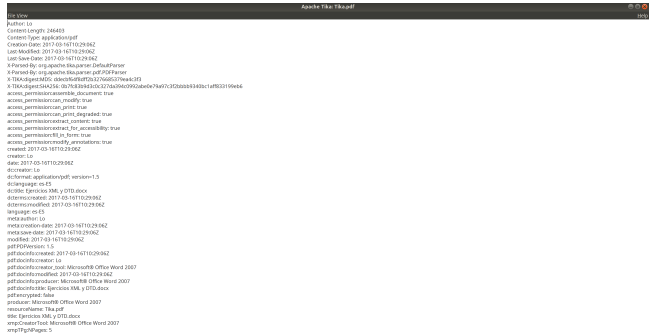
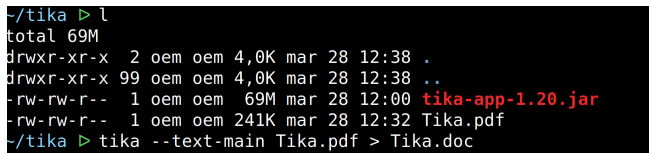


1. El autor es ‘Lo’.

1. El autor es ‘Lo’.

2. `tika -text-main Tika.pdf > Tika.doc`

3. tika -metadata https://www.uca.es



4.
 - a) echo 'Archivo de texto plano listo para comprimir.' > archivo.txt
 - b) cat archivo.txt
 - c) zip archivo.zip archivo.txt

d) tika -detect archivo.zip

```
/tika @ echo "Archivo de texto plano listo para comprimir" > archivo.txt
~/tika @ cat archivo.txt
Archivo de texto plano listo para comprimir
~/tika @ zip archivo.zip archivo.txt
  adding: archivo.txt (deflated 2%)
~/tika @ tika -detect archivo.zip
mar 28, 2019 12:42:32 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: J2KImageReader not loaded. JPEG2000 files will not be processed.
See https://pdfbox.apache.org/2.0/dependencies.html#jai-image-io
for optional dependencies.
mar 28, 2019 12:42:32 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: org.xerial's sqlite-jdbc is not loaded.
Please provide the jar on your classpath to parse sqlite files.
See tika-parsers/pom.xml for the correct version.
application/zip
```

Figura 4: Ejercicio 4.

e) tika -metadata archivo.zip

```
/tika @ tika -metadata archivo.zip
mar 28, 2019 12:43:04 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: J2KImageReader not loaded. JPEG2000 files will not be processed.
See https://pdfbox.apache.org/2.0/dependencies.html#jai-image-io
for optional dependencies.
mar 28, 2019 12:43:04 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: org.xerial's sqlite-jdbc is not loaded.
Please provide the jar on your classpath to parse sqlite files.
See tika-parsers/pom.xml for the correct version.
Content-Length: 215
Content-Type: application/zip
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.pkg.PackageParser
resourceName: archivo.zip
```

Figura 5: Ejercicio 4.

5. a) tika -metadata ImagenCorreo.jpg

```
/tika @ tika -metadata ImagenCorreo.jpg
mar 28, 2019 12:52:22 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: J2KImageReader not loaded. JPEG2000 files will not be processed.
See https://pdfbox.apache.org/2.0/dependencies.html#jai-image-io
for optional dependencies.
mar 28, 2019 12:52:22 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: org.xerial's sqlite-jdbc is not loaded.
Please provide the jar on your classpath to parse sqlite files.
See tika-parsers/pom.xml for the correct version.
Comments: CREATOR: gd-jpeg v1.0 (using IJG JPEG v62), quality = 85
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: 773560
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: jun mar 28 12:50:52 +01:00 2019
File Name: ImagenCorreo.jpg
File Size: 773560 bytes
Image Height: 1606 pixels
Image Width: 3520 pixels
JPEG Comment: CREATOR: gd-jpeg v1.0 (using IJG JPEG v62), quality = 85
Number of Components: 3
Number of Tables: 4 Huffman tables
Resolution Units: none
Thumbnail Height Pixels: 0
Thumbnail Width Pixels: 0
X Resolution: 1 dot
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.jpeg.JpegParser
Y Resolution: 1 dot
Comments: CREATOR: gd-jpeg v1.0 (using IJG JPEG v62), quality = 85
resourceName: ImagenCorreo.jpg
tiff:BitsPerSample: 8
tiff:ImageLength: 1606
tiff:ImageWidth: 3520
w:comments: CREATOR: gd-jpeg v1.0 (using IJG JPEG v62), quality = 85
```

Figura 6: Ejercicio 5.

b) tika -metadata ImagenWhatsapp.jpeg

```

$ tika -l tika --metadata ImageWhatsApp.jpeg
Mar 28, 2019 12:52:49 PM org.apache.tika.config.InitializableProblemHandlers$3 handleInitializableProblem
WARNING: TikaImageReader not loaded. JSDOMParser files will not be processed.
See https://pdfbox.apache.org/2.0/dependencies.html#aj-image-io
for optional dependencies.

Mar 28, 2019 12:52:49 PM org.apache.tika.config.InitializableProblemHandlers$3 handleInitializableProblem
WARNING: org.xml.sax.SaxException: not loaded.

Please provide the jar on your classpath to parse scribe files.

org.apache.pdfbox.pdmodel for the correct version.
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: 276260
Content Type: image/jpeg
Data Precision: 8 bits
File Modified Date: Tue Mar 28 12:50:37 +0100 2019
File Name: ImageWhatsApp.jpeg
File Size: 276260 bytes
Image Height: 720 pixels
Image Width: 1600 pixels
Number of Components: 3
Number of Tables: 4 Huffman tables
Resolution Units: none
Thumbnail Height Pixels: 0
Thumbnail Width Pixels: 0
0 Resolution: 1 dot
0 Parsed By: org.apache.tika.parser.DefaultParser
0 Parsed By: org.apache.tika.parser.jpeg.jpegParser
0 Resolution: 1 dot
resourceName: ImageWhatsApp.jpeg
riff:8bitsSample: 8
riff:ImageLength: 720
riff:ImageWidth: 1600

```

Figura 7: Ejercicio 5.

6. `tika -metadata https://www.uca.es > metadatosuca.txt`

```
Content-Encoding: UTF-8
Content-Language: es-ES
Content-Length: 97207
Content-Type: text/html; charset=UTF-8
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.html.HtmlParser
X-UA-Compatible: IE=EDGE
dc:title: Portal UCA { Portal principal de la Universidad de Cádiz
description:
generator: WPML ver:4.0.7 stt:1,4,2;
title: Portal UCA { Portal principal de la Universidad de Cádiz
viewport: width=device-width
```

7. a) `tika -text-main celebs_fixed.rdf > celebs_fixed.doc`

```

$ ./tika-b-tika --text-main celebs.fixed.rtf > celebs.fixed.doc
mar 28, 2019 1:18:43 PM org.apache.tika.config.InitializeableProblemHandler$3 handleInitializeableProblem
ADVERTENCIA: J2KImageReader not loaded. JPEG2000 files will not be processed.
See https://pdfbox.apache.org/2.0/dependencies.html#jai-image-io
for optional dependencies.
mar 28, 2019 1:18:43 PM org.apache.tika.config.InitializeableProblemHandler$3 handleInitializeableProblem
ADVERTENCIA: org.apache.xerces.impl.xpath.XPathNavigator not loaded.
Please provide the jar on your classpath to parse sqlite files.
See tika-parsers/pom.xml for the correct version.

```

Figura 8: Ejercicio 7.

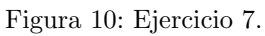
b) `tika -text-main celebs_fixed.rdf > celebs_fixed.pdf`

```

$ ./tika --text-main celebs fixed.pdf --> celebs fixed.pdf
mar 28, 2019 1:20:18 PM org.apache.tika.config.InitializeableProblemHandler$3 handleInitializeableProblem
ADVERTENCIA: J2KImageReader not loaded. J2PEGTools files will not be processed.
See https://pdfbox.apache.org/2.0/dependencies.html#ajl-image-io
for optional dependencies.
mar 28, 2019 1:20:18 PM org.apache.tika.config.InitializeableProblemHandler$3 handleInitializeableProblem
ADVERTENCIA: org.xerial:salticrabe is not loaded.
Please provide the jar on your classpath to parse salticrabe files.
See https://commons.apache.org/ for the correct version(s).

```

Figura 9: Ejercicio 7.



8.
 - a) `tika -metadata q.jpg > q.txt`
 - b) `tika -metadata r.jpg > q.txt`
 - c) `tika -metadata s.JPG > q.txt`

Figura 11: Ejercicio 8.

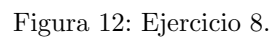


Figura 13: Ejercicio 9.

Figura 14: Ejercicio 9.

Page 5 de 9

```

~/tika -metadata uca.doc
mar 28, 2019 1:40:30 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: J2KImageReader not loaded. JPEG2000 files will not be processed.
See https://pdfbox.apache.org/2.0/dependencies.html#jai-image-io
for optional dependencies.

mar 28, 2019 1:40:30 PM org.apache.tika.config.InitializableProblemHandler$3 handleInitializableProblem
ADVERTENCIA: org.xerial's sqlite-jdbc is not loaded.
Please provide the jar on your classpath to parse sqlite files.
See tika-parsers/pom.xml for the correct version.
Content-Encoding: UTF-8
Content-Encoding: UTF-8
Content-Language: es-ES
Content-Language: es-ES
Content-Language: es-ES
Content-Length: 58837
Content-Length: 58322
Content-Length: 97207
Content-Type: application/xhtml+xml; charset=UTF-8
Content-Type-Hint: application/xhtml+xml; charset=UTF-8
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.html.HtmlParser
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.html.HtmlParser
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.html.HtmlParser
X-UA-Compatible: IE=edge
dc:title: Portal UCA - Portal principal de la Universidad de Cádiz
Description:
Generator: WPML ver:4.0.7 stt:1.4.2;
resourceName: uca.doc
title: Portal UCA - Portal principal de la Universidad de Cádiz
viewport: width=device-width

```

Figura 15: Ejercicio 9.

10. a) `tika -metadata primera.jpg > primera.txt`

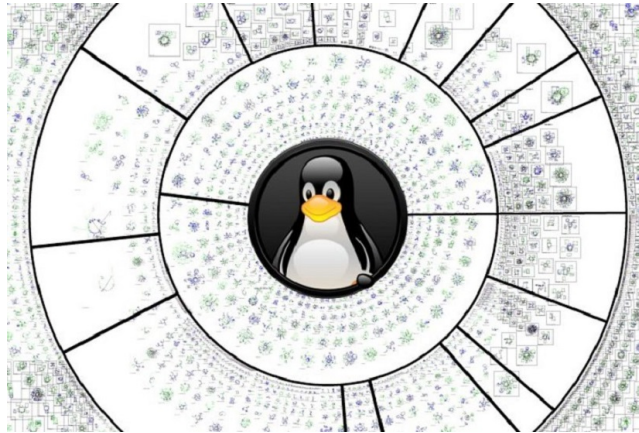


Figura 16: Ejercicio 10.

```

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: 163975
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: jue mar 28 13:47:27 +01:00 2019
File Name: primera.jpg
File Size: 163975 bytes
Image Height: 554 pixels
Image Width: 830 pixels

```

```
Number of Components: 3
Number of Tables: 2 Huffman tables
Resolution Units: inch
Thumbnail Height Pixels: 0
Thumbnail Width Pixels: 0
X Resolution: 72 dots
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.jpeg.JpegParser
Y Resolution: 72 dots
resourceName: primera.jpg
tiff:BitsPerSample: 8
tiff:ImageLength: 554
tiff:ImageWidth: 830
```

b) `tika -metadata segunda.png > segunda.txt`

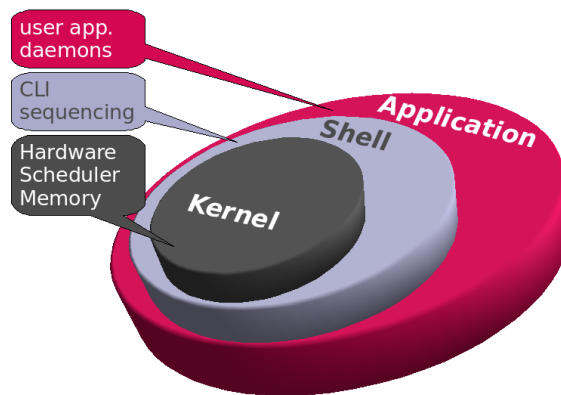


Figura 17: Ejercicio 10.

```
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: 163975
Content-Type: image/jpeg
Data Precision: 8 bits
File Modified Date: jue mar 28 13:47:27 +01:00 2019
File Name: primera.jpg
File Size: 163975 bytes
Image Height: 554 pixels
Image Width: 830 pixels
```

Number of Components: 3
 Number of Tables: 2 Huffman tables
 Resolution Units: inch
 Thumbnail Height Pixels: 0
 Thumbnail Width Pixels: 0
 X Resolution: 72 dots
 X-Parsed-By: org.apache.tika.parser.DefaultParser
 X-Parsed-By: org.apache.tika.parser.jpeg.JpegParser
 Y Resolution: 72 dots
 resourceName: primera.jpg
 tiff:BitsPerSample: 8
 tiff:ImageLength: 554
 tiff:ImageWidth: 830

c) tika -metadata tercera.png > tercera.txt

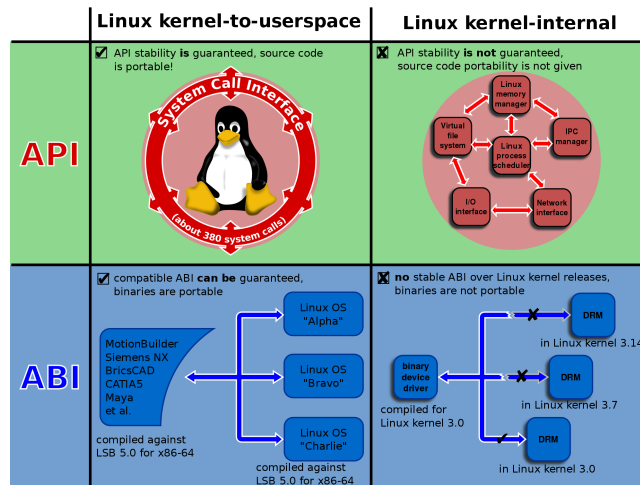


Figura 18: Ejercicio 10.

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: 163975
 Content-Type: image/jpeg
 Data Precision: 8 bits
 File Modified Date: jue mar 28 13:47:27 +01:00 2019
 File Name: primera.jpg
 File Size: 163975 bytes
 Image Height: 554 pixels
 Image Width: 830 pixels

Number of Components: 3
Number of Tables: 2 Huffman tables
Resolution Units: inch
Thumbnail Height Pixels: 0
Thumbnail Width Pixels: 0
X Resolution: 72 dots
X-Parsed-By: org.apache.tika.parser.DefaultParser
X-Parsed-By: org.apache.tika.parser.jpeg.JpegParser
Y Resolution: 72 dots
resourceName: primera.jpg
tiff:BitsPerSample: 8
tiff:ImageLength: 554
tiff:ImageWidth: 830