provide preprocessed files to use with the Mead summarizer. Instructions on how to prepare document can be found in the Mead user documentation.

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

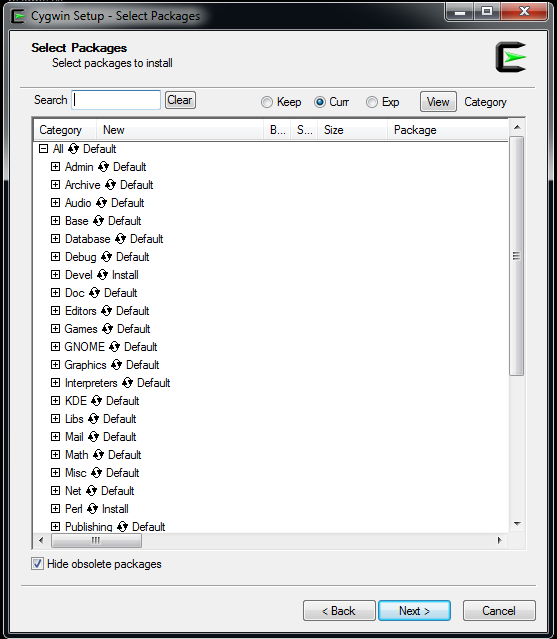
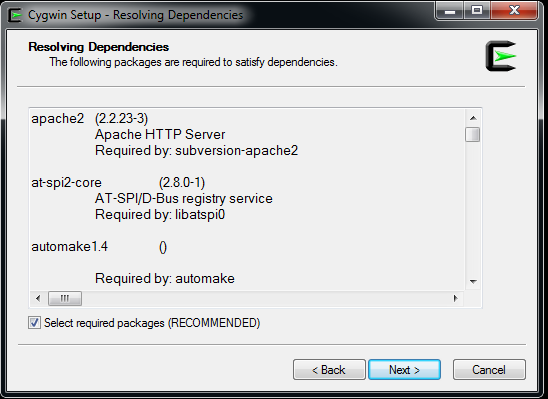
This document outlines how to install Mead on Windows.

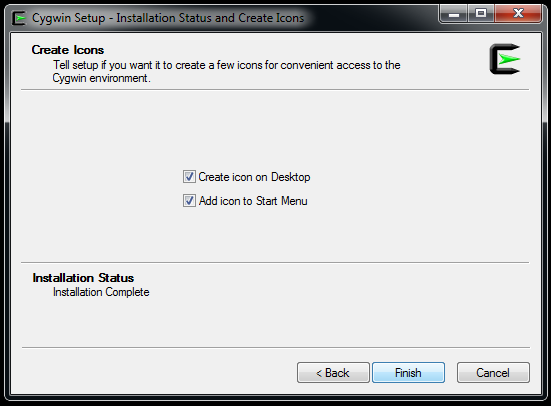
## Installation Overview:

1. Install Cygwin with the *devel* and *perl* packages
2. Copy and decompress MEAD-3.12.tar.gz onto c:
3. Run Install.pl

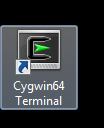
## Detailed Installation Instructions:

### Steps to Install Cygwin

1. Download Cygwin setup from <http://cygwin.com>
2. Run the setup program
3. Select and Install the **devel** and **perl** packages
4. Agree to any dependencies
5. On the final screen check Create **Icon on Desktop** and **Add icon to Start Menu**



### Steps to Mead

1. Download Mead from <http://www.summarization.com/mead/>
2. Decompress the file to C:\mead
3. Open Cygwin command Prompt (the desktop icon is shown below) 
4. Move to the mead directory. Type: cd /cygdrive/c/mead
5. Run the Installation script. Type: perl Install.PL

### Test Mead

1. Run the GA summarization. Type: bin/mead.pl GA3
2. 20 sentences should be returned as a summary of the GA3 text.

## Using Mead

1. Move to the mead directory. Type: cd /cygdrive/c/mead
2. To run mead on the GA3 data Type: bin/mead.pl GA3
   1. GA3 and other document are located into data directory under the mead directory: /cygdrive/c/mead/data/GA3

## Adding French Support

1. Download the French idf document. <http://Scaiano.com/edu/fridf>
2. Copy fridf into the mead/bin/

## Adding French Document For Summarization

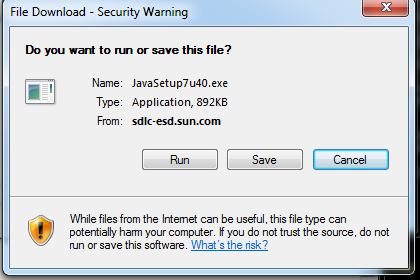
1. Download of French documents for summarization: <http://Scaiano.com/edu/mlp2011.zip>
2. Decompress them in c:\mean\data\mlp2011
   1. Note that we have already precprocessed the files for you and setup a basic configuration file.
   2. For instructions to prepare your own document set the section 18 of the Mead User Documentation
   3. For configuration instructions set sections 3 and 4
   4. The config files have already been set to use the French idf file (fridf) that you downloaded previously
3. To run mlp2011 (MultiLing Pilot 2011)
   1. Move to the mead directory. Type: cd /cygdrive/c/mead
   2. Type: bin/mead.pl -rc data/mlp2011/meadrc data/mlp2011/mlp000

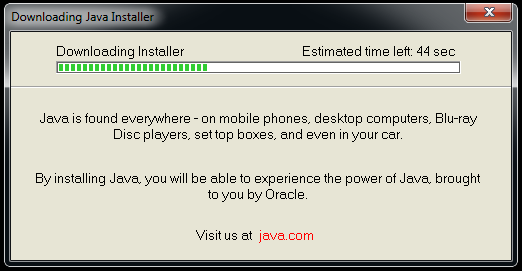
# Install Java

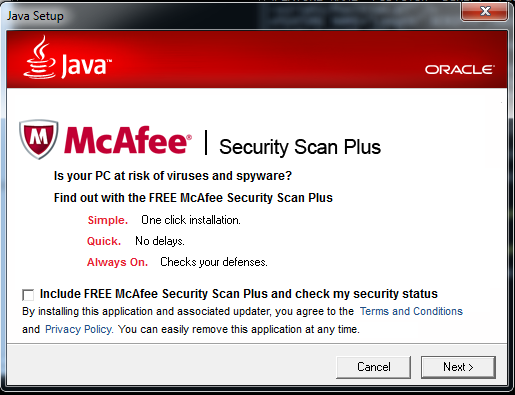
1. go to <http://www.java.com/>

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

1. Click *Free Java Download*
2. Macintosh HD:Users:martinscaiano:Desktop:Screen Shot 2013-09-21 at 2.55.07 PM.pngClick *Agree and Start Free Download*
3. Select *Run* if the option is presented



1. Click Install
2. Wait for the download to complete
3. Uncheck *Include Free McAfee…*  
   Don’t install more than you want or need. Most additional offers will not help you computer.



1. Wait for Installation to complete
2. Installation Complete

# Rouge

Rouge is an automatic evaluation for automatic summarization. It allows for easy though not intelligent comparison of various summaries to a set of human (or gold) summaries.

Rouge can be found at:

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