

DISSENY I INTERACCIÓ EN SISTEMES DE RECUPERACIÓ D'INFORMACIÓ

Mari Carmen Marcos

Secció de Ciències de la Documentació
Universitat Pompeu Fabra
(mcarmen.marcos@upf.edu)

Resum

Es pren com a marc la disciplina Interacció Persona-Ordinador (IPO) per presentar un estat de la qüestió sobre les interfícies d'usuari en sistemes de recuperació d'informació. S'hi estudia la interfície com un element clau de la interacció amb els sistemes informàtics, i es distingeixen dos moments principals del procés: la consulta a la base de dades i la presentació dels resultats. Es té en compte la diversitat d'usuaris pel que fa als seus coneixements previs sobre el sistema i la seva necessitat d'informació. S'hi presenten algunes tècniques d'accés a la informació, en què es destaquen el *browsing* o ullada, el *clustering* o conglomeració i les tècniques de visualització d'informació.

Paraules clau

Interacció persona-ordinador, Interfícies d'usuari, Sistemes de recuperació d'informació.

Resumen. *Diseño e interacción en sistemas de recuperación de información.*

Se toma como marco la disciplina Interacción Persona-Ordenador (IPO) para presentar un estado de la cuestión sobre las interfaces de usuario en sistemas de recuperación de información. Se estudia la interfaz como un elemento clave de la interacción con los sistemas informáticos, distinguiendo dos momentos principales del proceso: la consulta a la base de datos y la presentación de los resultados. Se atiende a la diversidad de usuarios en cuanto a sus conocimientos previos sobre el sistema y su necesidad de información. Se presentan algunas técnicas de acceso a la información, donde se destacan el *browsing* u ojeo, el *clustering* o agrupación y las técnicas de visualización de información.

Palabras clave

Interacción persona-ordenador, Interfaces de usuario, Sistemas de recuperación de información.

Abstract. *Design and interaction in information retrieval systems.*

Within the context of Human-Computer Interaction (HCI), this article reviews the state of the question regarding user interfaces in information retrieval systems. The interface is studied as a key element in interaction with computer systems, distinguishing two principal moments in the process: consulting the database and presenting results. Consideration is also given to user diversity in terms of previous knowledge of the system and need for information. Several techniques for accessing information are also presented, with particular attention to browsing, clustering and information visualisation techniques.

Keywords

Human-Computer Interaction, User interfaces, Information retrieval systems

Introducció

La recuperació d'informació es pot tractar des de diferents perspectives, que en la bibliografia sobre la matèria s'agrupen en dos corrents: l'algorítmic i el cognitiu. Aquestes denominacions fan paleses les diferents disciplines que interactuen i formen part de l'àmbit de la recuperació d'informació.

Tradicionalment, la recerca ha posat més èmfasi en l'aspecte algorítmic, tot i que el costat cognitiu, que tracta sobre el comportament de les persones en els processos de recuperació d'informació, també està tractat en la bibliografia. Els estudis sobre aquest tema demostren que els usuaris passen per diferents estats segons que van organitzant i articulant la seva necessitat d'informació (Marchionini, Komlodi, 1998). Alguns treballs en aquesta línia publicats en espanyol són el d'Altuna (1992), que fa un estat de la qüestió amb relació als OPAC, i el de Frías i Martín (1999), que presenten l'anàlisi transaccional com a tècnica per a la recollida de dades dels usuaris a través dels quals se n'estudia el comportament.

Una de les disciplines que en els últims anys ha agafat rellevància en aquest àmbit és l'anomenada Human-Computer Interaction (HCI), que en català es coneix amb el nom d'Interacció Persona-Ordinador (IPO).

Probablement, la recuperació d'informació és l'àrea de la documentació que es pot beneficiar més dels estudis sobre IPO per oferir als usuaris més bones interfícies que en facilitin la feina. El disseny de nous sistemes de recuperació d'informació passa per investigar noves tècniques que superin els inconvenients detectats. Això no obstant, són poques les recerques que s'han dut a terme des d'aquesta perspectiva.

La manera en què es presenta a l'usuari la informació en un sistema de recuperació d'informació (SRI) en fa variar la manera d'interactuar. En aquest sentit, s'estan investigant noves presentacions superiors a la consulta i la recuperació tradicional, i que permetin a l'usuari obtenir una imatge del sistema a partir de la qual localitzar el seu tema d'interès. Els estudis tendeixen cap a les presentacions que suporten l'hipertext i els mapes visuals de representació d'informació.

La disciplina Interacció Persona-Ordinador, tot i que està molt establerta, encara és poc utilitzada en l'àmbit de la Biblioteconomia i la Documentació, tal com fa palès el baix nombre de referències bibliogràfiques que hi ha sobre això a la nostra àrea. No volem pas dir amb això que no hi hagi un interès per part de la recerca en documentació sobre els aspectes d'interacció, sinó més aviat que aquest interès encara no s'ha desenvolupat d'una manera tan sistematitzada com en altres apartats.

En aquest article es parteix de la Interacció Persona-Ordinador, d'on es prenen els principis ja establerts en relació amb les interfícies d'usuari, per aplicar-los a la resolució del problema de la consulta a sistemes de recuperació d'informació. Amb això pretenem contribuir al desenvolupament d'estudis d'Interacció Persona-Ordinador en l'àmbit de la documentació, ja que ajudarà a avançar en el desenvolupament de sistemes més adaptats als usuaris.

1. El marc teòric de la interacció de les persones i els ordinadors: la IPO

1.1. Concepte i interdisciplinarietat de la IPO

La Interacció Persona-Ordinador (IPO) és la conjunció interdisciplinària de diverses ciències i tecnologies convertida avui en una disciplina per ella mateixa, atesa la seva trajectòria i el seu assentament en la recerca en les últimes dècades.

D'entre les diferents definicions que se n'han anat donant, en proposem una que considerem que en condensa d'altres de donades amb anterioritat per autors com Booth (1989), Marchionini (1995) o Shneiderman (1992): la IPO s'ocupa d'estudiar la creació de productes informàtics que ajudin en la realització de tasques als seus usuaris tenint en compte la facilitat d'ús, el temps d'execució, que evitin possibles errors i que, en conseqüència, els satisfacin.

Podem resumir els punts d'interès de la IPO seguint Booth (1989), el qual apunta que aquesta disciplina s'ocupa d'estudiar sobretot:

- El *hardware* i el *software* dels ordinadors i com intervenen en la interacció que es produeix quan les persones els fan servir.
- Els models mentals de les persones i el seu acostament al model mental reflectit en l'ordinador, que correspon al del programador.
- Les tasques per a les quals està preparat l'ordinador i com aquestes s'adapten a les necessitats dels seus usuaris.
- El disseny, centrat en les necessitats dels usuaris.
- L'impacte que produeix la presència i l'ús d'ordinadors en entorns socials i laborals.

Cal no oblidar que la IPO com a disciplina sorgeix del desenvolupament d'altres de molt establertes en la ciència, que es poden dividir en dos grups: aquelles que se centren en l'estudi de les persones que utilitzen els ordinadors, i les que s'ocupen dels factors tecnològics, és a dir, els que afecten els ordinadors. La recerca conjunta des d'aquests punts de vista ha portat a desenvolupar la disciplina que ara ens ocupa. Vegem quines són les àrees que influeixen de manera més directa en la IPO (la figura 1 presenta un esquema fet pel grup d'interès per a la interacció persona-ordinador SIGCHI, *Special Interest Group for Computer-Human Interaction*).

La filosofia, la psicologia, la fisiologia, la medicina i l'ergonomia —anomenada així a Europa i factors humans als Estats Units— són disciplines que ajuden a entendre millor la part humana de la IPO i, per tant, a dirigir el disseny dels sistemes cap als seus usuaris.

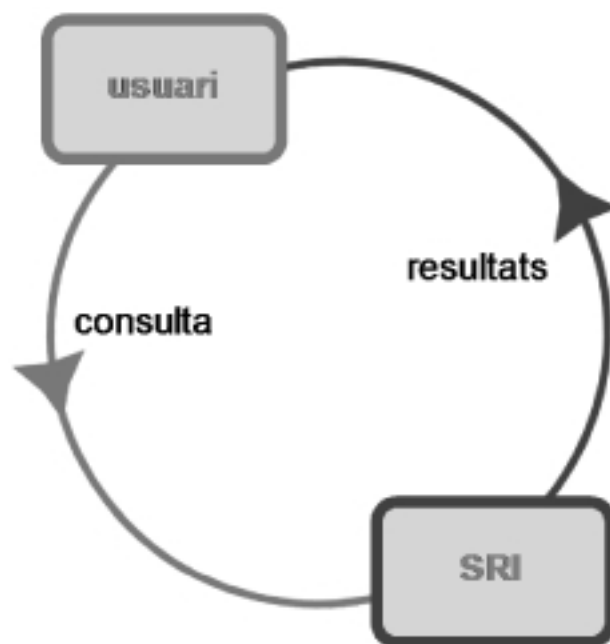


Figura 1. Elements de la IPO, pres de sigchi.org/cdg/cdg2.html

De totes aquestes ciències i disciplines, la més recent és l'ergonomia. Va néixer al segle xx molt vinculada a l'enginyeria i a la indústria, ja que mirava de dissenyar màquines de manera que fossin més còmodes de fer servir per part dels seus operaris; durant la Segona Guerra Mundial es va aplicar a l'àmbit militar i, al cap de poc temps, es va començar a aplicar en el civil. La ciència que avui aplica aquests coneixements és la informàtica.

A la taula 1 es recull una tria feta pel Grup d'Interès Especial sobre Interacció Persona-Ordinador (SIGCHI) de l'*American of Computing Machinery* (ACM) d'organismes d'interès internacional per a la IPO, com també algunes revistes especialitzades en aquest tema i els congressos més importants que se celebren en aquest àmbit.

TAULA 1. PRINCIPALS ORGANITZACIONS, REVISTES I CONGRESSOS SOBRE IPO.

DADES PRESES DE sigchi.org/cdg/cdg1.html, sigchi.org/cdg/cdgA.html

ORGANISMES INTERNACIONALS I NACIONALS PER A LA IPO

The ACM Special Interest Group on Computer Human Interaction

The British Computer Society Specialist Group on HCI

The European Association for Cognitive Ergonomics

The Human Factors Society Computer Systems Technical Group

The IFIP Technical Committee (TC 13) on Human-Computer Interaction

REVISTES ESPECIALITZADES EN IPO

ACM transactions on information systems (1983-)

Behavior and information technology (1983-)

Human factors (1959-)

Human-computer interaction (1983-)

Interacting with computers (1989-)

International journal of human-computer interaction (1989-)

International journal of man-machine studies (1969-)

SIGCHI bulletin (1982-)

CONGRESSOS ESPECIALITZATS EN IPO

ACM CHI Human Factors in Computing Systems Conference (1983-)

ACM User Interface Software Technology Conference (1988-)

Annual Meeting of the Human Factors Society (1957-)

BCS HCI SG Human-Computer Interaction Conference (1985-)

Conference on Computer Supported Cooperative Work (1986-)

European Conference on Cognitive Ergonomics

Hypertext Conference (1987-)

IFIP INTERACT Human Factors in Computing Conference (1984-)

International Conference on Human-Computer Interaction (1985-)

1.2. Recerca en IPO

Per la bibliografia publicada en els últims anys, s'observa que els estudis se centren especialment en aquestes àrees de recerca:

- Avenços en la teoria de la IPO. Tant des del punt de vista de la psicologia i de la sociologia com des de l'ergonomia, el disseny i la informàtica gràfica, la recerca intenta aconseguir models basats en experiències prèvies que serveixin, al seu torn, per desenvolupar recerques futures. Com tota disciplina científica, la IPO s'ocupa —tot i que no d'una manera tan explícita com altres matèries— de conformar aquella base teòrica necessària per establir paradigmes científics.
- Disseny de la informació en pantalla. S'hi estudien aspectes relatius al nombre d'elements que es presenten, la mida, la disposició, els colors utilitzats, la combinació de text, les imatges, els sons i els vídeos. Són aspectes d'ergonomia visual i de llegibilitat. L'ergonomia es tracta des dels aspectes humans més fisiològics (postures de l'esquelet en utilitzar l'ordinador) fins als relacionats amb l'entorn mediambiental (soroll, llum, temperatura) i organitzacional (posició de l'usuari dins de l'empresa, habilitat i experiència en el maneig d'ordinadors, satisfacció o insatisfacció amb les noves tecnologies). Es tenen en compte els usuaris amb necessitats especials en la interacció amb les màquines (infants, gent gran i discapacitats).
- Sistemes per a treball en grup i cooperatiu (CSCW). S'estudien els factors socials i organitzacionals del treball amb ordinadors. Hi ha força bibliografia sobre aquest tema; per exemple, Baecker (1995, capítol 11) o Grudin (1988, 1994). Pel que fa a aquest tema, Shakel (1997) creu que caldrà millorar els factors humans i els aspectes que afecten el treball en organitzacions, i afirma que cal treballar amb especialistes en aspectes socials i organitzacionals per integrar aquestes àrees en el disseny de sistemes d'informació. Portant aquestes idees a l'àmbit de la documentació i, més concretament, a un dels aspectes en què millor s'aprecien els avantatges de la col·laboració entre persones, trobem la recerca en catàlegs de biblioteca. En aquest sentit, Nichols (1998) considera que la recerca d'informació forma part de les activitats de les persones en què s'acostuma a donar una interacció amb altres companys. Entre aquestes interaccions hi ha la recomanació d'ítems interessants, compartir mètodes o tàctiques de recerca i explicacions informals sobre com utilitzar un sistema en particular.
- Formes d'organització de la informació. L'hipertext com a forma d'organització de la informació ha rebut un gran impuls després de la seva implantació de manera massiva en els documents web. Shakel (1997) afirma que a Internet caldrà que hi hagi una organització millor de la informació per fer augmentar la facilitat en la recuperació d'informació rellevant a la recerca. Les metadades, les ontologies i la prometedora web semàntica són el pròxim repte per a l'organització d'informació a la xarxa.
- Presentació de la informació en sistemes de realitat virtual i realitat augmentada. Les xarxes neuronals per mitjà de sistemes d'intel·ligència artificial i sistemes experts oferiran moltes possibilitats al desenvolupament d'aquests sistemes.
- L'admissió de la veu com a dispositiu d'entrada i sortida d'informació i els gestos com a dispositiu d'entrada facilitaran la interacció de les persones amb els sistemes informàtics.

1.3. La IPO a Espanya

La trajectòria de la interacció persona-ordinador a Espanya és relativament recent. Aquest fet es reflecteix en la poca recerca que s'havia fet fins fa poc sobre aquest tema i, en conseqüència, en la poca publicació que hi havia respecte a això en el nostre país. En els últims anys, el panorama ha canviat i avui dia és habitual trobar-la com a assignatura en els plans d'estudis de titulacions universitàries de les enginyeries informàtiques, per exemple, a Sevilla, Granada, Carlos III de Madrid, Ciudad Real, Albacete, València, Saragossa, Lleida, el País Basc, Astúries i Vigo (González, 2001; Lorés, 2002), i també en altres estudis com els de psicologia i fins i tot documentació, com és el cas del segon cicle d'aquests estudis a la Universitat Oberta de Catalunya.

Al mateix temps s'ha impulsat la bibliografia en espanyol en la qual es presenten avenços en recerca sobre IPO. Gràcies a la unió amb els diferents departaments universitaris que inclouen aquests estudis, la IPO s'ha fet un forat en la recerca del nostre país i avui ja en podem veure grans projectes, com és el cas del llibre digital *Introducción a la interacción persona-ordenador* (Lorés 2001, 2002; Abascal 2000), en el qual participen més d'una desena d'universitats espanyoles per dotar-lo de continguts. Cal dir que en vuit mesos ha superat les 13.000 visites i que és dels pocs materials en espanyol que indexa l'ACM. Aquest material docent es fa servir en diverses universitats per fer classe.

Al final del 1999 es va constituir l'Asociación para la Interacción Persona-Ordenador (AIPO) a l'Escuela Técnica Superior de Informática de la Universidad Autónoma de Madrid, amb l'objectiu de promoure i difondre la IPO, com també organitzar actes, jornades científicotècniques i altres tipus d'activitats relacionades amb aquesta i, paral·lelament, servir de vincle entre els professionals. L'AIPO es va constituir com una organització oberta a la comunitat universitària d'Espanya i Llatinoamèrica interessada en la IPO. De moment ha celebrat tres trobades nacionals en què s'han reunit experts en informàtica gràfica, psicologia cognitiva, sociologia, disseny, ergonomia, etc., i s'està preparant el quart.

Un altre projecte en marxa és el programa de doctorat interuniversitari en IPO que s'està construint amb la col·laboració d'aquestes universitats i que espera tenir una gran acollida en un camp que fins ara no oferia estudis de tercer cicle.

Des del febrer del 2003 disposa del domini www.aipo.es. Des d'aquí s'accedeix a les trobades anuals de l'AIPO, al llibre digital, a la informació sobre el doctorat, a enllaços d'interès, bibliografia i altra informació sobre interacció de les persones i els ordinadors.

Pel que fa a la docència d'IPO, Espanya segueix de prop les tasques que duu a terme el grup de treball ACM/IEEE-CS Joint Task Force on Computing Curricula. Aquest grup va redactar el 1991 un informe en el qual s'establien nou àrees temàtiques per cobrir la matèria de la disciplina d'informàtica, dins les quals hi havia la interacció persona-ordinador. Es continua treballant en aquesta guia curricular; de fet n'hi ha un esborrany del març del 2001 que manté aquesta disciplina com una de les que cal impartir en els estudis d'informàtica. L'informe es pot consultar a www.computer.org/education/cc2001/report.

2. La interfície d'usuari com a element clau de la interacció

2.1. La importància del disseny d'interfícies

En la nostra vida diària constantment interactuem amb persones i objectes. La manera en què ho fem amb aquests últims depèn de la nostra experiència passada amb aquelles mateixes coses o amb d'altres de semblants. Quan no coneixem un objecte, n'intuïm el funcionament tot

relacionant-ne l'aspecte amb el d'objectes que ja coneixem; si s'esdevé alguna cosa inesperada, pot donar lloc a una situació d'humor (de fet és un recurs molt utilitzat en la comèdia), però si passa quan estem treballant amb l'ordinador, el sentiment més habitual és el de frustració i empipament.

Don Norman (1988), en el seu llibre *Psychology of everyday things*, descriu situacions diàries en les quals es dona algun problema d'interacció a causa del mal disseny dels objectes. Amb relació a aquesta obra, cal destacar el canvi de títol que l'autor en va voler fer per a l'edició en rústica, que va ser *Design of everyday things* i que va decidir per una qüestió d'adaptació als lectors: el títol original s'adreçava a la comunitat investigadora, mentre que l'altre n'afavoria la venda de caràcter divulgatiu. Amb això volem assenyalar que Norman adapta la presentació del llibre als dos tipus de lectors, i això mateix ja és una lliçó: quan es tracta de disseny, cal adaptar el producte als usuaris potencials.

Les persones utilitzem els ordinadors per dur a terme tasques que altrament serien difícils d'executar; hi cerquem fer augmentar la productivitat, la precisió i la satisfacció en la nostra feina. Si fer tasques amb l'ordinador ha de ser complicat, no servirà de res recórrer-hi. Com que no tothom espera el mateix d'un sistema, caldrà crear un software flexible que beneficiï tothom; mentre que uns usuaris demanen a un editor de textos que tingui les funcions bàsiques, d'altres prefereixen una interfície gràfica amb multitud d'elements visuals que als primers els pot produir distracció. Per això, no hi pot haver un estil d'interfície òptim per a totes les persones i totes les tasques, però sí diferents tècniques que s'adaptin als diferents usuaris i treballs. El més adequat és que un producte pugui ser utilitzat per usuaris tant novells com freqüents i experts, cada un dins del seu nivell, i això es pot aconseguir conjugant diferents estils.

La interfície és la part del sistema a la qual l'usuari accedeix i, per tant, la que li donarà les claus necessàries perquè construeixi la seva representació mental del sistema, és a dir, el model mental. El seu disseny s'ha de plantejar amb tota cura a fi que el model que creï l'usuari sigui el correcte, és a dir, el mateix que el que té el creador del sistema. Aquesta és una de les raons, juntament amb d'altres com la llegibilitat, la usabilitat o el màrqueting, per les quals el disseny d'interfícies d'usuari esdevé una de les àrees amb més desenvolupament dins de la informàtica, si bé altres àrees del coneixement, com la psicologia, l'ergonomia i l'enginyeria de disseny, contribueixen des dels seus punts de vista a millorar el disseny de les interfícies de sistemes informàtics.

La interacció amb l'ordinador és similar a la que s'efectua amb la resta dels objectes que ens envolten. En aquesta, concretament, recorrem a les nostres experiències anteriors amb els elements del hardware, amb els sistemes operatius i amb els programes, i les conjuguem amb les nostres expectacions. El dissenyador d'interfícies ha de conèixer bé els usuaris del sistema que dissenya, les tasques que voldran dur a terme i la seva manera de fer servir l'ordinador. Aquesta interfície ha de donar a l'usuari les claus per conèixer la funció dels elements que la formen i li ha de permetre determinar-ne el comportament.

Seguint Laurel (1990), no hem de veure l'ordinador com una màquina que calcula, sinó com una màquina amb capacitat per representar accions en què els homes puguin participar. Per tant, el disseny d'una bona interfície ha d'aconseguir un bon equilibri entre «pensar» i «fer». Laurel (1990) vol agafar el concepte de conversa —el que anomenem interacció o diàleg entre l'home i la màquina— i deixar de banda la vella idea que qualsevol acció de l'home serà resposta per l'ordinador. Per això —i hi estem d'acord—, cal que hi hagi una coordinació, cal compartir una informació que permeti la conversa. En l'entorn persona-ordinador, aquest *common ground* és un espai en què els significats prenen forma a través de la col·laboració i l'aproximació successiva dels participants.

Segons Mayhew (1992), l'objectiu dels dissenyadors d'interfícies és facilitar a l'usuari el desenvolupament d'un model mental efectiu. Una manera molt habitual per aconseguir aquesta finalitat és fer servir metàfores. La metàfora d'ordinador queda descrita per Baecker (1995) amb la utilitat d'ajudar l'usuari a entendre un nou domini tot permetent-li manejar-lo amb els coneixements previs assimilats. Una de les més utilitzades en les interfícies gràfiques d'usuari (GUI) és la de l'escriptori, construïda a imatge de la taula de treball de l'usuari: documents, carpetes, maletí, paperera...

Com es pot desprendre del que s'ha dit fins ara, el fet de determinar la usabilitat d'un sistema serà una tasca difícil, ja que no es poden fixar unes pautes rígides, sinó que varien segons cada usuari, cada tasca que s'executi i les circumstàncies específiques en què això s'esdevé. La norma ISO 9241, a la seva part 11, explica els beneficis de mesurar la usabilitat en termes d'efectivitat (objectius que l'usuari pot assolir), eficiència (recursos gastats per assolir aquests objectius) i satisfacció (nivell amb què l'usuari considera que el producte és acceptable).

2.2. Concepte d'interfície

El diccionari *Webster's ninth new collegiate*, tal com recull Church (1999, p. 6), defineix interfície com «una superfície que forma un límit comú entre dos cossos, espais o fases; un lloc en què es troben sistemes independents i actuen i interactuen entre ells; o el mecanisme que fa possible la interacció o la comunicació».

La interfície queda definida per Baecker (1995) de manera breu com els dispositius d'entrada i sortida i el software que porten relacionat, i de manera àmplia com tot allò que conforma l'experiència de l'usuari amb ordinadors, incloent-hi la documentació, la formació i la persona. Dit d'una altra manera, és el lloc en què es troben, actuen i es comuniquen sistemes independents entre ells. La manera més senzilla d'explicar-ho seria dir que una interfície és el que l'usuari veu a la pantalla —si no es tenen en compte altres modes no visuals d'informació.

Alberico i Micco (1990) es refereixen a la interfície com a tot allò amb què interactua la persona en fer servir un sistema informàtic (software i hardware). El seu desenvolupament requereix coneixements de psicologia, enginyeria i programació, entre d'altres. Una interfície pot ser un grup de programes distribuïts entre diversos ordinadors; una part pot estar en un ordinador local, una altra, a l'ordinador remot que faci de servidor i una tercera, en una xarxa.

En la mateixa línia, Card, Moran i Newell (1983, p. 4) diuen que la comunicació entre una persona i un ordinador es dona en la interfície, i que aquesta està formada de totes les parts involucrades en aquesta comunicació, «els dispositius físics, com ara el teclat o la pantalla, com també els programes de l'ordinador per controlar la interacció».

Així mateix, Mandel (1997, p. 14) considera que la interfície «inclou tant el *hardware* de l'ordinador com el software que presenta informació a l'usuari i li permet interactuar amb la informació i l'ordinador». En aquesta definició, l'autor inclou els dispositius físics d'entrada i sortida, els elements sobre els quals actua l'usuari (punts on dur el ratolí, parts de la pantalla on tocar...), la informació amb la qual treballen els usuaris i la que els ajuda a manejar el sistema i l'ordinador.

Marchionini (1995, p. 38) planteja la definició des del punt de vista de la recuperació d'informació i es refereix a la interfície com «les representacions del coneixement en una base de dades i les eines, les regles i els mecanismes per accedir-hi i manipular-les [...]. Serveix com a intermediari entre l'usuari i la base de dades». Continua l'autor dient que «una interfície ha de proveir relacions robustes entre el contingut de la base de dades i les representacions conceptuals que manipulen els qui volen trobar informació».

Basant-se en aquesta última observació de Marchionini, Church (1999) determina que la interacció no es produeix entre l'home i la màquina, sinó entre l'home i la informació que aquesta forneix, per la qual cosa proposa canviar l'expressió d'interfície home-màquina o persona-ordinador per la d'interfície home-informació (*human-information interface*), que es definiria com «el límit o l'intermediari entre un usuari humà i la informació que cerca; una superfície o un límit en què un usuari contacta, interactua o es comunica amb les fonts d'informació per obtenir informació» (p. 19). Aquesta visió ens resultarà especialment útil en un entorn de recuperació d'informació, en què l'objectiu es troba en la informació, no pas en la màquina per ella mateixa. A més, aquesta nova expressió és més àmplia, ja que no implica que la interfície hagi de ser necessàriament un sistema informàtic.

Seguint amb les idees de Laurel (1990), considerem la interfície com un context compartit per a una acció en què tots dos —home i ordinador— són agents (definim agents com a «iniciadors d'una acció», seguint la *Poètica* d'Aristòtil). Quan no es dona aquest fons conceptual comú, apareixen missatges d'error i resultats inesperats perquè no hi ha hagut comunicació. Per resoldre aquests problemes tan habituals sorgeixen les metàfores en les interfícies, la missió de les quals és oferir a les persones un esquema conceptual que els resulti familiar. Però ni tan sols les millors metàfores funcionen sempre, ja que cada persona pot interpretar les diferents formes segons quin en sigui el referent en el món real. Tot i així, avui dia estan molt esteses a les interfícies gràfiques.

Laurel vol anar més enllà en el concepte de disseny d'interfícies i apunta la importància de la imaginació, ja que es tracta de representar alguna cosa que no existeix en el món real. En una representació a través d'un ordinador, l'usuari hi participa; no és el mateix que en la vida real, però sí que té efectes o conseqüències que s'esdevindrien en la realitat. La frase següent resumeix el concepte que té Laurel sobre les interfícies home-màquina: «Pensar sols en interfícies és quedar-se curt, dissenyar experiències home-màquina no és construir un escriptori millor. És crear mons imaginaris que tinguin una relació especial amb la realitat, mons en els quals puguem estendre, ampliar i enriquir la nostra capacitat de pensar, sentir i actuar» (Laurel, 1990, p. 33).

3. La interfície en els sistemes de recuperació d'informació (RI)

3.1. Dissenyar per a l'usuari

Podem dividir els tipus d'usuaris d'acord amb dues variables relacionades amb el seu coneixement previ:

- la concreció de la seva necessitat d'informació
- la seva experiència en el maneig del sistema de RI

Així, podem establir quatre tipus generals d'usuaris (vegeu la taula 2):

Els usuaris de tipus A parteixen d'una situació privilegiada, ja que coneixen amb claredat la seva necessitat d'informació, estan familiaritzats amb la col·lecció i tenen un bon maneig de les opcions de recerca que permet el sistema. Per a aquests usuaris (entre els quals es poden trobar bibliotecaris del mateix centre o d'altres centres), cal proveir mètodes ràpids de recerca que els permetin anar directament als documents objecte de recerca, tot i que no serà l'usuari majoritari del sistema.

TAULA 2. TIPOLOGIA D'USUARIS

	AMB NECESSITAT D'INFORMACIÓ BEN DEFINIDA	AMB NECESSITAT D'INFORMACIÓ POC DEFINIDA
Expert en el maneig de l'SRI	A	B
Novell en el maneig de l'SRI	C	D

Els usuaris de tipus B també coneixen bé la col·lecció i la forma de consulta de la base de dades, però no tant els objectius de la seva recerca. Serien usuaris habituals que cerquen informació sobre un tema determinat, però no estan segurs de quins documents els podrien interessar, potser perquè tenen poc coneixement de la matèria. Per a ells cal habilitar una funció de recerca similar al «passeig» per les prestatgeries, és a dir, una opció d'«ullada» temàtica que els eviti plantejar una consulta i, en canvi, els doni com a resultat documents propers a la seva necessitat d'informació. En aquests casos en què l'usuari no coneix amb exactitud què necessita trobar, anirà identificant entre els documents que obtingui aquells que consideri d'interès.

Els usuaris de tipus C i D no estan gaire familiaritzats amb la base de dades que faran servir, cosa que vol dir que no coneixen bé el sistema i encara menys el llenguatge de consulta que utilitza el sistema. Els de tipus C com a mínim tenen una idea clara de la seva necessitat d'informació, però tenen el problema de no saber formular la consulta al sistema. Per a ells cal habilitar mètodes que els permetin arribar als documents d'una manera directa però sense passar per una consulta que impliqui el maneig d'ordinadors. Una interrogació senzilla amb un o dos termes de recerca els podria donar un primer resultat sobre el qual després afinar.

Els usuaris de tipus D probablement han arribat al sistema d'una manera casual i voldran fer un cop d'ull al que conté, per si els pot interessar. No s'ha de deixar de banda aquest usuari potencial, sinó cobrir-ne la necessitat d'informació de la manera més efectiva possible. La manera idònia de donar-li a conèixer la col·lecció és oferir-li una visió general d'aquesta i permetre-li navegar pels diferents temes a fi que arribi a documents del seu interès. Com que no coneix el sistema de recerca, la ullada, també anomenada *browsing*, serà la millor opció per a aquest usuari casual, ja que no se li pot exigir que plantegi els termes de recerca en la seva consulta.

3.2. Tècniques d'accés a la informació

3.2.1. Ullada o *browsing*

La tècnica anomenada *browsing* en l'àmbit anglosaxó —que aquí traduïm com a ullada— apareix com a alternativa al tradicional *querying* (la consulta per interrogació) utilitzat en els sistemes de RI. És un mètode d'accés a la informació que consisteix a revisar amb la vista un espai amb el propòsit de reconèixer-hi objectes. Es pot dur a terme en espais d'una dimensió de forma seqüencial (per exemple, en una llista), o pot tenir lloc en un context estructurat que conté relacions jeràrquiques (per exemple en forma d'arbre) o bé semàntiques o associatives (per exemple, en forma de mapa). Es pot fer tant en contextos analògics (seria el cas de revisar les prestatgeries d'una biblioteca, per exemple) com en contextos digitals (podria ser en el si de documents d'un processador de textos, en presentacions multimèdia, a la web, etc.).

Lin (1997, p. 41) ho defineix com «un procés interactiu en el qual es pot visualitzar grans quantitats d'informació, percebre o trobar estructures o relacions, i seleccionar ítems tot

centrant-hi l'atenció visual». En el seu treball, Xia Lin, a més d'explicar en què consisteix aquesta tècnica perceptual, recull les utilitats que altres autors han cregut oportunes en l'àmbit de la recuperació d'informació:

- Thompson i Croft (1989) veuen el *browsing* com una bona estructura organitzacional, ja que els ítems relacionats estan els uns a prop d'altres.
- Motro (1986) i Marchionini (1987) creuen que és d'utilitat especial quan els usuaris no estan familiaritzats amb el contingut de la col·lecció.
- Belkin, Oddy i Brooks (1982a i 1982b) diuen que resulta útil als qui tenen dificultat per expressar la seva necessitat d'informació.
- Bates (1986) continua en aquesta línia opinant que resulta millor per als usuaris que cerquen un tipus d'informació que és més fàcil reconèixer que descriure.

Suposa un repte per a la RI i el disseny de sistemes d'informació, ja que, en primer lloc, cal identificar un punt d'entrada a la informació que posi en relació l'esquema mental de l'usuari amb l'organització de la informació, i això s'ha d'oferir per mitjà del disseny. En segon lloc, el mateix sistema ha de donar les claus a l'usuari sobre la rellevància de la informació amb relació a les seves necessitats d'informació.

En funció de la definició de l'objecte de recerca (és a dir, la necessitat d'informació) i la tàctica utilitzada, Marchionini (1995, p. 106) distingeix tres tipus d'ullada:

- Directa o específica, si l'usuari té clar el seu objectiu, el document que busca. Seria el cas, per exemple, de cercar un ítem en una llista.
- Semidirecta o predictiva, amb un objectiu general, és a dir, l'usuari busca informació sobre un tema, però no coneix els documents concrets que li seran d'utilitat, llavors prova amb termes més o menys generals sobre el seu objecte d'interès i en revisa els resultats. És un mètode que consumeix força temps, tot i que depèn del grau de generalitat o especificitat amb el qual s'iniciï.
- Indirecta o general, si l'usuari no té un objectiu definit, simplement revisa la col·lecció per recreació o curiositat.

Estem d'acord que la ullada evita als usuaris plantejar les recerques i els facilita aquesta tasca proporcionant matèries per les quals l'usuari pot navegar. Per tant, aquesta opció s'hauria d'oferir com a primer acostament a la col·lecció de manera que l'usuari tingués accés a un *overview* dels continguts als quals pot accedir en el sistema d'informació, i en els passos successius, de manera que es faciliti el reconeixement d'informació.

S'ha demostrat que el reconeixement d'informació produeix menys càrrega cognitiva que el plantejament de les consultes i permet a l'usuari navegar per la base de dades sense necessitat d'expressar amb un o diversos termes el tema del seu interès. Per tant, s'haurà d'oferir l'opció d'un accés a la informació per mitjà d'ullada a més de per mitjà d'interrogació (*querying*).

3.2.2. Conglomeració o *clustering*

Les anomenades tècniques de *clustering* intenten agrupar objectes semblants per mitjà d'algoritmes matemàtics. Aquestes tècniques aplicades a l'àmbit de la recuperació d'informació aconseguen crear de manera automàtica classificacions de documents tenint en compte cer-

tes similituds en el seu contingut. Per poder crear *clusters*, els documents es representen com a vectors de termes. Cal tenir en compte que la mida dels vectors és igual que la del vocabulari del conjunt recuperat. Per a cada document, el seu vector defineix un punt en un espai multidimensional. Les distàncies entre punts i la seva posició relativa són indicadors de la similitud entre els documents. Per mostrar els punts i les seves relacions a l'usuari, cal reduir el nombre de dimensions a una, dues o tres. Remetem a l'obra de Fernández i Moya (1998), on s'explica amb gran claredat el funcionament d'aquestes tècniques, i als articles de Can i Ozkarahan (1989), Salton i McGill (1983), Salton i Wong (1978), Salton (1989) i Rijsbergen (1979).

Les tècniques de *clustering* es fan servir per agrupar elements en funció de les seves semblances. De moment no són la pràctica més estesa en RI, tot i que s'està treballant en el desenvolupament de sistemes que ho incorporen basant-se en la idea que la coincidència de termes en documents indica la similitud entre aquells documents. Hi ha motius que en justifiquen l'ús: podem presentar a l'usuari els documents recuperats segons la seva similitud en comptes de fer-ho en forma de llista seqüencial, ja que si li interessa un document recuperat també li interessarà amb molta probabilitat un altre document que sigui similar a aquest.

Com a exemple de presentació de resultats amb agrupació en un cercador de la web, mostrem Vivísimo (figura 2).

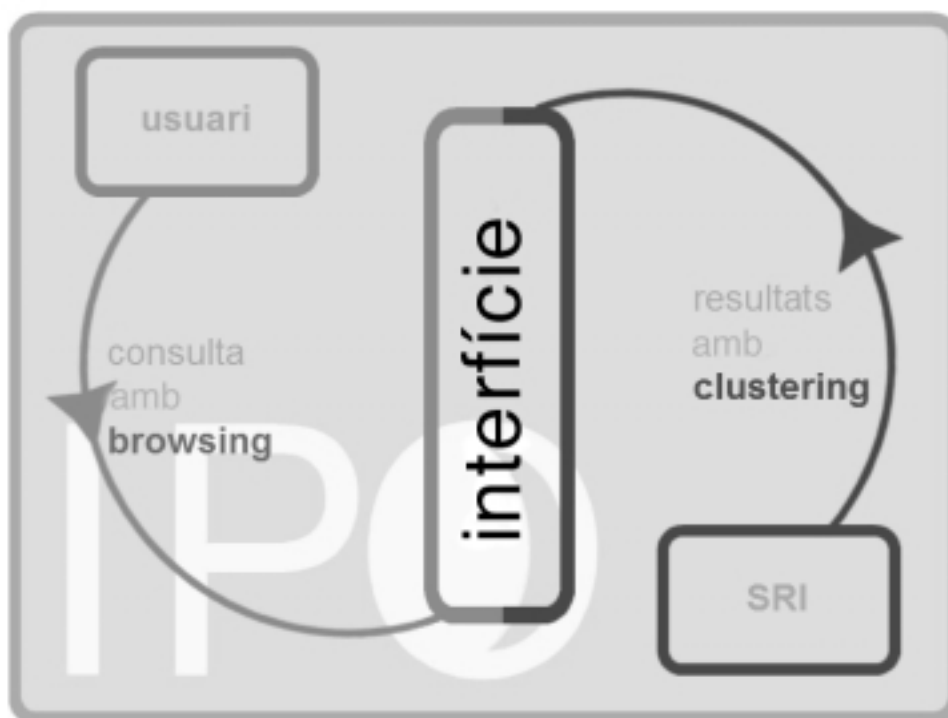


Figura 2. Aquest cercador presenta a la part esquerra els *clusters* que ha detectat dins dels resultats obtinguts, www.vivisimo.com

3.2.3. Visualització de la informació

Per acabar, insistirem en la presentació gràfica dels resultats, considerada avantatjosa davant de la textual, ja que evita la lectura seqüencial de la llista i redueix el temps de selecció de documents rellevants.

Un tret que fa suposar que aquestes presentacions amb visualització són avantatjoses és l'agrupació natural dels documents similars, que permet establir relacions que *a priori* no són evidents. Pel que sembla, els usuaris experts, és a dir, els habituats a utilitzar sistemes de cerca i recuperació d'informació textuals, es mouen amb més comoditat en aquests sistemes tradicionals, però aprenen molt ràpid a manejar els gràfics. En canvi, als usuaris novells en l'ús d'SRI, com que no tenen una experiència prèvia en el maneig d'aquest tipus de sistemes, els resulta més senzill d'entrada manejar els sistemes gràfics que no pas els textuals, perquè els permeten fer servir la intuïció per localitzar grups de documents rellevants.

Hi ha un gran camp de recerca en l'aplicació de la visualització a sistemes de recuperació d'informació. De fet, totes les fases del procés de recerca poden incorporar elements visuals, des de la presentació global del fons documental fins als atributs de cada document, passant pel mateix procés de recerca (tria de punts d'accés, combinació de termes en la consulta, presentació de resultats, reformulació de la consulta) (vegeu la taula 3).

TAULA 3. FASES DEL PROCÉS DE RECERCA D'INFORMACIÓ ON ES POT DONAR LA VISUALITZACIÓ

VISUALITZACIÓ EN LA FASE DE CONSULTA DE LA BASE DE DADES	
Visió global del conjunt complet de la col·lecció (<i>overview</i>)	Representació del procés de la recerca a la base de dades (incloent-hi la retroalimentació que es produeix després de l'obtenció dels resultats)
VISUALITZACIÓ EN LA FASE D'OBTENCIÓ DELS RESULTATS	
Representació del conjunt de documents obtinguts després d'una recerca (<i>preview</i>)	Representació dels atributs de documents concrets (<i>detail</i>)

A més, la ullada i l'agrupació presentats per mitjà de visualització faciliten a l'usuari la localització i l'associació de documents que altrament apareixerien de manera independent entre ells.

De moment, la majoria dels sistemes de recuperació d'informació que incorporen tècniques de visualització són prototips fets en laboratoris de recerca. D'altra banda, la recerca en visualització d'informació encara té molt de camp per estudiar i avaluar per obtenir resultats satisfactoris que puguin competir amb productes més tradicionals de recuperació d'informació.

3.3. Tres interfícies en el procés de recuperació d'informació

3.3.1. La interfície de consulta

La interfície de consulta és el primer contacte de l'usuari amb el sistema; per tant, ens trobem en la primera fase del procés de recerca. Com que hem de considerar un ventall ampli de possibles usuaris, hem de tenir-hi en compte des d'aquells amb una necessitat ben definida i un bon coneixement del maneig del sistema fins a aquells més casuals i amb necessitats d'informació vagues. En aquest darrer cas, les seves dificultats seran de diversa mena, ja que desconeixen tant el contingut de la col·lecció com el llenguatge d'indexació i d'interrogació del sistema.

Per tant, cal facilitar a l'usuari diferents camins per iniciar la recerca:

- Visió general de la col·lecció, per a qui no coneix el fons documental o vol fer un recorregut sense una destinació concreta. Des d'ella es pot aprofundir en temes més específics fins arribar als documents per mitjà de la ullada jeràrquica o associativa. Utilitza la tècnica de la ullada.

- Recerca simple, en què l'usuari introdueix un o dos termes amb els quals expressa la seva necessitat d'informació. Utilitza tècniques d'interrogació.
- Recerca avançada, que li permet fer ús de combinacions entre termes i limitar per alguns camps. Utilitza tècniques complexes d'interrogació.

Abadal (2002) planteja els elements que ha de contenir una pàgina d'aquest tipus, que, de manera resumida, són aquests: títol de la base de dades en què s'està buscant, sistema de recollida de la informació (formulari, llista desplegable...), acotació de la recerca a camps determinats, ús d'operadors, presentació d'índexs, missatges d'orientació, opcions per a la presentació dels resultats (quines dades es facilitaràn i en quin ordre apareixeràn els registres), botons per a execució de les accions, historial de recerques, idioma de la interfície.

3.3.1.1. Visió global de la col·lecció (*overview*)

Fins en aquest moment, la base de dades d'un SRI és una «caixa negra» per a l'usuari, el qual no pot conèixer per endavant quina informació hi ha si no és interrogant el sistema per aquells termes que consideri del seu interès.

En els casos en què l'usuari desconeix el contingut de la col·lecció, no pot estar segur de trobar-hi documents que en satisfacin la necessitat d'informació. Aquest problema pot deixar de ser-ho si no se li dóna una visió global dels temes que s'inclouen a la col·lecció, de manera

Directorio de sitios Web ¡Hemos organizado la web para ti!

<u>Arte y cultura</u> Literatura , Teatro , Museos , Guías ...	<u>Internet y ordenadores</u> WWW , Software , Chat , Redes ...
<u>Ciencia y tecnología</u> Astronomía , Biología , Ingeniería ...	<u>Materiales de consulta</u> Bibliotecas , Diccionarios ...
<u>Ciencia Sociales</u> Filosofía , Historia , Idiomas , Psicología ...	<u>Medios de comunicación</u> Radio , TV , Revistas , Periódicos ...
<u>Deportes y ocio</u> Deportes , Fútbol , Juegos , Turismo ...	<u>Política y gobierno</u> Elecciones , Boletines oficiales , Hacienda ...
<u>Economía y negocios</u> Empresas , Inmobiliarias , Empleo ...	<u>Salud</u> Medicina , Enfermedades , Embarazo ...
<u>Educación y formación</u> Primaria , Secundaria , Universidades ...	<u>Sociedad</u> Gastronomía , Religión , Para niños ...
<u>Espectáculos y diversión</u> Cine , Actores , Música , Humor , ¡Genial!...	<u>Zonas geográficas</u> Países , Europa , España , CC.AA. ...

Los mejores sitios web - Novedades

Figura 3. Els directoris de recerca de la web com Yahoo! classifiquen les pàgines en funció d'unes matèries predefinides amb un criteri temàtic propi, www.yahoo.com

que amb un primer cop d'ull sàpiga el que pot esperar de la base de dades, i després vagi aprofundint diversos nivells fins a localitzar el seu tema d'interès.

Per a això es poden fer servir les dues tècniques d'*overview*: jeràrquic i associatiu.

- Jeràrquic. Presentació de la col·lecció per mitjà d'una estructura jeràrquica temàtica, que es podria basar en les categories de classificació. Es pot fer de manera textual, com ho fan els directoris de la web (per exemple Yahoo!, figura 3) o amb una presentació gràfica com a WebBrain (figura 4).



Figura 4. WebBrain és un directori que mostra gràficament les relacions de jerarquia i associació que hi ha entre els temes de les pàgines que indexa, www.webbrain.com

- Associatiu. Presentació de la col·lecció per mitjà d'una estructura associativa fent servir les tècniques d'agrupació en funció de la temàtica dels documents (*clustering*). Es pot oferir de manera textual, com ho fan alguns multicercadors web, per exemple iBoogie (figura 5), o presentar aquesta informació en forma de mapa, per exemple, amb mapes autoorganitzats com els de Kohonen (figura 6).



Figura 5. iBoogie és un multicercador que presenta els grups en què se divideixen els resultats obtinguts, en funció de la coocurrència de termes, www.iboogie.com

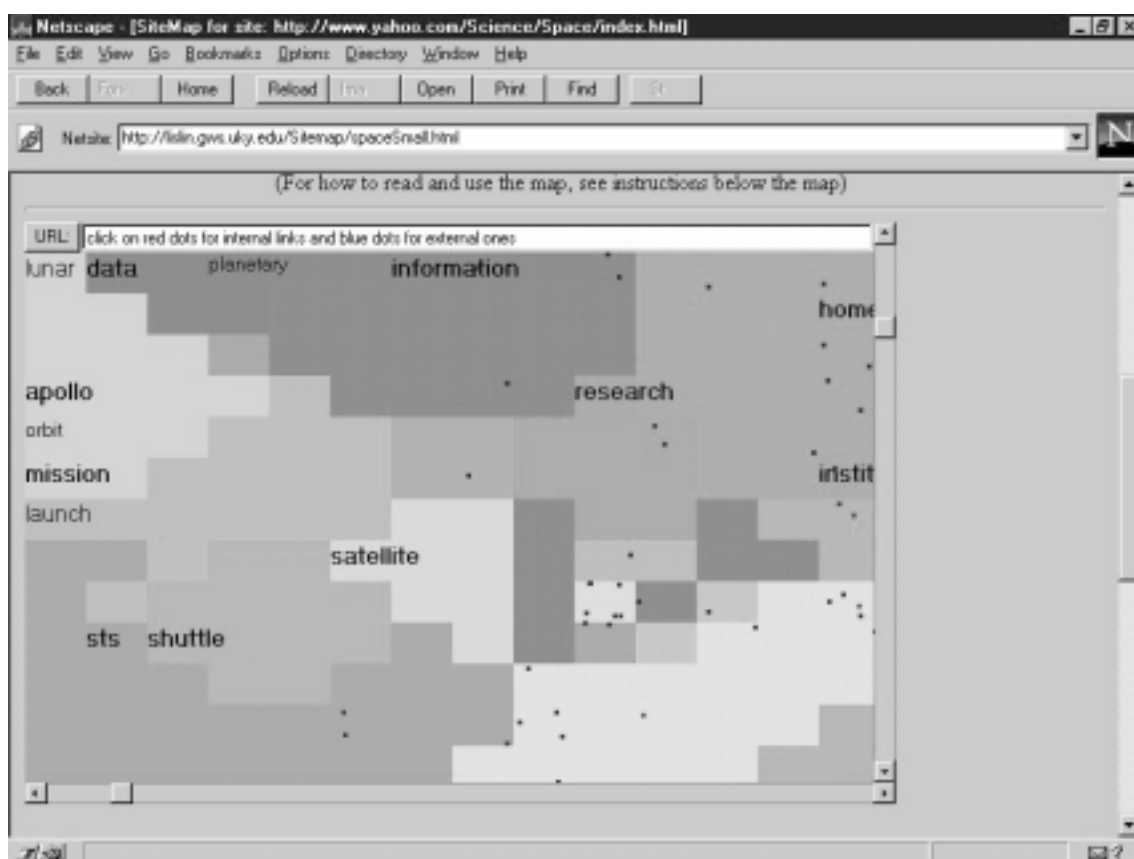


Figura 6. A SiteMap, Xia Lin (Drexel University) crea un mapa a partir de l'apartat d'astronomia i les ciències de l'espai del directori Yahoo! imatge presa de www.cybergeography.org/atlas/sitemap_large.jpg

De vegades l'usuari té una idea vaga de la seva necessitat d'informació. És conscient que necessita certa informació, però no sap quins seran els documents que li poden resoldre el problema i, a més, això el porta a no saber amb quins termes concretar la consulta (Belkin, Oddy, Brooks, 1982a, 1982b). Per a aquests casos és convenient recórrer a les tècniques d'ullada, bé des de l'*overview* o visió general que citàvem abans, bé des d'una recerca general que l'usuari pot introduir per mitjà de *querying* o interrogació al sistema. Des de qualsevol d'aquests dos punts d'accés, l'usuari obtindrà un conjunt de possibles temes en els quals aprofundir. El sistema ha de permetre fer un *zoom* per endinsar-se en les categories. És recomanable que es mantingui el context, a fi que l'usuari no perdi la noció de com ha arribat a un punt determinat i es desorienti. De la mateixa manera que en el cas de l'*overview*, la jerarquia es pot mostrar de manera textual o bé de manera gràfica. Tal com veiem en els exemples, en el cas d'optar per la representació gràfica o visual, es pot fer en forma d'arbre o en forma de mapa.

3.3.1.2. Cerca simple

La gran majoria dels usuaris, llevat que es tracti de professionals de la documentació o d'usuaris molt habituals amb una gran habilitat en el maneig de sistemes de recuperació d'informació, desconeixen el llenguatge d'interrogació que permet elaborar correctament el plantejament de recerca.

En els sistemes tradicionals de RI, l'usuari planteja al sistema la seva consulta tot introduint per mitjà del teclat les paraules que considera que podrien estar en els documents susceptibles d'interessar-li. Ens sembla important mantenir aquesta forma de recerca per diversos motius:

- Els usuaris estan acostumats a aquesta forma d'accés a la col·lecció (i molt més ara amb la generalització de l'ús dels cercadors de la web). Les figures 6 i 7 mostren la cerca simple en un OPAC i en un motor de recerca, respectivament, i s'hi pot observar que són molt semblants.
- Resulta d'utilitat per a aquells usuaris que volen indicar els criteris de recerca directament, sense fer el recorregut d'allò que és general a allò que és específic per ullada.

Consultar el Catálogo de la Biblioteca - Búsqueda sencilla

Figura 7. Cerca simple de l'OPAC de la biblioteca de la Universidad Carlos III de Madrid.

Figura 8. Cerca simple de Google.

En el cas que es faci servir una base de dades amb camps de vocabulari controlat, l'usuari no ha de saber necessàriament els termes que s'han fet servir en el sistema per descriure els documents que podrien resultar del seu interès. La manera de facilitar-li el coneixement dels termes que fa servir la col·lecció per descriure els seus documents és disposar d'una llista de termes acceptats o un tesaurus.

Ara bé, el fet de permetre-li la consulta del tesaurus podria suposar-li una complicació si no està acostumat a aquesta mena d'eines documentals. De tota manera, considerem que aquesta opció ha d'estar disponible per si realment té interès a consultar-lo. Una altra opció, més simple per a l'usuari, és que el sistema el remeti de manera automàtica des del terme no acceptat que ell tecleja fins als termes acceptats o relacionats amb aquest segons el tesaurus, de manera que se li suggereixin alternatives de consulta. És a dir, facilitar un tesaurus en línia que «tradueixi» els termes utilitzats per l'usuari a termes controlats. Com a exemple de siste-

REFERENCIAS CRUZADAS

Interface

ESTA ENTRADA NO SE UTILIZA ACTUALMENTE
 EN SU LUGAR SE UTILIZAN LAS SIGUIENTES ENTRADAS:

Entradas	Encabezamiento
44	► <u>Interficies</u>

Figura 9. L'OPAC de la biblioteca de la Universitat Politècnica de Catalunya remet al terme acceptat.

REFERENCIAS CRUZADAS

Entradas	Encabezamiento
305	► <u>Informàtica</u>
También encontrarás información bajo las siguientes:	
94	► <u>Cibernètica</u>
184	► <u>Llenguatges de programació</u>
160	► <u>Ordinadors</u>
12	► <u>Teledocumentació</u>
112	► <u>Bases de dades</u>
81	► <u>Disseny de sistemes [TÉRMINO ESPECÍFICO]</u>
1	► <u>Bancs de dades</u>
332	► <u>Intel·ligència artificial [TÉRMINO ESPECÍFICO]</u>

Figura 10. L'OPAC de la biblioteca de la Universitat Politècnica de Catalunya indica les matèries relacionades i permet efectuar des d'allí noves consultes.

ma de recuperació d'informació que ofereix aquesta possibilitat, mostrem l'OPAC de la biblioteca de la Universitat Politècnica de Catalunya, que remet l'usuari al terme correcte (figura 9) i ofereix matèries relacionades (figura 10).

Una altra possibilitat que ofereix el tesauro en línia és la de suggerir termes amb els quals ampliar temàticament la recerca (*broader terms*), restringir-la a temes més específics (*narrower term*) o continuar-la amb temes relacionats (*related terms*), i d'aquesta manera facilitar la reformulació de la consulta.

Per tant, amb aquesta forma d'accés, l'usuari pren la iniciativa i comença la recerca tot indicant les paraules que hauran d'aparèixer en el camp de matèries o descriptors.

3.3.1.3. Cerca avançada

Altres usuaris voldran acudir a una cerca avançada per establir els criteris de cerca amb més precisió. Pressuposem que no són experts en recuperació d'informació, si més no és evident que no ho seran en tots els SRI, per la qual cosa els hem de facilitar unes eines que els permetin construir la seva recerca d'una manera senzilla. Per fer-ho, disposarem de formularis en què es combinin els termes, però de manera que ells no hagin de fer anar operadors boo-

Consultar el Catálogo de la Biblioteca - Búsqueda Avanzada

Todos los campos	<input type="text"/>	Y
Autor	<input type="text"/>	Y
Título	<input type="text"/>	Y
Materia	<input type="text"/>	Y
Serie	<input type="text"/>	Y
Título de Revista	<input type="text"/>	
<input type="button" value="Buscar"/> <input type="button" value="Reestablecer"/>		
Bibliotecas	ALL	
idioma:	ANY	
formato:	ANY	
localización:	ANY	
Material:	ANY	
año:	<input type="text"/>	
ordenar por:	No	

Figura 11. Cerca avançada a l'OPAC de la biblioteca de la Universidad Carlos III de Madrid.

Google Sugerencias de Búsqueda | Todo acerca de Google

Búsqueda Avanzada

Buscar resultados con todas las palabras 10 resultados

con la frase exacta

con alguna de las palabras

sin las palabras

Idioma Producir páginas escritas en cualquier idioma

Formato de archivo Solamente devuelve resultados del formato de archivo cualquier formato

Fecha Mostrar páginas Web actualizadas durante en cualquier momento

Presencia Producir resultados en los que mis términos estén presentes en cualquier parte de la página

Dominios Solamente producir resultados del dominio o sitio Web

Ejemplos: .org, google.com [Más información](#)

Búsqueda Especifica a una Página

Similares Encontrar páginas similares a la página

Enlaces Encontrar páginas con enlaces a la página

Ejemplo: www.google.com/help.html

Figura 12. Cerca avançada en Google.

leans si no els coneixen bé, és a dir, que l'equació de recerca es farà de manera interna pel sistema i totalment transparent a l'usuari. Les figures 10 i 11 mostren, respectivament, aquesta opció en el catàleg de la Universidad Carlos III de Madrid i en Google, per continuar amb els exemples donats en el cas de la cerca simple i observar que tots dos recorren als formularis.

3.3.2. La interfície de resultats

Una vegada feta la consulta, el sistema ofereix a l'usuari un conjunt de documents que compleixen els criteris especificats. Tot sovint la informació que ofereix resulta incòmoda de manejar, ja que està en forma de llista, i a més resulta insuficient per avaluar la rellevància dels documents mostrats.

En aquesta llista de documents que compleixen els criteris de cerca establerts per l'usuari, cada document es mostra de manera independent en relació amb els altres, tot i que potser estiguin semànticament molt propers pel seu contingut temàtic. Considerem que és preferible agrupar els resultats establint relacions entre els documents recuperats, de manera que es presentin en funció de la seva semblança pel que fa al contingut.

En el cas de cerques per matèries, no sempre es podran establir subgrups, per tal com tots els documents compartiran molts termes en el seu text o diversos descriptors en el cas que estiguin indexats. Si el nombre de documents recuperats és baix, aquesta agrupació també perd en bona part el seu sentit. La manera de presentar aquestes relacions es pot fer amb les tècniques d'agrupament i es poden fer servir presentacions cartogràfiques, com els mapes de Kohonen, o bé recórrer a la metàfora del paisatge com ThemeView o de les constel·lacions com Galaxy, tots dos desenvolupats per SPIRE (www.pnl.gov/infoviz), o bé a la metàfora dels imants com a VIBE (www.itl.nist.gov/iaui/vvrg/emorse/papers/smc97.html) i WebVIBE (www2.sis.pitt.edu/~webvibe/webVibe/page02.html). Oferim l'exemple de KartOO, en què s'observa una metàfora cartogràfica per als resultats obtinguts en la cerca (figura 13).

Si s'ha fet servir el tradicional model booleà per plantejar una consulta, no és pertinent l'ordenació dels resultats per rellevància, ja que tots els registres compleixen la condició de cer-

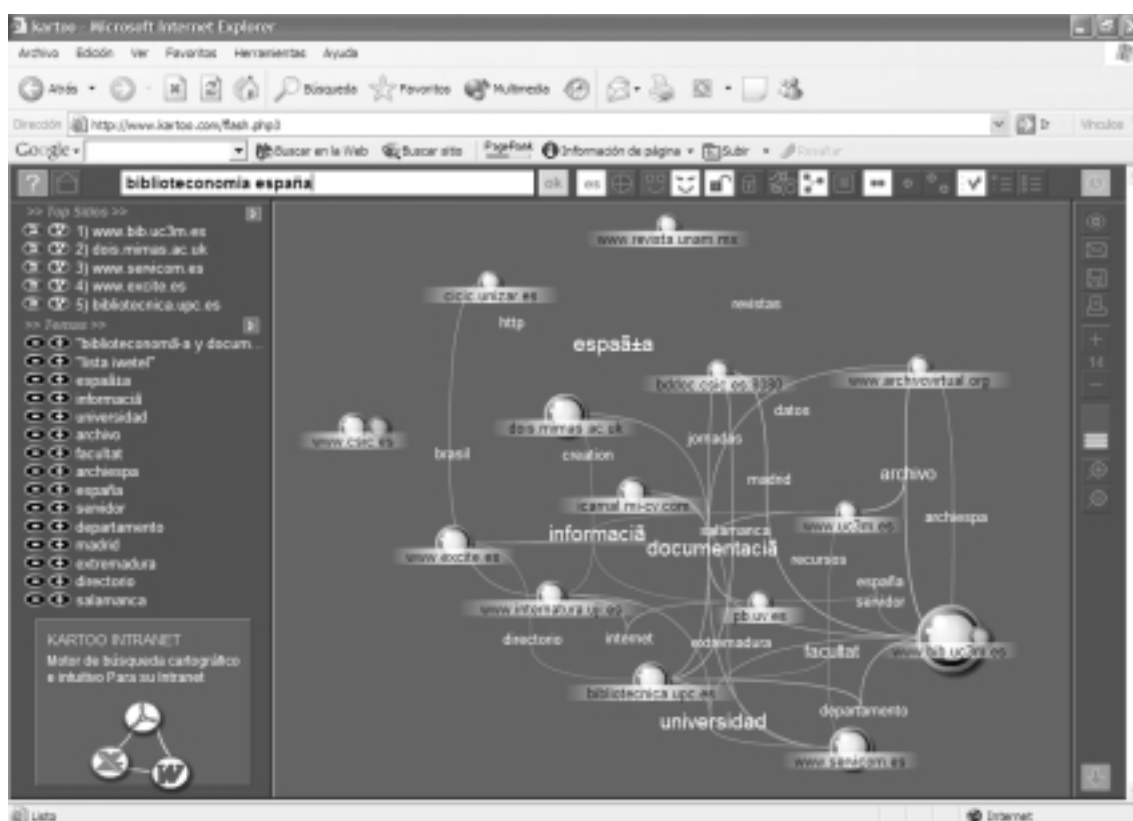


Figura 13. KartOO presenta els resultats en forma d'esferes la mida de les quals varia segons la rellevància que se'ls ha assignat. Els llocs web apareixen enllaçats entre ells segons la seva semblança de continguts www.kartoo.com

ca, però no hi ha nivells intermedis, per la qual cosa l'ordre establert vindrà donat principalment per altres criteris com l'alfabètic (cognom de l'autor) o el cronològic (data de publicació).

L'altra opció seria utilitzar un model de RI basat en la freqüència d'aparició dels termes en els documents, com fa el model vectorial, en què cada document coincideix en major o menor grau amb la consulta. D'aquesta manera, es pot establir un rànquing de rellevància dels resultats i presentar a l'usuari aquells documents que responen en major grau a la seva consulta. En el cas de fer servir models vectorials, poques vegades es donen a conèixer a l'usuari els criteris de rànquing.

Si pensem en el cas dels OPAC, hi trobem un altre problema afegit: el sistema ofereix no més la informació que tradicionalment apareix a la fitxa de catalogació. Aquesta informació que ofereix el registre bibliogràfic és insuficient perquè l'usuari pugui fer una valoració de la rellevància del document. Caldria fornir dades addicionals a les de la descripció catalogràfica, com ara l'índex del document, comentaris d'altres usuaris sobre el llibre, obres publicades pel mateix autor, etc.

Abadal (2002) també fa un recull d'elements que han de ser presents en les pàgines de resultats de les interfícies de recerca. De manera resumida, són aquests: títol de la base de dades que es consulta, consulta feta, nombre de documents recuperats i de documents visualitzats en pantalla en aquell moment, descripció dels documents amb possibilitat de selecció d'aquells que interessin a l'usuari, tipus de document de què es tracta, agrupació per categories, tria del format de presentació en pantalla, del nombre de registres que es visualitzen i de l'ordenació amb la qual es mostren, possibilitat de cerca de documents similars i de reformulació de la consulta sobre els resultats obtinguts, i navegació entre registres, entre d'altres.

3.3.3. La interfície de reformulació de la consulta

Amb la llista de resultats que compleixen el criteri de cerca es dona per acabada la consulta, per la qual cosa si l'usuari no en queda satisfet, haurà d'acudir de nou a la pantalla de consulta i formular una nova cerca. Aquest és el cas més habitual en els sistemes de recuperació d'informació.

Com que la cerca d'informació és un procés amb una continuïtat, no hem de permetre que l'usuari comenci el procés cada vegada que arriba a un conjunt de possibles resultats, sinó que aquest conjunt li serveixi per afinar la seva consulta i arribar a uns resultats òptims.

Per fer-ho, cal establir fórmules de retroalimentació en el sistema que siguin com més transparents millor per a l'usuari. Una tècnica consisteix que el sistema reformuli per ell mateix la consulta a partir dels documents que l'usuari marqui entre els obtinguts com a més rellevants a la seva necessitat d'informació. En sistemes a text complet, ho pot fer basant-se en la freqüència d'aparició de les paraules i tenint en compte en quines parts del text són; en bases de dades referencials haurà de tenir en compte alguns camps en llenguatge natural com el títol o el resum, però també haurà de considerar que hi hagi coincidència en els camps de les matèries i la classificació.

4. Reflexions finals

Aquest article tracta la interacció en els sistemes de recuperació d'informació i, per tant, pren com a disciplina marc la Interacció Persona-Ordinador (IPO), que s'ocupa d'estudiar la creació de productes informàtics que ajudin en la realització de tasques als usuaris. El panorama espanyol en l'estudi de la IPO és molt encoratjador, tot i la seva incorporació recent als plans d'estudi de les facultats espanyoles d'informàtica i en alguns estudis de psicologia i de documentació. En el cas concret de l'aplicació de la IPO al desenvolupament de productes per a documentació, la recerca encara és escassa, tot i que s'hi pot observar un creixement en els últims anys.

Una disciplina fonamental per als estudis d'IPO és la psicologia cognitiva, d'on es prenen les bases per a l'estudi del coneixement humà. La manera en què les persones percebem la informació, l'emmagatzemem, la mantenim a la memòria i la processem constitueix una clau important que cal tenir en compte en el disseny de sistemes més adequats als qui els utilitzaran. Un dels objectius principals de la IPO és crear sistemes que permetin a l'usuari formar una imatge mental correcta del seu funcionament, de manera que el model del dissenyador i el de l'usuari siguin el mateix, i d'aquesta manera evitar que es donin desajustos que provoquin una idea equivocada. La part del sistema que té la clau perquè el model mental que construeix l'usuari sigui la correcta és la interfície.

Cal aclarir que la interfície perfecta no existeix, perquè seria aquella que fos perfecta per a tots els usuaris, i atesa la complexitat de les persones i de les variades maneres d'actuar per fer front als problemes, no és possible assolir un sistema ideal per a tots ells. En canvi, sí que hem de fixar com a objectiu que la interfície aconseguixi que l'usuari faci servir el sistema d'una manera transparent, és a dir, que aquest es mogui amb desimboltura pensant el mínim possible i sense necessitat d'entendre com treballa el programa informàtic. Per tant, la interfície ideal és la que «desapareix» davant de l'usuari i li permet centrar-se en les tasques que fa i no en les que processa el sistema.

En plantejar el disseny d'interfícies per a sistemes de recerca d'informació, cal tenir en compte, entre altres coses, que el reconeixement d'informació produeix menys càrrega cognitiva que no pas el plantejament de les consultes i permet a l'usuari navegar per la base de da-

des sense necessitat d'expressar amb un o diversos termes el tema del seu interès. Per tant, s'haurà d'oferir l'opció d'un accés a la informació per mitjà d'ullada a més de per mitjà d'interrogació.

La visualització és una altra manera natural de les persones per entendre el seu entorn. Aplicada a la informació, les ajuda a formar una imatge mental de l'espai informatiu. Si la visualització té lloc en una interfície que té com a objectiu la recuperació d'informació, l'expressió creada per a aquests tipus de sistemes és «interfícies visuals de recuperació d'informació» (VIRI, *visual information retrieval interfaces*).

Hi ha un gran camp obert de recerca en l'aplicació de la visualització a sistemes de recuperació d'informació. De fet, totes les fases d'un procés de recerca poden incorporar elements visuals, des de la presentació global del fons documental fins als atributs de cada document, passant pel mateix procés de recerca (tria de punts d'accés, combinació de termes en la consulta, presentació de resultats, reformulació de la consulta). En canvi, el desenvolupament d'aquest tipus d'aplicacions encara és escàs en sistemes de recuperació d'informació, en part a causa d'una absència de recerca en aquest camp, i en part a causa de la dificultat que planteja trobar una metàfora adient per presentar grans quantitats d'informació en espais reduïts.

A més, la majoria dels sistemes que fan servir tècniques de visualització per mostrar la informació s'han desenvolupat per treballar amb bases de dades a text complet, o si més no amb un camp de resum, cosa que permet disposar d'una quantitat de termes suficient per fer un tractament automàtic del llenguatge i establir categories per similitud. En el cas de bases de dades referencials, la informació de contingut del document sols està en els camps de títol (amb els problemes d'ambigüitat que comporta el llenguatge natural) o en els normalitzats per a descripció (descriptors o matèries i codis de classificació), cosa que fa més complicat establir criteris de similitud.

Una altra tècnica que cal tenir en compte en la presentació de la informació és la conglomeració de documents, que es pot fer de manera automàtica per mitjà de tècniques de *clustering*. Aquesta presentació facilita a l'usuari la localització i l'associació de documents que altrament apareixerien de manera independent entre ells. La recuperació d'informació es comença a fer servir en la presentació de resultats de motors i multicercadors de la web. L'aplicació de la visualització per presentar la informació agrupada semànticament millora l'accés a aquesta per part dels usuaris, ja que agrupen els documents similars i això permet establir relacions que *a priori* no són evidents. Les tècniques de visualització permeten establir relacions semàntiques entre els documents de la col·lecció en funció de la seva ubicació en l'espai d'informació, característica que no mostren les habituals llistes de documents.

Bibliografia

- ABADAL, E. (2002). «Elementos para la evaluación de interfaces de consulta de bases de datos». *El profesional de la información*, 11, 5, p. 349-360.
- ABASCAL, J. *et al.* (2000). «Hacia un corpus docente virtual común en IPO: experiencia del Curso Virtual de Introducción a la Interacción Persona-Ordenador», en: *I Jornadas de Interacción Persona-Ordenador (Granada)*.
- ALBERICO, R.; MICCO, M. (1990). *Expert systems for reference and information retrieval*. London: Meckler.
- ALTUNA ESTEIBAR, B. (1992). «Comportamientos de uso y estrategias de búsqueda de los usuarios de catálogos automatizados: breve revisión de la investigación», en: *Miscelánea homenaje a Luis García Ejarque*. Madrid: Fesabid, p. 103-111.

- BAECKER, R.; RUDIN, J.; BUXTON, W. (1995). *Readings in human computer interaction: toward the year 2000*. San Francisco: Morgan Kaufmann.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. (1999). *Modern information retrieval*. Reading (EUA): Addison Wesley.
- BATES, M. (1986). «An exploratory paradigm for online information retrieval», en: BROOKS (ed.). *Intelligent information systems for the information society: proceedings of the 6th IRFIS Conference*. New York: North Holland, p. 91-99.
- (1989). «The design of browsing and berrypicking techniques for the online search interface». *Online review*, 13, 5, p. 407-424. www.gseis.ucla.edu/faculty/bates/berrypicking.html
- BELKIN, N. (2000). «Helping people find the way thy don't know». *Communications of the ACM*, 43, 8, p. 58-61.
- *et al.* (1990). «Taking account of user tasks, goals and behavior for the design of online public access catalogs». *ASIS'90: information in the year 2000: from research to applications: proceedings of the 53rd Annual Meeting of the American Society for Information Science*. Medford (EUA): Learned Information, p. 69-79.
- ; ODDY, R.; BROOKS, H. (1982a). «ASK for information retrieval. Part I: Background and theory». *Journal of documentation*, 38, 2, p. 61-71.
- ; —; —. (1982b). «ASK for information retrieval. Part II: Results of a design study». *Journal of documentation*, 38, 3, p. 145-164.
- BOOTH, P. (1989). *An introduction to human-computer interaction*. Hilldale (EUA): Erlbaum.
- CAN, F.; OZKARAHAN, E. (1989). «Dynamic cluster maintenance». *Journal of information processing & management*, 25, 3, p. 275-291.
- CARD, S.; MORAN, T.; NEWELL, A. (1983). *The psychology of human-computer interaction*. Hillsdale (EUA): Erlbaum.
- CHURCH, G. (1999). «The human-computer interface and information literacy: some basis and beyond». *Information technology & libraries*, 18, 1, p. 3-21.
- FERNÁNDEZ MOLINA, J. C.; MOYA ANEGÓN, F. de (1998). *Los catálogos de acceso público en línea: el futuro de la recuperación de información bibliográfica*. Málaga: Asociación Andaluza de Bibliotecarios.
- FERNÁNDEZ RUIZ, M. J.; ANGÓS ULLATE, J. M.; SALVADOR OLIVÁN, J. A. (2001). «Interfaces de usuario: diseño de la visualización de la información como medio para mejorar la gestión del conocimiento y los resultados obtenidos por el usuario», en: Congreso ISKO-España (5è, 2001: Alcalá de Henares). *La representación y organización del conocimiento: metodologías, modelos y aplicaciones*. Alcalá de Henares: Facultad de Documentación.
- FRÍAS MONTOYA, J. A.; MARTÍN RODRÍGUEZ, F. (1999). «El análisis transaccional como técnica de recogida de datos para el estudio del comportamiento de los usuarios del catálogo en línea», en: Congreso ISKO-España (4t, 1999: Granada). *Representación y organización del conocimiento en sus distintas perspectivas: su influencia en la recuperación de información*, p. 427-434.
- GONZÁLEZ ROMANO, J. (2001). *Assignatures d'Introducció a la Interacció Persona-Ordinador impartides a les universitats espanyoles: apèndix al projecte docent presentat a la Universitat de Sevilla al juny del 2001*. www.lsi.us.es/~mariano/assignaturas_ipo.pdf
- GRUDIN, J. (1988). «Why groupware applications fail: problems in design and evaluation». *Office: technology and people*, 4, 3, p. 245-264.
- (1994). «Groupware and social dynamics: eight challenges for developers». *Communications of the ACM*, 37, 1, p. 92-105.
- ISO 9241: *ergonomic requirements for office work with visual display terminals (VDTs)*, 1992-2000. www.iso.ch

- LAUREL, B. (ed.). (1990). *The art of human-computer interface design*. Reading (EUA): Addison-Wesley.
- LIN, X. (1997). «Map displays for information retrieval». *Journal of the American Society for Information Science*, 48, p. 40-54.
- LORÉS, J. et al. (2001). «Un corpus docente virtual, común y ubicuo en IPO», en: *II Congreso Internacional de Interacción Persona-Ordenador*. Salamanca.
- . (2002 act.). *La interacción persona-ordenador*. griho.udl.es/ipo [llibre digital].
- MANDEL, T. (1997). *The elements of user interface design*. New York: Wiley.
- MARCHIONINI, G. (1995). *Information seeking in electronic environments*. Cambridge: University Press.
- . (1987). «An invitation to browsing: designing full-text systems for novice users». *Canadian journal of information science*, 12, 3-4, p. 69-79.
- . (1989). «Information seeking strategies of novices using a full-text electronic encyclopedia». *Journal of the American Society for Information Science*, 40, 1, p. 54-66.
- ; KOMLODI, A. (1998). «Design of interfaces for information seeking». *Annual review of information science and technology (ARIST)*, 33, p. 89-130.
- MARCOS, M. C. (2001). «HCI (human-computer interaction): concepto y desarrollo». *El profesional de la información*, 10, 6, p. 4-16.
- . (2002). «Presente, pasado y futuro de la investigación en interacción hombre-ordenador aplicada a la documentación», en: MORÁN, M. A.; RODRÍGUEZ LÓPEZ, M. C. (coords.). *La Documentación para la investigación: homenaje a José Antonio Martín Fuertes*. León: Universidad, p. 365-385.
- . (2003). «Interacción persona-ordenador en las interfaces de recuperación de información», en: Jornadas Españolas de Documentación (8es: 2003: Barcelona). *Los sistemas de información en las organizaciones: eficacia y transparencia (6-8 febrero 2003, Barcelona)*. [Madrid]: Fesabid, p. 463-476.
- MAYHEW, D. (1992). *Principles and guidelines in software user interface design*. Englewood Cliffs: Prentice Hall.
- MISCH, F. (ed.) (1988). *Webster's ninth new collegiate dictionary*. Springfield: Merriam-Webster.
- MOTRO, A. (1986). «BAROQUE: A browser for relational databases». *ACM transactions on office information system*, 4, p. 164-181.
- NICHOLS, D. (1998). «Implicit rating and filtering», en: *Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering (Budapest, Hungary, 10-12 November 1997)*. www.comp.lancs.ac.uk/computing/research/cseg/projects/ariadne/docs/delos5.html; www.comp.lancs.ac.uk/computing/research/cseg/projects/ariadne/docs/delos5.pdf, p. 31-36.
- NORMAN, D. (1988a). *The design of everyday things*. New York: Doubleday.
- . (1988b). *The psychology of everyday things*. New York: Basic Books.
- RIJSBERGEN, C. van (1979). *Information retrieval*. [2a ed.] Londres: Butterworth.
- SALTON, G. (1979). «Progress report on automatic information retrieval», en: *Second International Conference on Information Storage and Retrieval (Dallas, Texas, 27-28 September)*.
- ; MCGILL, M. J. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.
- ; WONG, A. (1978). «Generation and search of clustered files». *ACM transactions on database systems*, 3, 4, p. 321-346.
- SHAKEL, B. (1997). «Human-Computer Interaction: whence and whither?». *Journal of the American Society for Information Science*, 48, 11, p. 965-986.
- SHNEIDERMAN, B. (1992). *Designing the user interface: strategies for effective Human-Computer Interaction*. Reading (EUA): Addison-Wesley.
- THOMPSON, R.; CROFT, W. (1989). «Support for browsing in an intelligent text retrieval system». *International journal of man-machine studies*, 30, p. 636-668.