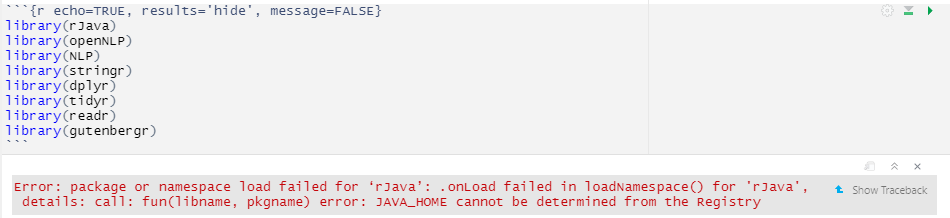
## JAVA and NER

Prologue: We will be doing Natural Language Processing using openNLP. This is a package that uses NLP libraries written in Java. In order to use these, you will need to install Java on your system. You can do so by clicking here: <https://java.com/en/download/>

You can opt to install it before your next step, or if you are not sure whether you have Java on your computer go through the setup below. It will throw an error when trying to launch **rJava**. At that point, you’ll have to install it anyway.

1. Install all rJava, openNLP, and NLP
2. Try loading libraries



If you get this threatening error message, you don’t have Java properly installed.

rJava is a package that connects the R language with the Java language. In order for this to work, you need to have Java installed on your system.

Go to: <https://www.java.com/en/download/manual.jsp>

1. Make sure you install the 64-bit version for Windows, here: <https://javadl.oracle.com/webapps/download/AutoDL?BundleId=245060_d3c52aa6bfa54d3ca74e617f18309292>
2. Load the libraries again.
   1. You might get the following error:



* 1. It makes no functional difference. If you want to fix it you have to “set Java on your PATH”. This basically tells all pieces of software where Java is installed on your computer so they can use java.exe.
     1. If it doesn’t work trying setting the path in R manually using this command in console:

Sys.setenv(JAVA\_HOME="C:\\Program Files\\Java\\jre1.8.0\_301\\")

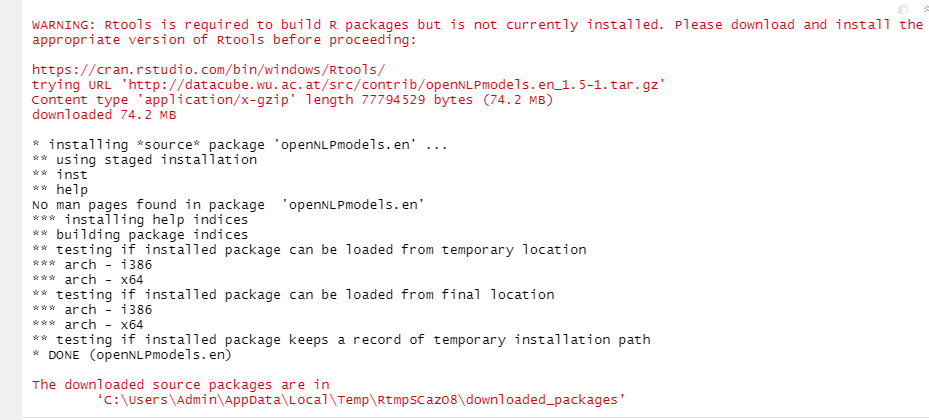
or wherever your Java directory is.

* + 1. This fix is temporary. In order to permanently add Java to your PATH you will have to go into advanced system settings. There’s a good overview here. Just remember that the installed java version is **jre1.8.0\_301** Replace whatever version they give with this one.
    2. <https://www.codejava.net/java-core/how-to-set-java-home-environment-variable-on-windows-10>
    3. Again, this is entirely optional

1. Install the NLP Models



1. You should get the message below. If you do not, RStudio needs to refresh the RJava installation. Restart R.



1. You should be good to go at this point. Follow the instructions in the Rmarkdown document. Between reading and processing time, this should take about 30 minutes.