Extraer páginas de un PDF en MacOs X desde el shell

Antecedentes

Que si esta, que si la otra,... un baile de aplicaciones algunas de ellas muy engañosas con el concepto **gratis** cuando en realidad son aplicaciones capadas que buscan tu dinero.

Opensource que no es lo mismo y es mejor. En mi artículo Extraer páginas de un PDF en Linux desde el shell escribí sobre algo parecido para Linux. Hoy le toca el turno a Mac Os.

Tras un incidente con la falta de usabilidad de Dominio.es en el que te piden un DNI en el proceso de validación de un cambio de titular, tuve que extraer del PDF que me envió el cliente, las dos imágenes del DNI, y después, juntarlas en una sola imagen (no admiten más de una), para terminar con un, **se me olvide decirte que la imagen no puede tener más de 600px de ancho) realice el trabajo en el shell

Programa Poppler

Poppler es un programa de rendering de PDF basado en xpdf

Usaremos Homebrew para instalarlo.

```
brew install poppler
```

Con Poppler, extremos las imágenes a través de su binario **pdfimages**

```
pdfimages archivo.pdf prefijo_salida
```

- `archivo.pdf` es el nombre de tu archivo
- `prefijo_salida` es el prefijo que tendrán tus imágenes extraidas

Si lo hacemos así las imágenes saldrán en formato .ppm así que tendríamos que convertirlos a JPG pro imperativo de dominios.es.

Programa ImageMagick

Usaremos uno de los programas de ImageMagick, así que si no lo tenemos (es un super programa, tanto en Linux como en Mac) lo instalaremos.

```
brew install imagemagick
```

Una vez instalado de todos los binarios o programas que instala usaremos varios.

`convert` para convertirlo a otro formato

```
convert prefijo_salida-0000.ppm prefijo_salida-0000.jpg
```

Como tenemos mas de una imagen que bueno un comando

```
for img in *.ppm; do convert "$img" "${img%.ppm}.jpg"; done
```

Tip

Si antes de hacer nada leemos el man de **poppler** directamente podemos hacer la extracción a jpg.

Usamos el parametro `-j`

```
pdfimages -j archivo.pdf prefijo_salida
```

Crear un único archivo

Bueno, cuando vas a subir el fichero descubres que sólo quieren un fichero.

Nos toca usar convert otra vez para añadirlos.

```
convert prefijo_salida-0000.jpg prefijo_salida-0001.jpg +append resultado_final.jpg
```

Ay... que ahora no lo quieren de max de 600px de ancho.

```
convert prefijo_salida-0000.jpg prefijo_salida-0001.jpg -append resultado_final.jpg
```

Observa que `+append` se convirtió en `-append` para crearlo en vertical uno tras otro.

Pero, no. No comprobé que el ancho de cada imagen fuera de menos de 600px, y resulto que tenía 610.

Bueno, ya es el último.

```
convert resultado_final.jpg -resize 600x resultado_final_redimensionado.jpg
```

Finales

Espero que lo hallas disfrutado. Ahora te dejo que tengo que seguir con mi proyecto, Baytuka

Aviso

Esta documentación y su contenido, no implica que funcione en tu caso o determinados casos. También implica que tienes conocimientos sobre lo que trata, y que en cualquier caso tienes copias de seguridad. El contenido el contenido se entrega, tal y como está, sin que ello implique ningún obligación ni responsabilidad por parte de Castris

Si necesitas soporte profesional puedes contratar con Castris soporte profesional.

Revision #1
Created 2024-05-07 10:18:40 UTC by Abkrim
Updated 2024-05-07 10:18:59 UTC by Abkrim