

## Práctica 2 - Tika

### Ejercicio 1: Abrir un archivo pdf en el modo ventana y decir quién es el autor del archivo, si tiene.

Primero ejecutamos el modo ventana abriendo un “cmd” en la misma pestaña que contiene el fichero .jar y ejecutando la orden: “java -jar tika-app-2.9.1.jar”. Con el modo ventana abierto, vamos a File > Open... y seleccionamos el fichero que deseamos (en este caso selecciono el documento que realice para la anterior práctica) ver Figura 1 y 2.

Como podemos ver en la Figura 3, en el apartado “dc:creator” obtenemos el autor del documento.



Figura 1 Modo ventana de Tika (apartado open)

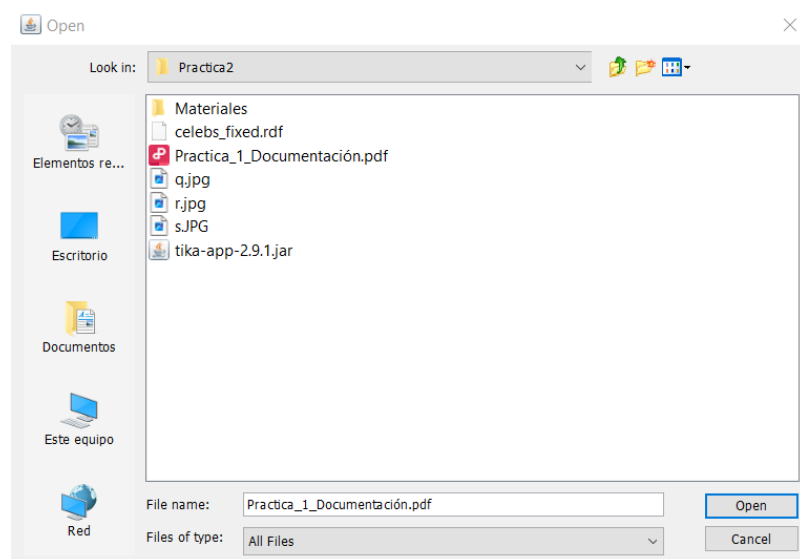


Figura 2 Selección del fichero

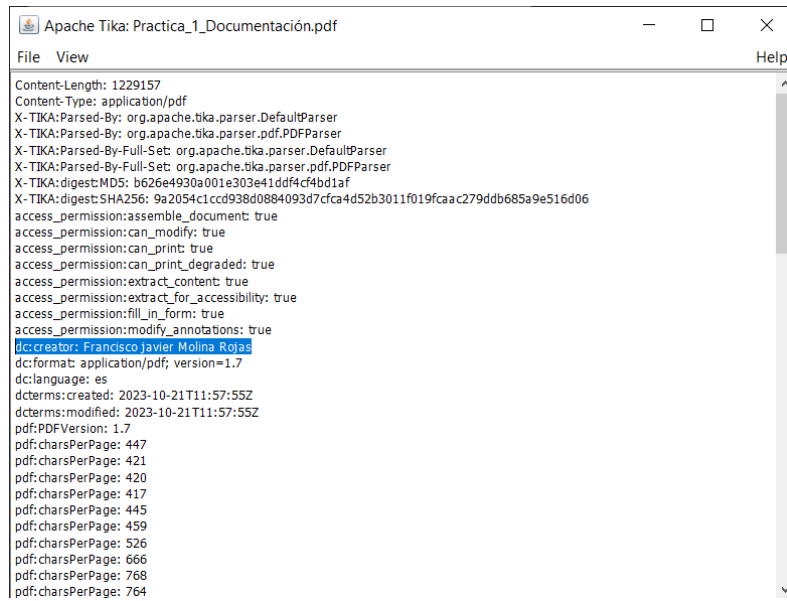


Figura 3 Creador del documento

## Ejercicio 2: Almacenar el contenido de un archivo .pdf en un archivo .doc.

Para obtener el contenido de un fichero ejecutaremos el fichero jar con la misma orden que en el apartado anterior, pero a esta le agregaremos la flag “--text” junto al fichero del que queramos obtener el contenido. Si dejásemos esto, la consola se llenaría con el contenido del fichero, así que usaremos el operador > para redirigir el flujo de salida a un fichero llamado test.doc, ver Figura 4.

En la Figura 5 y 6 podemos ver el contenido del fichero .pdf y del fichero .doc resultante.

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --text Practica_1_Documentación.pdf > test2.doc
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 4 Comando usado para obtener el contenido del fichero .pdf

```
Francisco Javier Molina Rojas
Recuperación de la Información
Práctica 1 - Expresiones Regulares en Java

Ejercicio 1. Comprobar si una cadena empieza por "abc".
Expresión regular usada: "abc.*"

Figura 1 Código del ejercicio 1

Figura 2 Ejemplo de ejecución del ejercicio 1

Como podemos ver en la salida, cuando la cadena empieza por abc, nos
devuelve un
resultado correcto (true) y cuando no, devuelve un resultado incorrecto
(false).
```

Figura 5 Contenido de test2.doc

Francisco Javier Molina Rojas

## Práctica 1 - Expresiones Regulares en Java

**Ejercicio 1.** Comprobar si una cadena empieza por "abc".

Expresión regular usada: `^abc.*`



```
1 ejercicio1.java X
2
3 ejercicio1.java > ejercicio1
4
5 import java.util.regex.Matcher;
6 import java.util.regex.Pattern;
7
8 public class ejercicio1 {
9     public static void main(String[] args) {
10         String cadena = "hahcdikndpiu29u3ue8ie3u19";
11         Pattern pat = Pattern.compile("^abc.*");
12         Matcher mat = pat.matcher(cadena);
13         System.out.println("Para la cadena: " + cadena + " el resultado es: " + mat.matches());
14     }
15 }
```

Figura 1 Código del ejercicio 1

Figura 6 Contenido del fichero .pdf

### Ejercicio 3: Ver los metadatos de un archivo que esté subido en una página web.

Para este ejercicio, usaremos el siguiente enlace que contiene un pdf:

<https://esingenieria.uca.es/wp-content/uploads/2023/11/Convocatoria-elecciones-Delegado-curso-23-24.report.pdf>, ver Figura 7.

En este caso además del comando “java -jar tika-app-2.9.1.jar” le añadimos la flag “--metadata” y la web, obteniendo como resultado los metadatos de la misma como podemos ver en la Figura 8.

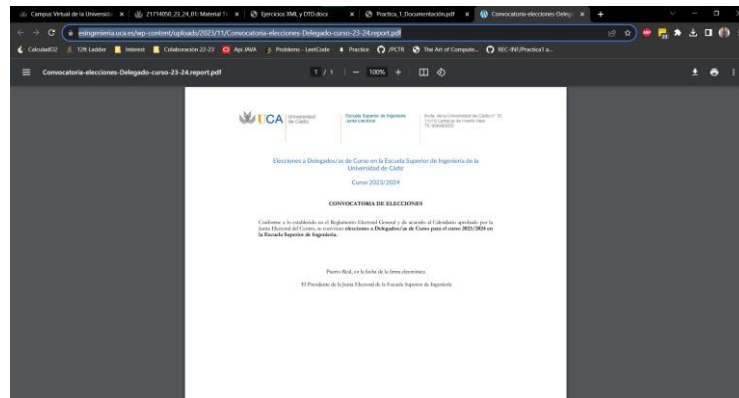


Figura 7 Archivo web

```
C:\Users\javier\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata https://esingenieria.uca.es/wp-content/uploads/2023/11/Convocatoria-elecciones-Delegado-curso-23-24.report.pdf
Content-Length: 333462
Content-Type: application/pdf
X-TIKA:Parsed-By: org.apache.tika.parser.DefaultParser
X-TIKA:Parsed-By: org.apache.tika.parser.pdf.PDFParser
access_permission:assemble_document: true
access_permission:can_modify: true
access_permission:can_print: true
access_permission:can_print_degraded: true
access_permission:extract_content: true
access_permission:extract_for_accessibility: true
access_permission:fill_in_form: true
access_permission:modify_annotations: true
dc:format: application/pdf; version=1.4
dcterms:created: 2023-11-02T13:24:21Z
dcterms:modified: 2023-11-02T13:24:21Z
pdf:PDFVersion: 1.4
pdf:charsPerPage: 1033
pdf:docinfo:created: 2023-11-02T13:24:21Z
pdf:docinfo:modified: 2023-11-02T13:24:21Z
pdf:docinfo:producer: iText® 5.5.13.3 #2000-2022 iText Group NV (AGPL-version)
pdf:encrypted: false
pdf:hasCollection: false
pdf:hasMarkedContent: false
pdf:hasXFA: false
pdf:hasXMP: false
pdf:producer: iText® 5.5.13.3 #2000-2022 iText Group NV (AGPL-version)
pdf:uncompressedCharsPerPage: 0
resourceName: Convocatoria-elecciones-Delegado-curso-23-24.report.pdf
xmpTPg:10Pages: 1
C:\Users\javier\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 8 Resultado del comando

**Ejercicio 4: Comprimir un archivo de texto sencillo y abrir con Tika el archivo comprimido.**

Creamos un fichero .txt con contenido en el mismo directorio donde guardamos en fichero pika, ver Figura 9.

Posteriormente, lo comprimimos el archivo en un .zip, ver Figura 10.

Por último, realizamos el comando “java -jar tika-app-2.9.1.jar” junto el archivo .zip, ver Figura 11.

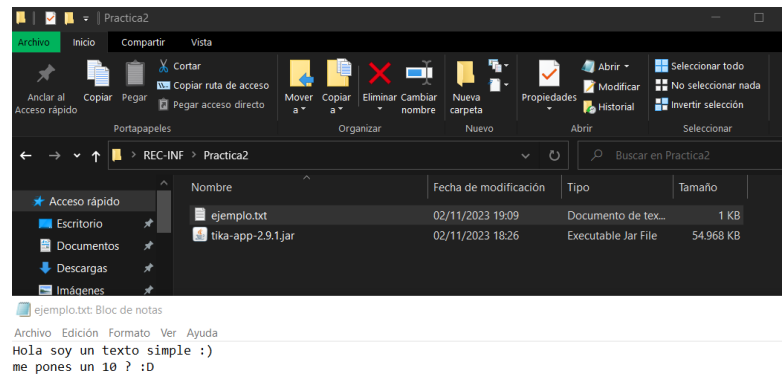


Figura 9 Fichero .txt

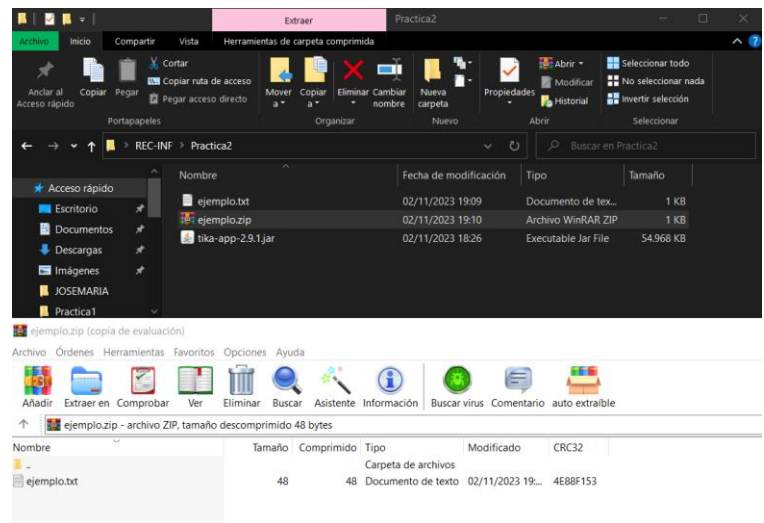


Figura 10 Fichero .txt comprimido

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar ejemplo.zip
<?xml version="1.0" encoding="UTF-8"?><html xmlns="http://www.w3.org/1999/xhtml">
<head>
<meta name="X-TIKA:Parsed-By" content="org.apache.tika.parser.DefaultParser"/>
<meta name="X-TIKA:Parsed-By" content="org.apache.tika.parser.pkg.PackageParser"/>
<meta name="resourceName" content="ejemplo.zip"/>
<meta name="Content-Length" content="204"/>
<meta name="Content-Type" content="application/zip"/>
<title/>
</head>
<body><div class="embedded" id="ejemplo.txt"/>
<div class="package-entry"><h1>ejemplo.txt</h1>
<p>Hola soy un texto simple :)&#13;
me pones un 10 ? :D</p>

</div>
</body></html>
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 11 Resultado del comando

**Ejercicio 5: Pasar por correo y Whatsapp una foto y comparar los metadatos comprobando las diferencias.**

Descargamos y almacenamos las fotos en el mismo repositorio que almacena el fichero .jar, ver Figura 12.

Posteriormente, realizamos dos comandos con “java -jar tika-app-2.9.1.jar” junto con el flag “--metadata” y cada una de las dos imágenes.

Finalmente, como vemos en las Figuras 13 y 14, la foto enviada por WhatsApp tiene menos metadatos que la enviada por correo. Esto es debido a la compresión que realiza WhatsApp al mandar cualquier tipo de foto.

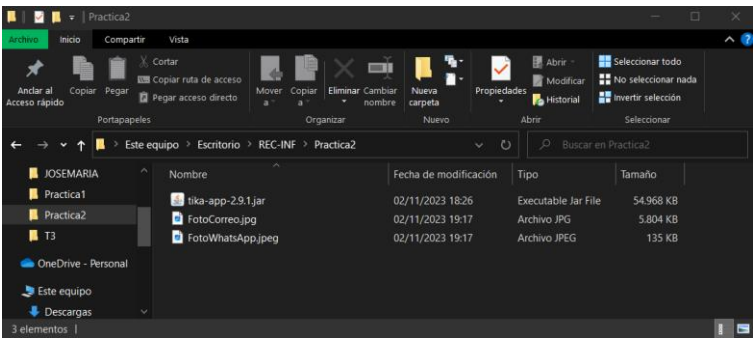


Figura 12 Fotos en el directorio

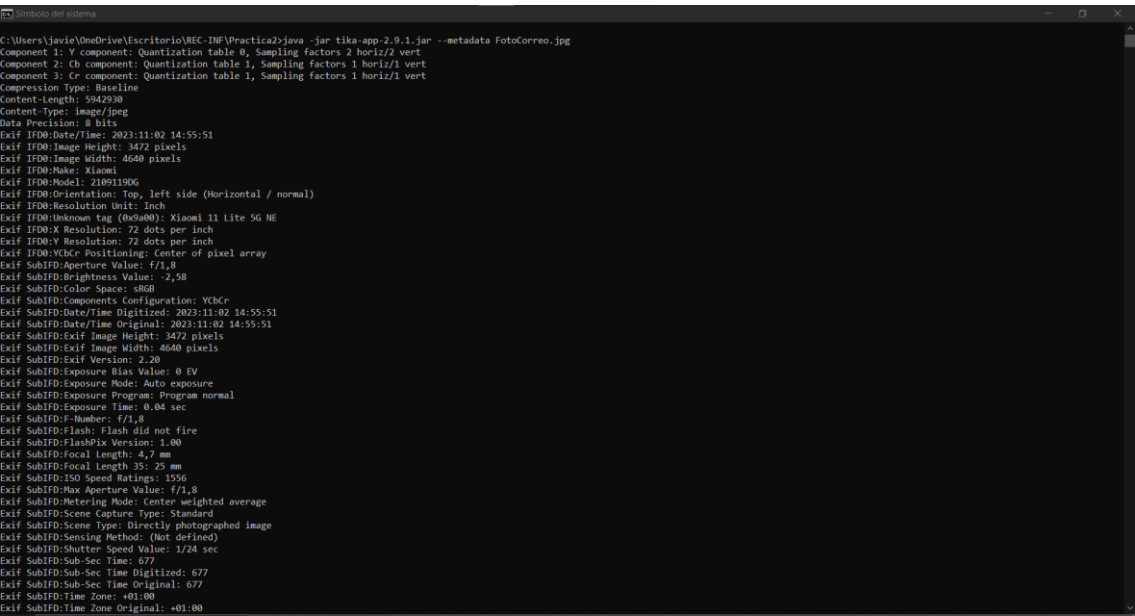


Figura 13 Metadatos de la foto por correo



```
Símbolo del sistema
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata FotoWhatsApp.jpeg
Component 1: Y component: Quantization table 0, Sampling factors 2 horiz/2 vert
Component 2: Cb component: Quantization table 1, Sampling factors 1 horiz/1 vert
Component 3: Cr component: Quantization table 1, Sampling factors 1 horiz/1 vert
Compression Type: Progressive, Huffman
Content-Length: 137922
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: jue nov 02 19:17:41 +01:00 2023
File Name: FotoWhatsApp.jpeg
File Size: 137922 bytes
Image Height: 1532 pixels
Image Width: 2048 pixels
Number of Components: 3
Number of Tables: 2 Huffman tables
Resolution Units: none
Thumbnail Height Pixels: 0
Thumbnail Width Pixels: 0
Version: 1.1
X-Resolution: 1 dot
X-TIKA:Parsed-By: org.apache.tika.parser.DefaultParser
X-TIKA:Parsed-By: org.apache.tika.parser.image.JpegParser
Y-Resolution: 1 dot
resourceName: FotoWhatsApp.jpeg
tiff:BitsPerSample: 8
tiff:ImageLength: 1532
tiff:ImageWidth: 2048
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 14 Metadatos de la foto por WhatsApp

**Ejercicio 6: Ver los metadatos de <http://www.uca.es/es/> y guardarlo en un archivo.txt.**

Primero, realizamos el comando “java -jar tika-app-2.9.1.jar” junto con el flag “--metadata” y con la url desada, además de usar el operador > para redirigir el flujo de salida a un fichero llamada “archivo.txt”, ver Figura 15.

Por último, podemos ver los metadatos de la url en dicho fichero, ver Figura 16.

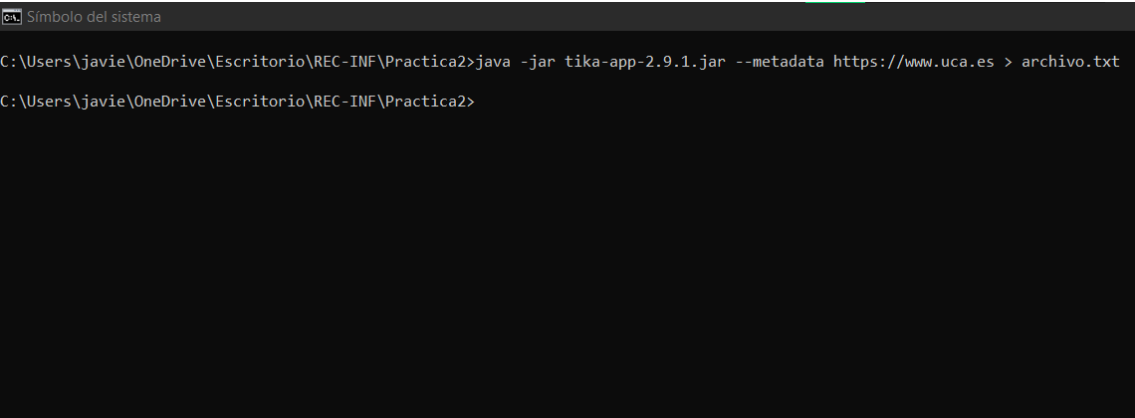


Figura 15 Comando realizado

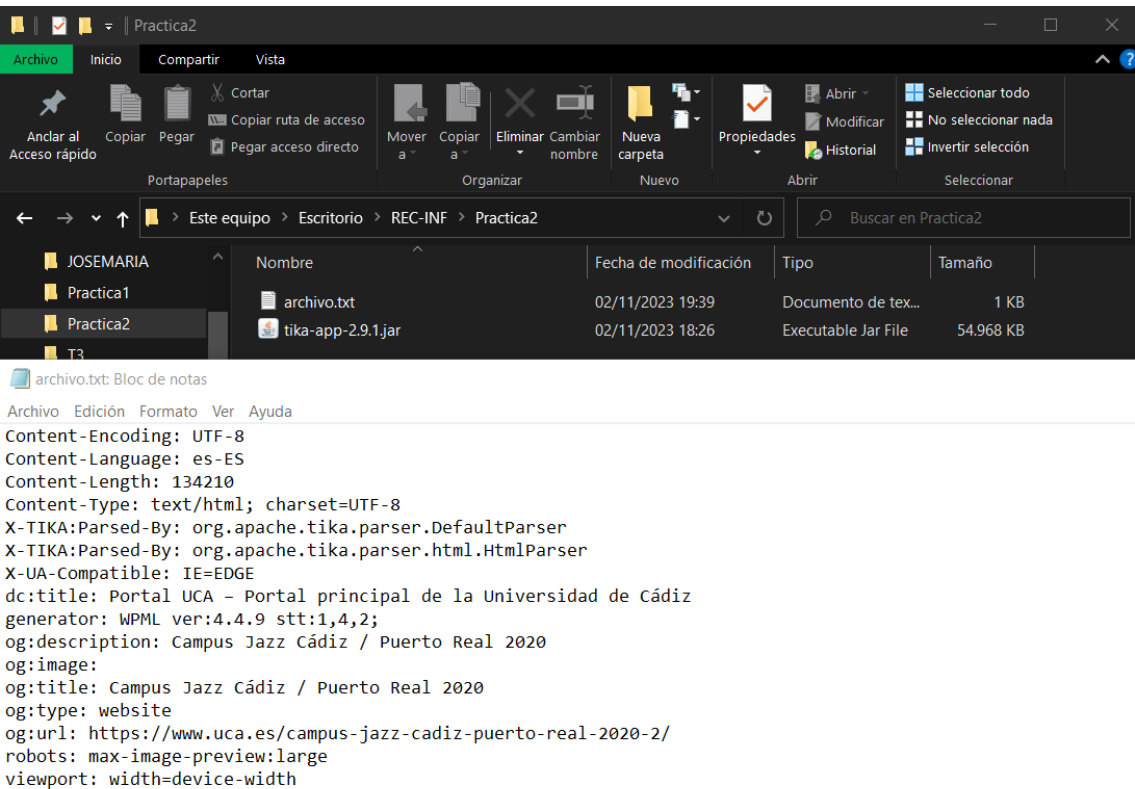


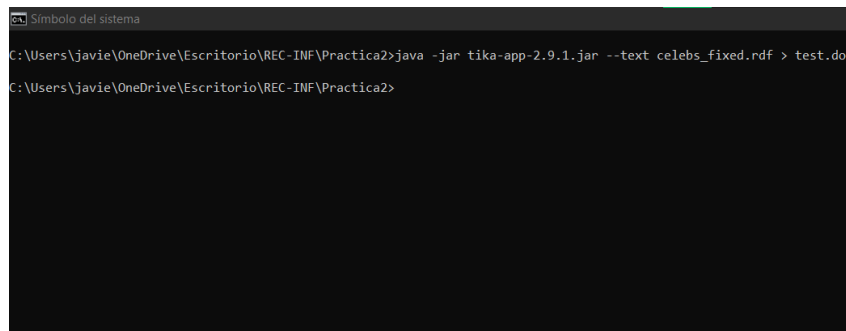
Figura 16 Contenido de archivo.txt

### Ejercicio 7: Pasar un archivo .rdf a .doc. ¿Pasar el archivo .doc o .rdf al formato .pdf dará error?

Primero realizaremos el paso de .rdf a .doc, como podemos ver en la Figura 17, con el comando “java -jar tika-app-2.9.1.jar”, el flag “--text” y el operador > para redirigir el flujo a un fichero test.doc.

Si vemos el contenido de este fichero obtendríamos lo que se ve en la Figura 18.

Si intentamos pasar este .rdf a .pdf, como podemos ver en la Figura 19, con el comando comando “java -jar tika-app-2.9.1.jar”, el flag “--text” y el operador > para redirigir el flujo a un fichero test.pdf, podremos ver en la Figura 20 que al intentar acceder a su información, este estará corrupto y no permitirá acceder al mismo.



```
Símbolo del sistema
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --text celebs_fixed.rdf > test.doc
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 17 Comando para pasar de .rdf a .doc

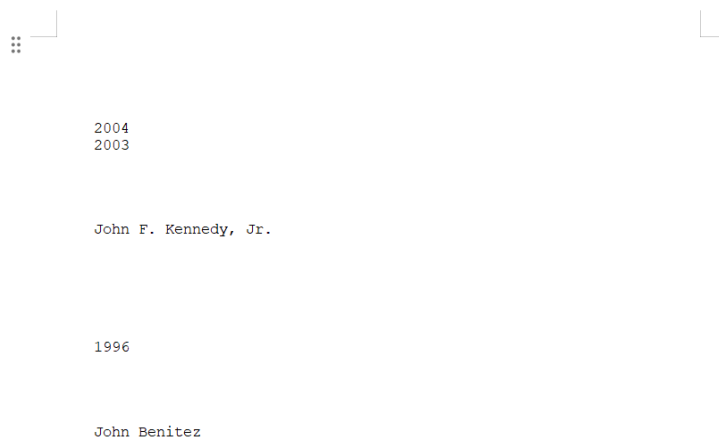


Figura 18 Contenido de test.doc

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --text celebs_fixed.rdf > test.pdf
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 19 Comando para pasar de .rdf a .pdf

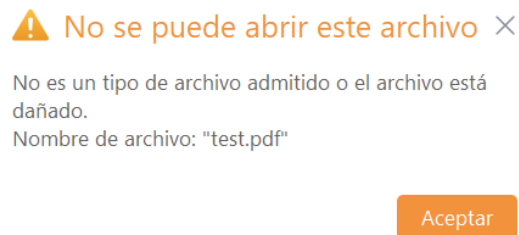


Figura 20 Fichero test.pdf corrupto

## Ejercicio 8: Descargar las 3 imágenes que se proporcionan en la carpeta “Material Tika” y decir qué imagen (o imágenes) se sacó (sacaron) con un producto perteneciente a apple computer inc.

Como podemos ver en las Figuras 21, 22 y 23 (metadatos de cada imagen), solo la imagen q y r están realizadas con un producto perteneciente a Apple.

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata s.JPG
AE Warning: AE Good
Auto Bracketing: Off
Blur Warning: No Blur Warning
Color Saturation: Normal
Component 1: Y component: Quantization table 0, Sampling factors 2 horiz/1 vert
Component 2: Cb component: Quantization table 1, Sampling factors 1 horiz/1 vert
Component 3: Cr component: Quantization table 2, Sampling factors 1 horiz/1 vert
Compression Type: Baseline
Content-Length: 2554023
Content-Type: image/jpeg
Data Precision: 8 bits
Dynamic Range: Standard
Exif IFD0:Copyright:
Exif IFD0:Date/Time: 2006:04:25 14:42:14
Exif IFD0:Make: FUJIFILM
Exif IFD0:Model: FinePix A500
Exif IFD0:Orientation: Top, left side (Horizontal / normal)
Exif IFD0:Resolution Unit: Inch
Exif IFD0:Software: Digital Camera FinePix A500 Ver1.03
Exif IFD0:X Resolution: 72 dots per inch
Exif IFD0:Y Resolution: 72 dots per inch
Exif IFD0:YCbCr Positioning: Datum point
Exif SubIFD:Aperture Value: f/4,9
Exif SubIFD:Brightness Value: 8,99
Exif SubIFD:Color Space: sRGB
Exif SubIFD:Components Configuration: YCbCr
```

Figura 21 Metadatos de la imagen s

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata q.jpg
Component 1: Y component: Quantization table 0, Sampling factors 2 horiz/2 vert
Component 2: Cb component: Quantization table 1, Sampling factors 1 horiz/1 vert
Component 3: Cr component: Quantization table 1, Sampling factors 1 horiz/1 vert
Compression Type: Progressive, Huffman
Content-Length: 66370
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: jue nov 02 19:51:06 +01:00 2023
File Name: q.jpg
File Size: 66370 bytes
ICC:Blue Colorant: (0,1431, 0,0606, 0,7141)
ICC:Blue TRC: 0.0, 0.0030976, 0.0069734, 0.0132296, 0.0217594, 0.0328832, 0.0467231, 0.0634623, 0.0832685, 0.1061, 0.5378653, 0.6034638, 0.6734874, 0.7480125, 0.8272068, 0.9109026, 1.0
ICC:CMM Type: Icms
ICC:Class: Display Device
ICC:Color space: RGB
ICC:Green Colorant: (0,3851, 0,7169, 0,0971)
ICC:Green TRC: 0.0, 0.0030976, 0.0069734, 0.0132296, 0.0217594, 0.0328832, 0.0467231, 0.0634623, 0.0832685, 0.1061, 0.5378653, 0.6034638, 0.6734874, 0.7480125, 0.8272068, 0.9109026, 1.0
ICC:Media Black Point: (0,0121, 0,0125, 0,0103)
ICC:Media White Point: (0,9642, 1, 0,8249)
ICC:Primary Platform: Apple Computer, Inc.
ICC:Profile Connection Space: XYZ
ICC:Profile Copyright: FB
ICC:Profile Date/Time: 2012:01:25 03:41:57
ICC:Profile Description: c2
ICC:Profile Size: 524
ICC:Red Colorant: (0,4361, 0,2225, 0,0139)
```

Figura 22 Metadatos de la imagen q

```
Seleccionar Símbolo del sistema

C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata r.jpg
Component 1: Unknown (0) component: Quantization table 0, Sampling factors 2 horiz/2 vert
Component 2: Y component: Quantization table 1, Sampling factors 1 horiz/1 vert
Component 3: Cb component: Quantization table 1, Sampling factors 1 horiz/1 vert
Compression Type: Progressive, Huffman
Content-Length: 128142
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: jue nov 02 19:51:06 +01:00 2023
File Name: r.jpg
File Size: 128142 bytes
ICC:Blue Colorant: (0,1431, 0,0606, 0,7141)
ICC:Blue TRC: 0.0, 0.0030976, 0.0069734, 0.0132296, 0.0217594, 0.0328832, 0.0467231, 0.0634623, 0.083
61, 0.5378653, 0.6034638, 0.6734874, 0.7480125, 0.8272068, 0.9109026, 1.0
ICC:CMM Type: lcms
ICC:Class: Display Device
ICC:Color space: RGB
ICC:Green Colorant: (0,3851, 0,7169, 0,0971)
ICC:Green TRC: 0.0, 0.0030976, 0.0069734, 0.0132296, 0.0217594, 0.0328832, 0.0467231, 0.0634623, 0.08
461, 0.5378653, 0.6034638, 0.6734874, 0.7480125, 0.8272068, 0.9109026, 1.0
ICC:Media Black Point: (0,0121, 0,0125, 0,0103)
ICC:Media White Point: (0,9642, 1, 0,8249)
ICC:Primary Platform: Apple Computer, Inc.
ICC:Profile Connection Space: XYZ
ICC:Profile Copyright: FB
ICC:Profile Date/Time: 2012:01:25 03:41:57
```

Figura 23 Metadatos de la imagen r

**Ejercicio 9: Describir el procedimiento seguido para guardar el contenido de una web cualquiera en un archivo html, este convertirlo en doc y comprobar los metadatos de este último.**

Para pasar el contenido de la web a un archivo html, usaremos el comando “java -jar tika-app-2.9.1.jar” y el operador > para redirigir el flujo de salida hacia un archivo llamado webEsi.html, ver Figura 24.

Como podemos ver en la Figura 25 se obtiene el formato html de la web.

Para pasar el archivo html a doc, usaremos el comando “java -jar tika-app-2.9.1.jar” y el operador > para redirigir el flujo de salida hacia un archivo llamado test.doc, ver Figura 26.

En la Figura 26, podemos ver el contenido de test.doc.

Por ultimo para visualizar los metadatos de este ultimo fichero, usaremos el comando “java -jar tika-app-2.9.1.jar”, el flag “--metadata” y el fichero test.doc, ver Figura 27

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar https://esingenieria.uca.es/ > webEsi.html
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 24 Comando pasar de web a html

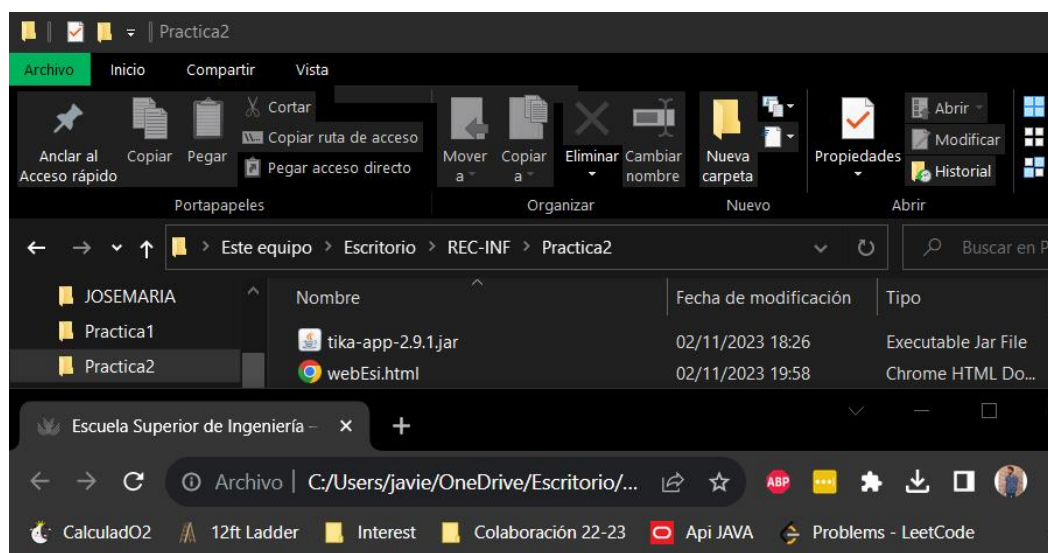


Figura 25 Muestra del html

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2> java -jar tika-app-2.9.1.jar webEsi.html > test.doc  
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 26 Comando pasar de html a doc

```
<?xml version="1.0" encoding="UTF-8"?><html  
xmlns="http://www.w3.org/1999/xhtml" lang="es-ES">  
<head>  
<link rel="icon" type="image/png" href="https://esingenieria.uca.es/wp-  
content/themes/theme_main_uca/images/logos/favicon-32x32.png"/>  
<link rel="icon" type="image/png" href="https://esingenieria.uca.es/wp-  
content/themes/theme_main_uca/images/logos/favicon-16x16.png"/>  
<link rel="dns-prefetch" href="//fonts.googleapis.com"/>  
<link rel="stylesheet" href="https://esingenieria.uca.es/wp-  
includes/css/classic-themes.min.css" type="text/css" media="all"/>  
<link rel="stylesheet" href="https://esingenieria.uca.es/wp-  
content/plugins/foobox-image-  
lightbox/free/css/foobox.free.min.css?ver=2.7.8" type="text/css"  
media="all"/>  
<link rel="stylesheet" href="https://esingenieria.uca.es/wp-  
content/plugins/uca-gcalendar/uca-gcalendar.css?ver=1.0.13"  
type="text/css" media="all"/>  
<link rel="stylesheet" href="//esingenieria.uca.es/wp-  
content/plugins/sitepress-multilingual-cms/templates/language-  
switchers/legacy-list-horizontal/style.css?ver=1" type="text/css"  
media="all"/>  
<link rel="stylesheet"  
href="https://fonts.googleapis.com/css?family=Lato%3A100%2C300%2C400"  
type="text/css" media="all"/>  
<link rel="stylesheet" href="https://esingenieria.uca.es/wp-  
content/themes/theme_main_uca/bootstrap/css/bootstrap.min.css?ver=1.3.0
```

Figura 27 Contenido test.doc

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2> java -jar tika-app-2.9.1.jar --metadata test.doc  
Content-Encoding: UTF-8  
Content-Encoding: UTF-8  
Content-Encoding: UTF-8  
Content-Language: es-ES  
Content-Language: es-ES  
Content-Language: es-ES  
Content-Length: 33186  
Content-Length: 32663  
Content-Length: 70849  
Content-Type: application/xhtml+xml; charset=UTF-8  
Content-Type-Hint: application/xhtml+xml; charset=UTF-8  
X-TIKA:Parsed-By: org.apache.tika.parser.DefaultParser  
X-TIKA:Parsed-By: org.apache.tika.parser.html.HtmlParser  
X-TIKA:Parsed-By: org.apache.tika.parser.DefaultParser  
X-TIKA:Parsed-By: org.apache.tika.parser.html.HtmlParser  
X-TIKA:Parsed-By: org.apache.tika.parser.DefaultParser  
X-TIKA:Parsed-By: org.apache.tika.parser.html.HtmlParser  
X-UA-Compatible: IE=EDGE  
dc:title: Escuela Superior de Ingenier|ja ÔÇô Sitio web de la Universidad de C|ídiz  
generator: WPML ver:4.0.7 stt:1,2;  
og:description: Grados  
og:title: Grados  
og:type: website  
og:url: https://esingenieria.uca.es/grados/  
resourceName: test.doc  
resourceName: webEsi.html  
robots: max-image-preview:large  
viewport: width=device-width  
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 28 Metadatos de test.doc



**Ejercicio 10: Descargar tres imágenes, a elección del alumno, de tres sitios diferentes donde los usuarios comparten imágenes como pueden ser: facebook, instagram, flickr, twitter... y comentar las diferencias que encontramos en los metadatos.**

Como podemos ver en las Figuras 29 y 30 (pertenecientes a Twitter), la Figura 31 (perteneciente a Facebook) y la Figura 32 (perteneciente a Pinterest); las principales diferencias que poseen son: Las fotos de Twitter y Facebook poseen información relativa al RGB, cosa que la de Pinterest no posee. Además, las fotos de Facebook y Twitter poseen una longitud superior a la de Pinterest.

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata FotoTwitter.jfif
Component 1: Y component: Quantization table 0, Sampling factors 2 horiz/2 vert
Component 2: Cb component: Quantization table 1, Sampling factors 1 horiz/1 vert
Component 3: Cr component: Quantization table 1, Sampling factors 1 horiz/1 vert
Compression Type: Progressive, Huffman
Content-Length: 63332
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: vie nov 03 09:12:14 401:00 2023
File Name: FotoTwitter.jfif
File Size: 63332 bytes
ICC-Blue Colorant: (0.1431, 0.0606, 0.7141)
ICC-Blue TNC: (0.0, 0.0000763, 0.0001526, 0.0002289, 0.0003052, 0.0003815, 0.0004578, 0.0005341, 0.0006104, 0.0006867, 0.000763, 0.0008392, 0.0009155, 0.0009918, 0.0010681, 0.0011444, 0.0012207, 0.001297, 0.0013733, 0.0014496, 0.0015259, 0.0016022, 0.0016785, 0.0017548, 0.0018311, 0.0019074, 0.0019837, 0.00206, 0.0021363, 0.0022126, 0.0022889, 0.0023652, 0.0024415, 0.0025178, 0.0025941, 0.0026704, 0.0027467, 0.002823, 0.0028993, 0.0029756, 0.0030519, 0.0031282, 0.0032045, 0.0032808, 0.0033571, 0.0034334, 0.0035097, 0.003586, 0.0036623, 0.0037386, 0.0038149, 0.0038912, 0.0039675, 0.0040438, 0.0041201, 0.0041964, 0.0042727, 0.004349, 0.0044253, 0.0045016, 0.0045779, 0.0046542, 0.0047305, 0.0048068, 0.0048831, 0.0049594, 0.0050357, 0.005112, 0.0051883, 0.0052646, 0.0053409, 0.0054172, 0.0054935, 0.0055698, 0.0056461, 0.0057224, 0.0057987, 0.005875, 0.0059513, 0.0060276, 0.0061039, 0.0061802, 0.0062565, 0.0063328, 0.0064091, 0.0064854, 0.0065617, 0.006638, 0.0067143, 0.0067906, 0.0068669, 0.0069432, 0.0070195, 0.0070958, 0.0071721, 0.0072484, 0.0073247, 0.007401, 0.0074773, 0.0075536, 0.0076299, 0.0077062, 0.0077825, 0.0078588, 0.0079351, 0.0080114, 0.0080877, 0.008164, 0.0082403, 0.0083166, 0.0083929, 0.0084692, 0.0085455, 0.0086218, 0.0086981, 0.0087744, 0.0088507, 0.008927, 0.0090033, 0.0090796, 0.0091559, 0.0092322, 0.0093085, 0.0093848, 0.0094611, 0.0095374, 0.0096137, 0.00969, 0.0097663, 0.0098426, 0.0099189, 0.0099952, 0.0100715, 0.0101478, 0.0102241, 0.0103, 0.0103763, 0.0104526, 0.0105289, 0.0106052, 0.0106815, 0.0107578, 0.0108341, 0.0109104, 0.0109867, 0.011063, 0.0111393, 0.0112156, 0.0112919, 0.0113682, 0.0114445, 0.0115208, 0.0115971, 0.0116734, 0.0117497, 0.011826, 0.0119023, 0.0119786, 0.0120549, 0.0121312, 0.0122075, 0.0122838, 0.0123601, 0.0124364, 0.0125127, 0.012589, 0.0126653, 0.0127416, 0.0128179, 0.0128942, 0.0129705, 0.0130468, 0.0131231, 0.0131994, 0.0132757, 0.013352, 0.0134283, 0.0135046, 0.0135809, 0.0136572, 0.0137335, 0.0138098, 0.0138861, 0.0139624, 0.0140387, 0.014115, 0.0141913, 0.0142676, 0.0143439, 0.0144202, 0.0144965, 0.0145728, 0.0146491, 0.0147254, 0.0148017, 0.014878, 0.0149543, 0.0150306, 0.0151069, 0.0151832, 0.0152595, 0.0153358, 0.0154121, 0.0154884, 0.0155647, 0.015641, 0.0157173, 0.0157936, 0.0158699, 0.0159462, 0.0160225, 0.0160988, 0.0161751, 0.0162514, 0.0163277, 0.016404, 0.0164803, 0.0165566, 0.0166329, 0.0167092, 0.0167855, 0.0168618, 0.0169381, 0.0170144, 0.0170907, 0.017167, 0.0172433, 0.0173196, 0.0173959, 0.0174722, 0.0175485, 0.0176248, 0.0177011, 0.0177774, 0.0178537, 0.0179299, 0.0180062, 0.0180825, 0.0181588, 0.0182351, 0.0183114, 0.0183877, 0.018464, 0.0185403, 0.0186166, 0.0186929, 0.0187692, 0.0188455, 0.0189218, 0.0189981, 0.0190744, 0.0191507, 0.019227, 0.0193033, 0.0193796, 0.0194559, 0.0195322, 0.0196085, 0.0196848, 0.0197611, 0.0198374, 0.0199137, 0.01999, 0.0200663, 0.0201426, 0.0202189, 0.0202952, 0.0203715, 0.0204478, 0.0205241, 0.0206, 0.0206763, 0.0207526, 0.0208289, 0.0209052, 0.0209815, 0.0210578, 0.0211341, 0.0212104, 0.0212867, 0.021363, 0.0214393, 0.0215156, 0.0215919, 0.0216682, 0.0217445, 0.0218208, 0.0218971, 0.0219734, 0.0220497, 0.022126, 0.0222023, 0.0222786, 0.0223549, 0.0224312, 0.0225075, 0.0225838, 0.0226601, 0.0227364, 0.0228127, 0.022889, 0.0229653, 0.0230416, 0.0231179, 0.0231942, 0.0232705, 0.0233468, 0.0234231, 0.0234994, 0.0235757, 0.023652, 0.0237283, 0.0238046, 0.0238809, 0.0239572, 0.0240335, 0.0241098, 0.0241861, 0.0242624, 0.0243387, 0.024415, 0.0244913, 0.0245676, 0.0246439, 0.0247202, 0.0247965, 0.0248728, 0.0249491, 0.0250254, 0.0251017, 0.025178, 0.0252543, 0.0253306, 0.0254069, 0.0254832, 0.0255595, 0.0256358, 0.0257121, 0.0257884, 0.0258647, 0.025941, 0.0260173, 0.0260936, 0.0261699, 0.0262462, 0.0263225, 0.0263988, 0.0264751, 0.0265514, 0.0266277, 0.026704, 0.0267803, 0.0268566, 0.0269329, 0.0270092, 0.0270855, 0.0271618, 0.0272381, 0.0273144, 0.0273907, 0.027467, 0.0275433, 0.0276196, 0.0276959, 0.0277722, 0.0278485, 0.0279248, 0.0280011, 0.0280774, 0.0281537, 0.02823, 0.0283063, 0.0283826, 0.0284589, 0.0285352, 0.0286115, 0.0286878, 0.0287641, 0.0288404, 0.0289167, 0.028993, 0.0290693, 0.0291456, 0.0292219, 0.0292982, 0.0293745, 0.0294508, 0.0295271, 0.0296034, 0.0296797, 0.029756, 0.0298323, 0.0299086, 0.0299849, 0.0300612, 0.0301375, 0.0302138, 0.0302901, 0.0303664, 0.0304427, 0.030519, 0.0305953, 0.0306716, 0.0307479, 0.0308242, 0.0309, 0.0309763, 0.0310526, 0.0311289, 0.0312052, 0.0312815, 0.0313578, 0.0314341, 0.0315104, 0.0315867, 0.031663, 0.0317393, 0.0318156, 0.0318919, 0.0319682, 0.0320445, 0.0321208, 0.0321971, 0.0322734, 0.0323497, 0.032426, 0.0325023, 0.0325786, 0.0326549, 0.0327312, 0.0328075, 0.0328838, 0.0329601, 0.0330364, 0.0331127, 0.033189, 0.0332653, 0.0333416, 0.0334179, 0.0334942, 0.0335705, 0.0336468, 0.0337231, 0.0337994, 0.0338757, 0.033952, 0.0340283, 0.0341046, 0.0341809, 0.0342572, 0.0343335, 0.0344098, 0.0344861, 0.0345624, 0.0346387, 0.034715, 0.0347913, 0.0348676, 0.0349439, 0.0350202, 0.0350965, 0.0351728, 0.0352491, 0.0353254, 0.0354017, 0.035478, 0.0355543, 0.0356306, 0.0357069, 0.0357832, 0.0358595, 0.0359358, 0.0360121, 0.0360884, 0.0361647, 0.036241, 0.0363173, 0.0363936, 0.0364699, 0.0365462, 0.0366225, 0.0366988, 0.0367751, 0.0368514, 0.0369277, 0.037004, 0.0370803, 0.0371566, 0.0372329, 0.0373092, 0.0373855, 0.0374618, 0.0375381, 0.0376144, 0.0376907, 0.037767, 0.0378433, 0.0379196, 0.0379959, 0.0380722, 0.0381485, 0.0382248, 0.0383011, 0.0383774, 0.0384537, 0.03853, 0.0386063, 0.0386826, 0.0387589, 0.0388352, 0.0389115, 0.0389878, 0.0390641, 0.0391404, 0.0392167, 0.039293, 0.0393693, 0.0394456, 0.0395219, 0.0395982, 0.0396745, 0.0397508, 0.0398271, 0.0399034, 0.0399797, 0.040056, 0.0401323, 0.0402086, 0.0402849, 0.0403612, 0.0404375, 0.0405138, 0.0405901, 0.0406664, 0.0407427, 0.040819, 0.0408953, 0.0409716, 0.0410479, 0.0411242, 0.0412, 0.0412763, 0.0413526, 0.0414289, 0.0415052, 0.0415815, 0.0416578, 0.0417341, 0.0418104, 0.0418867, 0.041963, 0.0420393, 0.0421156, 0.0421919, 0.0422682, 0.0423445, 0.0424208, 0.0424971, 0.0425734, 0.0426497, 0.042726, 0.0428023, 0.0428786, 0.0429549, 0.0430312, 0.0431075, 0.0431838, 0.0432601, 0.0433364, 0.0434127, 0.043489, 0.0435653, 0.0436416, 0.0437179, 0.0437942, 0.0438705, 0.0439468, 0.0440231, 0.0440994, 0.0441757, 0.044252, 0.0443283, 0.0444046, 0.0444809, 0.0445572, 0.0446335, 0.0447098, 0.0447861, 0.0448624, 0.0449387, 0.045015, 0.0450913, 0.0451676, 0.0452439, 0.0453202, 0.0453965, 0.0454728, 0.0455491, 0.0456254, 0.0457017, 0.045778, 0.0458543, 0.0459306, 0.0460069, 0.0460832, 0.0461595, 0.0462358, 0.0463121, 0.0463884, 0.0464647, 0.046541, 0.0466173, 0.0466936, 0.0467699, 0.0468462, 0.0469225, 0.0469988, 0.0470751, 0.0471514, 0.0472277, 0.047304, 0.0473803, 0.0474566, 0.0475329, 0.0476092, 0.0476855, 0.0477618, 0.0478381, 0.0479144, 0.0479907, 0.048067, 0.0481433, 0.0482196, 0.0482959, 0.0483722, 0.0484485, 0.0485248, 0.0486011, 0.0486774, 0.0487537, 0.04883, 0.0489063, 0.0489826, 0.0490589, 0.0491352, 0.0492115, 0.0492878, 0.0493641, 0.0494404, 0.0495167, 0.049593, 0.0496693, 0.0497456, 0.0498219, 0.0498982, 0.0499745, 0.0500508, 0.0501271, 0.0502034, 0.0502797, 0.050356, 0.0504323, 0.0505086, 0.0505849, 0.0506612, 0.0507375, 0.0508138, 0.0508901, 0.0509664, 0.0510427, 0.051119, 0.0511953, 0.0512716, 0.0513479, 0.0514242, 0.0515, 0.0515763, 0.0516526, 0.0517289, 0.0518052, 0.0518815, 0.0519578, 0.0520341, 0.0521104, 0.0521867, 0.052263, 0.0523393, 0.0524156, 0.0524919, 0.0525682, 0.0526445, 0.0527208, 0.0527971, 0.0528734, 0.0529497, 0.053026, 0.0531023, 0.0531786, 0.0532549, 0.0533312, 0.0534075, 0.0534838, 0.0535601, 0.0536364, 0.0537127, 0.053789, 0.0538653, 0.0539416, 0.0540179, 0.0540942, 0.0541705, 0.0542468, 0.0543231, 0.0543994, 0.0544757, 0.054552, 0.0546283, 0.0547046, 0.0547809, 0.0548572, 0.0549335, 0.0550098, 0.0550861, 0.0551624, 0.0552387, 0.055315, 0.0553913, 0.0554676, 0.0555439, 0.0556202, 0.0556965, 0.0557728, 0.0558491, 0.0559254, 0.0560017, 0.056078, 0.0561543, 0.0562306, 0.0563069, 0.0563832, 0.0564595, 0.0565358, 0.0566121, 0.0566884, 0.0567647, 0.056841, 0.0569173, 0.0569936, 0.0570699, 0.0571462, 0.0572225, 0.0572988, 0.0573751, 0.0574514, 0.0575277, 0.057604, 0.0576803, 0.0577566, 0.0578329, 0.0579092, 0.0579855, 0.0580618, 0.0581381, 0.0582144, 0.0582907, 0.058367, 0.0584433, 0.0585196, 0.0585959, 0.0586722, 0.0587485, 0.0588248, 0.0589011, 0.0589774, 0.0590537, 0.05913, 0.0592063, 0.0592826, 0.0593589, 0.0594352, 0.0595115, 0.0595878, 0.0596641, 0.0597404, 0.0598167, 0.059893, 0.0599693, 0.0600456, 0.0601219, 0.0601982, 0.0602745, 0.0603508, 0.0604271, 0.0605034, 0.0605797, 0.060656, 0.0607323, 0.0608086, 0.0608849, 0.0609612, 0.0610375, 0.0611138, 0.0611901, 0.0612664, 0.0613427, 0.061419, 0.0614953, 0.0615716, 0.0616479, 0.0617242, 0.0618, 0.0618763, 0.0619526, 0.0620289, 0.0621052, 0.0621815, 0.0622578, 0.0623341, 0.0624104, 0.0624867, 0.062563, 0.0626393, 0.0627156, 0.0627919, 0.0628682, 0.0629445, 0.0630208, 0.0630971, 0.0631734, 0.0632497, 0.063326, 0.0634023, 0.0634786, 0.0635549, 0.0636312, 0.0637075, 0.0637838, 0.0638601, 0.0639364, 0.0640127, 0.064089, 0.0641653, 0.0642416, 0.0643179, 0.0643942, 0.0644705, 0.0645468, 0.0646231, 0.0646994, 0.0647757, 0.064852, 0.0649283, 0.0650046, 0.0650809, 0.0651572, 0.0652335, 0.0653098, 0.0653861, 0.0654624, 0.0655387, 0.065615, 0.0656913, 0.0657676, 0.0658439, 0.0659202, 0.0659965, 0.0660728, 0.0661491, 0.0662254, 0.0663017, 0.066378, 0.0664543, 0.0665306, 0.0666069, 0.0666832, 0.0667595, 0.0668358, 0.0669121, 0.0669884, 0.0670647, 0.067141, 0.0672173, 0.0672936, 0.0673699, 0.0674462, 0.0675225, 0.0675988, 0.0676751, 0.0677514, 0.0678277, 0.067904, 0.0679803, 0.0680566, 0.0681329, 0.0682092, 0.0682855, 0.0683618, 0.0684381, 0.0685144, 0.0685907, 0.068667, 0.0687433, 0.0688196, 0.0688959, 0.0689722, 0.0690485, 0.0691248, 0.0692011, 0.0692774, 0.0693537, 0.06943, 0.0695063, 0.0695826, 0.0696589, 0.0697352, 0.0698115, 0.0698878, 0.0699641, 0.0700404, 0.0701167, 0.070193, 0.0702693, 0.0703456, 0.0704219, 0.0704982, 0.0705745, 0.0706508, 0.0707271, 0.0708034, 0.0708797, 0.070956, 0.0710323, 0.0711086, 0.0711849, 0.0712612, 0.0713375, 0.0714138, 0.0714901, 0.0715664, 0.0716427, 0.071719, 0.0717953, 0.0718716, 0.0719479, 0.0720242, 0.0721, 0.0721763, 0.0722526, 0.0723289, 0.0724052, 0.0724815, 0.0725578, 0.0726341, 0.0727104, 0.0727867, 0.072863, 0.0729393, 0.0730156, 0.0730919, 0.0731682, 0.0732445, 0.0733208, 0.0733971, 0.0734734, 0.0735497, 0.073626, 0.0737023, 0.0737786, 0.0738549, 0.0739312, 0.0740075, 0.0740838, 0.0741601, 0.0742364, 0.0743127, 0.074389, 0.0744653, 0.0745416, 0.0746179, 0.0746942, 0.0747705, 0.0748468, 0.0749231, 0.0749994, 0.0750757, 0.075152, 0.0752283, 0.0753046, 0.0753809, 0.0754572, 0.0755335, 0.0756098, 0.0756861, 0.0757624, 0.0758387, 0.075915, 0.0759913, 0.0760676, 0.0761439, 0.0762202, 0.0762965, 0.0763728, 0.0764491, 0.0765254, 0.0766017, 0.076678, 0.0767543, 0.0768306, 0.0769069, 0.0769832, 0.0770595, 0.0771358, 0.0772121, 0.0772884, 0.0773647, 0.077441, 0.0775173, 0.0775936, 0.0776699, 0.0777462, 0.0778225, 0.0778988, 0.0779751, 0.0780514, 0.0781277, 0.078204, 0.0782803, 0.0783566, 0.0784329, 0.0785092, 0.0785855, 0.0786618, 0.0787381, 0.0788144, 0.0788907, 0.078967, 0.0790433, 0.0791196, 0.0791959, 0.0792722, 0.0793485, 0.0794248, 0.0795011, 0.0795774, 0.0796537, 0.079729, 0.0798053, 0.0798816, 0.0799579, 0.0800342, 0.0801105, 0.0801868, 0.0802631, 0.0803394, 0.0804157, 0.080492, 0.0805683, 0.0806446, 0.0807209, 0.0807972, 0.0808735, 0.0809498, 0.0810261, 0.0811024, 0.0811787, 0.081255, 0.0813313, 0.0814076, 0.0814839, 0.0815602, 0.0816365, 0.0817128, 0.0817891, 0.0818654, 0.0819417, 0.082018, 0.0820943, 0.0821706, 0.0822469, 0.0823232, 0.0823995, 0.0824758, 0.0825521, 0.0826284, 0.0827047, 0.082781, 0.0828573, 0.0829336, 0.0830099, 0.0830862, 0.0831625, 0.0832388,
```

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata FotoFacebook.jpg
Component 1: Y component: Quantization table 0, Sampling factors 2 horiz/2 vert
Component 2: Cb component: Quantization table 1, Sampling factors 1 horiz/1 vert
Component 3: Cr component: Quantization table 1, Sampling factors 1 horiz/1 vert
Compression Type: Baseline
Content-Length: 207472
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: vie nov 03 09:18:14 +01:00 2023
File Name: FotoFacebook.jpg
File Size: 207472 bytes
ICC:Blue Colorant: (0.1431, 0.0686, 0.7139)
ICC:Blue TRC: 0.0, 0.0018921, 0.0037842, 0.0062867, 0.0095979, 0.0137179, 0.0186923, 0.0246281, 0.0315251, 0.0394293, 0.0483711, 0.0584115, 0.0695659, 0.0818799, 0.0953689, 0.1100786, 0.1260243, 0.1432517, 0.1617762, 0.1816129, 0.2028077, 0.225391, 0.2493629, 0.2747692, 0.3016098, 0.3299306, 0.3597314, 0.3910735, 0.4239261, 0.4583505, 0.4943465, 0.5319295, 0.5711452, 0.6119936, 0.65449, 0.6986648, 0.7445335, 0.7921111, 0.8413977, 0.8924544, 0.9452201, 1.0
ICC:Cyan Type: Icms
ICC:Class: Display Device
ICC:Color space: RGB
ICC:Device manufacturer: saws
ICC:Device model: ctrl
ICC:Green Colorant: (0.3051, 0.7169, 0.0971)
ICC:Green TRC: 0.0, 0.0018921, 0.0037842, 0.0062867, 0.0095979, 0.0137179, 0.0186923, 0.0246281, 0.0315251, 0.0394293, 0.0483711, 0.0584115, 0.0695659, 0.0818799, 0.0953689, 0.1100786, 0.1260243, 0.1432517, 0.1617762, 0.1816129, 0.2028077, 0.225391, 0.2493629, 0.2747692, 0.3016098, 0.3299306, 0.3597314, 0.3910735, 0.4239261, 0.4583505, 0.4943465, 0.5319295, 0.5711452, 0.6119936, 0.65449, 0.6986648, 0.7445335, 0.7921111, 0.8413977, 0.8924544, 0.9452201, 1.0
ICC:Media White Point: (0.9505, 1, 1.089)
ICC:Primary Platform: Microsoft Corporation
ICC:Profile Connection Space: XYZ
ICC:Profile Copyright: CCB
ICC:Profile Date/Time: 2018-03-20 09:14:29
ICC:Profile Description: uRGB
ICC:Profile Size: 456
ICC:Red Colorant: (0.436, 0.2224, 0.0139)
ICC:Red TRC: 0.0, 0.0018921, 0.0037842, 0.0062867, 0.0095979, 0.0137179, 0.0186923, 0.0246281, 0.0315251, 0.0394293, 0.0483711, 0.0584115, 0.0695659, 0.0818799, 0.0953689, 0.1100786, 0.1260243, 0.1432517, 0.1617762, 0.1816129, 0.2028077, 0.225391, 0.2493629, 0.2747692, 0.3016098, 0.3299306, 0.3597314, 0.3910735, 0.4239261, 0.4583505, 0.4943465, 0.5319295, 0.5711452, 0.6119936, 0.65449, 0.6986648, 0.7445335, 0.7921111, 0.8413977, 0.8924544, 0.9452201, 1.0
ICC:Signature: acsp
ICC:Tag Count: 9
ICC:Version: 2.1.0
ICC:XYZ values: 0.964 1 0.825
Image Height: 1241 pixels
Image Width: 877 pixels
Number of Components: 3
Number of Tables: 4 Huffman tables
Resolution Units: none
Thumbnail Height Pixels: 0
Thumbnail Width Pixels: 0
Version: 1.1
X Resolution: 1 dot
X-TIKA:Parsed-By: org.apache.tika.parser.DefaultParser
```

Figura 31 Metadatos foto Facebook

```
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>java -jar tika-app-2.9.1.jar --metadata FotoPinterest.jpg
Component 1: Y component: Quantization table 0, Sampling factors 2 horiz/2 vert
Component 2: Cb component: Quantization table 1, Sampling factors 1 horiz/1 vert
Component 3: Cr component: Quantization table 1, Sampling factors 1 horiz/1 vert
Compression Type: Progressive, Huffman
Content-Length: 76956
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: vie nov 03 09:22:01 +01:00 2023
File Name: FotoPinterest.jpg
File Size: 76956 bytes
Image Height: 564 pixels
Image Width: 564 pixels
Number of Components: 3
Number of Tables: 2 Huffman tables
Resolution Units: inch
Thumbnail Height Pixels: 0
Thumbnail Width Pixels: 0
Version: 1.1
X Resolution: 72 dots
X-TIKA:Parsed-By: org.apache.tika.parser.DefaultParser
X-TIKA:Parsed-By: org.apache.tika.parser.image.JpegParser
Y Resolution: 72 dots
resourceName: FotoPinterest.jpg
tiff:BitsPerSample: 8
tiff:ImageLength: 564
tiff:ImageWidth: 564
C:\Users\javie\OneDrive\Escritorio\REC-INF\Practica2>
```

Figura 32 Metadatos foto Pinterest

**Ejercicio 11: Se puede trabajar con Tika desde Eclipse. Describe los pasos que han de realizarse para poder crear un proyecto Tika en Eclipse en el que se extraiga el contenido de un fichero PDF.**

Primero, una vez instalado eclipse, creamos un proyecto java y desmarcamos la opción “create module-info.java file”, ver Figura 33.

Después creamos en el directorio “src” un fichero clase con el nombre “ExtraerPdf”, ver Figura 34 y 35.

Posteriormente, añadimos en el build path el fichero .jar de Tika para añadir sus librerías, ver Figura 36 y 37.

Con esto ya podríamos crear nuestro script (explicación a continuación) : Se instancia un objeto de la clase File de java con la referencia al fichero deseado (path hacia el fichero), se instancia un objeto de la clase Tika, definimos una variable String que almacene el contenido del fichero y para obtener este, usamos el método `parseToString()` de la clase Tika, que recibe un fichero de entrada y devuelve su contenido. Por ultimo se imprime por pantalla el contenido. Para una visualización del código ver Figura 38.

Por último, ejecutamos el código de “ExtraerPdf.java” y obtendremos en la consola de salida el contenido del pdf, ver Figura 39.

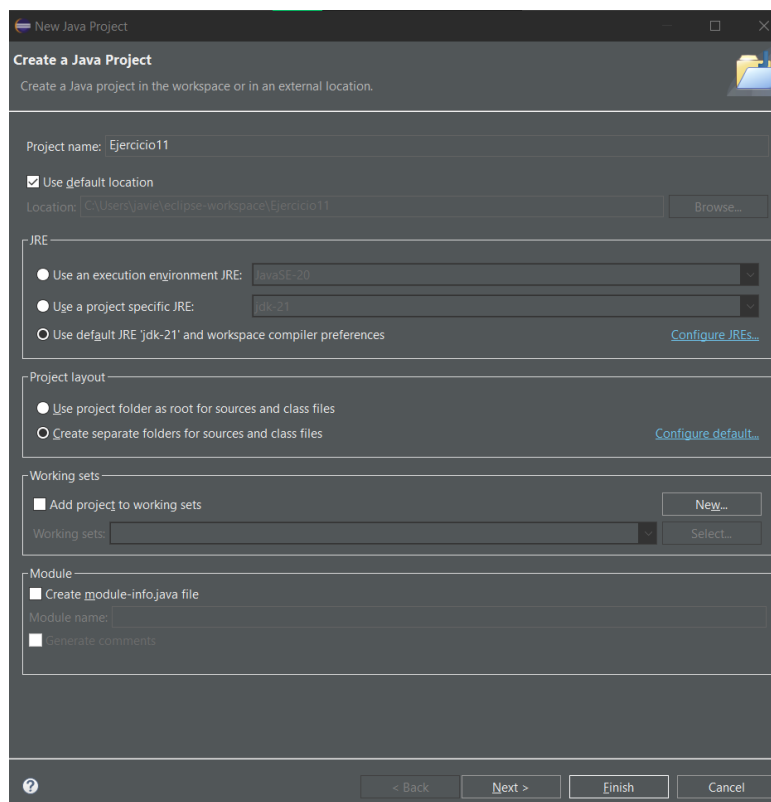


Figura 33 Creación del proyecto

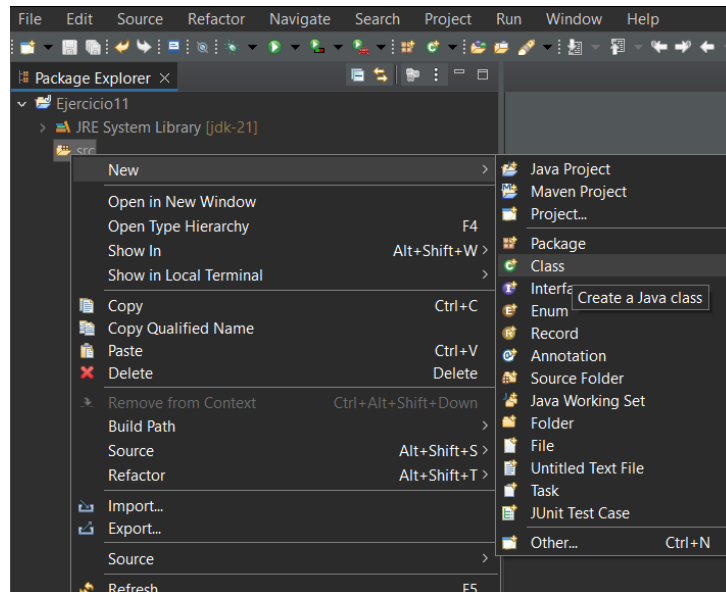


Figura 34 Creacion del fichero class I

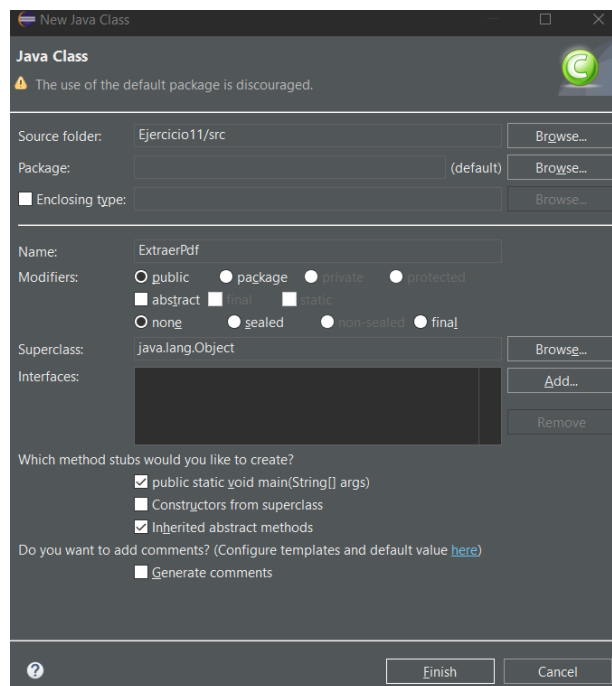


Figura 35 Creacion del fichero class II

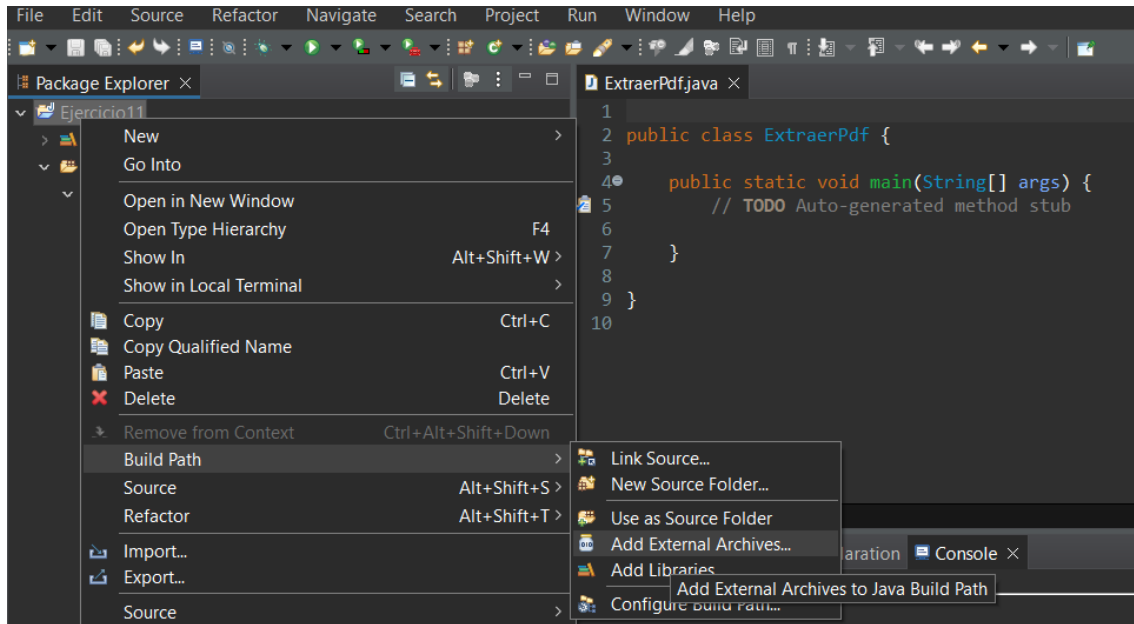


Figura 36 Añadir el fichero jar al Build Path I

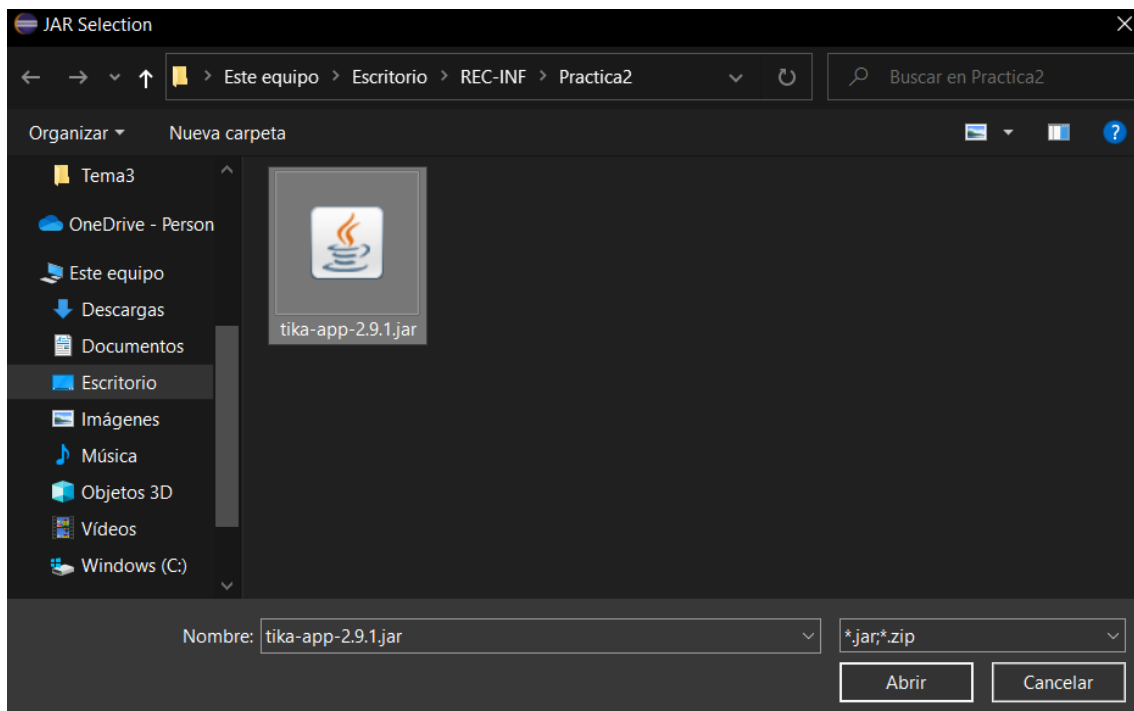


Figura 37 Añadir el fichero jar al Build Path II

```
ExtraerPdf.java ×
1 import org.apache.tika.Tika;
2 import org.apache.tika.exception.TikaException;
3
4 import java.io.File;
5 import java.io.IOException;
6
7 public class ExtraerPdf {
8
9     public static void main(String[] args) throws IOException, TikaException {
10         File ficheroPdf = new File("C:\\Users\\javie\\OneDrive\\Escritorio\\REC-INF\\Practica1\\Practica_1_Documentación.pdf");
11         Tika t = new Tika();
12         String Contenido = t.parseToString(ficheroPdf);
13         System.out.println(Contenido);
14     }
15 }
16
17
18
19
20
```

Figura 38 Código de ExtraerPdf.java

```
Problems Javadoc Declaration Console ×
<terminated> ExtraerPdf [Java Application] C:\Program Files\Java\jdk-21\bin\javaw.exe (4 nov 2023, 14:52:50 – 14:52:52) [pid: 4624]

Francisco Javier Molina Rojas

Recuperación de la Información

Práctica 1 - Expresiones Regulares en Java

Ejercicio 1. Comprobar si una cadena empieza por "abc".
Expresión regular usada: "^abc.*"

Figura 1 Código del ejercicio 1
```

Figura 39 Salida por consola (contenido del pdf)