



B Universitat de Barcelona

CRAI Unitat de Serveis als Usuaris

MANUAL DE DIGITALITZACIÓ

CENTRE DE RECURSOS PER A L'APRENTATGE I LA INVESTIGACIÓ

1 - Guia per escaneig de documents

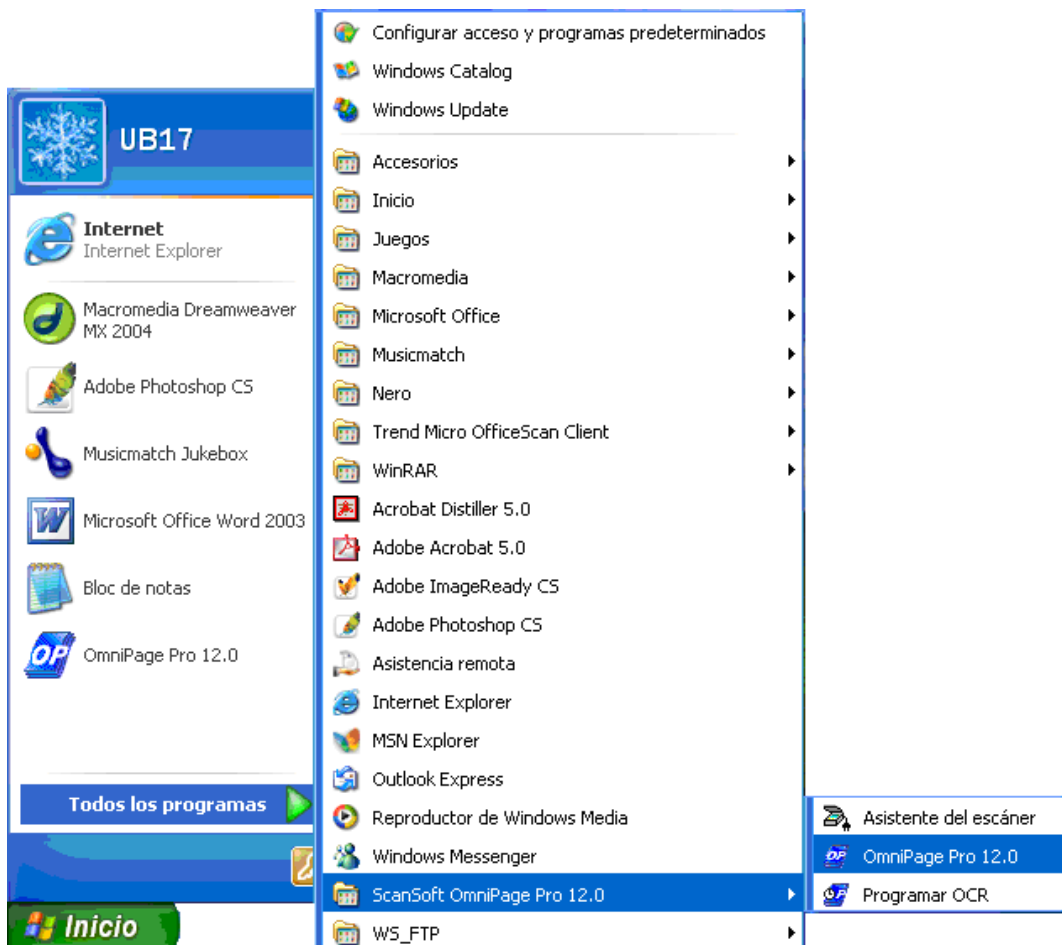
amb el programa Omnipage Pro 12.0

Aquesta guia suposa que es disposa d'escàner i del programa Omnipage Pro 12.0 correctament instal·lats.

L'aplicació Omnipage Pro 12.0, tot i que té altres utilitats, és un programa dissenyat principalment per escanejar documents de text i guardar-los en un format editable. Tot això es realitza mitjançant l'eina OCR (Optical Character Recognition). Aquesta eina llegeix imatges i les reconeix en text que pot ser editable en un processador de textos tipus Word.

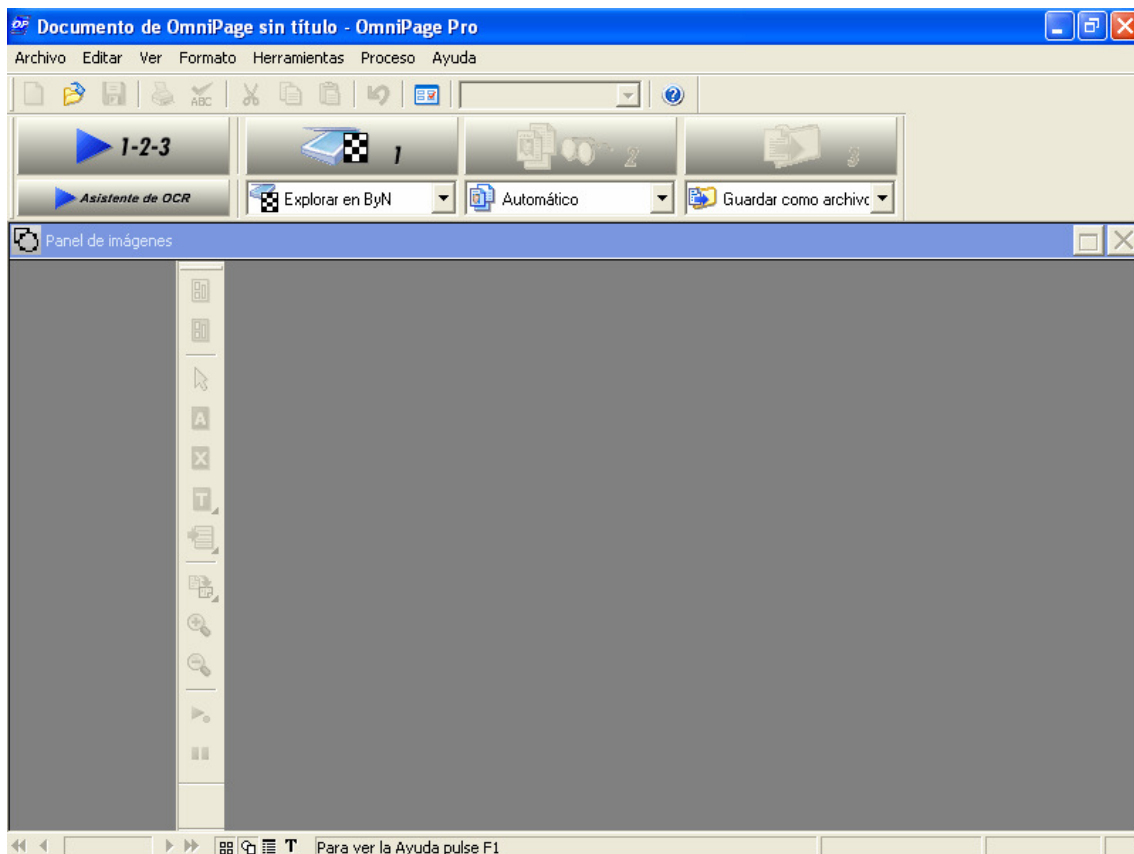
Els passos que han de seguir-se per escanejar un document de text són els següents:

A. Primer de tot executar el programa, que es troba sota la carpeta *Scansoft Omnipage Pro 12.0* si no s'ha indicat el contrari en la instal·lació.



CRAI Unitat de Serveis als Usuaris

En iniciar el programa apareixerà la següent pantalla:



B. Es pot posar ara un full a l'escàner.

C. En la pantalla de l'Omnipage poden distingir-se dues opcions que es troben ressaltades.



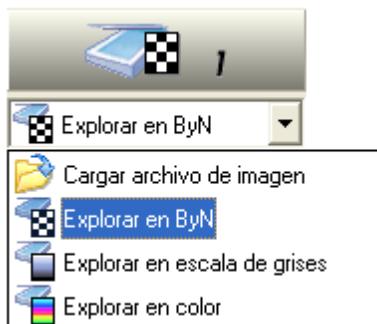
Aquest botó portaria a:

1. Escanejar la pàgina
2. Reconèixer la imatge
3. Guardar el document

Tots aquests passos s'executen de forma automàtica.

És, en canvi, més recomanable realitzar els tres passos de forma manual ja que es poden així escollir unes opcions interessants.

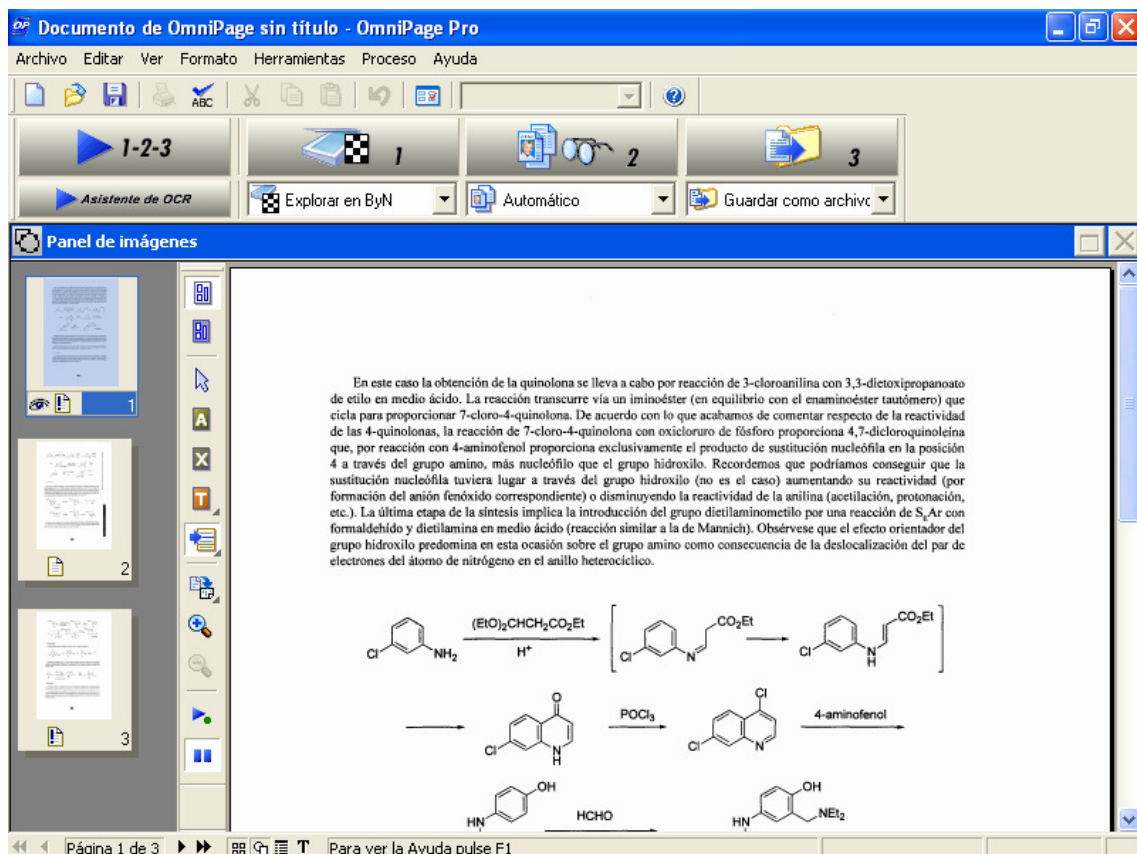
CRAI Unitat de Serveis als Usuaris



El següent botó, en desplegar la llengüeta, dóna diverses opcions. Es pot explorar:

1. En només blanc i negre: recomanable quan el que s'escanegi sigui tan sols text.
2. En escala de grisos: si a més de text conté algun gràfic en blanc i negre.
3. En color: si hi ha imatges en color i volen guardar-se d'aquesta manera.
4. Un arxiu d'imatge: si ja es disposa de la imatge digitalitzada i es vol reconèixer el text.

S'han posat uns fulls a l'escàner a tall d'exemple. Si s'escaneja en només blanc i negre i es pressiona la icona número 1, obtindrem una pàgina del tipus:



Documento de OmniPage sin título - OmniPage Pro

Archivo Editar Ver Formato Herramientas Proceso Ayuda

1-2-3

Asistente de OCR

Explorar en ByN Automático Guardar como archiv

Panel de imágenes

En este caso la obtención de la quinolona se lleva a cabo por reacción de 3-cloroanilina con 3,3-dietoxipropanoato de etilo en medio ácido. La reacción transcurre vía un iminoéster (en equilibrio con el enamoéster tautómero) que cicla para proporcionar 7-cloro-4-quinolona. De acuerdo con lo que acabamos de comentar respecto de la reactividad de las 4-quinolonas, la reacción de 7-cloro-4-quinolona con oxícloruro de fósforo proporciona 4,7-dicloroquinoleína que, por reacción con 4-aminofenol proporciona exclusivamente el producto de sustitución nucleófila en la posición 4 a través del grupo amino, más nucleófilo que el grupo hidroxilo. Recordemos que podríamos conseguir que la sustitución nucleófila tuviera lugar a través del grupo hidroxilo (no es el caso) aumentando su reactividad (por formación del anión fenóxido correspondiente) o disminuyendo la reactividad de la anilina (acetilación, protonación, etc.). La última etapa de la síntesis implica la introducción del grupo dietilaminometilo por una reacción de S_NAr con formaldehído y dietilamina en medio ácido (reacción similar a la de Mannich). Obsérvese que el efecto orientador del grupo hidroxilo predomina en esta ocasión sobre el grupo amino como consecuencia de la deslocalización del par de electrones del átomo de nitrógeno en el anillo heterocíclico.

Clc1ccc(N)cc1
 $\xrightarrow{H^+}$
 $(EtO)_2CHCH_2CO_2Et$
 \rightarrow
Clc1ccc(NC(=O)OCC)cc1
 \rightarrow
Clc1ccc2c(c1)cnc2
 $\xrightarrow{POCl_3}$
Clc1ccc2c(c1)cnc2
 $\xrightarrow{4\text{-aminofenol}}$
Clc1ccc2c(c1)cnc2
 \rightarrow
Clc1ccc2c(c1)cnc2

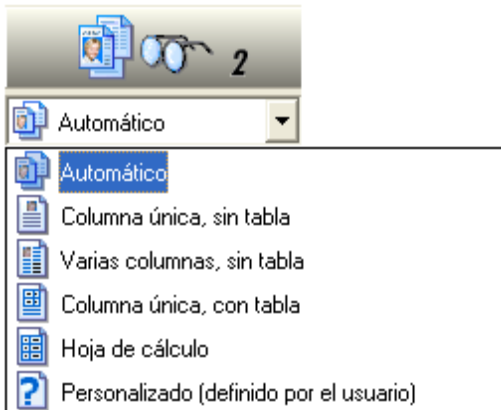
Clc1ccc2c(c1)cnc2
 \xrightarrow{HCHO}
Clc1ccc2c(c1)cnc2

Página 1 de 3

Para ver la Ayuda pulse F1

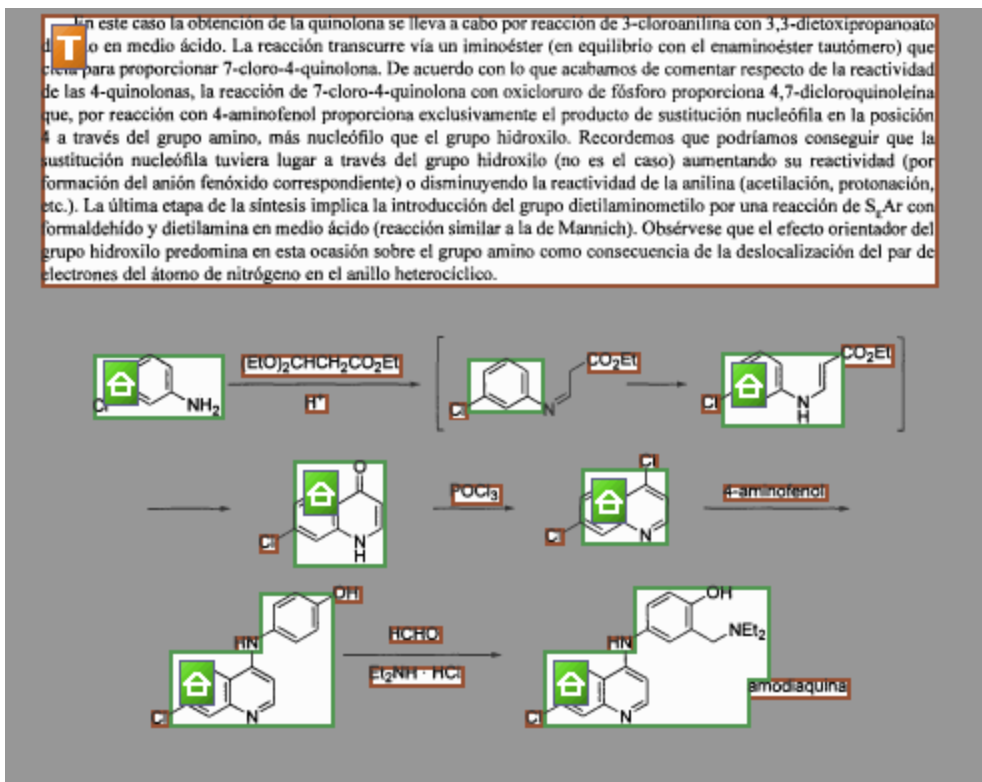
CRAI Unitat de Serveis als Usuaris

D. S'activen ara, altres possibilitats.



Aquesta és la icona de reconeixement. Reconeix la imatge escanejada en text, imatge o taula. L'opció que es recomana fer servir és *Automático*.

Abans de prémer aquesta icona però, es recomana especificar el tipus de zones que es tenen. Si no es fa així, un reconeixement directe donaria com a resultat el següent:



Les seccions verdes són reconegudes com a imatge i les vermelles com a text.

Donat el resultat de fer la selecció automàtica es recomana seleccionar manualment les seccions. Això es portarà a terme mitjançant les icones del centre de la pantalla.

CRAI Unitat de Serveis als Usuaris



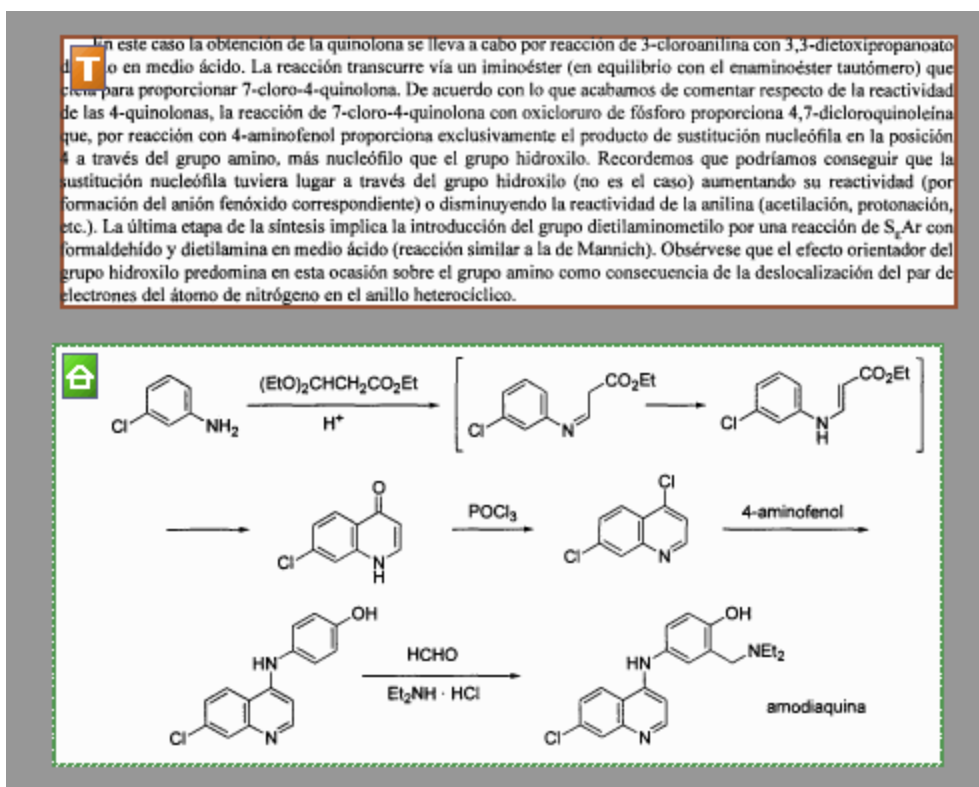
Mitjançant la icona vermella es poden seleccionar les seccions que són text.

Amb la icona blava les seccions que són taules. Informació que està continguda en files i columnes.

Amb la verda, aquelles que són imatges.

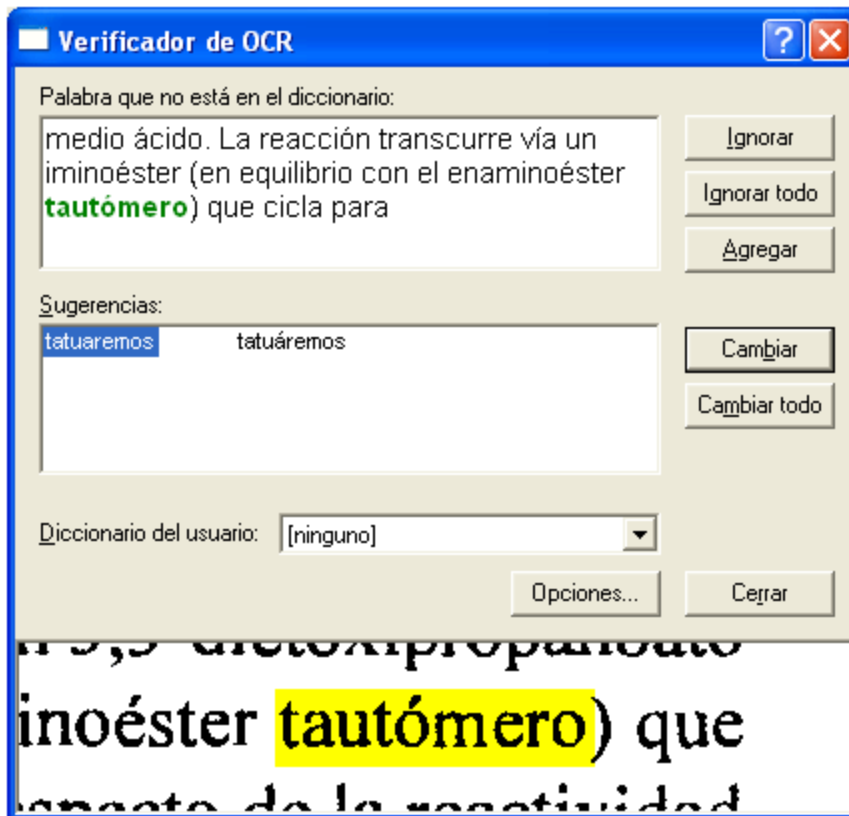
Per canviar d'icona s'ha de pitjar a la fletxa grisa que es troba en la part inferior a la dreta del botó.

Fent la selecció manualment podria fer-se una selecció del tipus.



Un cop seleccionats els diferents tipus de zones de cada pàgina es premerà el botó de reconeixement (*Automático*). En aquest cas es tenen dues seccions, una de text i l'altre d'imatge. Les zones en gris no seran reconegudes, les zones vermelles (text) i blaves (taules) obriran una eina tipus diccionari, les zones en verd (imatge) seran directament exportades sense un tractament previ.

L'eina diccionari obre una finestra com la següent.

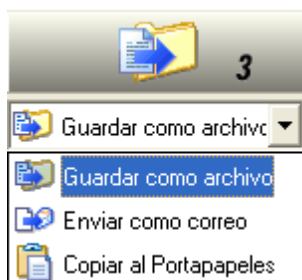


L'Omnipage té incorporat un diccionari en castellà. Un dels avantatges d'aquesta eina és quan el document no té molt bona resolució. Les paraules que no han estat ben reconegudes no es troben contingudes en el diccionari i aquest les mostra una a una i ofereix la possibilitat de modificar-les manualment (en la part ressaltada en verd) o automàticament oferint altres paraules semblants (si les hi ha).

Sempre es pot cancel·lar aquesta opció tancant la finestra del diccionari.

Un dels problemes que presenta Omnipage és que tan sols conté un diccionari en castellà i per tant quan es vulguin reconèixer documents en altres llengües el diccionari es parará pràcticament a cada paraula; el que pot fer-se és tancar l'eina del diccionari. Potser llavors hi haurà paraules que han estat mal reconegudes però que podran ser editables en qualsevol programa de text.

D. Després de fer la selecció i el reconeixement de les seccions s'ha de guardar el document en un arxiu en el format desitjat mitjançant el botó número 3.



En prémer-lo el programa obrirà una finestra de diàleg que permet guardar el document en algun dels formats següents:

Excel [*.xls]

Text [*.txt]

PDF [*.pdf]: es pot llegir amb l'Acrobat Reader o editar-se amb l'Adobe Acrobat.

Html [*.htm]: pot llegir-se amb qualsevol navegador d'Internet o editar-se en qualsevol editor de pàgines web tipus Dreamweaver o Frontpage.

Microsoft Word [*.doc]

RTF [*.rtf]: editable en qualsevol processador de text tipus Word o Wordpad.

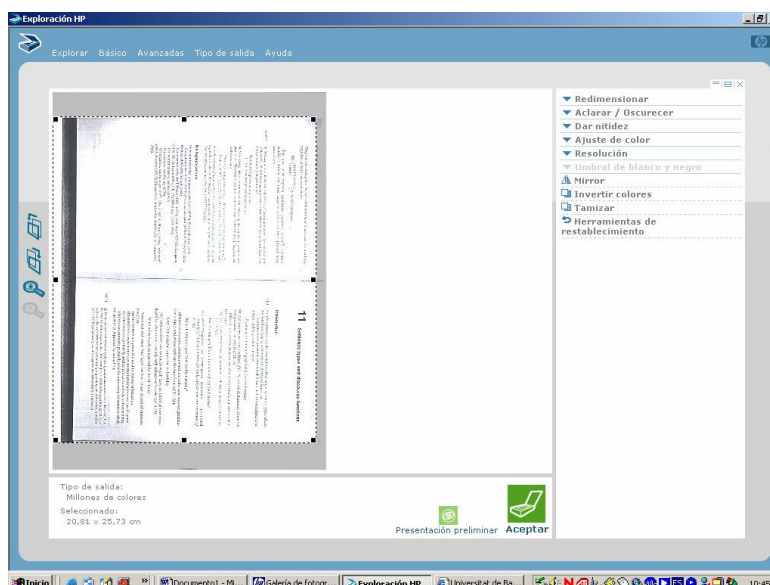
2- Guia per escaneig de documents

amb el programa HPDirector

MANUAL DIGITALITZACIÓ¹

- Documents text a .pdf

1. Obrir el programa de l'escàner
2. *Herramientas* i *Explorar*
3. El programa farà una exploració del document
4. S'obrirà la *Pantalla de Exploración*
5. Amb els requadres que marquen el document podem seleccionar la part que volem escanejar i després fer els canvis necessaris per obtenir el resultat amb les diferents funcions de la pantalla d'exploració.



Les diferents funcions² de la pantalla serien: *Explorar* (per canviar tipus de suport) en aquest cas el que surt per defecte; *Básicos* (per rotar la imatge, redimensionar) sobretot farem servir *Rotar imagen*; *Avanzadas* (per ajustar color, variar blanc i negre, la resolució) en el nostre cas al ser blanc i negre utilitzarem *Umbral de blanco y negro* que s'activa des de *Tipo de salida*, amb aquesta opció podrem treure foscors del full; *Tipo de salida* (les diferents formes en color, gris i b/n) per la nostra opció *Blanco y negro (1 bit)*.

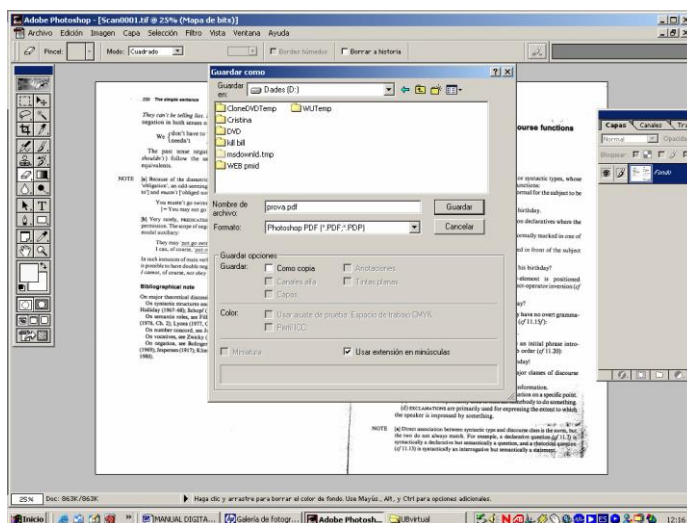
6. Una vegada fets els canvis necessaris cliquem *Aceptar*
7. Apareixerà en la *Galería de fotografías* el document com a imatge .tif que hem digitalitzat
8. Seleccionem la imatge i cliquem l'opció de *Programas* per enviar aquesta imatge a Photoshop. Seleccionar el programa *Photoshop*

¹ Depenent del programa podrà variar la terminologia o algun procés. Programa d'escàner HP scanjet 5500c

² Aquestes mateixes funcions estan també a la barra d'eines del costat

9. El programa de l'escàner envia la imatge a Photoshop i des d'allà podrem modificar la imatge (netejar-la amb la goma o eliminar coses seleccionant-les i suprimint, modificar mides i resolucions, etc.) i canviar el format a .pdf

10. Amb els canvis ja fets, si són necessaris, canviarem el format a .pdf fent el següent: *Archivo, Guardar como*, posem un nom a l'arxiu, on el guardem i a *Formato* seleccionem *Photoshop PDF, Guardar*, i l'arxiu .tif haurà passat a .pdf.



Recordeu que els .tif es quedaran guardats al disc dur de l'ordinador i haurem d'esborrar-los cada cert temps per buidar-lo. Normalment estan a *Mis imàgenes*.

Es farà un document .pdf per full. Si el professor vol un arxiu de diversos fulls, s'haurà de crear un amb el Adobe Acrobat insertant tots els documents en un de sol (*Documento, Insertar*, arxiu a inserir i a on).

- Diapositives

Seràn imatges amb bona qualitat. Una imatge per diapositiva, amb alta resolució (1200) en .tif

1. Connectar l'adaptador a l'escàner per poder escanejar les diapositives
2. Obrir el programa de l'escàner
3. Ficar les diapositives que volem escanejar, l'adaptador normalment té capacitat x 3 diapositives cada cop
4. *Herramientas* i *Explorar*
5. El programa farà una exploració de l'adaptador
6. S'obrirà la *Pantalla de Exploración*
7. A *Explorar* canviar l'opció a *Diapositivas* i automàticament tornarà a explorar les diapositives, donant com a resultat l'escaneig de les diapositives i no de tot l'adaptador.
8. Amb el requadres que marquen les diapositives podem seleccionar la que volem escanejar o totes.


CRAI Unitat de Serveis als Usuaris

9. Fer els canvis necessaris com en el cas de document de text (amb les funcions ja esmentades). Un canvi probable serà el cas de les resolucions (augmentarà el pes però també la qualitat, amb 1200 és suficient).
10. Una vegada fets els canvis necessaris cliquem *Aceptar*
11. Apareixerà en la *Galería de fotografías* el document com a imatge .tif que hem digitalitzat.
12. Seleccionem la imatge de les diapositives i cliquem l'opció de *Programas* per enviar aquesta imatge a Photoshop. Seleccionar el programa *Photoshop*
13. El programa de l'escàner envia la imatge a Photoshop i des d'allà haurem de crear diferents arxius per cadascuna de les diapositives si hem de digitalitzar més d'una, a més de poder modificar-les com a la de text (netejar, eliminar coses seleccionant-les i suprimint, modificar mides i resolucions, etc.)
14. Per guardar una a una i canviar el format (cosa poc recomanable en el cas de diapositives perquè es busca qualitat i el .tif és de les millors) fer el mateix que per .pdf: *Archivo, Guardar como*, posem un nom a l'arxiu on el guardem i a *Formato* seleccionem el que volem, si no el que hi ha per defecte .tif, *Guardar*, i fer això per cadascuna.

Recordeu que els .tif, resultants de la digitalització (anomenats normalment Scanxxx) es quedaran guardats al disc dur de l'ordinador i haurem d'esborrar-los cada cert temps per buidar-lo. Normalment estan a *Mis imágenes*.

3 – Altres recursos a la web

<http://www.library.cornell.edu/preservation/tutorial-spanish/contents.html>



Llevando la Teoría a la Práctica
Tutorial de Digitalización de Imágenes

Tabla de contenidos

Cómo utilizar este tutorial

Imprimir este tutorial

Busqueda del tutorial ➔

Visión en:

- ➔ inglés
- ➔ francés

Prefacio

1. Terminología básica
2. Selección
3. Conversión
4. Control de calidad
5. Metadatos
6. Infraestructura técnica
 - A. Cadena de digitalización
 - B. Creación de imágenes
 - C. Gestión de archivos
 - D. Entrega
7. Presentación
8. Preservación digital
9. Gestión
10. Formación continua

© 2000-2003 Biblioteca de la Universidad de Cornell / Departamento de Investigación

http://www.hugorodriguez.com/cursos/escanear_01.htm

Curso de escáner



Prohibida la reproducción sin el consentimiento escrito del autor.
Todos los artículos, cursos, fotografías y gráficos de esta página tienen Copyright y están registradas en el registro de propiedad intelectual de Barcelona.
Redactor: si te inspiras en el contenido de esta página para tu propio artículo, te agradecería que lo hagas constar y me lo comuniques.

© Hugo Rodríguez 2004

¿Todavía no te entiendes con tu escáner?. ¿Te ocurre a menudo que los negativos salen de todos los colores menos del que realmente son?. ¿Los retratos parecen de marcianos verdes?. ¿Las copias en papel no tienen extraños artefactos?. ¿Falta detalle en las sombras de tus diapositivas?. Eso es que no aprovechas correctamente las prestaciones de tu escáner. Aprende con este curso a usarlo correctamente y a exprimir al máximo tu escáner. No sólo eso, entenderás porqué sale bien de una forma y no de otras. Y si al final del curso, todavía no consigues escanear con calidad, entonces no será porque no lo hagas bien; es que...tu escáner no da más de sí!.



http://www.hugorodriguez.com/cursos/curso-idigital_01.htm

Iniciación a la imagen digital



Prohibida la reproducción sin el consentimiento escrito del autor.
Todos los artículos, cursos, fotografías y gráficos de esta página tienen Copyright y están registradas en el registro de propiedad intelectual de Barcelona.
Redactor: si te inspiras en el contenido de esta página para tu propio artículo, te agradecería que lo hagas constar y me lo comuniques.

© Hugo Rodríguez 2003

LA IMAGEN DIGITAL Y EL CODIGO BINARIO

La imagen digital está formada por un conjunto definido de puntos llamados píxeles. La imagen tradicional, sobre material sensible, también está formada por puntos o granos de plata metálica; la gran diferencia estriba en que en la fotografía tradicional, los granos son irregulares y están situados de forma aleatoria mientras que en la digital forman una matriz con filas y columnas. Cuanto mayor sea el número de filas y columnas, mayor será el detalle de la imagen digital y mayor será su tamaño de archivo.

Cada píxel de una imagen almacena la información de su tono o luminosidad, donde el tono negro es el valor 0 y el blanco el valor más alto (normalmente 255 en escala de grises), pero en formato binario. El formato binario es una notación numérica como puede ser la que utilizamos normalmente, la decimal, con la diferencia de que la binaria es la única que entienden los ordenadores.

Cuando nos referimos a la cifra "19" tan sólo necesitamos dos cifras en sistema decimal, el 1 y el 9, pero en formato binario hacen falta más. Esto es debido a la forma que tienen los ordenadores de almacenar la información que es a base de microinterruptores, que básicamente son interruptores como cualquiera de los que tenemos en las lámparas, con sólo 2 posiciones: apagado y encendido. La posición apagado se hace corresponder con un 0 y la de encendido con un 1. Así pues, en formato binario solo hay 2 cifras, el 0 y el 1 y todos los números se componen de estas dos cifras.

Normalmente las cifras binarias se componen de un número total de ceros y unos que es potencia de 2, como 8, 16 o 32. A cada uno de estos unos y ceros se les llama Bit y a un conjunto de 8 bits se le llama Byte. Si un número en formato binario tiene menos de 8 bits como es 11001, se suele completar hasta alcanzar 1 Byte como mínimo: 11001000.

