

Gestió documental amb microordinadors: característiques, estructura i tecnologia dels sistemes de gestió documental

LLUÍS CODINA BONILLA*
ERNEST ABADAL FALGUERAS**

Resum: Els sistemes de gestió documental (SGD) són programes que permeten la creació i explotació de bases de dades documentals. Estan orientats a la gestió de documentació científica i tècnica (informació de tipus textual poc estructurada), en oposició als sistemes de gestió de bases de dades (SGBD) dirigits a la gestió de dades empresarials (informació de tipus factual molt estructurada). Es presenta una tipologia dels SGD, així com les característiques, tecnologia i estructura bàsiques d'aquests. Finalment, es comenten les noves tendències en tecnologia de sistemes documentals. S'adjunta també una relació dels principals SGD que poden trobar-se al mercat espanyol.

1. Les bases de dades i la recuperació de la informació

En una època com l'actual, caracteritzada per un creixement quasi exponencial de la informació, els problemes de recuperació són els més citats pels gestors d'informació. Els mètodes tradicionals (creació d'inventaris, índexs o catàlegs manuals) no poden afrontar tots sols la gestió i control de grans volums de dades i d'informació. Es tracta, com és ben sabut,

* Universitat Pompeu Fabra.

** Escola Universitària "Jordi Rubió i Balaguer" de Biblioteconomia i Documentació.

d'instruments de costosa elaboració que proporcionen una recuperació incompleta, poc versàtil i de resposta lenta.

La microinformàtica està participant molt activament en la resolució d'aquest tipus de problemes i disposa de tot un vessant aplicat al camp del tractament de la informació i dels documents. Els microordinadors han anat augmentant les prestacions (potència, velocitat, capacitat d'emmagatzematge, etc.) de forma paral·lela a la reducció del seu preu. Les actuals característiques converteixen els microordinadors en uns valuosíssims instruments per processar la informació a nivells professionals.¹ Per altra banda, també s'han desenvolupat programes i aplicacions, orientats a facilitar la gestió dels documents o de les dades, així com per automatitzar diversos processos tècnics de les disciplines encarregades del tractament de la informació (arxivística, biblioteconomia i documentació).

D'altra banda, els sistemes de gestió de bases de dades (SGBD) són, actualment, la forma més eficient d'organitzar i estructurar grans volums d'informació, ja que en permeten la gestió d'una forma completament automatitzada. Un SGBD és un conjunt de programes que permeten la creació, actualització i accés de base de dades.² Sobre aquest tipus de programes versarà l'eix d'aquest article, el qual en mostrarà les principals característiques, estructura i tecnologia.

2. Els programes de recuperació de la informació

No tots els SGBD són iguals, ja que dessota aquesta denominació genèrica queda inclosa una diversitat notable de programes amb estructura, nivell i aplicacions diferents. Un dels criteris de classificació més importants és el que fa referència a l'aplicació a la qual s'orienta el programa, al tipus de problema que intenta resoldre.

Així doncs, segons aquest criteri, existirien dues grans famílies de programes:

- Sistemes d'orientació general
- Sistemes orientats a la gestió documental

1. Les principals tendències i línies d'evolució en l'àmbit de la microinformàtica ja van ser comentades en un article anterior. (vegeu Codina, 1992).

2. El terme *base de dades* s'utilitza sovint amb dos significats diferents sense especificar, en la majoria dels casos, a quin d'ells es fa referència. D'aquesta manera, es posaria la mateixa etiqueta de base de dades tant a *Medline*, *Índice Médico Español* o *Iberlex* (conjunt de dades organitzades lògicament i gravades en un suport legible per màquina), com a *Basis*, *Knosys*, o *CDS/ISIS* (programes per crear i explotar bases de dades) quan, en estricta aplicació, tan sols es podria utilitzar amb propietat aquest nom per fer referència al primer grup d'entitats que, algunes vegades, es posen al servei del públic a través de les xarxes de telecomunicacions o distribuïdes en suport òptic. D'altra banda, tant *Basis*, com *Knosys* o *CDS/ISIS* serien exemples de *sistemes de gestió de bases de dades*.

2.1. La informació no estructurada i els Sistemes d'orientació general

Els sistemes d'orientació general són una subclasse dels SGBD, però, per raons històriques, tenen el mateix nom que la classe general. Així, doncs, existeix un malentès, també històric, sobre aquests tipus de programes, ja que hom tendeix a creure, fins i tot en ambients professionals, que només hi ha un representant d'aquesta espècie de programes. Com veurem, això no correspon a la realitat.

En tot cas, els SGBD d'orientació general, són coneguts com SGBD *tout court*. Aquests sistemes permeten la gestió de dades i informacions de molt diversa naturalesa, encara que estan especialment ben adaptats a les necessitats de la gestió empresarial i administrativa en general. En aquest sentit, actualment la majoria de les empreses disposen d'un o més SGBD com a nucli dels seus processos de gestió.

El concepte d'*informació-no-estructurada* és útil per caracteritzar el tipus de documents que sol tractar la Documentació, a diferència del tipus de documents o del tipus d'informació que solen tractar els sistemes comptables, administratius, etc.

L'oposició informació estructurada *versus* informació no estructurada significa que alguns documents o dades són de naturalesa i extensió previsible i estadísticament regular. En canvi, la informació no estructurada és aquella intrínsecament imprevisible pel que fa a l'extensió, naturalesa i format. Per exemple, el format d'una lletra de canvi, d'un albarà o d'una factura sempre serà el mateix. Tanmateix, no hi ha dos articles de revista ni dos llibres que tinguin la mateixa estructura. D'altra banda, aquest concepte fa referència també al fet que els documents estructurats resulten altament previsibles i podem anticipar quants n'hi haurà d'una sèrie al cap d'un any, quines dades contenen, quants i quan es necessitaran, qui els necessitarà, de quina forma ens els demanaran, etc. És impossible, però, determinar la major part d'aquestes mateixes coses relatives al fons típic d'una biblioteca universitària, o del centre de documentació d'un diari, per no parlar del fons d'una fonoteca o d'una iconoteca, etc.

Així doncs, els SGBD estan pensats per al tractament d'informació molt estructurada (p.e. dades comptables) i de tipus textual (lletres i xifres). També acostumen a disposar de facilitats per efectuar càlculs i tractaments estadístics bàsics amb els valors numèrics que gestionen. No obstant això, són poc aptes per a la gestió d'informacions textuales llargues i no estructurades (p.e. articles de revista).

Com que foren els primers sistemes que es van desenvolupar i els que tenen un major nombre d'instal·lacions a tot el món (amb molta diferència), se'ls sol anomenar *sistemes de gestió de bases de dades* (SGBD) o *data base management systems (DBMS)*, sense més ni més, com si no hi hagués cap altre tipus de SGBD.

Per altra part, a causa de l'èxit del model relacional per estructurar les dades (comparat amb els models clàssics jeràrquics o en xarxa), quasi tots els SGBD per a microordinadors són relacionals, així com bona part dels SGBD per a minis i grans ordinadors.

2.2. Sistemes de gestió documental (SGD)

Els sistemes de gestió documental (SGD) estan concebuts per a la gestió de documents no estructurats, com ara documents científics (articles, ponències, tesis, etc.), tècnics (informes, patents, etc.) o culturals (articles de diaris, fotografies, etc.), en lloc de simples dades o informació estructurada.

Els SGD permeten, per tant, la gestió de fons documentals científics/tècnics/culturals de qualsevol naturalesa, i això inclou la gestió de qualsevol col·lecció de textos, imatges i fonogrames.

Per tot això, és lògic que les principals aplicacions dels SGD es trobin en la gestió de fons documentals (de qualsevol institució, organisme o empresa), en l'organització de bibliografies, en la informatització de catàlegs de biblioteques, així com en la gestió de fitxers informàtics produïts per altres programes que continguin text. Aquestes aplicacions són, estadísticament, marginals i no és estrany que molts professionals de la informàtica en desconeguin l'existència. D'aquesta "marginalitat" provenen els tradicionals malentesos entre documentalistes i informàtics. Sempre que el departament de documentació o la biblioteca d'una empresa han de negociar la informatització del seu fons documental o de qualsevol altre dels seus processos tècnics, el departament d'informàtica intentarà imposar un SGBD, i tendirà a ignorar l'especificitat dels SGD.

Tot i que els sistemes documentals tradicionals estan orientats a tractar informació exclusivament de tipus textual (referencial o font), també existeixen programes que poden gestionar i incorporar informació multimèdia (text, gràfics, imatges i so) a la seva base de dades. *Gestionar*, en aquests casos, significa:

- *Descriure* la informació no textual
- *Recuperar* i mostrar la informació no textual

En aquest tipus de programes, el principi de funcionament sol ser el següent: la descripció és de tipus textual, i va associada a un o més documents no textuais (imatges, per exemple). Els documents no textuais estan digitalitzats en un suport informàtic, la qual cosa permet que el programa els pugui mostrar en pantalla com a resposta a una recerca d'informació.

Això significa que, en aquests programes, allò que es recupera directament és un registre textual i, indirectament, la imatge o el so associat a

aquest registre. D'aquí prové l'expressió anglosaxona *text retrieval systems*, que posa l'èmfasi en el fet que es tracta de recuperació d'informació textual, encara que siguin sistemes per a la gestió de documents no textuais.

Per tant, d'acord amb l'estat actual de la tecnologia microinformàtica, la gestió d'un fons documental mitjançant microordinadors actualment es pot fer, segons el seu contingut, de dues maneres diferents:

- a) Amb registres textuais que contenen una representació abreujada (document secundari) del document original (document primari). En aquest cas, la base de dades només gestiona referències, que remeten a la consulta del document original. Aquest document original pot ser textual o no textual. En ambdós casos, es pot considerar que es tracta de bases de dades referencials.
- b) Amb registres que contenen tot el document, a més a més d'una referència d'aquest, composta per una sèrie de camps descriptius i d'altres de contingut (com p.e. el resum o els descriptors). Els documents poden ser textuais o no textuais. En el primer cas, els registres consisteixen en tot el document sencer, no en una simple representació, i es parla de bases de dades de text complet. En el segon cas, els registres poden consistir en imatges, per exemple; imatges que es troben contingudes a, i formen part de, la pròpia base de dades. En ambdós casos es pot considerar que es tracta de bases de dades font.

En aquest punt, cal assenyalar que hi ha una certa ambigüitat, difícil d'evitar, quan es diu que una base de dades "és una base d'imatges". Pot tractar-se d'una base de dades textual sobre imatges, o bé d'una base de dades amb imatges. L'ambigüitat podria eliminar-se si es fes servir sempre una terminologia rigorosa, o si sempre es diferenciés clarament entre el domini o la temàtica d'una base de dades (imatges, per exemple) i el tipus d'informació que contenen els seus registres.

En tot cas, des d'un punt de vista tecnològic, la majoria dels SGD de la generació actual es caracteritzen pels elements bàsics següents :

- *Disposen d'una tècnica d'indexació basada en l'ús del fitxer invertit.*
El fitxer invertit s'anomena així per oposició al fitxer directe, que és el conjunt de tots els registres introduïts a la base de dades, i que estan disposats de forma seqüencial (seguint l'ordre cronològic en el qual van ser donats d'alta). Per altra part, el fitxer invertit o fitxer invers, és el fitxer de cerca i consta de la relació alfabètica de tots els termes continguts en camps indexables dels registres. Es podria comparar a l'índex analític d'una publicació.

- *Utilitzen l'àlgebra booleana per efectuar les operacions de recuperació.*
L'àlgebra de Boole constitueix la base del llenguatge d'interrogació de la majoria dels SGD per construir operacions de recerca de certa complexitat (vegeu.3.3.2.).
- *Utilitzen registres d'extensió indefinida basats en el model d'una fitxa bibliogràfica, en comptes d'una estructura tabular (pròpia dels sistemes de gestió de bases de dades relacionals).*
Els registres basats en el model textual o en el model de la fitxa bibliogràfica tradicional s'oposen als registres basats en el model tabular en el sentit que, els primers, segueixen intuïtivament l'estructura d'una típica fitxa bibliogràfica (vegeu figura 2), mentre que el model tabular és una estructura matemàtica (vegeu figura 1) que s'utilitza en els sistemes relacionals, en la qual cada fila és un registre i cada columna és un camp. Aquesta estructura matemàtica de dades no admet grups de repetició (per exemple, dos autors, diverses paraules clau, etc.), és a dir, que les files només poden tenir valors atòmics i, a més a més, exigeix una extensió fixa de les columnes.

L'estructura relacional de la taula fa que el model relacional estàndard no s'adapti bé a les necessitats típiques de la gestió documental. Com podem observar a la figura 1, quan hi ha dos autors en un mateix llibre, s'han de crear dues files (o dos registres) per tal de recollir aquest fet, degut a la imposició del model tabular relacional. En canvi, el model de fitxa, no presenta cap restricció en aquest sentit.

Figura 1:

Registres bibliogràfics representats en una estructura tabular. Cada fila és un registre, i cada camp és una columna.

Núm. registre	Autor	Títol	Data
10	SALTON	Recuperació d'informació	1986
10	MCGILL	Recuperació d'informació	1986

Figura 2:

Un registre bibliogràfic representat en una fitxa. Cada camp és una zona d'informació que pot contenir grups de repetició (autors, per exemple).

NÚM. REG.	: 10
AUTOR	: SALTON; MCGILL
TÍTOL	: Recuperació d'informació
DATA	: 1986

En bona part per la pròpia naturalesa del model relacional –com ja hem vist– i, en part també, per raons històriques, els sistemes relacionals es van adaptar exclusivament a les aplicacions de gestió de dades de les empreses (i no de gestió de documents) i, en conseqüència, fins ara, la qualificació *relacional* equivalia a dir “no vàlid per a aplicacions documentals”, perquè els programes convencionals que incorporen un sistema relacional ni suporten les cerques de tipus documental ni aporten la més mínima eina típica de la gestió documental.

Recentment, en canvi, han sortit sistemes capaços de combinar les dues coses: prestacions documentals i model relacional (vegeu 4.) en el si de programes mixtos que combinen dos motors de recuperació d'informació complementaris entre ells: un motor relacional i un motor documental.

A banda d'això, els sistemes de gestió de bases de dades (SGBD) tenen uns trets diferents dels sistemes de gestió documental (SGD) també per raó de la seva orientació cap a la solució de problemes diferents.

Quant a la seves aplicacions i característiques funcionals, la taula 1 recull les diferències principals entre els SGD i els SGBD:

Taula 1. Principals diferències entre un SGD i un SGBD	
SGD (gestió documental)	SGBD (orientació general)
Orientats a la gestió de documents: llibres, articles de revista, fotografies, etc.	Orientats al tractament de dades empresarials (administratives i de gestió): comptabilitat, nòmines, etc.
Preparats per a informació poc o gens estructurada: textos d'extensió no previsible, amb format o sense.	Preparats per a informació molt estructurada: textos breus i xifres, ambdós amb un format predeterminat.
Proporcionen respostes a usuaris personals: intenten seleccionar documents pertinents a una necessitat d'informació.	Proporcionen fluxos de dades a altres programes informàtics (p.e. programa de confecció de nòmines, llistats comptables, etc.)
Es troben presents a determinats sectors econòmics, en els quals cal gestionar fons documentals científics o tècnics com a part de les seves activitats (p.e. empreses periodístiques, hospitals, laboratoris farmacèutics, etc.). Necessaris en serveis públics de lectura i investigació: biblioteques i centres de documentació de tota mena.	Es poden trobar a qualsevol tipus d'empresa, pública o privada, i de qualsevol sector econòmic. És una de les aplicacions bàsiques de l'ofimàtica, juntament amb el processador de textos i el full de càlcul.

Amb el temps, el model clàssic de SGD, tal com s'acaba de sintetitzar, s'ha anat diversificant i ha provocat l'aparició de productes molt específics i adequats per a determinades aplicacions (p.e. confecció de bibliografies, gestió de documents multimèdia, etc.) i no tant per a d'altres. A més a més, han sorgit diverses estratègies en la gestió del coneixement i de la informació, com els hipertextos o els sistemes d'indexació de fitxers. A continuació es presenta una breu descripció de cada una d'aquestes orientacions documentals particulars.

2.2.1. Sistemes clàssics

Amb la denominació de *sistemes clàssics* es coneixen aquells sistemes (vegeu Sieverts, 1991) que agruparien els SGD més tradicionals.³ Són els que s'ajustarien més a la descripció genèrica abans apuntada, i disposen de totes les característiques estructurals i tecnològiques indicades al punt 2.2. Presenten una notable semblança funcional amb els programes que utilitzen els grans distribuïdors de bases de dades ASCII per gestionar les seves bases de dades documentals.

Els sistemes clàssics es caracteritzen pels següents trets diferencials:

- Organitzen la informació en registres d'estructura homogènia, de forma que cada document queda representat mitjançant fitxes o registres amb els mateixos camps. Aquests camps posseeixen uns atributs que marquen la seva forma d'indexació. Alguns programes només admeten un tipus de registre, i només deixen tenir una base de dades oberta cada vegada, encara que es poden crear tantes bases de dades com es necessiti. Els sistemes de nivell més alt, en canvi, permeten tenir diverses bases de dades obertes alhora.
- Proporcionen als documentalistes un conjunt d'eines que permeten controlar la indexació dels documents, per exemple, editant el fitxer de paraules buides i el fitxer invertit. Els sistemes de més alt nivell permeten la inclusió d'un thesaurus en línia que pot assistir el documentalista en el moment de la indexació de documents, i també l'usuari final en el moment de la recuperació d'informació.
- Incorporen funcions de creació de bases de dades, altes i modificacions integrades en el propi programa.

De fet, són els sistemes amb més instal·lacions arreu del món, i són els programes més típicament documentals, en el sentit que permeten la gestió automatitzada de qualsevol fons documental mitjançant fitxes referencials o

3. Les denominacions *clàssic* i *tradicional* no comporten cap judici de valor. Només significa que, estadísticament, suposen la majoria dels sistemes instal·lats i que es tracta de la tecnologia més madura.

amb text complet. El sistema d'alt nivell proporciona també eines a l'administrador de la base, tals com establiment de diversos nivells de seguretat, estadístiques d'ús, controls d'accés, etc.

A diferència d'altres sistemes, donen per suposat que la càrrega de dades es farà, sobretot, des del teclat o important fitxers amb un format uniforme i predeterminat. Fins ara, la majoria donava també per suposat que les fitxes serien de tipus referencial. Darrerament, possibiliten el treball amb text complet i alguns programes admeten registres textuais amb desenes o centenars de pàgines.

2.2.2. Sistemes de gestió de bibliografies

La principal característica dels sistemes de gestió de bibliografies rau en el fet que no tan sols gestionen documents, sinó que, a partir d'aquests, són capaços de generar bibliografies segons diversos estils (p.e. MLA, ISO, etc.).

Acostumen a ser programes d'utilització senzilla i adreçats al sector acadèmic i científic. Estan pensats per a un usuari personal que s'ocuparà de realitzar totes les funcions de la base de dades (des de la creació fins a les exportacions o cerques). El seu ús, per tant, ha de ser molt simple perquè la formació d'aquesta classe d'usuaris no és comparable a la dels professionals de la informació i documentació. (Per aquest motiu, p.e., les possibilitats de cerca no són tan sofisticades com en els sistemes clàssics.)

Els registres de la base de dades són aprofitats moltes vegades per complementar treballs científics (ja siguin articles de revista, o informes de tota mena). És per aquest motiu que un dels aspectes més elaborats el constitueix les sortides d'informació: es poden editar llistes bibliogràfiques segons els estils marcats per diverses publicacions periòdiques.

Altres facilitats que acostumen a permetre són:

- barrejar diversos models de registre (articles, llibres, notes de lectura, etc.) en una mateixa base de dades
- importar registres de grans bases de dades (distribuïdes en línia o en CD-ROM) i transferir-los directament a les bases de dades pròpies creades amb el programa.

2.2.3. Sistemes d'indexació

El seu tret diferencial consisteix en la generació d'índexs a partir de fitxers de text preexistents creats mitjançant algun processador de textos. Per què cal incorporar un editor de textos en una base de dades si en el mercat ja hi ha processadors boníssims? ¿No és millor aprofitar-los i fer que el programa de base de dades s'ocupi tan sols del procés d'indexació? Aquesta seria una mica la filosofia comercial latent en els sistemes d'indexació.

Aquest tipus de programes, per tant, no disposa de mòdul d'alimentació de la base de dades (altes, baixes i modificacions de registres) i fonamenten

l'actualització de la base de dades en la importació de fitxers externs, provinents generalment d'un processador de textos. Això implica, per altra banda, que per a qualsevol alteració (modificació o supressió) en el fitxer principal s'ha de recórrer a un programa extern (el processador de textos).

Disposen de filtres que els permeten interpretar el format de dades dels programes més populars per a microordinadors, i per tant, poden crear índex de tots els fitxers de dades que hi hagi en un disc dur en una xarxa d'ordinadors. Quan es recupera el document, es recupera en el format nadiu amb el qual va ser creat. Sovint fan servir el sistema de fitxers invertit, però no necessàriament.

2.2.4. Hipertextos

Els sistemes de gestió d'hipertextos (SGH) són programes que permeten la creació i explotació d'hipertextos. Aquests hipertextos (o hiperdocuments) es componen de grups de registres, homogenis o heterogenis, i cada registre pot ser un document independent (macrotextos) o parts del mateix document (microtextos).

Cada registre (*node*, en terminologia d'hipertext) pot tenir una estructura diferent, i a més, incloure text, imatge o so, indistintament. La xarxa formada per aquest conjunt de registres pot estar unida per *lligams* (*links*), que pot establir l'autor de l'hipertext; o es poden crear de forma automàtica, per exemple vinculant tots els registres que tinguin determinades paraules clau o d'altres característiques formals o de contingut.

L'establiment d'un lligam suposa que és possible passar del registre A al registre B mitjançant una marca visible en el registre A, anomenada *botó*. Aquesta marca pot adquirir dues formes: una paraula en negreta, o una icona que representa, d'una manera gràfica, la funció que compleix de permetre la navegació entre registres.

La xarxa formada pels nodes o registres pot estar organitzada de forma seqüencial, jeràrquica, o en forma d'arbre lògic; però, en tot cas, de manera que es pugui anar d'un registre a l'altre mitjançant una navegació de node a node, activant els botons corresponents.

L'estructura de l'hiperdokument pot ser tan simple o tan complexa com vulgui l'autor, i el procés de creació és una operació summament complexa, semblant en molts aspectes al procés de disseny i creació d'un llibre. Quan es decideix crear un hipertext s'ha d'estudiar quina informació en formarà part, quina extensió tindran els nodes de text, quines imatges il·lustraran el text, etc. També cal dissenyar l'estructura de l'hipertext pensant en les formes d'accés a la informació, i en la necessitat d'utilitzar índexs analítics, sumaris, etc.

Normalment, els SGH no suporten formes de recuperació d'informació gaire sofisticades, però, en canvi, solen proporcionar moltes formes de navegació entre els nodes de la xarxa.

Molts hipertextos es distribueixen amb un SGH que només permet operacions de lectura. En particular, aquest tipus de SGH són molt útils per a la distribució de llibres electrònics, com ara enciclopèdies o manuals tècnics. El suport ideal, en aquests casos, sol ser el CD-ROM i d'aquesta manera és com s'estan distribuint diverses enciclopèdies electròniques al món anglosaxó (Grolier i Compton's, per exemple). En aquestes enciclopèdies, cada article és un node de l'hiperdokument, i aquest node pot estar vinculat a diverses imatges que il·lustren l'article, o a un enregistrament sonor, per exemple. A més a més, s'han predefinit diverses formes de navegar per la informació, seguint una estructura de menús que tenen les diverses opcions il·lustrades amb icones i que ofereixen diverses possibilitats a cada nivell.

A efectes conceptuals cal tornar sobre la divisió bàsica entre un programa que permet crear i llegir hipertextos, i que nosaltres hem proposat d'anomenar sistema de gestió d'hipertextos (SGH), i el conjunt de dades o de documents que formen l'hipertext en si mateix, i que alguns autors denominen hiperdokument o hipertext indistintament. A efectes conceptuals, es considera que un hiperdokument és una forma específica de base de dades o, si es vol, una extensió de les bases de dades convencionals (vegeu Fiderio, 1988).

2.2.5. Sistemes d'orientació personal o *personal information systems*

Al món anglosaxó té molt èxit una classe de programes que presenta una combinació de prestacions diferents i que se solen denominar *personal information systems* (PIM). Aquests gestors personals d'informació (GPI) estan orientats a solucionar problemes d'informació personal en el sentit que presuposen que l'administrador de la base i l'usuari final seran la mateixa persona. Per aquest motiu solen ser programes molt flexibles i de fàcil utilització, però amb poques o cap possibilitat d'establir controls de seguretat i accés, atès que normalment la mateixa persona que crea la base de dades en serà l'únic usuari.

Una altra característica típica dels GPI és que no imposen la necessitat de treballar amb camps o amb formats de registre prefixats, i així cada registre pot tenir un format diferent o, fins i tot, no tenir-ne (*free-format*).

Per altra banda, solen estar mancats d'instruments de control de la indexació, com ara fitxers de paraules buides i no contempen la possibilitat d'editar el fitxer invertit, etc.

Funcionalment, solen ser una combinació de gestor de bases de dades en format lliure, gestor de disc dur, útils per visualitzar fitxers de diversos programes (full de càlcul, processador de textos, bases de dades, etc.) i sistemes d'indexació d'arxius. Alguns d'ells incorporen petites prestacions pròpies de sistemes de gestió d'hipertextos.

Finalment, s'ha de dir que, com que la casuística (realment endimoniada) d'aquests sistemes és tan diversa, alguns no s'ajusten del tot a cap dels trets

mencionats, i tenen com a única característica que permet situar-los dins la categoria dels GPI, el fet que no presenten cap de les eines pròpies de la figura d'un administrador d'una base de dades documental (assignació de nivells de seguretat, administració d'usuaris, control de la política d'indexació, etc.).

2.3. Sistemes compostos

Diverses aplicacions de gestió pròpies del món documental tenen com a nucli un o més sistemes de gestió de bases de dades. Les més importants són les següents:

– *Sistemes d'automatització de biblioteques.*

Aquests sistemes solen tenir un mòdul de recuperació d'informació (o OPAC) que consisteix, de fet, en un sistema de gestió documental. També solen tenir un sistema de gestió de bases de dades, de tipus relacional, per gestionar les dades sobre els préstecs, les reclamacions, les adquisicions, etc. Tot aquest conjunt d'aplicacions diverses queden ocultes a l'usuari final sota una capa de software que actua com a interfície amb l'usuari, i que uneix les aplicacions en un únic sistema de menús, de manera que els usuaris veuen un únic programa.

– *Sistemes de gestió d'arxius.*

Aquests sistemes solen tenir un nucli consistent en un sistema de gestió de bases de dades, que permet la gestió de sèries documentals, el control d'espurgaments, dels trasllats, etc. Com que aquests sistemes no necessiten gestionar documentació científic-tècnica, sinó, més aviat, documentació administrativa (quasi sempre agrupada en sèries i no en documents individuals) no solen incorporar un sistema de gestió documental pròpiament dit, sinó que el seu motor de recuperació d'informació està basat en un sistema de tipus relacional.

– *Sistemes de gestió electrònica de documents.*

Una altra aplicació recent que inclou al seu nucli un sistema de gestió de bases de dades (de vegades de tipus documental i de vegades de tipus relacional), són els *sistemes de gestió electrònica de documents* (SGED). Aquests són programes que vinculen la imatge facsimil del document a la seva descripció textual de manera que, una vegada recuperat el registre (o conjunt de registres), es pugui consultar el document original en pantalla i/o es pugui reproduir amb una impressora làser. L'objectiu d'aquests sistemes és evitar els documents en paper i, en tot cas, evitar els desplaçaments a l'arxiu per recuperar el document original. Els SGED incorporen, al voltant d'un SGD o d'un SGBD, una sèrie d'eines de hardware i de software que permeten di-

gitalitzar documents, comprimir imatges, imprimir-les, etc. Funcionalment incorporen totes les prestacions d'un sistema clàssic més les següents funcions:

- Control d'escàner, per tal de poder digitalitzar imatges de documents bidimensionals i incorporar-les a la base de dades.
- Control de càmeres de vídeo, per tal de poder digitalitzar imatges d'objectes tridimensionals, amb el mateix objectiu anterior.
- Control de discos òptics, per tal d'emmagatzemar els grans volums d'informació que consumeix la imatge.
- Possibilitat d'associar imatges a registres textuais.

3. Estructura funcional dels SGD

A banda de les orientacions concretes, quasi tots els SGD comparteixen una sèrie d'elements estructurals per tal de permetre la creació i explotació de les dades i, per tant, malgrat les diferències específiques comentades, tots els SGD solen presentar un grup de funcions comunes. Aquestes funcions es podrien agrupar en els següents apartats:

- Mòdul d'administració de la base de dades
- Mòdul de manteniment
- Mòdul d'indexació i recuperació de la informació
- Mòdul de difusió de la informació

3.1. Mòdul d'administració de la base de dades

Comprèn totes aquelles operacions relacionades amb el disseny i creació de bases de dades. Es tracta d'unes operacions que cal realitzar en el moment de crear la base de dades i que, normalment, no han de ser alterades. Les més importants són les següents:

– *Definició del model de registre*

D'acord amb el tipus d'informació que s'ha de gestionar, i segons les necessitats de recuperació dels usuaris als quals se servirà, aquest mòdul permet concretar quins seran els camps que formaran part del registre, així com els noms d'aquests, longitud, etiqueta, tipus de dades que suporta, criteri d'indexació i restriccions.

– *Determinació de la política d'indexació*

Les eines per fixar les polítiques d'indexació solen gravitar al voltant de dos instruments principals:

- La creació d'un diccionari de paraules buides, o diccionari negatiu, que marca quines paraules no aniran al fitxer invertit i no seran, per tant, punts d'accés al document, per tal d'evitar un creixement desmesurat del fitxer invertit i reduir, per tant, l'índex d'ocupació de disc, però també, per poder produir llistes de descriptors més fàcils de manipular i consultar. Un fitxer de paraules buides sol permetre la inclusió de 300 a 1.000 termes.
- Els fitxers de validació, o fitxers d'autoritats. Alguns SGD permeten definir taules de valors obligatoris, és a dir, que alguns camps només poden contenir termes que estiguin presents en una llista de paraules clau o descriptors autoritzats, o bé que estiguin dins d'un cert rang de valors. Això permet fixar polítiques d'indexació amb llenguatges controlats, i/o evita que certs camps es puguin omplir amb valors erronis.

– *Establiment de nivells de seguretat i gestió d'usuari*

Per tal d'assegurar la privacitat de l'accés a la base de dades, els programes documentals solen permetre la inclusió de paraules de pas que impedeixin la utilització de les dades per part d'usuaris no autoritzats. Poden existir diversos nivells de seguretat: control del manteniment (sols podran realitzar altes, baixes i modificacions de registres els usuaris autoritzats expressament), accés a la base de dades, control d'impressions, etc. D'aquesta forma, poden existir usuaris autoritzats als nivells 0 i 1 de la base de dades (p.e. consulta i manteniment) però no al nivell 3 (p.e. impressió), o a l'inrevés.

3.2. Mòdul de manteniment

Aquest mòdul comprèn les operacions relacionades amb l'alimentació de la base de dades, o sigui, amb les altes, les modificacions i l'eliminació de registres. Les principals funcions que controla són les següents:

– *Introducció de registres (altes)*

Les dades s'han d'introduir seguint l'estructura i ordre de camps preestablerta en el moment de crear la base de dades, en el cas dels sistemes clàssics. En alguns programes serà possible utilitzar diversos models de registre o fulls d'entrada de dades (un per a articles de revista, un altre per a llibres, o fotografies, etc.). La introducció de registres es pot realitzar, normalment, per diferents vies: teclat i importació de fitxers (els dos casos més freqüents), i digitalització d'objectes bidimensionals (amb escàners) o tridimensionals (amb càmeres de vídeo).

Alguns programes incorporen facilitats d'edició, com per exemple, operacions de blocs, que ajuden a l'hora de donar altes a registres amb molt text; així com la possibilitat de crear macro-comandaments que permeten evitar

operacions molt repetitives, com haver d'introduir cada vegada el nom d'una mateixa publicació, la data del registre, etc.

- Modificació de registres

Alguns programes permeten la realització de modificacions globals en tots els registres de la base de dades. D'aquesta manera, indicant tan sols una vegada el terme o grup de termes a substituir s'aconseguirà el canvi global. En molts casos, a més, es poden afegir condicions: "substituir el terme A pel terme B només quan es trobi en el camp C".

Normalment, les modificacions es realitzen tenint en pantalla el(s) registre(s) a modificar o suprimir. En aquest sentit, hi hauria un paral·lelisme amb les modificacions manuals: primer cal anar al catàleg o fitxer, localitzar la fitxa que cal esmenar i, tot seguit, modificar-la. Per facilitar les modificacions alguns programes disposen d'un senzill editor de textos propi. En els sistemes d'indexació, en canvi, l'operació és més complexa a causa de la necessitat de traspasar els registres a un processador de textos extern.

3.3. Mòdul d'indexació i recuperació de la informació

3.3.1. La indexació

La qualitat de les eines per fixar polítiques d'indexació és un paràmetre crític de tot SGD, juntament amb les possibilitats del seu llenguatge d'interrogació, i ambdós formen el nucli del seu motor de recuperació de la informació.

Des del punt de vista de la màquina, el procés d'indexació consisteix en la lectura dels registres (fitxer lineal) per formar els índexs que serviran per a la recuperació de la informació (fitxer invers). D'aquesta manera, totes aquelles paraules o cadenes de caràcters que es troben en un camp indexat (marcat per a índex), i que no constin en el fitxer de paraules buides seran traspasades automàticament a l'índex invers i ordenades alfabèticament. Així doncs, tota paraula no buida serà un punt d'accés al registre.

Aquesta és la tècnica bàsica de la indexació automàtica. Alguns sistemes permeten refinar aquest mètode i hi incorporen l'anàlisi de la freqüència estadística dels termes en un registre determinat, comparat amb la freqüència dels mateixos termes al conjunt de la base, i atorguen un pes d'acord amb aquest valor estadístic, de tal manera que només formen part del fitxer invertit aquelles paraules que superen un cert llindar.

Des del punt de vista de les persones, el procés d'indexació consisteix en la determinació, a través d'un procés intel·lectual, del contingut temàtic d'un document i en l'assignació de les paraules clau i/o dels descriptors que millor el representin. Alguns sistemes permeten combinar els dos sistemes bàsics d'indexació, l'automàtic i l'intel·lectual. Aquest últim es pot implementar o bé mitjançant camps que s'han de validar contra un fitxer de paraules autoritzades o contra un thesau-

rus, o bé amb camps que admeten separadors de descriptors, de forma que es pugui considerar un sol descriptor un grup de paraules. D'altra banda, la majoria d'SGDs permeten que l'usuari decideixi quins camps vol indexar i quins no.⁴

En tot cas, els termes d'indexació solen agrupar-se en un o més fitxers invertits. Aquests fitxers contenen la relació alfabètica dels termes, els quals estan associats als números de registre en els quals aquests es troben. Alguns SGD disposen de fitxers inversos en els quals consta, a més, la posició concreta del terme dins del camp. Actualment, el fitxer invers acostuma a ocupar entre un 60-90% de l'espai del fitxer lineal (la proporció depèn del programa concret, del número de camps que es volen indexar i del fitxer de paraules buides).

– *Actualització*

Tots els registres afegits al fitxer lineal solen seguir el procés d'inversió comentat, consistent en la seva exploració i extracció dels termes que formaran part del fitxer invertit o índex. Alguns programes permeten realitzar la inversió després de la introducció de grups de registres (*batch*), a elecció de l'usuari. D'altres programes realitzen la inversió de forma automàtica després de cada alta o modificació d'un registre.

– *Reindexació*

Quasi tots els SGD permeten reindexar tota la base de dades, per tal de crear nous índexs si, per algun motiu s'ha modificat la política d'indexació, el fitxer de paraules buides, o s'ha alterat l'estructura dels camps.

3.3.2. La recuperació d'informació

El mòdul de recuperació de la informació, juntament amb el sistema d'indexació, és el nucli de tot SGD, i de la qualitat i prestacions d'aquest, depèn la qualitat de tota la base de dades. En aquest nucli, estretament relacionat amb el sistema de creació dels índexs, i en la forma en la qual s'implementa de cara a l'usuari final, rau la principal diferència d'un SGD respecte a un SGBD.

Generalment, els SGD, incorporen les següents formes de recuperació de la informació:

– *Consulta a través dels índexs*

Alguns sistemes permeten la consulta en línia del fitxer invertit. D'aquesta manera és possible saber, primer, quants documents estan associats a una determinada paraula o a un descriptor i, segon, recuperar automàticament tots els registres que contenen aquesta paraula o descriptor. És també molt

4. Els que no permeten aquesta opció condicionen molt l'aplicació que es vol desenvolupar, perquè obligaran a indexar tots els camps i produiran índexs més voluminosos i amb més soroll.

útil per a usuaris novells que desitgin realitzar consultes simples (un autor, un títol, una matèria, etc.). Els índexs, com a relació alfabètica de tots els termes significatius d'una base de dades, permeten obtenir-ne una visió general del contingut i, per tant, actuen com a sistema de representació del coneixement contingut a una base de dades.

- *Consulta assistida per menús*

Es tracta d'un sistema de cerca simplificada en què l'usuari no necessita conèixer la sintaxi del llenguatge d'interrogació de la base, sinó tan sols respondre a les preguntes que el sistema li proposa ("vols buscar un autor, un títol, o una matèria?"), i, a continuació, indicar quin és el terme buscat. Habitualment, aquesta funció s'implementa també mostrant un formulari que representa els camps de la base de dades, i l'usuari ha de incloure un terme de recerca als camps que consideri oportú. En els dos casos, el sistema de recuperació calcula l'equació de recerca que correspon a l'elecció feta per l'usuari.

- *Consulta al fitxer lineal*

Encara que no es tracti estrictament d'una forma de cerca, aquesta opció la permeten molts SGD i consisteix a poder avançar i retrocedir per entre els registres de la base de dades, disposant d'una visió general del seu contingut i estructura. Els anglosaxons utilitzen el terme *browse* (fullejar) per referir-se a aquesta opció.

Els sistemes de gestió d'hipertextos s'han especialitzat molt en aquestes funcions i permeten operacions de *browse* molt sofisticades, basades en la xarxa de lligams establerts entre els nodes de la base (registres).

- *Definició d'equacions de cerca*

Quasi tots els SGD (i notablement els sistemes clàssics) disposen d'un llenguatge d'interrogació més o menys sofisticat, que segueix una sintaxi pròpia (no estan normalitzats) i utilitza una sèrie d'operadors per combinar els termes de la cerca.

Aquest llenguatge és necessari per a consultes complexes que combinin més d'un concepte o termes de cerca. Segons el programa de què es tracti disposarà de determinats operadors o instruccions, però a la base de quasi tots els llenguatges d'interrogació sol trobar-s'hi l'àlgebra de Boole.

El llenguatge d'interrogació inclou el conjunt d'instruccions i operadors que permet posar en contacte les consultes dels usuaris (formulades en llenguatge natural i, per tant, intel·ligibles per al sistema) amb la lògica interna d'emmagatzematge del programa. Per aconseguir això, és necessari formalitzar la consulta fins a convertir-la en el que s'anomena equació de cerca, que està estructurada seguint uns paràmetres que el sistema és capaç d'entendre i als quals pot donar resposta.

Els principals elements i operadors del llenguatge d'interrogació d'un SGD són els següents:

- Operadors booleans:

Unió, intersecció i negació de conjunts.

- Operador de proximitat:

Utilitzat per trobar registres en els quals un terme es troba en un nombre determinat de paraules d'un altre terme.

- Cerca limitada a un camp concret:

Permet acotar a quin(s) camp(s) volem buscar un terme.

- Cerca per rang:

Permet buscar els registres que es troben a partir d'un valor de referència o entre dos valors. Els principals rangs a establir són: "major que", "menor que", "entre .. i ..".

- Truncaments i màscares:

Alguns programes permeten truncament a la dreta, a l'esquerra i al mig de la paraula. Alguns programes permeten l'ús de màscares que substitueixen tants caràcters com màscares emprades.

- Semblança fonètica:

Amplia la recuperació d'un terme a tots aquells que tinguin una expressió fonètica similar. Provoca molt soroll en la recuperació però pot ser útil en determinats casos.

- Cerca amb l'ús de thesaurus en línia o amb sinònims:

Alguns programes poden traduir el terme de la recerca als termes autoritzats al thesaurus, o ampliar-la buscant també pels termes més genèrics, o pels sinònims, etc.

- Combinació de conjunts de cerca precedents:

És freqüent la necessitat de combinar equacions de recerca precedents amb noves línies de recerca, i alguns programes que permeten aquesta opció assignen números a cada subconjunt recuperat.

- Combinació de diversos operadors:

Alguns programes permeten combinar diversos termes i operadors en una sola operació de recerca, particularment amb l'ús de parèntesis, per tal d'indicar quin operador afecta a cada terme i amb quina procedència s'ha de calcular la recerca.

D'altres possibilitats complementàries consisteixen a localitzar informació no indexada (mitjançant una exploració seqüencial dels registres), guardar estratègies de cerca per a proporcionar serveis de DSI (difusió selectiva de la informació) i ressaltar els termes de cerca en el registre per tal de veure ràpidament el context dels termes buscats.

3.4. Mòdul de difusió de la informació

Comprèn tant els aspectes relacionats amb la sortida de registres de la base de dades (exportacions) com les operacions que s'ocupen de la incorporació i adaptació de fitxers externs de registres (importacions). Ambdós processos són fonamentals en un SGD, ja que asseguren la difusió i possibilitat d'intercanvi de les seves dades amb l'exterior del sistema.

3.4.1. Exportacions

En una base de dades no n'hi ha prou amb recuperar i visualitzar per pantalla els registres. Normalment els usuaris necessiten, a més, poder disposar físicament del conjunt de referències que han seleccionat. Per aquest motiu és important que el programa permeti la transferència d'informació des del sistema cap a l'exterior. Per a la denominació d'aquestes operacions sol emprar-se l'expressió *exportació de dades*, incorporant un terme de l'àmbit econòmic.

Aquest mòdul ha de contemplar la possibilitat d'exportar la informació a fitxers informàtics amb diversos formats. És imprescindible que es pugui exportar en format ASCII i que s'hi puguin incloure marques de final de registre i de final de camp, així com que es puguin ordenar els registres per un o més camps abans de fer l'operació d'exportació.

- El format

La forma o estil segons els quals es visualitzaran o imprimiran els registres depèn del format utilitzat. Disposar d'un llenguatge de format de registres més o menys complex és una important característica diferencial entre els diferents SGD.

El format assenyala dos grans tipus d'informacions al sistema:

- Els camps que hi han d'aparèixer i l'ordre en què es visualitzaran.
- Els caràcters (signes de puntuació, etc.) i literals (expressions textuais) que han d'aparèixer entre un i altre camps.

A partir d'un mateix registre podrien configurar-se, p.e., estils diversos com els que mostra la figura següent.

Figura 3. Diversos formats d'exportació d'un mateix registre bibliogràfic.

Format A	Autor: F.W. Lancaster Títol: Toward paperless information systems Ciutat: New York Editorial: Academic Press Any: 1978 Pàgines: 216 p.
Format B	F.W. Lancaster. <i>Toward paperless information systems</i> . New York : Academic Press, 1978. 216 p.
Format C	LANCASTER, F.W. Toward paperless information systems / F.W. Lancaster. – New York : Academic Press, 1978. – 216 p.

- El suport

Les exportacions de registres es poden realitzar bàsicament per dues vies: paper (les referències s'envien a la impressora per tal de traspassar-les sobre paper) o suport magnètic (els registres exportats quedaran emmagatzemats en un disquet o al disc dur per tal de poder ser manipulats posteriorment per un altre programa diferent).

3.4.2. Importacions

Aquest terme fa referència a l'operació inversa a la descrita anteriorment, o sigui, a la incorporació de dades des de l'exterior cap al sistema. És una de les operacions més complexes de la gestió d'una base de dades.

La possibilitat d'intercanviar fitxers d'informació processats per diferents programes és una opció cada cop més necessària en els SGD. La importació de dades pot ser de diverses classes: d'un processador de textos, d'un altre SGD, o d'una base de dades en línia.

4. Tendències

Els últims vint anys, la tecnologia documental s'ha mantingut pràcticament sense canvis, perquè els sistemes actuals presenten la mateixa tecnologia

bàsica que els sistemes dels anys 70. Ni tan sols els hipertextos no són tan innovadors com molts els consideren, perquè existeixen també des dels anys 70 i, més que incorporar autèntiques innovacions tècniques, representen una nova forma de concebre l'ús de tècniques preexistents, per tal de posar-les al servei de la gestió de documents multimèdia. Allò que ha resultat innovador en els hipertextos és la seva filosofia, consistent a concebre l'ordinador com una màquina literària, i les grans possibilitats d'estructurar els documents com xarxes amb nodes d'informació i lligams conceptuals entre ells.

En tot cas, en els últims dos o tres anys, s'han estat produint un seguit d'innovacions en els SGD que consisteixen, o bé en la incorporació (encara molt tímida) d'algunes teories de la recuperació d'informació, tal com sistemes de ponderació i formes de recuperació no booleanes; o bé en intents d'integrar els sistemes documentals en arquitectures d'informació corporatives, on han de poder conviure amb sistemes relacionals.

Concretament, les principals innovacions es poden agrupar en els apartats següents:

- *Incorporació de tècniques per calcular el grau de pertinença dels documents recuperats.*

Alguns programes poden realitzar una estimació de la probabilitat de pertinença de cada document recuperat, i poden mostrar, en primer lloc, els documents més pertinents, és a dir, els que s'ajusten més a les condicions de la consulta de l'usuari, obviant el principal inconvenient de la lògica booleana del tot o res, segons la qual, un document o és pertinent o no ho és, sense possibilitat d'establir un grau de pertinença.

- *Incorporació de prestacions pròpies d'hipertextos.*

Una de les opcions que comencen a incloure cada vegada més programes és la possibilitat de considerar els registres de la base de dades com a nodes d'una xarxa de documents vinculats entre ells. D'aquesta manera, és possible disposar de noves formes per moure's entre el conjunt de registres (*navegació* és un dels noms emprats per designar aquest tipus de consulta), a més de les formes determinades per l'àlgebra de Boole.

- *Incorporació de tècniques especials de construcció d'estratègies de recuperació.*

Amb la introducció d'aquests recursos és possible assignar pes (ponderació) als termes de la pregunta, i també crear equacions de cerca a partir de documents que l'usuari indica com a representatius de la seva necessitat d'informació.

- *Combinació de tecnologies relacionals i tecnologies documentals en programes mixtos.*

Aquesta tendència és particularment important i obrirà, possiblement, una nova dimensió als programes de gestió documental. La tecnologia relacional gaudeix d'un alt grau de formalisme matemàtic i, per tant, les bases de dades construïdes amb aquesta tecnologia són especialment consistents. Al mateix temps, l'èxit del model relacional en el món de la gestió empresarial fa que hi hagi un format internacional (anomenat SQL, sigles d'*structured query language*) que cada vegada és incorporat per un major nombre de programes. L'existència d'una norma comuna permet que els programes que saben "llegir" SQL puguin explotar qualsevol base de dades amb format SQL, independentment del programa amb el qual s'han creat les dades.

Alguns programes documentals, doncs, estan combinant la consistència, la seguretat i l'estandarització pròpies de l'àlgebra relacional amb aquelles eines de la documentació que resulten imprescindibles per a una gestió professional, com ara l'ús de sistemes de fitxer invertit i els sistemes de control de la indexació.

En aquesta nova estirp de programes documentals, que podria ser una de les tecnologies dominants en el futur, els registres posseeixen una estructura tabular, per tal de respectar el model de taula propi dels sistemes relacionals, però disposen també d'un nou tipus de camp, el contingut del qual és el nom d'un document no tabular, ja sigui un text o una imatge. És a dir, a la taula hi ha dos tipus de camps: els camps convencionals que s'omplen amb valors prefixats, i els camps de tipus punter, que contenen la identificació i l'adreça del fitxer documental. A aquests camps, la literatura tècnica anglosaxona els anomena BLOBs, per *binary large objects* (grans objectes binaris).

D'aquesta manera, quan l'administrador de la base defineix una nova base de dades, pot decidir si els camps seran relacionals o documentals. En el segon cas, aleshores, es tractaran amb tècniques pròpies dels sistemes documentals clàssics (ja comentades) de forma independent al tractament relacional convencional que té la taula. Es tracta, per tant, d'una extensió de les prestacions relacionals clàssiques, per tal d'adaptar-la a les necessitats de la documentació, de tal manera que en aquests tipus de programes conviuen les dues lògiques de treball, la lògica relacional i la lògica documental. Aquesta última s'incorpora amb totes les conseqüències, incloent, per tant, eines tan purament documentals com thesaurus, gestió de paraules claus, gestió de fitxers invertits, gestió de diccionari de paraules buides, etc.

Normalment, aquests tipus de programes estan preparats per treballar en entorns d'arquitectura client/servidor, la qual cosa els fa especialment idonis per treballar en xarxes, així com per explotar bases de dades distribuïdes. A més a més, aquests programes, com que van adreçats especialment a la gestió documental, incorporen eines específiques per als documentalistes, com la gestió de thesaurus, l'administració d'usuari, l'edició de lèxics, etc.

- *Adaptació dels SGD a sistemes operatius que suporten interfícies gràfiques d'usuari (IGU)*

Darrerament, la microinformàtica està derivant cap a sistemes operatius que incorporen interfícies gràfiques d'usuari (IGU), tal com Windows i OS/2, sent Windows una de les més implementades, gràcies a la seva compatibilitat total amb el sistema MS-DOS.

Per aquesta raó, diversos SGD que existien només en versió MS-DOS s'estan adaptant a l'entorn Windows, i cal esperar que el mateix succeirà amb el sistema operatiu OS/2. Això significa, en general, programes documentals molt més versàtils, amb capacitats multimèdia (Windows i OS/2 són essencialment multimèdia) i amb corbes d'aprenentatge molt reduïdes.

La utilització d'una IGU facilita la gestió de documents no textuais, ja que tots els tipus d'informacions que suporta l'IGU així com tots els tipus de perifèrics controlats pel sistema operatiu, passen a ser suportats per qualsevol programa que hagi estat escrit per aquest sistema operatiu.

5. Els SGD al mercat espanyol

La taula 2 presenta una relació dels principals SGD que es poden trobar al mercat espanyol. La inclusió de programes en aquesta llista s'ha fet atenent als criteris següents:

- que es tracti de programes per a microordinadors compatibles (és a dir, que disposen d'un microprocessador Intel i que, en el vessant lògic, utilitzen el sistema operatiu MS-DOS, o un clònic del MS-DOS –p.e. DR-DOS–, o sistemes compatibles amb MS-DOS –com l'OS/2 o l'entorn Windows–)
- disposició d'un servei de suport tècnic o distribuïdor a l'Estat espanyol.

Per tal de donar una idea del nivell professional dels programes s'utilitza una triple classificació: nivell personal, nivell mig i alt nivell. El nivell es refereix bàsicament als aspectes següents:

- disponibilitat d'eines per a l'administrador de la base (d'indexació i de seguretat);
- possibilitat de mantenir oberta més d'una base de dades, volum d'informació que poden suportar, i possibilitats multiusuari (en xarxa o en sistemes operatius multiusuari).

El programes de *nivell personal* (gestors personals d'informació i els gestors de bibliografies) solen tenir molt poques d'aquestes eines, per les raons apuntades al punt 2.2.5.

Els programes de *nivell mig* disposen de certes funcions de seguretat, sovint del tipus tot o res, és a dir, que els usuaris que tenen el password d'accés ho poden fer tot, i els que no el tenen no poden fer res, o només poden llegir informació. Inclouen també algunes possibilitats de control de la inde-

xació, com declarar camps no indexats, camps que s'han de validar contra un fitxer de termes autoritzats, o es poden establir llistes de sinònims.

Per altra part, els programes d'*alt nivell* proporcionen eines completes d'administració de la base, a efectes de la definició de diversos tipus d'usuari amb diversos privilegis. Permeten tot tipus de control d'indexació incloent l'ús de thesaurus capaços de complir normatives ISO. També solen estar adaptats a diverses plataformes de hardware i de software, incloent sistemes operatius multiusuari.

(Els sistemes de gestió electrònica de documents solen ser superconjunts dels sistemes clàssics, és a dir, tenen totes les prestacions d'un sistema clàssic més la possibilitat de gestionar imatge.)

Finalment cal advertir que, malgrat que els autors han intentat aplicar aquests paràmetres d'una forma objectiva, és impossible evitar del tot un cert factor de subjectivitat i, per tant, aquesta classificació no pot comprometre als distribuïdors o productors del programes esmentats, els quals poden tenir, molt legitimament, un criteri molt diferent sobre el seu propi producte. En tot cas, només una demostració efectuada per personal autoritzat i un període de prova en condicions realistes d'ús, poden donar una idea ajustada de les autèntiques prestacions d'un programa.

Taula 2. Principals SGD al mercat espanyol		
Nom	Tipus	Distribuïdor
Archivist	Gestió bibliografies i gestor d'informació personal, amb algunes prestacions documentals molt originals.	Glahn (93) 451 45 46
AskSam	Gestor d'informació personal, amb prestacions d'hipertext, i amb llenguatge de programació. Té algunes prestacions de sistema documental clàssic de nivell mig.	BSI Multimedia (93) 266 31 20
BRS/Search	Sistema documental clàssic de molt alt nivell, amb prestacions de gestió electrònica de documents. Existeix en versions per a diversos sistemes operatius, incloent MS-DOS i Windows.	Eurotelemática (93) 310 77 66

Nom	Tipus	Distribuïdor
CDS/ISIS	Sistema documental clàssic d'alt nivell (en distribució només per a organismes sense ànim de lucre). Disposa d'un potent llenguatge per dissenyar formats. Per a diversos sistemes operatius, incloent MS-DOS i algunes versions d'Unix.	Unesco (Distribuïdor regional per a Espanya: CINDOC, (91) 563 54 82
Clarity	Sistema de gestió electrònica de documents, amb una orientació documental clàssica, de nivell mig.	Micronet (93) 439 15 40
Concord Textmaster	Sistema d'indexació, amb un complet motor de recuperació d'informació d'orientació documental. N'existeixen versions personals i professionals.	Lagrange Group (93) 322 15 80
Docu-Imagen	Sistema de gestió electrònica de documents d'alt nivell, de tipus documental / relacional, amb arquitectura client / servidor i funcions de seguiment d'expedients i correu electrònic. Per a entorn Windows.	Bull Espanya (93) 330 66 11
Guide	Sistema de gestió d'hipertextos, amb possibilitats de gestionar documents multimèdia. Per a entorn Windows.	BSI Multimedia (93) 266 31 20
HyperMap -Doc	Sistema de gestió electrònica de documents d'alt nivell, amb tecnologia relacional / documental, i arquitectura client/servidor. Per a entorn Windows.	Cognivision Research (93) 323 63 22
Inmagic	Sistema documental clàssic de nivell mig, amb un llenguatge de programació per gestionar les sortides d'informació sobre paper.	DOC-6 (93) 414 06 79

Nom	Tipus	Distribuidor
InvesDoc	Sistema de gestió electrònica de documents d'alt nivell, amb tecnologia relacional / documental, i arquitectura client / servidor. Per a entorn Windows.	Investrónica (93) 478 78 18
Knosys	Sistema documental situat entre un gestor personal d'Informació i un sistema clàssic de nivell mig.	Micronet (93) 439 15 40
Reference Manager	Gestió de bibliografies (amb un gran nombre de formats predefinits) i Gestor Personal d'Informació.	DOC-6 (93) 414 06 79
Smartext	Sistema de gestió d'hipertextos, molt orientat a documents multimèdia i amb un sistema de recuperació d'informació de tipus documental. Per a entorn Windows	Lotus Ibérica (93) 419 01 04
Taurus	Sistema de gestió electrònica de documents d'alt nivell, de tipus documental / relacional, amb arquitectura client / servidor per a l'entorn Windows.	Chemdata (91) 435 22 02
Texto-Logotel Texto-GED	Sistema documental clàssic de molt alt nivell, amb la possibilitat d'incorporar un llenguatge de programació propi, Logotel. Funciona amb diversos sistemes operatius, com MS-DOS i Unix. Texto-Ged és la versió Texto per a Windows, amb capacitat de gestió electrònica de documents.	Chemdata (91) 435 22 02

ANNEX:

Mentre es realitzava aquest article, els autors han tingut coneixement d'altres programes, encara que només a nivell de documentació comercial.⁵

Nom	Tipus	Distribuïdor
Archidoc	Sistema de gestió electrònica de documents.	Infodoc (91) 577 73 83
Infodoc Information Retrieval System	Sistema de gestió documental de tipus multimèdia	Infodoc (91) 577 73 83
Optosaurus	Sistema de gestió electrònica de documents.	Estado Sólido (91) 308 01 49
PC-Search	Sistema d'indexació i recuperació	Chemdata (91) 435 22 02
Personal Librarian	Sistema de gestió documental per a l'entorn Windows.	CSI (91) 265 34 55

6. Bibliografia

Blair, D.C. *Language and representation in information retrieval*. Amsterdam [etc.]: Elsevier, 1990.

Canals, I. "Introducción al hipertexto", *Revista Española de Documentación Científica*, 13.2 (1990). p. 685-709.

Citroen, L.C. "Micro-computer software for information retrieval: how to make a well-founded choice", *Program*, 23.2 (1989) 141-150.

5. Malgrat la crisi general del mercat de la informàtica, els SGD i, particularment, els SGED constitueixen un dels sectors que estan experimentant un major creixement. Per aquest motiu, fins que aquest àmbit no s'estabilitzi es fa molt difícil presentar un llistat exhaustiu d'aquestes aplicacions.

Clayton, M. *Gestión de automatización de bibliotecas*. Salamanca; Madrid : Fundación Germán Sánchez Ruipérez, 1991. 331 p. (Biblioteca del libro)

Codina, L. "Gestión electrónica de documentos", *Binary*, mayo (1992), p. 63-75.

Codina, L. "Nous sistemes operatius per a microordinadors i les seves possibles implicacions per a la gestió documental", *Item*, 10, (1992) p. 64-78

Codina, L. "Bases de datos documentales para microordenadores", *Terceiras Jornadas Españolas de Documentación Automatizada*, Palma: UIB, 1990.

Fidel, R. *Database design for information retrieval: a conceptual approach*. New York [etc]: John Wiley, 1987. 232 p.

Fiderio, J. "A gran vision". *Byte*, october (1988).

Fondin, H. "L'évolution des logiciels documentaires". *Les marchés de l'information documentaire*. J. Treffel, prés., A. Vuillemin, dir. Paris: ADBS, 1991. 313 p.

Gerrie, B. "Small scale databases", a *Small scale databases*. Melbour-ne: Academic Press, 1986. 1-15.

"Informatique et recherche documentaires", *Archimag*, 35, (juin 1990) p. 30-39.

K. Parsaye... (et al.) *Intelligent databases: object-oriented, deductive hyper-media technologies*. New York: John Wiley & Sons, 1989.

Kemp, A. *Computer-based knowledge retrieval*. London: Aslib, 1988. 399 p.

Khoshafian, S. et al. *Intelligent offices: object-oriented multi-media information management in client/server architectures*. New York: John Wiley & Sons, 1992.

Nieuwenhuysen, P. "Evaluating text storage and retrieval software", *The electronic library*,

Bertrand, R. (dir). *Micro-ordinateur et traitement de l' information*. Paris: A Jour, 1991. 216 p. (Nouvelles Technologies Documentaires.)

Rada, R. *Hypertext: from text to expertext*. New York: McGrawHill, 1991.

Salton, G. *Automatic text processing: the transformation, analysis, and retrieval of information by computers*. Reading: Addison-Wesley, 1989.

Small scale databases. Melbourne: Academic Press, 1986.

Sieverts, E.G. i M. Hofstede. "Software for information storage and retrieval tested, evaluated and compared. Part 1 - General introduction", *The Electronic Library*, 9.3 (1991): 145-154.

E. G. Sieverts... (et al.) "Software for information storage and retrieval tested, evaluated and compared. Part 2 - Classical retrieval systems", *The Electronic Library*, 9 .6 (1991): 301-316.

E. G. Sieverts... (et al.) "Software for information storage and retrieval tested, evaluated and compared: Part 4 - Indexing and full-text retrieval programs", *The Electronic Library*, 10.4 (1992): 195-206.

Tenopir, C., i G. Lundeen. *Managing your information: how to design and create a textual database on your micro-computer*. New York; London: Neal-Schuman, 1988. 226 p.

"Text indexing", *Electronic documents*, 1.3 (1992): 1-32.