

L'evolució de la distribució de bases de dades en línia en els darrers anys des de la perspectiva catalana

VICENÇ MELÉNDEZ*

Resum Es descriu l'evolució de la distribució de bases de dades en línia i es delimita i caracteritza la informació professional o especialitzada, en relació amb altres tipus d'informació que també són accessibles en línia. S'analitzen els factors que han influït i influeixen en la distribució de bases de dades a escala internacional; en especial, l'aparició dels suports òptics CD-ROM i dels serveis videotex. S'assenyalen les tendències observades en l'activitat dels distribuïdors més importants. Finalment es descriuen les situacions espanyola i catalana en aquest context.

La distribució de bases de dades en línia durant els anys 80

En la dècada dels 80 s'ha consolidat l'activitat de distribució de bases de dades en línia que va sorgir, amb caràcter públic, els anys 70. El 1973 es van obrir a Espanya dos serveis d'accés a una dotzena de bases de dades bibliogràfiques carregades al Servei de Recuperació d'Informació de l'Agència Europea de l'Espai: un a l'INTA i l'altre al CIDC (actualment Institut d'Estadística de Catalunya). L'accés a aquests sistemes ha passat a ser, al llarg d'aquest temps, un procediment completament normal de consulta de la informació.

Aquests anys s'han caracteritzat pel manteniment del predomini dels serveis de distribució de bases de dades dels Estats Units en un mercat internacional sense duanes ni fronteres. Predomini que es va iniciar quan es va fer possible, a meitat dels anys 70, l'accés des de la xarxa telefònica commutada als nous distribuïdors

* SOCADI

dels Estats Units, Dialog i SDC, els quals van començar oferint un conjunt de fonts d'informació, de tipus bibliogràfic, produïdes per agències governamentals que constituïen l'espina dorsal (15) d'aquests serveis.

La transmissió de dades als Estats Units era més senzilla que als països europeus, a la majoria dels quals no hi havia possibilitat d'arribar-hi més que per truca-da directa o bé, en el cas d'Esalrs, mitjançant una xarxa punt a punt i uns equips de cost elevat.

Durant aquest període s'ha accentuat el creixement del nombre i la varietat de bases de dades accessibles i s'han diversificat les temàtiques que a l'inici eren majoritàriament tècniques (quadres 1 i 2).

Quadre 1. Estadístiques de bases de dades

Directory Issue	Number of Databases*	Number of Database Producers	Number of Online Services	Number of Gateways
1979/80	400	221	59	
1980/81	600	240	93	
1981/82	965	512	170	
1982/83	1.350	718	213	
1983/84	1.878	927	272	
1984/85	2.453	1.189	362	
1986	2.901	1.379	454	35
1987	3.369	1.568	528	44
1988	3.699	1.685	555	59
1989	4.062	1.813	600	75
1990 (current)	4.465	1.950	645	88

Font: *Directory of online databases*. New York: Cuadra Associates.

Quadre 2. Bases de dades de Dialog per tipus (en %)

Bibliogràfiques	58,4
Catàleg i directoris	19,4
Text complet	15,7
Documents no bibliogràfics	2,0
Estadístiques	1,7
Propietats Físico-químiques	58,4
Numèriques no estadístiques	1,0
Processos en curs	1,0

Font: Database catàleg. Dialog. 1.1.90. Agregació pròpia sobre 200 bd.

En el panorama europeu han sorgit distribuïdors petits en un nombre excessiu en relació amb els usuaris potencials, sostinguts pels poders públics de cada país en molts casos i amb una projecció internacional escassa. Durant la dècada s'ha normalitzat l'accés entre xarxes de commutació de paquets dels diferents països, després que les Comunitats Europees, en un intent de donar més homogeneïtat al mercat europeu, posessin en funcionament una xarxa provisional (Euronet) que unifica l'accés a tots els distribuïdors europeus. Des de les Comunitats també es va intentar la implantació d'un llenguatge comú d'interrogació, ara adoptat per la ISO, el Common Command Language (4).

Segons un estudi de l'any 1989 de l'Information Market Observatory (IMO) de les Comunitats Europees citat a la revista *Infotecture* (8), el nombre de distribuïdors a Europa era de 227, amb un creixement del 8%, i de 300 als Estats Units on es manté estable des de 1986. Entre els 10 primers distribuïdors del món, 8 eren dels Estats Units. Els 5 primers distribuïdors d'aquest país tenien el 42% de les bases accessibles allí, mentre que a Europa els 5 primers n'agrupaven només el 26%.

La fragmentació política i la diversitat lingüística han estat i són dificultats per al desenvolupament de serveis de distribució que competeixin amb els nord-americans.

Al llarg d'aquesta dècada s'ha vist també com les dimensions del mercat de les bases de dades en línia en el seu vessant més professional, eren inferiors a les previstes i, per altra banda, que es tracta d'un mercat molt concentrat.

Segons la revista *Monitor* (14) i citant la publicació *Information Market Indicators*, el mercat de les bases de dades textuais i bibliogràfiques va ser, el 1988, de 456 milions de dòlars, dels quals 110 corresponien a Dialog i 228 a Mead data Central. Setze bases de dades acumulaven el 61% del consum total. *Monitor* indicava també que el mercat europeu podria ser el 15% del dels Estats Units amb

uns 68 milions de dòlars, la qual cosa seria compatible amb les xifres de facturació dels principals distribuïdors europeus. Les dades serien les següents: Dialog (Europa) 16,5; Data-Star 15; Esa-Irs 13; Questel 8, i Stn 7.

Se citen distintes causes per explicar el baix grau de desenvolupament de la informació especialitzada o professional respecte d'altres tipus d'informació: la natura d'aquesta informació, que sovint no presenta un caràcter d'urgència i aplicabilitat propi d'altres tipus d'informació com ara la informació en temps real (de cotitzacions de valors, tipus de canvi, d'interès, etc.) o la informació de gran públic, que es canalitza per les xarxes de videotex; la dificultat de cerca d'informació especialitzada, que sovint requereix un professional per a dur-la a terme; les característiques dels sistemes d'interrogació emprats ara com ara, que obliguen a definir el tema cercat amb molt de rigor i no ajuden prou l'usuari a poder realimentar la seva cerca a partir de la informació rellevant que localitza; la tarificació penalitzadora de l'usuari final, basada en el temps de consulta més que en la quantitat d'informació.

Per contra, altres tipus de serveis de distribució d'informació en línia, en àrees econòmiques i també en àrees d'interès general que han anat sorgint amb posterioritat, han tingut un creixement superior (vegeu quadre 3). Malgrat que es tracti de serveis amb finalitat i usuaris diferents, no hi ha unes fronteres clares entre ells i així podem veure com, per exemple, Dialog ofereix, cada cop més, informació econòmica quasi en temps real.

Quadre 3. Nombre d'usuaris i % de vendes, per tipus de serveis

	% 1988	31/12 1988	31/12 1987	31/12 1986	31/12 1985	31/12 1984	% vendes 1988
Interès general	38,38	1.018.654	860.000	677.000	505.025	326.100	5,10
% d'increment			18,45	27,03	34,05	54,87	
Empresa i Finances	28,73	762.594	667.182	558.618	468.165	377.069	37,20
% d'increment			14,30	19,43	19,32	24,16	
Ciència i Tècnica	15,24	404.529	363.491	345.841	328.619	229.237	17,40
% d'increment			11,29	5,10	5,24	43,35	
Crèdit i Solvència	8,83	234.338	205.100	157.701	140.848	111.500	21,40
% d'increment			14,26	30,06	11,97	26,32	
Línies aèries (inf.)	7,07	187.655	170.448	149.300	114.700	69.000	11,10
% d'increment			10,10	14,16	30,17	66,23	
Prensa	1,75	46.468	45.661	45.897	42.112	38.700	7,80
% d'increment			1,77	-0,51	8,99	8,82	
Total	100,00	2.654.238	2.311.882	1.934.357	1.599.469	1.151.606	100,00
% d'increment			14,81	19,52	20,94	38,89	

Font. Digital Information Group. *Information Industry's Annual Report 1989/90.*

Factors que influeixen en la distribució

Al llarg de la dècada passada s'han produït desenvolupaments que han influït i influeixen en la distribució de bases de dades en línia.

D'una manera continuada, els equips informàtics han anat augmentant de potència mentre que els preus es mantenen o baixaven. Gràcies a això, la informàtica documental ha deixat de ser patrimoni d'uns pocs grans centres oficials. Això ha fet possible l'augment continuat del nombre de bases de dades no solament bibliogràfiques sinó també en text complet, directoris i d'altres tipus d'informació de natura similar que veiem actualment. A escala espanyola ha donat lloc a l'aparició del fenomen de l'autodistribució de les bases de dades (centres que no tan sols produeixen bases sinó que addicionalment hi donen l'accés al públic). Aquest procés d'augment de potència, pel que fa als equips microordinadors d'ús individual, ha posat els fonaments per a la consulta de fonts d'informació en local (sense accedir a un distribuïdor remot). Però el desenvolupament que ho ha permès ha estat el CD-ROM, suport d'edició de caràcter òptic, de gran capacitat d'emmagatzematge.

Des de l'inici de la dècada es parlava del gran impacte que tindrien els suports d'emmagatzematge òptic en documentació, però no han estat els videodiscos ni els discos òptics digitals els que finalment han tingut aquest impacte sinó el disc compacte per a àudio, amb la seva aplicació per a dades. Aquest petit disc, de 12 cm de diàmetre i capacitat d'emmagatzematge superior a mig gigabyte, s'ha fet imprescindible en la distribució de bases des de la seva aparició l'any 1985. El seu èxit es basa en el fet que es tracta d'un suport barat, durador i amb una gran capacitat. Però sobretot té una característica fonamental que és el grau d'estandardització assolit. Els creadors, Philips i Sony, es posaren d'acord en les seves característiques físiques, dimensions, velocitat de funcionament, etc., a fi que un mateix disc pogués consultar-se en qualsevol lector. Amb posterioritat a la seva aparició, s'acordà l'estàndard que defineix l'ús dels CD-ROM des de diferents sistemes operatius.

L'Associació Americana d'Editors de Discos Òptics dona les dades del quadre 4 pel que fa al desenvolupament del CD-ROM.

Quadre 4. CD-ROM

	1989	1990
Títols impresos	1.418	3.024
Discos venuts	867.849	1.784.718
Lectors Instal·lats	544.588	1.290.162
Lectors venuts	381.190	766.324
Xifra de negocis en lectors*	204.967	409.934
Xifra de negocis en discos*	647.783	1.174.166

* en milers de dòlars

Font: *Infotecture*, 8 Nov. 1990, citant l'Associació Americana d'Editors de Discos Òptics.

Tot això ha remodelat el cicle de difusió de la informació (vegeu quadre 5). Els usuaris s'han anat traslladant des de la consulta de bases de dades en paper a la consulta en línia a causa dels avantatges que comporta la utilització dels operadors booleans quant a rapidesa, precisió i exhaustivitat. Amb l'aparició del CD-ROM, part dels usuaris de les versions en paper s'han traslladat directament al CD-ROM, igual que usuaris en línia que consulten algunes fonts de manera intensiva (p.e. grans biblioteques i centres de documentació d'empreses). L'accés en línia es manté per a fonts que tenen un ús mitjà o baix i per a aquelles fonts que, pel grau d'actualització, són especialment indicades per a aquest mitjà. La consulta del CD-ROM no depèn del bon o mal funcionament de les telecomunicacions, pot ser utilitzat més fàcilment per l'usuari final, ja que no es paga per temps, i s'adquireix com una publicació impresa per quantitat d'informació i no consulta a consulta.

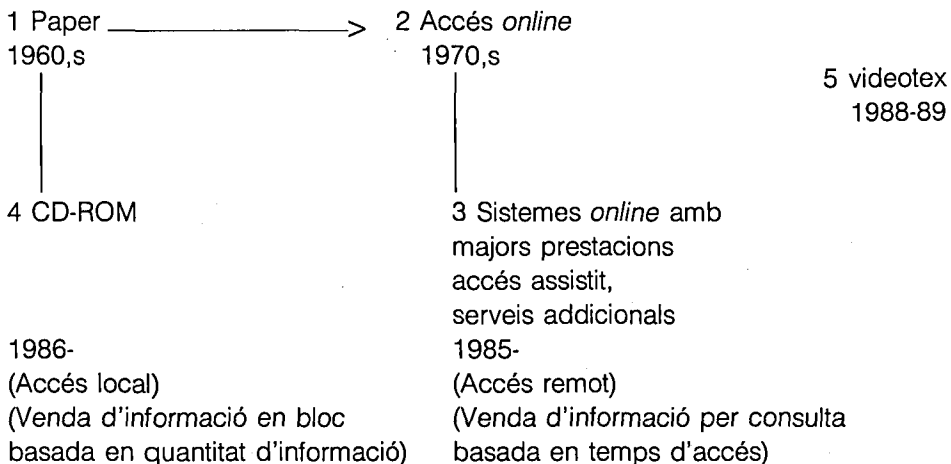
Per tal de frenar aquesta migració i també per a ampliar el nombre d'usuaris entre els no experts, els distribuïdors instal·len sistemes d'interrogació assistida, presten serveis addicionals, etc.

Des del punt de vista de les telecomunicacions, a Europa, les empreses que se n'encarreguen han desenvolupat de manera important, encara que desigual segons els països, xarxes videotex pensades per a la llar, l'esbarjo, la comunicació i les transaccions, és a dir per a un públic molt més ampli. Progressivament, les aplicacions professionals s'hi han anat obrint camí i competeixen amb la distribució clàssica o 'ascii'. A França hi ha accessibles per minitel bases tan importants com Kompass, Pascal o Medline. L'ús del terminal i dels serveis videotex és senzill, i la facturació es fa mitjançant el rebut telefònic (servei quiosc), a preus per minut diferents d'acord amb el valor dels serveis oferts, els quals s'agrupen segons vies d'entrada (números als quals truca l'usuari) a la xarxa videotex. El distribuïdor d'informació rep els ingressos de la companyia de telèfons. Tot i ésser

el videotex un sistema més car i menys eficient per a consultar informació especialitzada, resulta curiós comprovar que els avantatges de poder utilitzar les bases de dades amb un procediment senzill, conegut, sense entrebancs referents a contractes, pagament de factures, etc, units al fet que el cobrament corresponent a la consulta d'informació és dut a terme per la companyia de telèfons en el rebut normal, determinen un consum important d'informació. Es posa de manifest d'aquesta manera la importància del marketing en matèria de serveis d'informació.

Així doncs, la distribució de bases de dades en línia es veu influïda per l'ús del CD-ROM en general, i en els països europeus on el videotex té importància, aquest també pressiona sobre el mercat de les bases de dades professionals.

Quadre 5. Evolució dels procediments de difusió de la informació especialitzada



Segons G. Anderla (3), als Estats Units el CD-ROM comptaria actualment amb una cinquena part del mercat i l'en línia amb el 80%. A França, la situació seria diferent: l'any 1989, el CD-ROM podria haver tingut cap al 6%, i la repartició del mercat entre serveis en línia tradicionals i el minitel hauria passat d'una relació del 90% al 7% (l'any 1985) a una altra del 54% al 40%. Segons Anderla, recuperar fins a 5 referències a una base de dades com Pascal en un distribuïdor convencional costa menys de 2 minuts i té un preu d'aproximadament 5 dòlars, mentre que per mitjà del minitel requereix 10 minuts i té un preu de 19 dòlars.

reix accés a informació professional, tant en 'ascii' com en videotex, espanyola i internacional. Això s'ofereix junt amb serveis de comunicació: conferència i correu electrònic.

En el camp de la informació econòmica cal destacar el nou intent d'EFE per a oferir serveis en temps real a Espanya en competència amb els grans d'aquest sector com son Reuter i Telerate.

Sembla que hi ha un marge entre tots aquests serveis i vies de distribució de la informació per a un distribuïdor 'ascii' tradicional, sempre i quan es tingui en compte aquest nou context que s'està creant. Un distribuïdor d'aquesta naturalesa comptaria amb una oferta d'informació i amb una demanda potencial suficients. Milloraria la situació actual, ja que reduiria els costos de distribució per separat que actualment s'estan produint (amb equips de promoció, de formació, informàtics, etc, diferents). Significaria una oferta més atractiva per al públic, cosa que podria fer augmentar la demanda i en conseqüència els ingressos dels productors. Addicionalment en diferenciar-se els papers de productor i de distribuïdor de bases, es podrien millorar tant les mateixes bases com els serveis oferts. Altres beneficis d'ordre general serien: l'impacte sobre els hàbits d'ús de la informació per part dels usuaris, l'estímul a la creació de bases de dades i la popularització dels sistemes telemàtics. A fi de no patir la fragmentació dels serveis 'ascii' europeus, caldria assegurar l'articulació d'aquest distribuïdor a nivell europeu.

La situació a Catalunya

Hem mencionat, a l'inici, el paper de les bases de dades bibliogràfiques generades pels serveis o agències governamentals, que van permetre la consolidació dels distribuïdors Dialog i SDC.

De manera anàloga els serveis de l'Administració central han donat lloc als distribuïdors, els quals es concentren a Madrid. Respecte al CD-ROM, la situació és semblant, tant pel que fa als productes desenvolupats com a les empreses que hi treballen.

L'Administració catalana està generant també bases de dades de manera similar. A Catalunya hi ha, però addicionalment, iniciatives d'institucions, en general sense ànim de lucre, que han desenvolupat petites bases de dades: d'educació, de disseny, sobre temes de construcció, de temes polítics, de terminologia, del sector electrònic, de medicaments, etc. Aquestes bases no tenen sortida adequada cap al públic, la qual cosa perjudica el seu desenvolupament.

Per altra banda, el videotex s'està desenvolupant molt bé a Catalunya i és un sector on la iniciativa privada pot cercar serveis rendibles ja que l'equipament de

hard i soft és més assequible i les oportunitats de generar demanda són més favorables per a una gamma d'informacions de consulta senzilla, que és la més indicada per al videotex. Segons fonts de Telefónica, l'ús del videotex, si es descompten els serveis de les pàgines blanques, grogues i blaves, té lloc fonamentalment a Madrid i Barcelona en parts similars. Al gener i febrer d'enguany el nombre d'hores de consulta a la província de Barcelona va ser de 20.000 hores, i a tot a Espanya va ascendir a 82.000 hores durant el mes de març. Abadal (1) situa en un 31% el percentatge de centres servidors catalans sobre el total espanyol (i més del 50% en relació amb les bases de dades disponibles). Aquest percentatge, segons la informació que apareix a la *Guía Videotex* (9), es mantindria actualment. La informació més especialitzada tindria, segons fonts de Telefónica, una consulta significativa.

Comentaris finals

El punt bàsic que possibilita l'existència d'una indústria de la informació és la disponibilitat de bones fonts. Sense això no es pot parlar de distribució d'informació. Cal que s'incentivi la producció de bases de dades en tots els mitjans de difusió d'informació avui existents.

Paral·lelament cal promoure l'ús dels mitjans de consulta en suport electrònic. Un bon exemple d'això és la xarxa de valor afegit Spritel sorgida al País Basc.

Respecte a la situació de la distribució de bases de dades especialitzades, es pot detectar una manca de vertebració de l'oferta d'informació a escala espanyola. La unificació d'aquesta oferta pot ser viable si s'articula convenientment dins el mercat europeu.

Abril 1991

Bibliografia

- (1) E. Abadal i M.J. Recoder, «La Industria de las Bases de Datos en Catalunya. Perspectivas para la década de los 90.» *Terceras Jornadas Españolas de Documentación Automatizada*, Palma de Mallorca, 1990, 433-449.
- (2) Concha Alvaro, «La Distribución de Bases de Datos en España. Situación Actual y Perspectivas de Futuro.» *Terceras Jornadas Españolas de Documentación Automatizada*, Palma de Mallorca, 1990, 519-533.
- (3) Georges Anderla, «The Minitel Challenges Must Make Online Hosts Reassess.» *Information World Review*, January 1991, p. 10.
- (4) Tomàs Baiget, *Estudio y Anteproyecto de Distribución de Bases de Datos en España*, CIDC, Barcelona, 1989, Parte 1, 207p.

- (5) Tomàs Baiget i Vicenç Meléndez, «Las Bases de Datos al Servicio de la Empresa. Situación Actual y Perspectivas.», *Jornadas Internacionales sobre Información Electrónica y Empresa*, San Sebastián, 1990, Ponencia XV, 6 p.
- (6) Germán Beascochea, «Diseño de los Servicios Videotex: la Búsqueda de Fuentes de Información con Contenidos de Utilidad.», *Jornadas Internacionales sobre Información Electrónica y Empresa*, San Sebastián, 1990, Ponencia XXI, 9 p.
- (7) *Digital Information Group, Information Industry Factbook. The Information Industry's Annual Report*. 1989/90 Edition, Stamford, CT: Digital Information Group, 1989, 532 p.
- (8) *Infotecture*, 15 febrer 1991.
- (9) *Guía Videotex*, març 1991.
- (10) Institut d'Estadística de Catalunya. *Directori de Centres d'Accés a Bases de Dades*. Espanya. Barcelona: Institut d'Estadística de Catalunya, 1991 (en premsa).
- (11) Institut d'Estadística de Catalunya. *Estadística d'Ús de Bases de Dades a Espanya*. Barcelona: Institut d'Estadística de Catalunya, 1991 (en premsa).
- (12) Link, *The European Electronic Information Industry 1988-1992*. Link Resources, 1989, 90 p.
- (13) Vicenç Meléndez, «Avances Recientes y Tendencias Previsibles de Bases de Datos. Distribución de Bases de Datos.» *Encuentro sobre Bases de Datos de la Administración Pública*, Madrid, 1990, Consejo Superior de Informática.
- (14) *Monitor*, juny 1990.
- (15) Diane H. Smith, «Online Government Databases into the Maelstrom», *Database*, June 1988, 56-62.
- (16) Enric Teixidó, «Les Empreses de Serveis Videotex», *Universitat Tècnica d'Estiu de Catalunya*, Sitges, 1990, vol. 2, 1-15.