

# Instalação Apache Tika

# Download

Os arquivos estão disponíveis em [tika.apache.org/download.html](https://tika.apache.org/download.html)

- Tika-1.19.1-src.zip

Arquivos fonte para serem incluídos em um projeto Java

- Tika-app-1.19.1.jar

Aplicação Java para ser usado por linha de comando

# Dependências

Java

# Build do código fonte usando Maven

A execução do seguinte comando na pasta base criará as fontes e instalará os artefatos resultantes em seu repositório local do Maven

```
mvn install
```

# Exemplo de Uso - Extração de texto

```
import org.apache.tika.Tika;

public class SimpleTextExtractor {
    public static void main(String[] args) throws Exception {
        // Create a Tika instance with the default configuration
        Tika tika = new Tika();

        // Parse all given files and print out the extracted
        // text content
        for (String file : args) {
            String text = tika.parseToString(new File(file));
            System.out.print(text);
        }
    }
}
```

# Exemplo de Uso - Detecção de Tipo

```
import org.apache.tika.Tika;

public class SimpleTypeDetector {
    public static void main(String[] args) throws Exception {
        Tika tika = new Tika();

        for (String file : args) {
            String type = tika.detect(new File(file));
            System.out.println(file + ": " + type);
        }
    }
}
```

# Exemplo de Uso - Linha de comando

```
Java -jar Tika-app-1.19.1.jar --text exemplo.pdf
```

```
Java -jar Tika-app-1.19.1.jar --text-main http://www.google.com
```

```
Java -jar Tika-app-1.19.1.jar --metadata exemplo.pdf
```

```
Java -jar Tika-app-1.19.1.jar --detect file
```

```
Java -jar Tika-app-1.19.1.jar --language file.doc
```