

텍스트 분석을 위한 R 패키지 설치 가이드

<https://mrchypark.github.io/textR/installation>

[pdf버전] [문의하기] [의견 및 오류 신고]

스타누르기는 콘텐츠 제작자를 춤추게 합니다.

박찬엽

2018년 10월 23일

windows 사용자

windows 10 64bit, windows 7 32bit에서 테스트하였습니다.

설치를 위한 패키지 설치

아래 코드를 R 콘솔에서 실행하세요.

```
# stringi 패키지를 설치합니다.  
# 미리 컴파일된 버전을 설치하는 방법입니다.  
install.packages(  
  "https://github.com/gagolews/stringi/raw/master/devel/winbuilder/3.5/stringi_1.2.4.zip",  
  repos = NULL  
)  
  
## 패키지 버전을 확인해주세요. 1.2.4로 나오면 잘 설치된 것입니다.  
packageVersion("stringi")
```

```
## [1] `1.2.4`
```

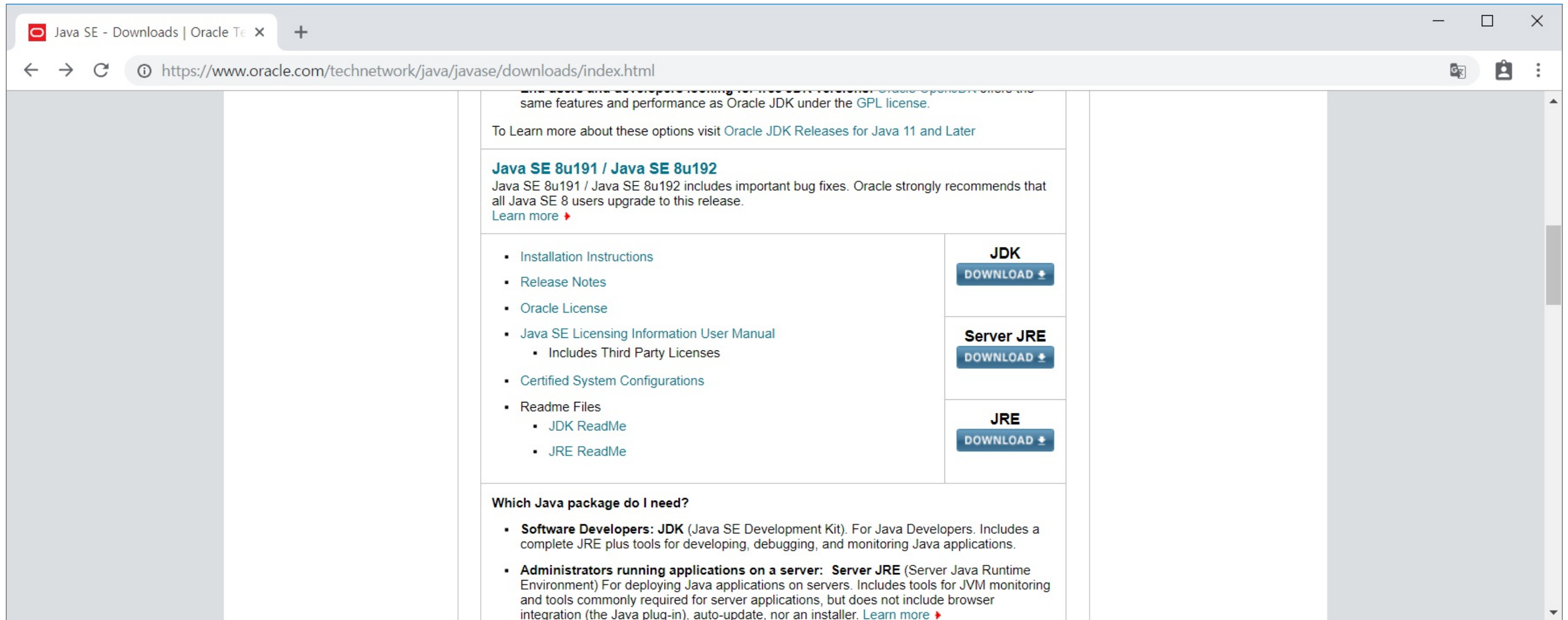
jdk 8 설치

<https://www.oracle.com/technetwork/java/javase/downloads/index.html> 방문



jdk 8 설치 (계속)

아래로 스크롤 하여 **Java SE 8uXXX** 에서 JDK의 다운로드 클릭



The screenshot shows the Oracle Java SE Downloads page for Java SE 8u191 / Java SE 8u192. The browser address bar shows the URL: <https://www.oracle.com/technetwork/java/javase/downloads/index.html>. The page content includes a list of links for installation instructions, release notes, Oracle license, and manuals. On the right side, there are three download buttons: **JDK** DOWNLOAD, **Server JRE** DOWNLOAD, and **JRE** DOWNLOAD. Below these, a section titled "Which Java package do I need?" provides guidance for software developers and administrators.

Java SE 8u191 / Java SE 8u192
Java SE 8u191 / Java SE 8u192 includes important bug fixes. Oracle strongly recommends that all Java SE 8 users upgrade to this release.
[Learn more](#)

- Installation Instructions
- Release Notes
- Oracle License
- Java SE Licensing Information User Manual
 - Includes Third Party Licenses
- Certified System Configurations
- Readme Files
 - JDK ReadMe
 - JRE ReadMe

JDK
DOWNLOAD

Server JRE
DOWNLOAD

JRE
DOWNLOAD

Which Java package do I need?

- Software Developers: JDK** (Java SE Development Kit). For Java Developers. Includes a complete JRE plus tools for developing, debugging, and monitoring Java applications.
- Administrators running applications on a server: Server JRE** (Server Java Runtime Environment) For deploying Java applications on servers. Includes tools for JVM monitoring and tools commonly required for server applications, but does not include browser integration (the Java plug-in), auto-update, nor an installer. [Learn more](#)

jdk 8 설치 (계속)

jdk 8 다운로드를 위한 화면

The screenshot shows the Oracle Java SE Development Kit 8 Downloads page. The browser address bar displays the URL: <https://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html>. The page features the Oracle logo and a navigation menu. The main content area is titled "Java SE Development Kit 8 Downloads" and includes a description of the JDK, a list of links for downloading the JDK, and a section for "Java SDKs and Tools". The page also includes a sidebar with links to Java SE, Java EE, Java ME, Java SE Subscription, Java Embedded, Java Card, Java TV, Community, and Java Magazine. The footer contains the text: "You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this

Oracle Technology Network / Java / Java SE / Downloads

Overview Downloads Documentation Community Technologies Training

Java SE Development Kit 8 Downloads

Thank you for downloading this release of the Java™ Platform, Standard Edition Development Kit (JDK™). The JDK is a development environment for building applications, applets, and components using the Java programming language.

The JDK includes tools useful for developing and testing programs written in the Java programming language and running on the Java platform.

See also:

- Java Developer Newsletter: From your Oracle account, select **Subscriptions**, expand **Technology**, and subscribe to **Java**.
- Java Developer Day hands-on workshops (free) and other events
- Java Magazine

JDK 8u191 [checksum](#)
JDK 8u192 [checksum](#)

Java SE Development Kit 8u191

You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this

Java SDKs and Tools

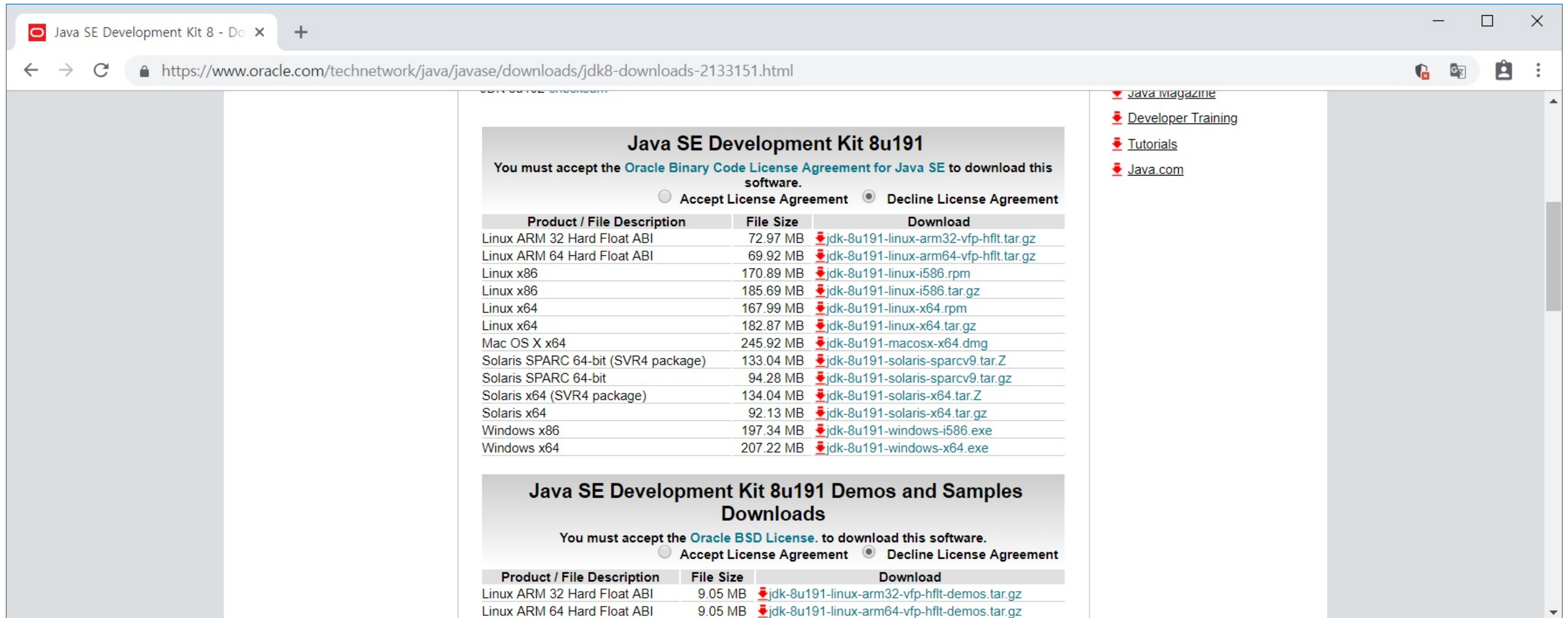
- [Java SE](#)
- [Java EE and Glassfish](#)
- [Java ME](#)
- [Java Card](#)
- [NetBeans IDE](#)
- [Java Mission Control](#)

Java Resources

- [Java APIs](#)
- [Technical Articles](#)
- [Demos and Videos](#)
- [Forums](#)
- [Java Magazine](#)
- [Developer Training](#)
- [Tutorials](#)
- [Java.com](#)

jdk 8 설치 (계속)

Java SE Development Kit 8uXXX 위치에 **Accept License Agreement** 클릭



Java SE Development Kit 8u191

You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this software.

☐ Accept License Agreement ☒ Decline License Agreement

Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	72.97 MB	jdk-8u191-linux-arm32-vfp-hflt.tar.gz
Linux ARM 64 Hard Float ABI	69.92 MB	jdk-8u191-linux-arm64-vfp-hflt.tar.gz
Linux x86	170.89 MB	jdk-8u191-linux-i586.rpm
Linux x86	185.69 MB	jdk-8u191-linux-i586.tar.gz
Linux x64	167.99 MB	jdk-8u191-linux-x64.rpm
Linux x64	182.87 MB	jdk-8u191-linux-x64.tar.gz
Mac OS X x64	245.92 MB	jdk-8u191-macosx-x64.dmg
Solaris SPARC 64-bit (SVR4 package)	133.04 MB	jdk-8u191-solaris-sparcv9.tar.Z
Solaris SPARC 64-bit	94.28 MB	jdk-8u191-solaris-sparcv9.tar.gz
Solaris x64 (SVR4 package)	134.04 MB	jdk-8u191-solaris-x64.tar.Z
Solaris x64	92.13 MB	jdk-8u191-solaris-x64.tar.gz
Windows x86	197.34 MB	jdk-8u191-windows-i586.exe
Windows x64	207.22 MB	jdk-8u191-windows-x64.exe

Java SE Development Kit 8u191 Demos and Samples Downloads

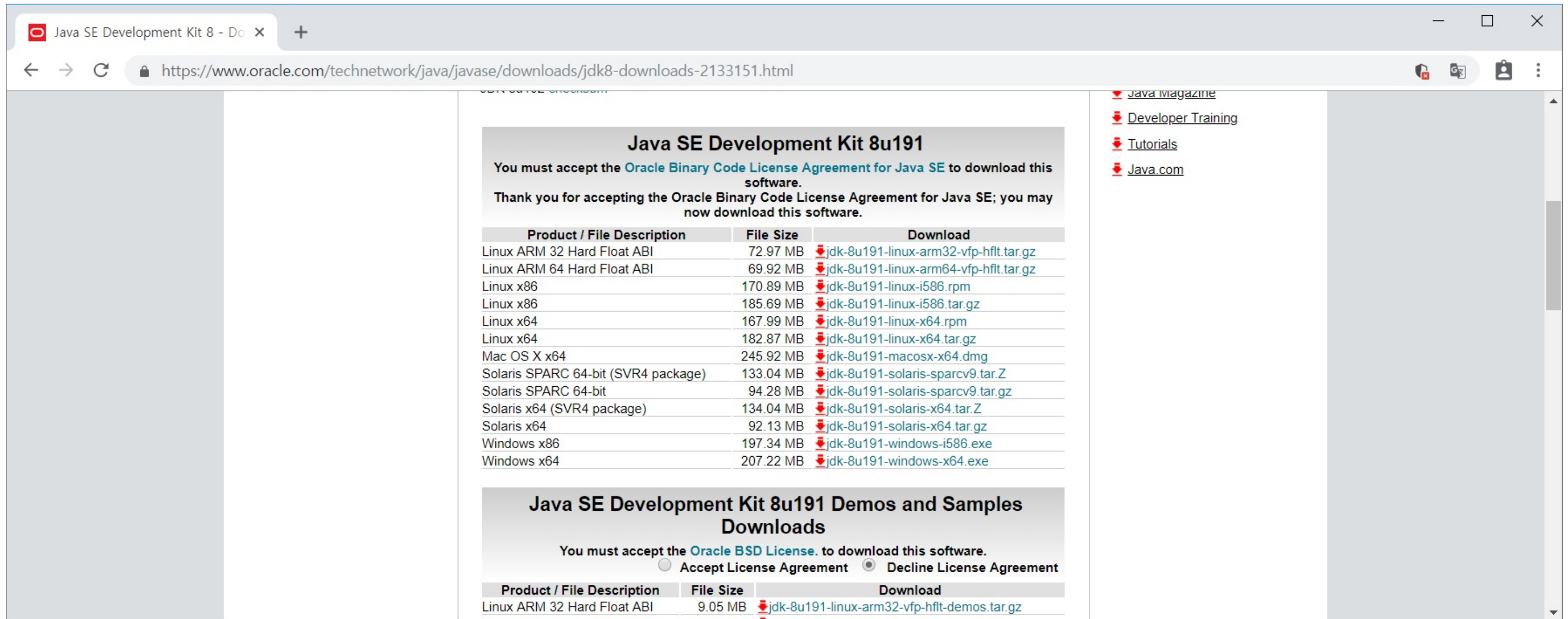
You must accept the [Oracle BSD License](#) to download this software.

☐ Accept License Agreement ☒ Decline License Agreement

Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	9.05 MB	jdk-8u191-linux-arm32-vfp-hflt-demos.tar.gz
Linux ARM 64 Hard Float ABI	9.05 MB	jdk-8u191-linux-arm64-vfp-hflt-demos.tar.gz

jdk 8 설치 (계속)

x64가 64bit, x86이 32bit를 의미하므로 컴퓨터의 bit에 맞는 버전을 다운로드



Java SE Development Kit 8u191

You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this software.

Thank you for accepting the Oracle Binary Code License Agreement for Java SE; you may now download this software.

Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	72.97 MB	jdk-8u191-linux-arm32-vfp-hflt.tar.gz
Linux ARM 64 Hard Float ABI	69.92 MB	jdk-8u191-linux-arm64-vfp-hflt.tar.gz
Linux x86	170.89 MB	jdk-8u191-linux-i586.rpm
Linux x86	185.69 MB	jdk-8u191-linux-i586.tar.gz
Linux x64	167.99 MB	jdk-8u191-linux-x64.rpm
Linux x64	182.87 MB	jdk-8u191-linux-x64.tar.gz
Mac OS X x64	245.92 MB	jdk-8u191-macosx-x64.dmg
Solaris SPARC 64-bit (SVR4 package)	133.04 MB	jdk-8u191-solaris-sparcv9.tar.Z
Solaris SPARC 64-bit	94.28 MB	jdk-8u191-solaris-sparcv9.tar.gz
Solaris x64 (SVR4 package)	134.04 MB	jdk-8u191-solaris-x64.tar.Z
Solaris x64	92.13 MB	jdk-8u191-solaris-x64.tar.gz
Windows x86	197.34 MB	jdk-8u191-windows-i586.exe
Windows x64	207.22 MB	jdk-8u191-windows-x64.exe

Java SE Development Kit 8u191 Demos and Samples Downloads

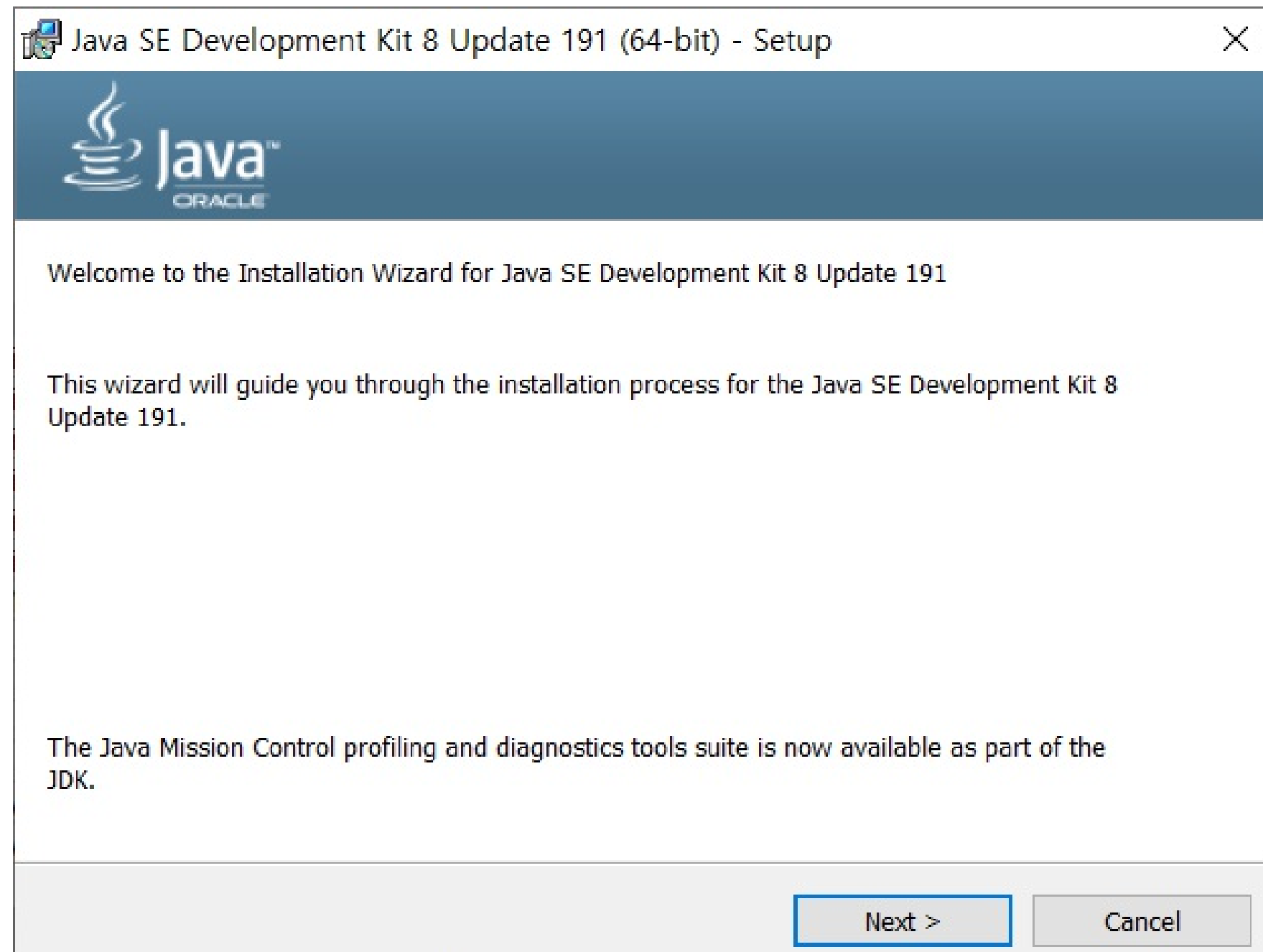
You must accept the [Oracle BSD License](#) to download this software.

☐ Accept License Agreement ☒ Decline License Agreement

Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	9.05 MB	jdk-8u191-linux-arm32-vfp-hflt-demos.tar.gz

jdk 8 설치 (계속)

설치 파일을 실행 후 다른 설정없이 **Next**를 계속 클릭하여 설치를 진행



KoNLP 패키지 설치 및 테스트

KoNLP는 가장 유명한 형태소 분석기입니다.

```
install.packages("KoNLP")  
library(KoNLP)  
  
test <- "한글 테스트 입니다."  
extractNoun(test)
```

맨 마지막 코드로 아래 결과가 나와야 합니다.

```
[1] "한글"    "테스트"
```

RcppMeCab 패키지 설치 및 테스트

RcppMeCab은 빠른 형태소 분석기입니다.

```
install.packages("RmecabKo")
install.packages("RcppMeCab")
RmecabKo::install_mecab("c:/mecab") # c:에 권한이 없다면 "d:/mecab"으로 설정
library(RcppMeCab)
test <- "한글 테스트 입니다."
pos(test)
```

아래 결과가 나와야 합니다.

```
$`한글 테스트 입니다.`
[1] "한글/NNG"      "테스트/NNG"    "입니다/VCP+EF" ". /SF"
```

만약에 글자가 깨진다면

글자가 깨진다면 아래 코드로 시도해 보세요.

```
library(RcppMeCab)
test <- "한글 테스트 입니다."
# iconv 함수는 인코딩을 변경하는 함수입니다.
test <- iconv(test, to = "UTF-8")
pos(test)
```

추가 패키지 설치

tidytext 패키지는 텍스트를 tidy하게 다룰 수 있게 해줍니다.

presidentSpeechKr 패키지는 역대 대통령 연설문 텍스트를 사용할 수 있게 해주는 패키지입니다.

```
install.packages("remotes")  
install.packages("tidytext")  
remotes::install_github("forkonlp/presidentSpeechKr")  
  
library(tidytext)  
library(presidentSpeechKr)
```

위 패키지들은 패키지를 불러오는 것이 동작하는 것을 확인해주세요.

macOS 사용자

터미널에서 설치해야 하는 도구들

1. JDK8

동영상 가이드를 참고하여 설치해주세요.

<https://www.youtube.com/watch?v=v8xZWbIASc0>

```
java -version
```

버전이 1.8.XXXXX로 표시되면 잘 설치된 것입니다.

터미널에서 설치해야 하는 도구들

2. mecab-ko

형태소 분석기를 설치합니다.

공식 문서를 참고하세요.

<https://bitbucket.org/eunjeon/mecab-ko-dic>

맥의 터미널에서 아래 명령어를 수행합니다.

```
wget https://bitbucket.org/eunjeon/mecab-ko/downloads/mecab-0.996-ko-0.9.2.tar.gz
tar zxfv mecab-0.996-ko-0.9.2.tar.gz
cd mecab-0.996-ko-0.9.2
./configure
make
make check
# 아래 실행 후 컴퓨터 비밀번호 입력
sudo make install
```

터미널에서 설치해야 하는 도구들

3. mecab-ko-dic

형태소 분석기의 사전을 설치합니다.

공식 문서를 참고하세요.

<https://bitbucket.org/eunjeon/mecab-ko-dic>

맥의 터미널에서 아래 명령어를 수행합니다.

```
wget https://bitbucket.org/eunjeon/mecab-ko-dic/downloads/mecab-ko-dic-2.1.1-20180720.tar.gz
tar zxfv mecab-ko-dic-2.1.1-20180720.tar.gz
cd mecab-ko-dic-2.1.1-20180720
./configure
make
make check
# 아래 실행 후 컴퓨터 비밀번호 입력
sudo make install
```

KoNLP 패키지 설치 및 테스트

KoNLP는 가장 유명한 형태소 분석기입니다.

```
install.packages("KoNLP")  
library(KoNLP)  
  
test <- "한글 테스트 입니다."  
extractNoun(test)
```

아래 결과가 나오면 잘 동작하는 것입니다.

```
[1] "한글"    "테스트"
```


RcppMeCab 패키지 설치 및 테스트

RcppMeCab은 빠른 형태소 분석기입니다.

```
install.packages("remotes")  
library(remotes)  
install_github("junhewk/RcppMeCab")  
library(RcppMeCab)  
  
test <- "한글 테스트 입니다."  
pos(test)
```

아래 결과가 나와야 합니다.

```
$ `한글 테스트 입니다.`  
[1] "한글/NNG"      "테스트/NNG"    "입니다/VCP+EF" ". /SF"
```

추가 패키지 설치

tidytext 패키지는 텍스트를 tidy하게 다룰 수 있게 해줍니다.

presidentSpeechKr 패키지는 역대 대통령 연설문 텍스트를 사용할 수 있게 해주는 패키지입니다.

```
install.packages("tidytext")  
install_github("forkonlp/presidentSpeechKr")  
  
library(tidytext)  
library(presidentSpeechKr)
```

위 패키지들은 패키지를 불러오는 것이 동작하는 것을 확인해주세요.

과제

1. 앞의 가이드에 따라 총 4개 패키지를 설치해 주세요.
2. 아래 코드를 R 콘솔에서 실행하여 스크린샷을 제출해 주세요.

```
library(KoNLP)
library(tidytext)
library(presidentSpeechKr)
library(RcppMeCab)

test <- "한글 테스트 입니다."
extractNoun(test)
pos(test)
```