

# ANEXO I

# LA HEMEROTECA

*Internet en la Prensa Diaria*

Historia, Sociedad, Tecnología y Crecimiento de la Red

Dirección

el Periódico Mércas, 14 de junio de 1995 **SOCIEDAD**

■ COMUNICACIÓN

## Internet capta una 'infopista' catalana con 120 rutas por todo el mundo

Una empresa agrupa todos los recursos en catalán existentes en la red

**Fecha** 24/06/99 01.40

**Tema** El Comentario / El capital intelectual -- Retevisión ficha a decenas de directivos de Telefónica

La Brujula . Net. El diario de Internet  
[ + 57.000 suscriptores diarios ; versión reducida y anticipada ]  
Suscripción gratuita y bajas en <http://labrujula.net/suscribir/>  
La actualidad completa del día a día en <http://labrujula.net>

El Comentario ( N ° 256 )  
LAS OPERADORAS DESCUBREN EL CAPITAL INTELECTUAL:  
RETEVISIÓN FICHA A DECENAS DE DIRECTIVOS DE TELEFÓNICA

Por cerca de diez millones de euros  
**BT ultima la compra de Arrakis, uno de los líderes de acceso a Internet**

Listo Mi PC

## ANEXO I

### HEMEROTECA DE INTERNET

Grupos temáticos por los que se han clasificados todos los artículos.

#### 1.- INTERNET

- 1.1 Estadísticas
- 1.2 Usuarios
- 1.3 Noticias del Sector-Internet
- 1.4 Conferencias - Congresos
- 1.5 Novedades

#### 2.- OPERADORES

- 2.1 Sector Clásico Telecomunicaciones
- 2.2 Nacional
- 2.3 Internacional
- 2.4 Fusiones empresariales
- 2.5 Empresas (excepto móviles)

#### 3.-TELEFONÍA MOVIL

- 3.1 Ofertas y Precios
- 3.2 Despliegues y fechas
- 3.3 Operadores
- 3.4 Tecnologías WAP, GPRS, UMTS

#### 4.-TECNOLOGÍA

DISPOSITIVOS–HARDWARE (Novedades)  
SOFTWARE AVANZADO

- 4.1 PDAs,
- 4.2 Nuevos Dispositivos
- 4.3 Web-TV
- 4.4 Terminales Móviles

#### 5.-REGULACIÓN

ACUERDOS ALCANZADOS  
LIBERALIZACIÓN  
PRECIOS IMPUESTOS POR EL EJECUTIVO

- 5.1 CMT
- 5.2 MISCELANEA



Recopilación de los artículos de internet procedentes de periódicos generalistas y revistas del sector utilizados como material de apoyo en la realización de esta tesis  
Foto: Jordi H.Prat

**Tamaño de la Muestra:**

Si de los **2.103** Artículos de prensa diaria y revistas especializadas, Escogemos **600** Artículos Monográficos de Internet.

Después de un proceso de escaneado y de su posterior tratamiento y procesado<sup>1</sup>, para convertirlos en texto, se han encontrado un total de **40.322** palabras distintas.

Si se refina un poco más la búsqueda y se acumulan las palabras equivalentes [Internet] = [internet.] se ha obtenido un total de **25.592** palabras distintas.

Si se eliminan los artículos y las preposiciones en aras a buscar cuales son los Sustantivos que mayormente se repiten en las definiciones de internet, y a su vez cortamos la “cola” del histograma<sup>2</sup> nos quedamos finalmente con **804** palabras.

PALABRA	FRECUENCIA
DE	26.115
LA	12.935
QUE	9.603
EL	9.499
EN	9.371
A	7.787
Y	7.076
LOS	5.804
LAS	4.194
UN	3.756
POR	3.629
SE	3.546
DEL	3.119
INTERNET	3.079

A continuación se presentan las **100** primeras acompañadas de su frecuencia de aparición.

Es interesante destacar que el operador **Telefónica** se encuentra en tercera posición absoluta, cuando realizamos el análisis sobre el total de la muestra. Descendiendo a posiciones muy inferiores a medida que realizamos el mismo ejercicio únicamente sobre artículos aparecidos en un determinado año. Ello nos indica la representatividad de Telefónica en la red, que no se inicia hasta mediados de 1996.

La siguiente palabra es **Millones**. Término presente en la mayoría de descripciones de la red: millones de usuarios, millones de ordenadores...

El término **Acceso** es algo que también preocupa a los periodistas en los temas analizados. Sin duda es el gran tema a resolver, debido a la lentitud de las conexiones. La progresiva apertura a la competencia, ha hecho que se hablara mucho más de estos temas.

Le siguen **Empresas** y **Pesetas** ambas expresiones también muy utilizadas en el contexto en que se mueve la prensa. Puesto que acostumbra a destacar las grandes inversiones que las organizaciones *dicen* realizar. Se demuestra que en muchos casos únicamente es el papel que aguanta estas afirmaciones. Intentando utilizar a la prensa como medio para posicionar dichas empresas.

Este tipo de información es la que se encuentra en la Base de Datos denominada *Hemeroteca*. Que puede ser consultada de forma electrónica, para poder realizar consultas sobre los artículos publicados entre los años 1994 y 2002. A continuación se enumeran en función de su frecuencia de repetición las **100** palabras que más aparecen.

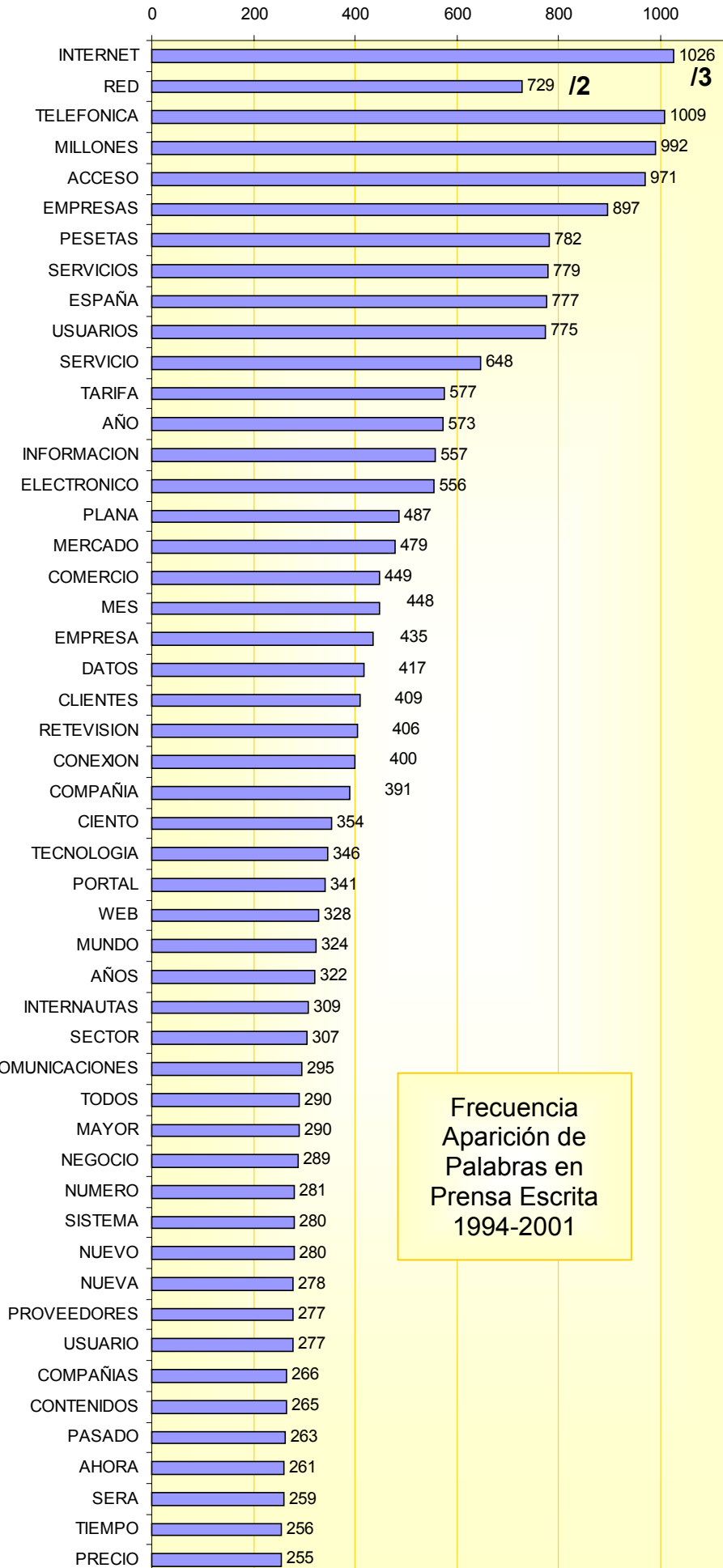
<sup>1</sup> Herramientas realizadas por Javier López Moliner, que describe en su proyecto de fin de carrera.

<sup>2</sup> Eliminando todas aquellas palabras cuya frecuencia de aparición sea menor a 20.

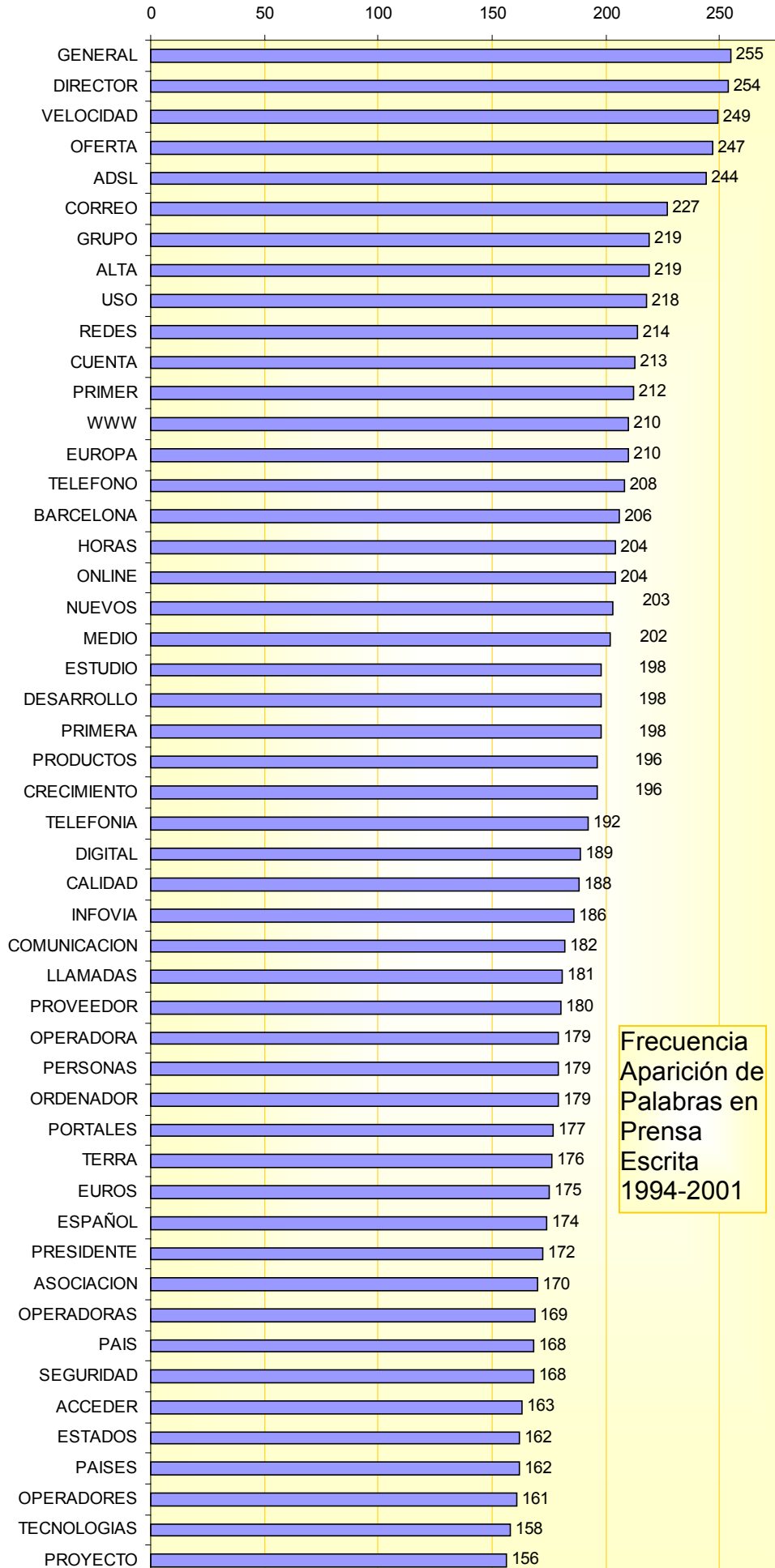
Frecuencia de Aparición de las palabras, en los 600 artículos monográficos de internet analizados.

En primer lugar aparece como es lógico INTERNET con 3.079 iteraciones, seguida de las palabras RED con 1458 y TELEFÓNICA con 1009, seguida del vocablo MILLONES (992) y de ACCESO: 971

Se han eliminado los artículos y las preposiciones.



Frecuencia Aparición de Palabras en Prensa Escrita 1994-2001

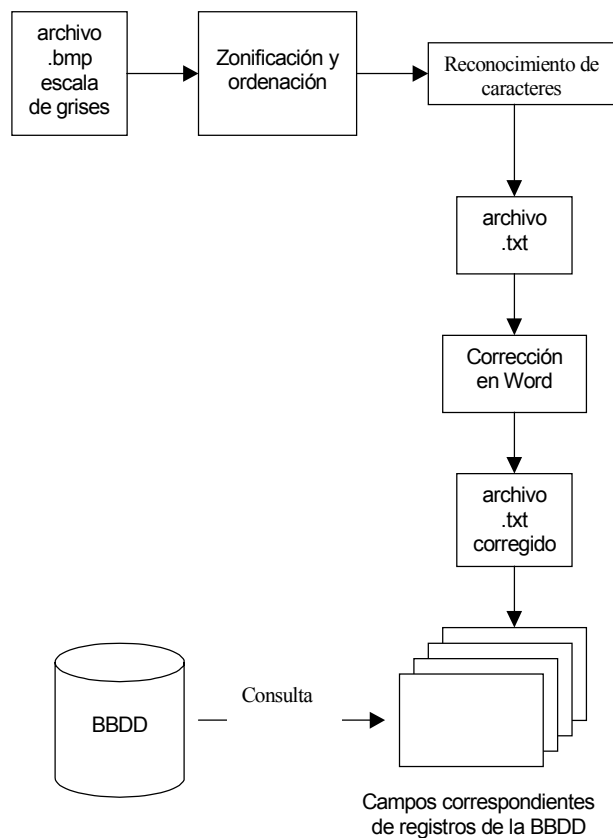


Frecuencia  
Aparición de  
Palabras en  
Prensa  
Escrita  
1994-2001

El proceso seguido para su obtención ha sido mediante el desarrollo de una aplicación de Gestión Documental<sup>3</sup>

Que nos permite obtener en formato texto, los artículos que se van escaneando, de manera que podamos clasificarlos y almacenarlos en la Base de Datos Documental.

Una vez escaneado el documento, se realiza su proceso de conversión a texto, mediante un sistema OCR.

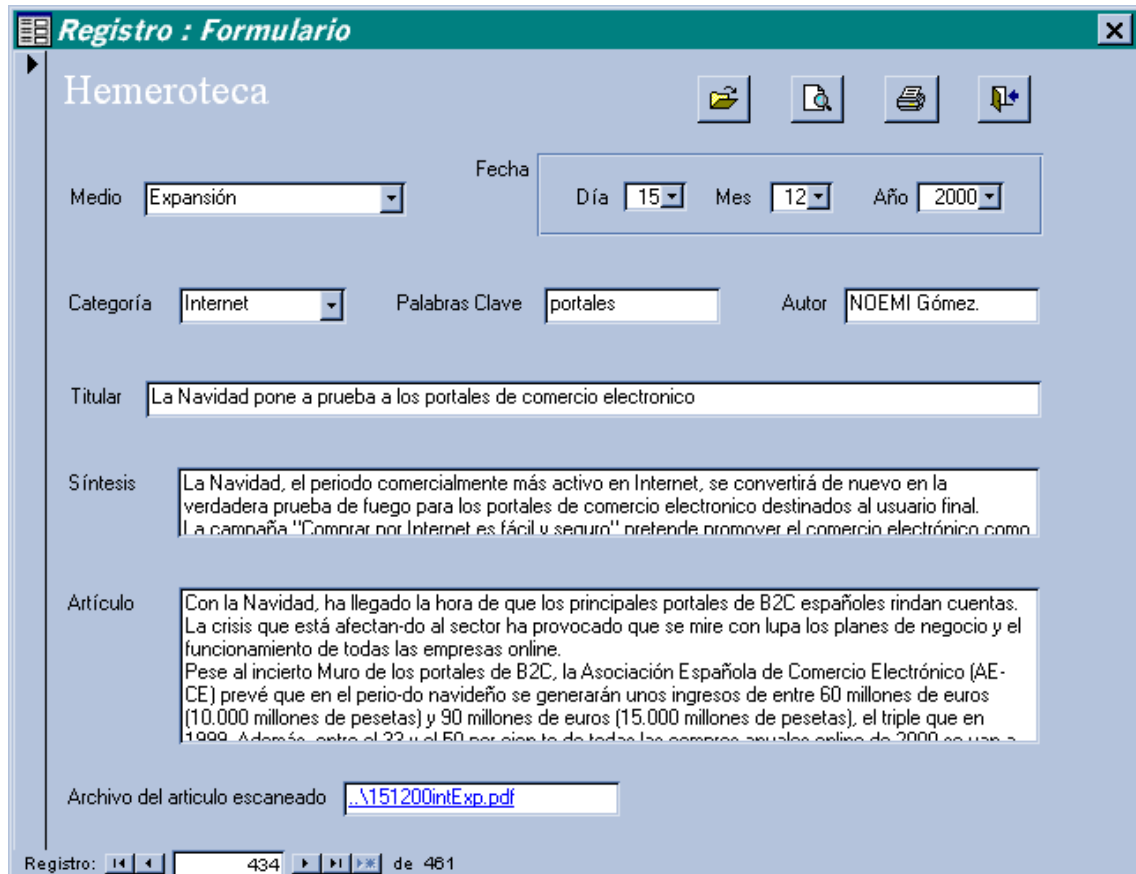


Proceso de Conversión de Imagen a Texto. Previamente se detectan las zonas del documento.

<sup>3</sup> Desarrollada en el Proyecto de Javier López Molinero.

La Base de Datos, está formada por los artículos de prensa, de los que para cada uno se cita la fuente, autor, Titular, resumen y el contenido en sí del artículo. Al incorporar la fecha de su publicación, nos permite realizar búsquedas por año, mes o día, conociendo un poco más detalladamente los hechos que sucedieron en determinada época.

La estructura es sencilla pero efectiva, veamos una muestra gráfica.



**Registro : Formulario**

Hemeroteca

Medio:  Fecha: Día  Mes  Año

Categoría:  Palabras Clave:  Autor:

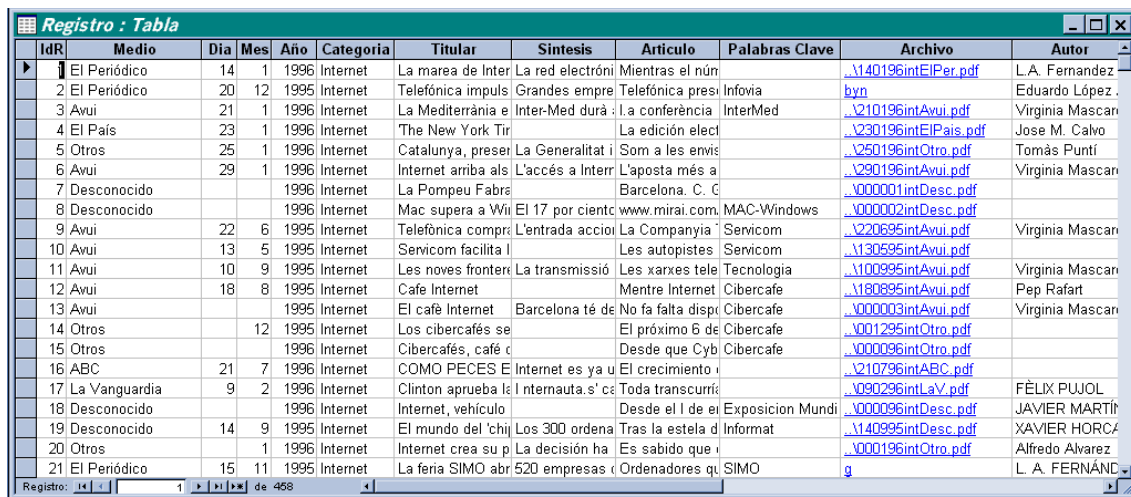
Titular:

Síntesis:

Artículo:

Archivo del artículo escaneado:

Registro: 434 de 461



**Registro : Tabla**

IdR	Medio	Día	Mes	Año	Categoría	Titular	Síntesis	Artículo	Palabras Clave	Archivo	Autor
1	El Periódico	14	1	1996	Internet	La marea de Inter	La red electróni	Mientras el núm		..\140196intEIPer.pdf	L.A. Fernandez
2	El Periódico	20	12	1995	Internet	Telefónica impuls	Grandes empre	Telefónica pres	Infovia	byn	Eduardo López .
3	Avui	21	1	1996	Internet	La Mediterrània e	Inter-Med durà	La conferència	InterMed	..\210196intAvui.pdf	Virginia Mascan
4	El País	23	1	1996	Internet	The New York Tir		La edición elect		..\230196intEIPais.pdf	Jose M. Calvo
5	Otros	25	1	1996	Internet	Catalunya, preser	La Generalitat i	Som a les emis		..\250196intOtro.pdf	Tomàs Puntí
6	Avui	29	1	1996	Internet	Internet arriba als	L'accés a Interr	L'aposta més a		..\290196intAvui.pdf	Virginia Mascan
7	Desconocido			1996	Internet	La Pompeu Fabra		Barcelona. C. G		..\000001intDesc.pdf	
8	Desconocido			1996	Internet	Mac supera a Wini	El 17 por cient	www.mirai.com,	MAC-Windows	..\000002intDesc.pdf	
9	Avui	22	6	1995	Internet	Telefónica compr	L'entrada accio	La Companyia	Servicom	..\220695intAvui.pdf	Virginia Mascan
10	Avui	13	5	1995	Internet	Servicom facilita l		Les autopistes	Servicom	..\130595intAvui.pdf	
11	Avui	10	9	1995	Internet	Les noves fronters	La transmissió	Les xarxes tele	Tecnologia	..\100995intAvui.pdf	Virginia Mascan
12	Avui	18	8	1995	Internet	Café Internet		Mentre Internet	Cibercafé	..\180895intAvui.pdf	Pep Rafart
13	Avui			1995	Internet	El café Internet	Barcelona té de	No fa falta disp	Cibercafé	..\000003intAvui.pdf	Virginia Mascan
14	Otros			12	1995	Internet	Los cibercafés se		El próximo 6 de	..\001295intOtro.pdf	
15	Otros			1996	Internet	Cibercafés, café d		Desde que Cyb	Cibercafé	..\000009intOtro.pdf	
16	ABC	21	7	1996	Internet	COMO PECES E	Internet es ya u	El crecimiento		..\210796intABC.pdf	
17	La Vanguardia	9	2	1996	Internet	Clinton aprueba la	Internauta.s' cs	Toda transcurri		..\090296intLaV.pdf	FÉLIX PUJOL
18	Desconocido			1996	Internet	Internet, vehiculo		Desde el I de e	Exposicion Mundi	..\000009intDesc.pdf	JAVIER MARTÍN
19	Desconocido	14	9	1995	Internet	El mundo del chi	Los 300 ordena	Tras la estela d	Informat	..\140995intDesc.pdf	XAVIER HORCA
20	Otros			1	1996	Internet	Internet crea su p	La decisión ha	Es sabido que	..\000196intOtro.pdf	Alfredo Alvarez
21	El Periódico	15	11	1995	Internet	La feria SIMO abr	520 empresas	Ordenadores qu	SIMO	g	L. A. FERNÁNDEZ

Registro: 1 de 458

Detalle interno de las tablas obtenidas en el estudio