
EL CD-ROM, EMMAGATZEMATGE ÒPTIC D'INFORMACIÓ

Lluís Soler i Carrascosa

La tecnologia CD-ROM pot convertir-se en un futur pròxim en un atractiu mitjà de transmissió d'informació per al món documental. Per bé que aquesta aplicació del disc òptic pot ser encara considerada com a experimental, i el seu veritable potencial encara no s'ha materialitzat, els editors estan començant a reconèixer el CD-ROM com una adequada via per distribuir grans quantitats d'informació

Desde molts anys ençà, el paper ha estat el mitjà normal per a transmetre informació. Fins i tot avui en dia, una gran quantitat de documents es produeixen a base de paper. Però cal dir que grans quantitats d'informació en aquest mitjà poden no tenir la vàlua real que tindrien si estiguessin sobre suport magnètic o òptic. Nombrosos usuaris tenen necessitat d'accés a certes informacions per prendre decisions, i desitgen la correcta informació, al moment, lloc i en format adequat i al millor preu possible. La coincidència d'alguna o totes aquestes condicions és el que determina la vàlua de la informació. Des que Gutenberg va inventar la impremta, la pàgina impresa ens ha servit suficientment, però actualment hi ha la necessitat de manipular bases de dades d'un volum considerable.

EL CD-ROM: NOVA TECNOLOGIA

Actualment, la vàlua de la informació ha estat descoberta amb les grans

bases de dades i en particular a través de la tecnologia CD-ROM (*compact disc, read only memory*). Amb aquesta nova tecnologia, l'accés a una informació és a l'abast del teclejar d'un dit. El CD-ROM és una ràpida i econòmica manera de localitzar una documentació; és portable, virtualment indestructible, relativament barat, fàcilment reproducible i, en general, el mitjà més eficaç de distribuir vastes quantitats d'informació digitalitzada. Aquesta aptitud es reflecteix en el relatiu creixement del nombre de productes CD-ROM que han sortit últimament al mercat.

Aquesta tecnologia va ser anunciada l'octubre de 1983, però val a dir que la seva primera presentació comercial fou al *Microsoft First International Conference on CD-ROM*, el març de 1985. Com a tecnologia seriosa que avui coneixem, s'estableix en el *CD-ROM Expo*, el setembre de 1988.

Com a petita descripció del CD-ROM, es pot dir que es tracta d'una tecnologia d'emmagatzematge. La capacitat d'un disc de CD-ROM permet desar milers de programes, accessibles com si es

tractés d'un simple disquet flexible. De fet, un disc òptic permet suplir l'espai equivalent a 1500 disquets de 360 Kb.

Una de les aplicacions actuals del CD-ROM està dirigida al món editorial. Exemples com la *Grolier Encyclopedia*, el 1985, han fet del CD-ROM un adequat mitjà per a l'edició electrònica. Bases de dades químiques, mèdiques, estudis biològics, astronòmics,...., han utilitzat aquesta tecnologia per obtenir un accés ràpid a tota la informació. Un disc òptic pot suplir l'espai equivalent a 250.000 pàgines de paper DIN A4, i amb avantatges considerables. Referent a pàgines de llibre, el nombre es convertiria en unes 60.000. Si considerem, com a mitjana, que un llibre pot tenir prop d'un 750 pàgines, pot concloure's que un disc òptic conté l'espai necessari per a desar-hi al voltant d'uns 80 llibres.

És també útil com a sistema per a distribuir informació. Institucions nord-americanes, com ara la CIA, el FBI o la Navy, han estat interessades a utilitzar el disc compacte per desar immenses quantitats d'informació, fer còpies sobre

CD-ROM i distribuir-les a totes les seves oficines.

El CD-ROM és, per tant, un mitjà ideal per a editar i/o distribuir informació relativament estable que no s'hagi d'actualitzar més sovint que un cop al mes. La informació que es pot emmagatzemar va des de la textual fins a la gràfica o sonora.

Quan fa aproximadament quatre anys que el CD-ROM va fer la seva aparició al món documentalista, aquella inicial excitació que es va produir ha passat a ser una postura d'espera envers aquesta nova tecnologia. El món dels documentalistes encara resta a l'espera de la inicial promesa del CD-ROM de ser una nova solució. La perspectiva de futur mostra un sistema de recerca d'informació de fàcil utilització pels usuaris, independent i a cost més reduït que la utilització de les bases de dades en línia. El disc compacte es va aproximant actualment cap a aquestes expectatives, però encara queda un llarg camí per recórrer. Dificultats per instruir els usuaris, problemes de *hardware* i *software* impedeixen, ara per ara, l'establiment definitiu d'aquesta tecnologia.

La creença d'alguns productors de CD-ROM que els seus sistemes són autoexplicatius i autosuficients és, segons la nostra experiència, errònia. És, per tant, necessari que tot producte CD-ROM tingui una documentació tècnica clarament redactada, així com una guia d'instal·lació i manteniment del sistema. No subministrar aquest manual pot incidir en l'èxit d'un producte. També és important el suport telefònic de l'usuari amb dificultats de funcionament, igual que l'existència d'instruccions de fàcil utilització i pantalles d'ajuda a l'usuari. És necessari recordar que els discs òptics han d'anar dirigits a documentalistes, i no a tècnics informàtics. Per tant, el producte CD-ROM ha de ser un sistema de fàcil utilització, amb processos d'instal·lació i operativitat adequats a gent no especialitzada. Per exemple, hi ha hagut casos en què la instal·lació d'un nou CD-ROM per part d'un usuari ha significat la impossibilitat d'utilitzar-ne un altre que feia servir anteriorment. Aquesta incompatibilitat és fàcilment resoluble per part d'un informàtic, però es converteix en un contratemps greu

per a gent no especialitzada. És per això necessari recordar que el documentalista és un usuari amb necessitat d'utilitzar diversos productes, i que, per tant, aquests no haurien de col·lidir, almenys de manera irresoluble.

La normalització del *hardware*, del *software* de recerca i l'operativitat d'aquest *software* constitueix un dels majors obstacles per a la consolidació de la tecnologia CD-ROM. Les reticències de gran nombre de documentalistes a la utilització del disc òptic es basen en la necessitat d'adquirir *hardware* i *software* especialitzat per acomodar a productes CD-ROM individualitzats. Una veritable normalització suposaria l'eliminació d'incompatibilitats entre diferents menes de lectors òptics i una major aproximació en la concepció dels programes de recerca d'informació. El CD-ROM té un enorme potencial per convertir-se en un element integral al servei de tota mena de documentalistes. Però val a dir que es fa necessari un esforç dels productors per normalitzar els seus productes, i crear un sentiment, en el món documentalista, de facilitat d'utilització i rendibilitat.

PROCÉS DE PRODUCCIÓ D'UN CD-ROM

El 1986 es va produir el primer CD-ROM al nostre país. Es tractava de l'edició electrònica en CD-ROM del *Diccionario de Medicina Marín*. Es va desenvolupar un software especialitzat per a la indexació de les dades i una interfície (*interface*) per a la recuperació de la informació.

La preparació de les dades, terme que descriu les tècniques de recollida i processament de dades per a ser enregistrades sobre CD-ROM, és una de les etapes més crítiques i que normalment ocupen més temps en el procés de producció d'un CD-ROM. La immensitat de dades a processar és subestimada dintre de tot el procés. Per fer correctament la producció del disc compacte és necessari el coneixement de l'estructura de les dades, unes bones eines per donar forma a les dades i indexar-les, el talent dels tècnics i infinita paciència. En la següent figura es presenten les

etapes necessàries per produir un CD-ROM.

Definició dels requeriments de l'usuari
Recollida de dades
Conversió de dades, disseny del programa i desenvolupament
Conversió de dades i manipulació
Indexació de dades
Indexació del sistema i avaluació
Format lògic
Pre-reproducció mestra/ reproducció mestra

Definició dels requeriments de l'usuari

Abans de començar qualsevol projecte, és necessari una anàlisi completa de la mena de producte, i de les característiques i funcions que es requereixen. Aquesta primera etapa és crucial i determina moltes de les futures decisions que es prendran. Les tasques que s'han d'efectuar en aquesta etapa inclouen anàlisi de mercat, anàlisi de competència, determinació dels canals de distribució, anàlisi de preus i magnitud del projecte, així com estimacions de temps i recursos necessaris.

El resultat d'aquesta primera etapa ha de ser un pla de com portar a la realitat el projecte. Si no s'ha aconseguit convèncer els responsables de la viabilitat o no del producte final, ara és el moment de parar el procés.

Recollida de dades

Les dades que han de ser situades sobre el CD-ROM poden estar originàriament en diferents formats. Hi poden estar com a text en pàgines impreses, fotografies, cintes de fotocomposició, fitxers de tractament de tex-

tos,... Aquesta etapa ha de preparar les dades per a ser manipulades dintre de l'ordinador, sobre base magnètica. Les activitats que poden ser necessàries en aquest moment inclouen:

1.- Entrada de text per teclat. És un procés relativament lent i penós, a causa del volum d'informació que ha de ser introduït al disc compacte. L'entrada de la informació a base magnètica mitjançant aquest mètode és cara, amb gran necessitat de mà d'obra i susceptible d'errors molt més freqüents com menys processos de detecció d'errors s'hi apliquin. Alguns productors de CD-ROM han utilitzat mà d'obra barata de països subdesenvolupats per fer aquest procés.

2.- Entrada de text per reconeixement òptic de caràcters (OCR, *optical character recognition*). Actualment s'han desenvolupat paquets de *software* que amb escàners de 200, 300 i 400 dpi (*dot per inch*: punts per polzada) permeten obtenir reconeixements de pàgines impreses de prop del 98,0-99,8 % segons quin sigui el tipus de lletra a llegir i quina la qualitat de la impressió. La velocitat d'entrada de text per aquest mètode és de 5 a 10 vegades superior, amb una fiabilitat major que l'entrada de text per teclat. El cost, per tant, de la digitalització d'informació via OCR és molt inferior a l'entrada per teclat del text, però pel fet que l'OCR pot produir defectes de reconeixement de caràcters, és necessari desenvolupar una sèrie d'eines de tractament posterior de la informació i correcció, per obtenir una rèplica sense errors de la informació original sobre base paper. Aquests paquets de *software* d'OCR s'han desenvolupat tant per a un entorn MACINTOSH com per a un d'IBM-PC. Algunes d'aquestes aplicacions són: MacinTEXT, ReadStar III Plus, TextSan, Read-It, Publish Pac per a un entorn MACINTOSH i AutoRead, ReadStar III Plus i TexRight per a un d'IBM-PC. Altres empreses han desenvolupat serveis d'OCR, per a usuaris que no desitgin adquirir el *hardware* i *software* necessaris, però que tinguin necessitat de fer servir aquesta tècnica. Alguns d'aquests serveis es poden trobar a Editorial Marín, S.A., *Servicios Auxiliares Lingüísticos*, S.A. o a *Centro de Textos Electrónicos*, S.A.

3.- Recuperació d'informació a partir de cintes magnètiques de fotocomposició o documents de tractament de textos. Es fa necessari usualment una transcodificació de la informació des del sistema on es trobaven inicialment les dades fins al sistema operatiu de treball.

4.- Procés de digitalització o escànnig de fotografies, documents,... Durant alguns anys els productors de CD-ROM han creat aplicacions prototípiques utilitzant el blanc i el negre per a les imatges. Actualment, hi ha un gran interès, per part de potencials usuaris, per crear sistemes amb text i imatge integrats.

Un primer aspecte a considerar és quina resolució cal utilitzar en la digitalització de les imatges. La resposta a aquesta pregunta és una conseqüència clara de la qualitat d'imatge que es desitgi. Com més qualitat d'imatge es vulgui obtenir, més resolució s'haurà d'aplicar, i més espai de disc serà necessari per a emmagatzemar la informació. Per exemple, una imatge digitalitzada a 200 dpi ocupa un espai aproximadament de mig megabyte. A 300 dpi, aquest espai es converteix en 1Mb, (Megabyte: un milió de bytes) i 2Mb si la resolució és de 400 dpi. Com es pot observar, el volum del fitxer generat per la imatge creix amb el quadrat de la resolució.

En resposta a la pregunta de quina resolució s'ha utilitzat, val a dir que cal considerar els perifèrics de què es disposa per a imprimir la informació. Els monitors poden mostrar imatges des de 40 dpi a 300 dpi. Les impressores ho poden fer de 100 dpi a 600 dpi. Per tant, quan s'han de fer servir altes resolucions es necessita un volum més gran per a emmagatzemar la informació i perifèrics més cars. Una més baixa resolució pot incidir de manera important en la qualitat de la imatge.

Normalment les imatges digitalitzades amb una resolució de 300 dpi solen ser satisfactòries en aquestes aplicacions.

Conversió de dades, disseny del programa i desenvolupament

Després que les dades han estat recollides i convertides a un format lleigible per l'ordinador, és probable que

■ **La creença d'alguns productors de CD-ROM que els seus sistemes són autoexplicatius i autosuficients és, segons la nostra experiència, errònia.**

aquest format sigui incompatible amb els programes d'indexació i recuperació d'informació. Aquestes dades, per tant, han de ser novament tractades i codificades per a ser llegides i indexades. Per exemple: els codis característics de les cintes de fotocomposició han de ser traduïts convenientment d'acord amb el programa de recuperació d'informació de què es disposi. És absolutament necessari tenir un coneixement complet de les dades de què es disposa abans de començar aquesta tasca. És freqüent cometre l'error de pensar que prenent una mostra de les dades i analitzant-la, aquella estructura és extrapolable al conjunt de dades total. És important assegurar-se que la mostra de dades del conjunt és representativa del total. La millor manera de no tenir complicacions en aquest aspecte és tenir totes les dades a mà, a l'hora de començar el procés.

El procés de disseny del programa d'indexació i recuperació d'informació és una part important del projecte, i pot ser, en alguns casos, el responsable de l'èxit o fracàs del producte CD-ROM. Actualment, hi ha la possibilitat de produir un producte CD-ROM a partir d'un *software* d'indexació i recuperació d'informació ja existent. Hi ha empreses a Espanya i a l'estranger que ofereixen el servei de producció del CD-ROM a partir de la base de dades del client. En són exemples: *BRS Information Technologies (BRS/Search)*, *Dataware Inc. (CD Answer)*, *Cambridge Scientific Abstracts (Compact Cambridge)*, *Reteaco Inc. (Findit)*,

Hewlett Packard/Meridian Data (HP LaserRETRIEVE), KnowledgeSet Corporation (KRS), Editorial Marín (LST), SilverPlatter Information Inc. (SilverPlatter Search and Retrieval Software).

De tota manera, hi ha una sèrie de consells importants a l'hora de desenvolupar, o escollir entre les possibilitats que ofereix el mercat, una interfície adequada per al producte CD-ROM en projecte. Dissenyar una interfície potent, efectiva i fàcil d'utilitzar és, al mateix temps, un art i una ciència. Els següents aspectes contribueixen a obtenir-ne una de bona:

- Arquitectura general (estructura de la informació).
- Disseny visual o gràfic.
- Contingut de les ajudes per a l'usuari.
- Programació.

Es poden definir sis etapes en el procés de disseny i desenvolupament:

1. Definició del problema.

És important identificar de manera clara el problema que cal solucionar i les característiques concretes d'una solució amb èxit.

2. Anàlisi de necessitats.

A qui va dirigit el producte?

De quina manera es farà ús d'aquest producte?

S'han de determinar les característiques del públic que ha de fer servir el producte i les necessitats que aquest ha de cobrir exactament.

3. Esbós de solució.

Aquest és el moment d'imaginar quina forma hauria de tenir aquesta interfície perquè produís el servei pel qual ha estat pensada.

4. Disseny.

S'ha d'idear la manera de poder plasmar a la pantalla la idea que s'ha desenvolupat. S'ha de proveir l'estructura general, però permetre, d'alguna manera, que el futur usuari pugui escollir certs formats. El control per l'usuari és un element crític de la interfície. En el disseny del programa cal considerar text, gràfics i so.

5. Desenvolupament.

És convenient programar un prototipus per sotmetre'l a totes aquelles proves necessàries per saber si el producte complirà les especificacions requerides. És important fer tests sobre aquest producte amb usuaris reals. Un cop les proves s'hagin acabat s'ha d'in-

■ Les dades emmagatzemades han de ser indexades perquè el procés de recuperació de la informació sigui ràpid

cloure allò que s'hagi après de l'usuari final en el prototipus, i acabar el producte inicial. El cicle de producció funciona de la següent manera: programar, fer diversos tests, modificar. Freqüents repeticions d'aquest cicle asseguruen l'èxit del producte final.

6. Producció.

Desenvolupament de la versió final per producció. Aquesta ha d'estar ja preparada per tests alfa i beta. S'anomena test alfa tota la sèrie de proves a què es sotmet dins de l'àrea pròpiament de programació, abans que els distribuïdors del producte no n'hi facin unes altres que es coneixen com a test beta.

Conversió de dades i Manipulació

Els processos que usualment es donen en aquesta etapa són els següents:

1. Conversió de la informació sobre base magnètica, com ara cintes de foto-composició, fitxers de tractament de textos,..., al format necessari per al sistema de recuperació d'informació.

2. Agrupament de la informació per indexar-la posteriorment.

3. Compresió d'imatges. Les imatges són emmagatzemades en format comprimit, per aprofitar al màxim l'espai del disc òptic. Depenent del contingut de la imatge, els algoritmes de compresió poden arribar a produir *ratios* de compresió des de 5 a 1 fins a casos extrems de 40 a 1. Amb aquests processos de compresió és possible emmagatzemar moltes més imatges i, per tant, aprofitar de manera millor l'espai del CD-ROM. També, la compresió

de la imatge fa que la seva lectura sigui molt més ràpida que en forma no comprimida, per bé que llavors es necessita un procés de descomposició de la imatge que pot afectar, segons el *hardware* disponible en cada cas, el temps total per mostrar la imatge a la pantalla.

4. Encriptació de dades.
5. Inserció automatitzada de formats de pantalla.
6. Correcció d'errors detectats.
7. Identificació de referències encruades.

Indexació de dades

Les dades emmagatzemades han de ser indexades perquè el procés de recuperació de la informació sigui ràpid. Les bases de dades necessiten diferents menes d'indexació per fer-les tan útils com sigui possible. Els processos que es fan en aquesta etapa són els següents:

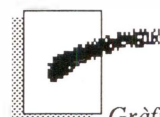
1. Inversió del text.
2. Indexació per camps.
3. Indexació d'imatge i so.
4. Resolució de referències encruades.
5. Compresió dels fitxers indexats resultants.

Indexació del sistema i avaluació.

En aquest moment, les dades ja preparades i el programa desenvolupat anteriorment han d'integrar-se per fer una simulació del sistema complet que permeti observar aspectes concrets, tant de les dades com del *software*, que caldria modificar per obtenir un funcionament millor.

Els tipus de processos que s'han d'efectuar són els següents:

1. Seguiment de l'aplicació.

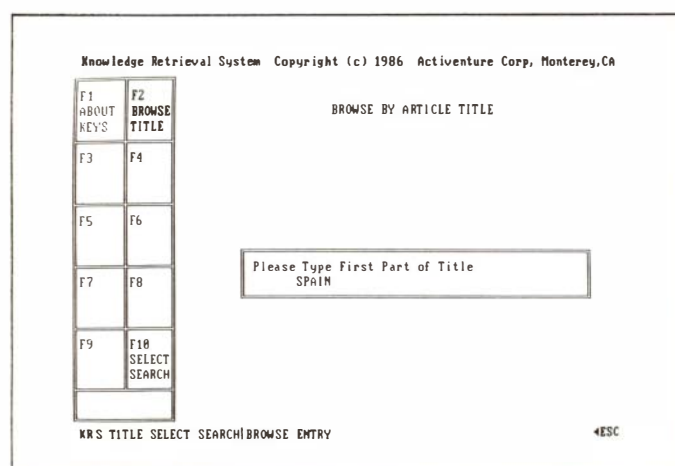


Gràfics de pantalla de l'enciclopèdia electrònica Grolier en CD-ROM (Knowledge Retrieval System) que conté 9 milions de paraules, 50.000 pàgines electròniques i 30.000 articles en un sol disc compacte.

RECERCA D'ARTICLES

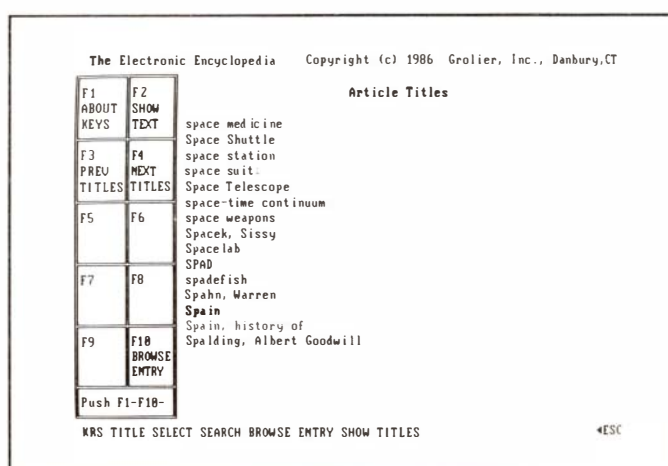
INTRODUCCIÓ DE LA VEU QUE ES VOL CERCAR

Per a la recerca d'una veu, solament s'han d'introduir els primers caràcters d'aquella.



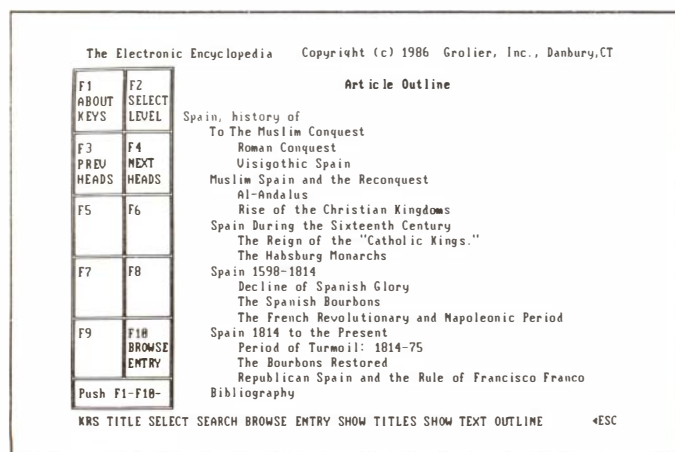
SELECCIÓ DE L'ARTICLE QUE ES VOL CONSULTAR

Les veus més semblants a aquella que es vol consultar i que l'usuari ha escrit també apareixen a la pantalla per tal de determinar la que es vol visualitzar.



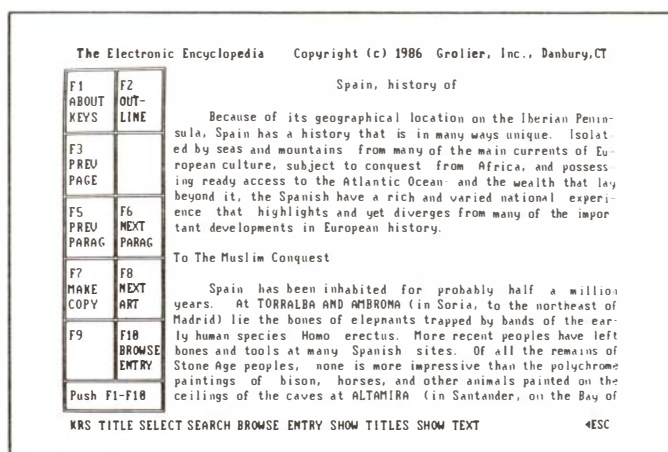
SUMARI DE L'ARTICLE CONSULTAT

La pantalla de l'ordinador ofereix l'estructura de l'article, i amb això s'aconsegueix una visió de la seva articulació amb la possibilitat d'accedir al corresponent apartat.



EL TEXT DE L'ARTICLE

La pantalla ofereix una de les pàgines electròniques de l'article. L'operador pot seleccionar l'avanç o bé el retrocés de paràgrafs des del començament de l'article fins al final.



RECERCA DE PARAULES

OBTENCIÓ DE LES PARAULES RELACIONADES

Després d'haver especificat la paraula que interessa a l'usuari, l'ordinador "explora" tot el contingut de l'enciclopèdia i indica el nombre de vegades que aquell terme es troba a l'obra, i també el nombre d'articles en els quals figuren les paraules objecte de la recerca. En l'exemple de pantalla, es demanen les paraules "Martin" i "Luther" però no "King", car es tracta d'un altre personatge.

Knowledge Retrieval System Copyright (c) 1986 Activenture Corp, Monterey, CA

F1 ABOUT KEYS	F2 SEARCH WORDS	Search Option Article Titles Bibliographies Fact Boxes Tables Article Text All Categories	TO FIND A TOPIC Enter search words at blinking cursor, then push SEARCH WORDS to find article titles. Then push SHOW TITLES to select an article.	Relation Option Negate Words In an Article In a Paragraph Words Apart 20 Exact Order
F3 LOOKUP WORDS	F4 SHOW TITLES		Push SHOW TITLES To See Titles There are 181 occurrences in 59 Articles	
F5 SEARCH OPTION	F6 RELATM OPTION	The word(s) MARTIN along with LUTHER but not KING along with along with		
F7 NEW QUERY	F8 SAVE QUERY			
F9	F10 SELECT SEARCH			

KRS TITLE SELECT SEARCH WORD SEARCH ←ESC

NOMS DE LES ENTRADES DELS ARTICLES QUE ACOMPLEIXEN LES CONDICIONS DE LA RECERCA

La pantalla ofereix els noms d'entrada dels corresponents articles en els quals es troben les paraules "Martin" i "Luther", però no la paraula "King". A partir d'aquí, es pot seleccionar l'article que es vulgui per consultar el text. En el cas de l'exemple, demanem "Henry VIII, King of England".

The Electronic Encyclopedia Copyright (c) 1986 Grolier, Inc., Danbury, CT

F1 ABOUT KEYS	F2 SHOW TEXT	Freq Article Title 1 Erfurt 2 Erikson, Erik 3 Europe, history of 1 European universities 1 Flacius Illyricus, Matthias 3 Frederick III, Elector of Saxony 1 German literature 1 Germany, history of 1 Grunewald, Matthias 1 Halle 1 Henry VIII, King of England 1 Huguenots 1 Hutten, Ulrich von 1 hymn 1 indulgences 1 James, Epistle of 1 Jimenez de Cisneros, Francisco 2 Karlstadt 1 Leo X, Pope
F3 PREV TITLES	F4 NEXT TITLES	
F5	F6	
F7	F8	
F9	F10 WORD SEARCH	

Push F1-F10
KRS TITLE SELECT SEARCH WORD SEARCH SHOW TITLES ←ESC

TEXT DE L'ARTICLE QUE ES VOL CONSULTAR

Després d'haver especificat les condicions de recerca, la pantalla ofereix la primera pàgina on apareixen les paraules objecte de la recerca que, fins i tot, es mostren amb un color especial per a destacar-les.

The Electronic Encyclopedia Copyright (c) 1986 Grolier, Inc., Danbury, CT

Henry VIII, King of England

In 1521, Henry had written a treatise against Martin Luther, for which Pope Leo X had awarded him the title "Defender of the Faith." Despite the organizational charges of his Reformation, Henry never adopted Protestant doctrines. The Latin Mass remained in use throughout his life, and theological changes were relatively minor, although both Cromwell and Thomas CRAWMER, Henry's archbishop of Canterbury, favored Protestant beliefs.

Foreign Affairs.

Several wars between England and France were fought during Henry's reign. He was personally in command of the English army that captured the towns of Therouanne and Tournai and defeated the French in the famous Battle of the Spurs (1513). In 1520, Henry met the French king FRANCIS I in a grandiose demonstration of friendship on the so-called Field of the Cloth of Gold, near Calais. Soon after, however, he joined Emperor Charles V in war (1522-27) against France.

F1 ABOUT KEYS	F2 OUT- LINE	Print Options Print Article Print Paragraph Page Breaks Page Length 66 Left Margin 10 Form Feed	Disk Options Set File Name TXT Text Form DOC Document PRN Print Form	Format Options Single Space Double Space Line Size 63 Right Justify Hyphenate Words Reduce Caps
F3 PREV PAGE	F4 NEXT PAGE			
F5 PREV PARAG	F6 NEXT PARAG	The Current Disk File is: LUTERO.DOC		
F7 MAKE COPY	F8 NEXT ART			
F9	F10 WORD SEARCH			

Push F1-F10-
KRS TITLE SELECT SEARCH WORD SEARCH SHOW TITLES SHOW TEXT ←ESC

OPCIÓ DE CÒPIES IMPRESES O BÉ EN DISC PER A UN PROCÉS POSTERIOR

En el curs de la consulta es pot imprimir l'article que interressi o bé determinats paràgrafs d'aquest mitjançant un quadre d'opcions.

Knowledge Retrieval System Copyright (c) 1986 Activenture Corp, Monterey, CA

F1 ABOUT KEYS	F2	Print Options Print Article Print Paragraph Page Breaks Page Length 66 Left Margin 10 Form Feed	Disk Options Set File Name TXT Text Form DOC Document PRN Print Form	Format Options Single Space Double Space Line Size 63 Right Justify Hyphenate Words Reduce Caps
F3 PRINT COPY	F4 DISK COPY			
F5 PRINT OPTION	F6 DISK OPTION	The Current Disk File is: LUTERO.DOC		
F7 FORMAT OPTION	F8			
F9	F10 SHOW TEXT			

Push F1-F10-
KRS TITLE SELECT SEARCH WORD SEARCH SHOW TITLES SHOW TEXT MAKE COPY ←ESC

2. Verificació de la integritat de les dades.

3. Optimització del lloc que ocupen les dades al disc.

4. Avaluació del rendiment del programa de recuperació d'informació.

5. Qualitat de les dades que conté l'aplicació.

És freqüent en aquesta etapa descobrir resultats inesperats referents tant a les dades com al *software*, que poden fer necessari refer alguna part del projecte. És important localitzar aquests problemes en aquesta etapa, abans que el CD-ROM hagi estat fabricat i enviat als usuaris.

Format Lògic

En aquesta etapa totes les formes en què es trobin les dades, text, imatge, bases de dades estructurades, so o qualsevol altra forma, convergeixen cap a una estructura unificada. Les activitats que es desenvolupen en aquesta etapa són:

1. Establiment de l'estructura lògica requerida pels fitxers. Actualment els formats més usuals són High Sierra i ISO9660 (el format High Sierra s'utilitza aproximadament en un 95 % dels discs fabricats).

2. Construcció de la imatge de disc.

3. Producció de la cinta de 9 pistes per polzada amb etiqueta ANSI.

Pre-reproducció mestra/ Reproducció mestra

Quan la cinta magnètica ha estat produïda, és enviada a una planta de reproducció mestra (*mastering*). Els tècnics processen les dades i apliquen mètodes de detecció d'errors i correcció de la informació. El sumari de les activitats del procés de pre-reproducció mestra és:

1. Verificació de la legibilitat de la cinta.

2. Càlcul del nivell 3, detecció d'errors i correcció de bits.

3. Escripció de les dades al disc magnètic.

4. Càlcul del nivell 1 i 2, detecció d'errors.

Quan els processos de detecció d'errors i correcció s'han efectuat, les dades són utilitzades per a produir el CD-

ROM. Es tracta d'un disc de policarbonat de plàstic de 12 cm de diàmetre i 1,2 mm de gruix. Per injecció en un motlle, la informació es representa digitalment en forma de "plans" i "sots" al llarg dels 5 km de pista espiral que recorre el disc de dintre cap a fora.

Aquest procés de pre-reproducció mestra/reproducció mestra el fan empreses especialitzades. Algunes d'aquestes companyies que poden fer el servei, quan ja tota la resta de processos abans esmentats s'ha portat a terme, són les següents: a Anglaterra, *Discronics Inc.*; a França, *Digipress* i *Moulage Plastique de l'Ouest*; al Japó, *Dai Nippon Printing Company Ltd.* i *Sanyo Electric Company*; als Estats Units, *Denon America*, *Discronics Inc.*, *3M Optical Recording Department*, *Philips and DuPont Optical* i *Sanyo Laser Products*, i a la República Federal d'Alemanya, *Philips and DuPont Optical* i *Sonopress GmbH*. En general tots aquests fabricants tenen condicions i preus particulars que depenen normalment del temps d'utilització del sistema i del volum d'informació processada.

Un cop el producte CD-ROM ha estat finalitzat, cal efectuar una sèrie de proves sobre la seva qualitat. Els tests a efectuar sobre el producte acabat poden ser els següents:

- Sobre la interfície, per estar segurs que treballa correctament a totes les configuracions de *hardware*.

- Sobre la consistència del so, efectes visuals i animació.

- Sobre el contingut del text.

- Sobre el funcionament conjunt del producte, per estar segurs que cap procés que pugui realitzar l'usuari no provoqui una incoherència de funcionament.

ACTITUD DELS EDITORS ENVERS EL CD-ROM

Actituds generals

A) Hi observen més obstacles que beneficis: crea la necessitat, per part de l'editor, d'investigar en una nova tecnologia, sense certesa de saber on arribarà. Al moment de produir un CD-ROM pot haver-hi greus problemes, que fan prendre una actitud molt prudent a l'editor.

B) Necessitat de crear nous productes: hi ha molt interès en la creació de nous productes, que poden ser extrets de conjunts de dades ja existents anteriorment. Però hi ha el problema de si seran o no competitius enfront de la versió impresa del producte o del servei en línia. Actualment es pot dir que els editors de CD-ROM estan a la recerca d'un *best-seller*: un producte que estimularà les vendes de lectors de CD-ROM, així com els programes Lotus 1-2-3 o dBase estimulen actualment les vendes d'ordinadors. Els *best-sellers* són productes que estableixen predisposició de consum per part dels usuaris, i que, per tant, justifiquen el desenvolupament i distribució de nous productes CD-ROM i obren el mercat a productes més especialitzats. Avui dia els *best-sellers* en CD-ROM poden ser definits com a productes que hagin arribat a vendre's en nombre superior a 1000 unitats. Donant per bona aquesta definició, es pot dir que al voltant d'uns 10 productes entren en aquest criteri. Entre ells hi ha *Books in Print*, *ERIC database*, *The Grolier Encyclopedia*, *PC-SIG* i *Microsoft Bookshelf*.

C) Necessitat de més facilitats tècniques: per a moltes aplicacions, els gràfics són una capacitat necessària. Però el tractament d'imatges és encara més complex que l'emmagatzematge de text sobre base òptica. En altres casos, per a algunes publicacions, un sol disc CD-ROM pot no ser suficient (pot ser necessari el multivolum i configuració múltiple dels fitxers).

Beneficis del CD-ROM per a l'editor

A) Possibilitats de nous mercats.

Pel fet que la gran majoria de productes CD-ROM desenvolupats han estat orientats a fins molt concrets, el mercat potencial ha estat molt limitat. Un cop els professionals dels camps de la medicina, la llei, l'enginyeria, ..., creguin en aquests productes, el volum de vendes creixerà de manera vertiginosa. No obstant això, cal dir que certs productes referents a aquestes àrees no han aconseguit despertar l'atenció de l'usuari. Molts d'aquests professionals necessiten dades actuals en els seus camps, per utilitzar-les immediatament als seus

treballs. Hi ha d'altres possibilitats com les següents:

- Dissenyar CD-ROMs per a països que no poden fer servir sistemes en línia. Alguns dels primers usuaris són Hong Kong i Austràlia. La Xina i l'Índia en serien usuaris potencials.

- Preparar bases de dades per a departaments acadèmics d'universitats.

- Distribuir bases de dades a biblioteques públiques.

- Utilitzar el disc òptic com a eina d'ensenyament per a potencials usuaris de serveis en línia.

B) Més control sobre les pròpies dades: es generen necessitats de digitalitzar dades pròpies i fer-ne ús, d'aquestes i d'altres de què es disposi, per ser sotmeses a processos de control i correcció d'errors, amb vista a ser utilitzades com a base de dades en CD-ROM o en una altra tecnologia.

C) Creació de nous productes: utilitzant la gran quantitat de dades de què disposen, els editors poden crear nous productes orientats a nous mercats i amb un valor afegit que en alguns casos pot ser considerable. Durant el 1990, els editors de CD-ROM haurien de donar preferència a nous productes amb les següents característiques:

- Èmfasi en la necessitat de dissenyar productes dirigits a usuaris professionals.

- Inclusió de gràfics (preferiblement amb colors).

- Inclusió de múltiples bases de dades en un sol disc compacte o múltiples bases de dades estructurades dintre d'un sol fitxer.

- Possibilitat d'executar els productes amb un *hardware* diferent de l'entorn IBM-PC. Fins ara l'ordinador IBM personal i compatible en general ha estat la forma més popular per a aplicacions de CD-ROM. Es donarà, durant els pròxims mesos, una forta competència amb el *hardware* d'Apple Macintosh, perquè el fet de tenir una interfície d'usuari del sistema operatiu, més atractiva comparada amb un entorn IBM-PC, resulta molt atractiu per a usuaris finals. Els CD-ROMs que es desenvoluparan hauran de funcionar també en un entorn Mac.

- Possibilitat d'utilitzar sense restriccions aquests productes CD-ROM en entorns de xarxes locals.

Les bases de dades amb orientació clínic, que serveixin d'ajuda en el diagnòstic diferencial, seran uns productes de gran importància. L'atlas del cos humà guiarà doctors i estudiants a través de l'anatomia humana amb gràfics i text integrat. Les bases de dades de referències que actualment estan en forma de documents, fins i tot sistemes en línia de serveis documentals prendran format com a productes CD-ROM.

En general, els editors de CD-ROM produiran aplicacions que ofereixin famílies de productes. Combinant una sèrie de bases de dades, hi ha la possibilitat de proveir l'usuari d'una informació més completa i valuosa, i fins i tot de més fàcil i ràpid accés. Aquesta idea és també molt atractiva per als mateixos usuaris, perquè elimina possibles inconvenients d'interferències entre productes de diferents editors.

Professionals dedicats al màrqueting, operadors de telèfon i, en general, tota aquella gent que ha de fer un ús freqüent de molts números de telèfon, trobaran, en el CD-ROM, resposta a alguns dels seus problemes. Diversos editors estan estudiant productes capços de recuperar amb rapidesa noms

■ **Combinant una sèrie de bases de dades, l'usuari pot proveir-se d'una informació més completa**

i números de telèfon. Probablement, diferents versions d'aquests productes estaran disponibles ja durant el 1990.

Una altra mena de productes correspondran a editors de diaris o revistes, que generaran productes CD-ROM amb l'acumulació de tot el material imprès durant tot l'any en una publicació. Aquests discs òptics seran rebuts per les persones que n'hagin estat subscriptors. D'aquesta manera tindran, al final del període, tot el material sobre base òptica per poder consultar qualsevol dada en qualsevol moment, i podran

substituir, amb el disc compacte, la versió impresa de la informació.

Obstacles del CD-ROM per als editors

A) Erosió de productes existents: la substitució de productes impresos per productes d'edició electrònica en CD-ROM pot ser un greu inconvenient per a alguns editors. Per a molts d'ells, les versions de les obres impreses són la major part de la font d'ingressos de l'empresa. Petites disminucions en percentatges de vendes poden significar pèrdues importants.

B) Pirateria de les dades: un segon inconvenient és el referent a la poc suggestiva idea per a l'editor de tenir 20 anys de treball i dades en un sol disc, amb la possibilitat que qualsevol usuari o una altra empresa pugui fer ús d'aquesta informació per a diferents objectius.

C) Temps il·limitat d'ús: alguns editors no estan d'acord amb el fet que la base de dades generada pugui ser utilitzada de manera il·limitada per part de l'usuari. Són, en canvi, partidaris que es paguin quantitats variables en funció de l'ús que se'n faci, tal com passa amb els serveis en línia.

La progressió del CD-ROM els últims anys es pot resumir de la següent manera: el 1987, els productes CD-ROM eren utilitzats en demostracions i conferències, però no eren normalment comercialitzats. El 1988, ha guanyat molts adeptes dintre dels documentalistes. El 1990, els productes CD-ROM han de ser utilitzats per professionals de diferents sectors.

Un primer pre-requisit, però, és la baixada de preus dels lectors de CD-ROM. Els preus dels lectors haurien de baixar fins a valors a la ratlla de les 60 a 70.000 pta, en lloc de les 200.000 actuals. Un altre aspecte important és la necessitat de poder treballar en xarxa, la qual cosa permetrà l'augment notable de les vendes.

ACTITUD DELS USUARIS ENVERS EL CD-ROM

Actituds generals

A) Poca experiència: molts usuaris

finals tenen poca experiència o pràcticament no en tenen cap, en la recerca documental. Les necessitats que solen tenir són relativament poc freqüents. Aquest fet posa a prova les capacitats de la interfície dissenyada per al producte CD-ROM.

B) Poc coneixement de la base de dades: les peculiars propietats d'una base de dades no són fàcils d'explicar a un usuari final.

Un programa d'ensinistrament en l'ús de la base de dades, per a cada cas en forma específica, és necessari per donar a l'usuari final el màxim benefici possible en el temps d'utilització del disc compacte.

Avantatges del CD-ROM per a l'usuari

A) Independència: tenint un accés personalitzat a la base de dades i essent possible redefinir o perfeccionar certes recerques, l'usuari té un sentit de llibertat en la utilització del sistema. D'aquesta manera, les seves recerques no depenen de l'actuació d'un intermediari que hagi d'entendre les seves necessitats.

B) Resultats al moment: un cop s'han trobat les referències, hi ha la possibilitat d'utilitzar els resultats al moment per efectuar noves recerques o donar definitivament per bones les que ja s'han fet.

C) La informació és lliure.

D) Major interacció: la possibilitat de visualitzar informació en diferents formats, així com la possibilitat de disposar de diccionaris, dona una estimulació intel·lectual que no és possible en sistemes de servei en línia.

Obstacles en la utilització del CD-ROM

Alguns d'aquests obstacles poden suprimir-se o en molts casos disminuir-se amb la utilització d'eines d'ensinistrament, i de programes d'ensenyament amb ajudes que permetin a l'usuari resoldre dificultats com ara les següents:

- Necessitat d'aprendre sobre la seva pròpia base de dades.

- Necessitat que l'usuari entengui que la recerca que aporta resultats per-

fectes és infreqüent.

- Mètodes de recerca inadequats que poden requerir més temps d'utilització dels terminals.

PROS I CONTRES D'AQUESTA TECNOLOGIA

Als usuaris d'informació, en realitat, no els importa de manera especial ni com ni on està desada la informació, mentre un sistema no demostrï grans avantatges enfront d'un altre de més

■ Un equip de CD-ROM té la possibilitat de combinar text i imatges

antic, si és que el sistema antic compleix els requeriments i funcions que se li demanen. Evidentment, els avantatges del CD-ROM són molts: capacitat d'emmagatzematge d'alta densitat, llegibilitat, interactivitat, baix cost de reproducció per a grans volums d'informació, economia de distribució, distribució local d'arxius d'informació, possibilitat de gran capacitat d'emmagatzematge per a ordinadors potents, *hardware* relativament barat, virtual indestructibilitat del suport òptic, integritat de les dades, manteniment de la base de dades fàcil (per als editors de bases de dades, l'actualització pot ser un procés lent i costós). Un equip de CD-ROM també té avantatges pel fet que, de manera fàcil, pot traslladar-se l'equip d'un lloc a un altre i pot ser utilitzat per diferents usuaris. Té la possibilitat de combinar text i imatges. Ofereix capacitat pràcticament il·limitada de recerques per l'usuari, amb possibilitat d'utilitzar potents eines de recuperació d'informació. Permet la recerca de termes que han de ser indexats i índexs per ser estructurats i combinats de maneres diferents per aconseguir una recuperació de la informació molt més ràpida. Els costos d'utilització per part de l'usuari

són fixos, i això tendeix a animar i fomentar aquesta utilització. Finalment, un aspecte que constitueix un important avantatge és el fet que aquests sistemes de CD-ROM tendeixen a ser cada cop més simples i de més fàcil utilització.

Entre els desavantatges de la tecnologia CD-ROM poden esmentar-se: Inconvenients de manca d'actualització suficient de les bases de dades, que normalment és menys freqüent que en bases de dades de serveis en línia; costos de preparació de dades relativament alts; l'accés a les dades pot ser en alguns casos relativament lent i ineficient comparat amb el temps d'accés de les dades de base magnètica o en sistemes en línia.

Cada indústria té la seva particular infraestructura, òrgans de comunicació, canals de distribució,... El mercat del CD-ROM ha mostrat un creixement, principalment en tres sectors: empreses, educació i usuaris particulars. El mercat de les empreses ha suportat el major nombre d'aplicacions i desenvolupaments (normalment en mercats verticals). Ara la tecnologia CD-ROM passa un període de tendència a la baixa, tant del hardware com del software necessaris. Els mercats de l'educació i dels usuaris particulars encara han de créixer de manera espectacular. ■

LLuís Soler i Carrascosa
és Enginyer químic

BIBLIOGRAFIA

OPTICA'87. The International Meeting for Optical Publishing and Storage (Amsterdam 14-16 abril 1987). Learned Information Ltd., Oxford, 1987.

The International Directory of Information Products on CD-ROM (1987). Mark Rasdall. Alan Armstrong & Associates Ltd., Londres, 1987.

MICROSOFT CD-ROM 1989-1990 Yearbook. Recollit per Salley Oberlin i Joyce Cox. Microsoft Press, Washington, 1989.

Designing Knowledge-Based Systems. T.R. Addis. Kogan Page Ltd., Londres, 1987.

Online Information 88. 12th International Online Information Meeting (Londres 6-8 desembre 1988). Learned Information Ltd., Oxford, 1988.