

COM BUSCAR INFORMACIÓ PER INTERNET?

Localitzar informació al Web requereix conèixer els instruments apropiats i seguir un procés o estratègia que es pot resumir en:

- Formular la cerca amb els termes que millor descriuen el problema d'informació
- En funció de la tipologia o complexitat de la cerca decideix quin instrument de cerca heu d'aplicar: directori o cercador (i dins d'aquests quin pot ser el més apropiat)
- Depenent dels resultats de la primera cerca, si aquesta ha estat molt rica seleccioneu els documents pertinents obtinguts. En canvi si ha estat pobre en resultats o infructuosa inicieu una nova cerca en un altre cercador o amb uns altres termes alternatius (tal vegada en els equivalents en anglès).
- Un cop identificats documents pertinents al nostre problema, ells mateixos ens aportaran més informació (enllaços a nous documents, terminologia específica) per ampliar, si s'escau, la consulta.

INSTRUMENTS DE CERCA D'INFORMACIÓ

Els motors de cerca són eines dirigides a la recuperació d'informació disponible a Internet. Bàsicament, un cercador es compon d'un *robot* que rastreja les pàgines de la xarxa cercant novetats i indexant-les, una base de dades a on emmagatzema una breu referència de la informació indexada i la seva adreça electrònica i una interfície que permet a l'usuari interrogar el sistema i rebre els resultats obtinguts.

Característiques

- Disposen d'un robot que explora la xarxa automàticament per localitzar i classificar recursos.
- Busquen per paraules clau.
- Estan compilats per màquines i no persones.
- La informació s'organitza de forma automàtica.
- Els resultats no es mostren ordenats per categories o matèries.
- Permeten cerques molt complexes.
- Indexen diferents parts de les pàgines web (títol, encapçalament, text complet, adreça).

Avantatges dels motors de cerca

- Permeten fer cerques a partir del llenguatge natural.
- Es busquen els termes a qualsevol part del document (pàgina web).
- La seva tipologia és variada: n'hi ha de generals, d'especialitzats i d'acadèmics.

Limitacions dels motors de cerca

- Poden provocar excés d'informació, amb resultats que sovint no tenen cap relació amb l'estratègia de cerca.
- La qualitat dels resultats obtinguts pot ser molt irregular.
- L'absència d'un llenguatge controlat dificulta l'accés a la informació (en el punt anterior aquest aspecte es presenta com un avantatge, ja que l'ús del llenguatge natural resulta més senzill per qui busca informació; però també dona com a resultat la dispersió de la informació degut a fenòmens com ara la sinonímia, la polisèmia, etc.)
- La indexació automàtica fa que els resultats siguin sovint decebedors.

Els diferents serveis de cerca d'informació a Internet es poden classificar en: directoris, cercadors, meta-cercadors (cercadors múltiples) i bases de dades.

- Directoris. Un índex o directori és una recopilació manual de documents que classifica adreces d'altres pàgines Web, organitzant-les temàticament. Existeixen nombrosos directoris, uns de generals com *Yahoo* (www.yahoo.com), uns altres són compilats per personal especialitzat com el de l'editorial *Enciclopedia Britannica* (<http://info.britannica.co.uk>). Finalment hi ha directoris que sols cobreixen un camp molt específic com *CENET Networks* (<http://downloads.zdnet.com/>) un directori de programari de lliure difusió que cobreix diversos àmbits i particularment té una bona secció de programari educatiu (en anglès). Un altre directori de l'àmbit lingüístic català és *CerCAT* (www.cercat.com).

Alguns directoris útils són:

Web Site Collation Project: www.websitecollationproject.com/

Internet Invisible: www.internetinvisible.com

Galaxy: www.galaxy.com/directory/

Lukor: www.lukor.com/directorio

Google Directori: <http://www.google.es/dirhp?hl=es>



- Cercadors. Un cercador (en anglès "search engine") és un programa que fa cerques dins una base de dades. En el cas que ens ocupa, la base de dades és de documents Web. La base de dades es construeix gràcies a uns programes anomenats "robots". Els robots són programes encarregats de fer buidats periòdics de les pàgines Web. També se'ls anomena "aranyes" pel fet que viatgen per la "teranyina". La seva finalitat és la de visitar els ordinadors que emmagatzemen informació d'accés públic, indexar-la i enviar aquest índex a l'ordinador matriu on es sumen tots els índexs alimentant de forma automàtica una gran base de dades documental. Fer una cerca a la Web vol dir fer una consulta a una base de dades documental creada per una empresa o entitat que manté el cercador (n'hi ha mes d'un miler). El motor de cerca rep la interrogació de l'usuari que consisteix en una o més paraules, fa la cerca a la seva base de dades i extreu una llista ordenada de documents que compleixen del tot o en part la interrogació. L'ordre d'aparició dels resultats depèn d'una puntuació que associa el programa a cada document quan fa la cerca, i en cada cas varia. Existeixen nombrosos cercadors a la Web que es diferencien per:
 - la quantitat de vegades que refresquen la informació del seu índex (tinguem en compte que la informació a la Web és molt dinàmica: hi ha altes i baixes de pàgines contínuament)
 - la quantitat d'informació que emmagatzemen de cada pàgina Web que el robot visita. Hi ha cercadors que sols es fixen en el títol, d'altres en els descriptors (si n'hi ha) i d'altres en tot el text.

- l'àmbit geogràfic que cobreixen: hi ha cercadors d'abast universal i d'altres que sols recullen informació d'ordinadors situats dins un domini geogràfic concret. Els cercadors generals més importants a nivell mundial són:

 www.excite.com	 www.lycos.com
 www.altavista.com	 www.clusty.com
 www.alltheweb.com	 http://www.bing.com/ (Microsoft)

En castellà destaquen:

 www.ozu.es	 www.sol.es
 www.yahoo.es	 www.lycos.es

I en català:

 www.som-hi.com	 www.cercat.com
---	---

 NOSALTRES CLASSIFICATS www.nosaltres.com	
 www.que.cat	 ct.yahoo.com

En aquesta pàgina tens tots els buscadors espanyols i catalans:
http://www.tingloop.com/buscadores/buscadores_espanoles.htm

I en aquesta altra tens un bon llistat de cercadors:
<http://www.diba.es/altreswebs/cercadors.asp>

I com a curiositat podeu consultar la relació que hi ha entre els diferents buscadors: <http://www.search-this.com/search-engine-decoder/>

- Meta-cercadors o cercadors múltiples: Cerquen informació simultàniament dins diferents cercadors (normalment els podem escollir) i van classificant les respostes a mesura que van arribant. Els més eficients eliminen els duplicats. Solen ser de menors prestacions que els cercadors individuals (més lents, amb menys recursos per recerques avançades) però simplifiquen la realització de cerques múltiples cosa molt útil en un primer nivell de cerca. Destaca Askjeeves (www.askjeeves.com) que accepta formular les preguntes en llenguatge natural. També tenim el Metacrawler (www.metacrawler.com) i el Cyber411 que fa servir 15 buscadors a la vegada (www.cyber411.com).

 www.metacrawler.com	 www.askjeeves.com	 www.cyber411.com
 www.buscamultiple.com	 www.dogpile.com	 www.ez2find.com
 eu.ixquick.com	 www.search.com	 http://www.webcrawler.com/
 www.soople.com		 www.Freeality.com

- Bases de dades:

Les bases de dades bibliogràfiques organitzen i estructuren la informació amb l'objectiu de facilitar-ne la recuperació, habitualment per autor, paraules clau, matèries... Les bases de dades són habitualment productes comercials en què un editor decideix quin tema tractarà i quina informació recopilarà sobre aquest tema. També existeixen, però, moltes bases de dades creades per institucions professionals o científiques que permeten l'accés lliure a la informació recopilada.

La majoria de bases de dades contenen referències (o citacions) bibliogràfiques d'articles publicats a revistes generals, especialitzades, diaris i congressos. Tot i així, també hi ha bases de dades que ofereixen informació bibliogràfica d'altres tipus de documents com llibres o capítols de llibre (*Iconda*), tesis (*Inist*), patents (*Derwent*), lleis (*Aranzadi*), notícies de diari (*Factiva*), etc. La informació continguda ha passat per una revisió. És a dir, s'ha fet una selecció de quines publicacions són les més importants en una especialitat temàtica.

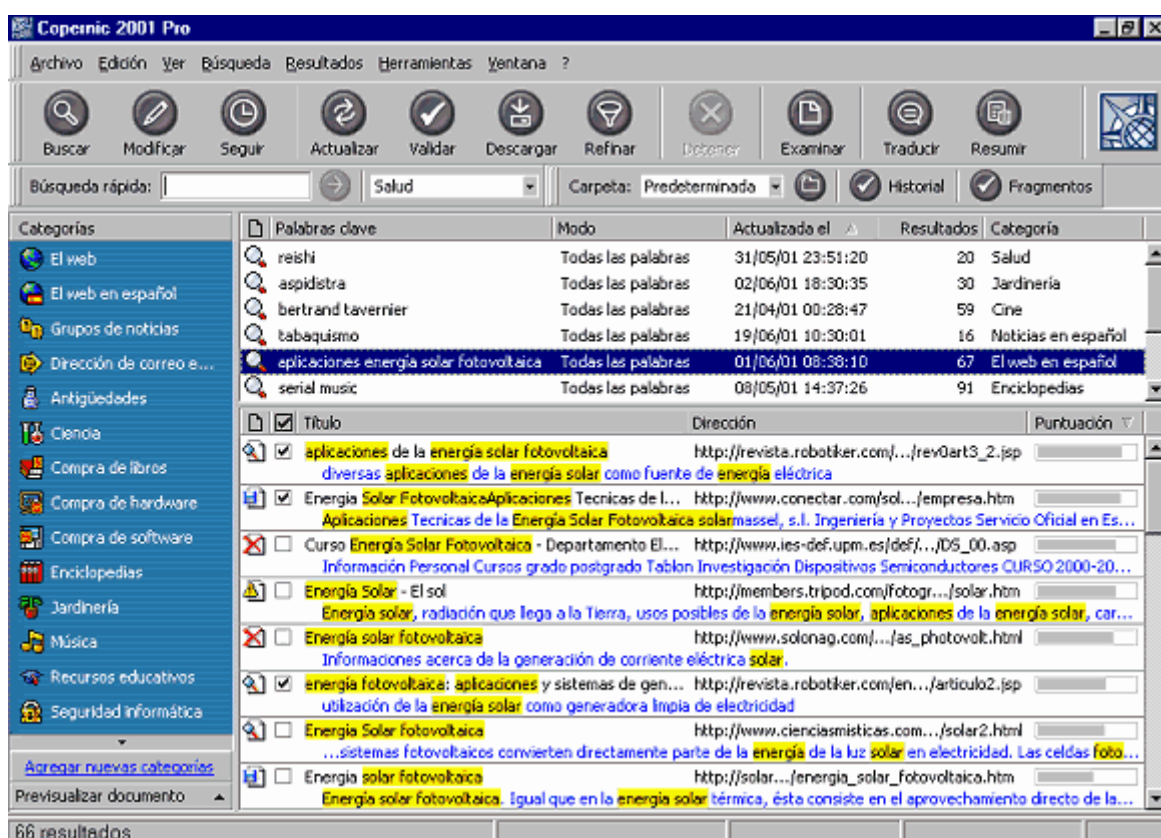
Algunes són multidisciplinàries (Ex.: *Web of Knowledge*). Altres són especialitzades en temes concrets (Ex.: *World Textiles*, especialitzada en indústria i tecnologia tèxtil; *Urbadoc*, especialitzada en urbanisme...). Utilitzeu les bases de dades bibliogràfiques quan vulgueu trobar informació actualitzada i especialitzada del vostre tema en revistes generals, revistes científiques, diaris, congressos, etc.

Direct Search www.freepint.com/gary/direct.htm	Elaborada per Gary Price, autor de "List of Lists". Es una col·lecció de més de 800 llocs, amb opció de cerca, que conté una àmplia gama de temàtiques.
A Collection of Search Engines www.leiden.edu	Una base gratuïta de bibliografies i bases de dades sobre bibliografia.
Complete Planet www.completeplanet.com	Descobreix més de 70.000 bases de dades possibles.
Librarians' Index to the Internet www.lii.org	<i>Librarians' Internet Index</i> (LII) és un web on setmanalment pots veure i rebre novetats sobre bases de dades..
Infomine infomine.ucr.edu/Main.html	Col·lecció de més de 20.000 enllaços acadèmics comentats sobre temes de diferents disciplines universitàries.
About.com www.about.com	Col·lecció acadèmica de webs sobre diversos temes, cada un supervisat per un "guia". La qualitat d'aquest directori no és uniforme, però val la pena consultar-lo. S'hi pot comprar i hi ha un xat; tot plegat està presentat amb un èmfasi informal, que pot resultar empador.
Internet Scout Report Archives http://scout.wisc.edu/archives/	Conté els arxius de la col·lecció més antiga dels webs seleccionats.
Martindale's The Reference Desk http://www.martindalecenter.com/	Col·lecció que utilitza multimèdia. S'hi poden trobar enllaços per a ciències de la salut.
Price's List of Lists http://www.specialissues.com/lol/	Centenars dels "millors", dels "més" i altres llistes i estadístiques, algunes d'útils i altres de trivials.
Computer-Mediated Communication (CMC) Information Sources www.december.com/cmc/info	També coneguda com "La llista de desembre", una àmplia col·lecció d'enllaços sobre ordinadors, Internet i tecnologia de les comunicacions. Inclou premsa, webs, projectes i més coses.
Suite 101.com www.suite101.com	És un ampli directori en format de portal, construït per 1.000 experts voluntaris. No cal ser-ne membre per accedir a la majoria de les pàgines. La qualitat no és uniforme, però és excel·lent i acadèmica, depenent del tema. Feu clic en l'autor d'una secció per veure la seva preparació professional.
Wikipedia www.wikipedia.org	La base de dades més gran del món. Està en 236 idiomes i qualsevol persona pot editar els seus articles. Com qualsevol hi pot escriure té problemes de fiabilitat.

FenixDoc http://biblioteca.upc.es/Fenixdoc/	La base de dades catalana més àmplia sobre producció científica. Pertany a la UPC.
Findarticles: http://findarticles.com/	Articles publicats en més de 3000 diaris i revistes.

- El cercador intel·ligent.

Copernic, producte de l'empresa canadenca *Copernic Technologies*, és un agent cercador que transmet una equació de cerca a un conjunt de cercadors de manera simultània, en recupera les referències pertinents i les ordena segons el grau mitjà de rellevància obtingut de cadascun dels cercadors. L'empresa creadora de *Copernic* es va constituir l'any 1996 amb el propòsit de desenvolupar solucions innovadores en tecnologies d'agents de cerca per tal d'aconseguir un accés i una gestió eficients de la gran massa d'informació present a Internet. El programa es va desenvolupar l'any següent en una versió experimental, i es va posar a l'abast del públic el novembre de 1998. Des d'aleshores s'han llançat versions amb elements nous, la majoria dels quals, però, només estan disponibles en les edicions comercials. L'edició gratuïta ofereix els serveis bàsics de consulta i, amb la intenció de persuadir els possibles compradors de les edicions comercials, dóna una idea força aproximada de les possibilitats del programa complet.



Copernic Basic, l'edició gratuïta, permet consultar simultàniament prop de 80 cercadors importants agrupats en set categories: "El web" més una categoria opcional relacionada amb un idioma o un país, "Grupos de noticias", "Direcciones de correo electrónico", "Compra de libros", "Compra de hardware" i "Compra de software". Malgrat que és l'edició reduïda d'un programa comercial, ofereix tantes o més possibilitats que les versions completes d'altres cercadors múltiples com ara *Lexibot*, *NetAttaché Pro* o *Internet EZ Search*.

El programa pot obtenir-se en www.copernic.com

PORTALS TEMÀTICS

Els portals es poden definir com llocs web que inclouen una base de dades de recursos electrònics classificats per tema. D'altra banda, molts d'ells es caracteritzen per oferir altres serveis en línia al seus usuaris com, per exemple: cercadors, grups de discussió, correu electrònic, notícies, fòrums, xat, llistes de distribució, compres, receptes, música, jocs... Aquests serveis sorgeixen com una alternativa als catàlegs i a les bases de dades per organitzar i recuperar els recursos electrònics d'accés remot.

Característiques

- Inclouen bases de dades de recursos electrònics d'accés obert i de pagament.
- Mostren els recursos electrònics organitzats per matèries.
- Ofereixen una sèrie de serveis en línia d'interès per als seus usuaris (cercadors, grups de discussió, correu electrònic, notícies, fòrums, xat, llistes de distribució, compres, receptes, música, jocs...)
- Molts d'ells permeten cercar recursos electrònics navegant per l'estructura jeràrquica de temes, de forma que en seleccionar un tema et mostra un altre grup de temes més específics que contenen, cadascun d'ells, recursos electrònics de l'especialitat.

Utilitza els portals temàtics...

- Per cercar recursos electrònics (llocs web, associacions, documents diversos...) especialitzats i de qualitat;
- Per accedir a un llistat seleccionat de recursos sobre una matèria;
- Per contactar amb altres persones de l'especialitat mitjançant llistes de distribució, directoris, fòrums, etc.;
- Per accedir a altres serveis relacionats amb el teu àmbit de cerca (agenda de cursos, congressos, novetats, etc.).

Inconvenients dels portals

- Obtenim menys resultats perquè la base de dades és més petita que la d'un cercador automàtic.
- No permeten cercar el contingut de les pàgines d'un lloc web, perquè només n'indexen la pàgina principal.
- Estan menys actualitzats perquè el procés de selecció i descripció dels recursos electrònics és manual.
- Donat que la descripció i el manteniment de la base de dades és manual, comporten un major cost de creació i manteniment.
- Les cerques són més complexes ja que requereixen diversos passos, però poden ser més precises sota termes específics.

Els portals actuals es poden dividir en dos tipus, de caràcter general (*Gateways* o *passarel·les*) o de caràcter temàtic (*Virtual libraries*). Mentre que els portals general són portals destinats al públic en general i acostumen a ser de baixa qualitat, els portals temàtics estan destinats a especialistes i disposen de continguts d'alta qualitat.

A continuació us dono diferents portals temàtics agrupats per categories temàtiques:

Genèrics:	-eresmas: www.eresmas.com -Terra: www.terra.es -World Online : www.worldonline.net -Ya: www.ya.com -Portal de Lleida: www.lleida.org
Pluridisciplinaris	- BUBL: http://bubl.ac.uk - Infomine: http://infomine.ucr.edu/
Ciència i tecnologia	- Chemedica: http://www.chemedia.com/ - Science Gateway: http://www.sciencegateway.org/ - Tecnociencia: http://www.tecnociencia.es

Arquitectura	- SoloArquitectura: http://www.soloarquitectura.com - Arquitectura: http://www.arquitectura.com - Construred: http://www.construred.com/ - Arq.com: http://www.arq.com.mx
Enginyeria	- The Engineer online: http://www.theengineer.co.uk - Intute: http://www.intute.ac.uk/sciences/ - eFunda: http://www.efunda.com/ - Engineers Edge: http://www.engineersedge.com/
Enginyeria aeronàutica	- AERADE: http://aerade.cranfield.ac.uk/
Electrònica	- Electrical and Electronics Engineering Virtual Library : http://webdiee.cem.itesm.mx/wwwvlee/
Enginyeria informàtica	- Intute Science, Engineering and Technology http://www.intute.ac.uk/sciences/
Enginyeria mecànica	- ICrank.com http://icrank.com/
Química	- Chemdex http://www.chemdex.org/ - ChemIndustry http://chemindustry.com/ - ChemSpy.com http://www.chemspy.com/
Telecomunicacions	- AboutAI.net http://www.aboutai.net/ - Goodman's Marks http://www.gbmarks.com/index.html - Nanotechnology-L http://www.nano.org.uk/links.htm
Física	- Physicsworld http://physicsworld.com/cws/home
Matemàtiques	- The Math Forum http://mathforum.org/library/ - The Mathematical Atlas http://www.math.niu.edu/~rusin/known-math/ - Math on the web http://www.ams.org/mathweb/ - Math-Net http://www.math-net.de/
Medi ambient	- Ciencias ambientales http://www.cienciasambientales.com/ - Earth Trends http://earthtrends.wri.org/ - ELDIS http://www.eldis.org/ - EnviroLink http://www.envirolink.org/
Medicina	- Intute. Health and Life Sciences http://www.intute.ac.uk/healthandlifesciences/medicine/
Literatura	- Lletra: http://lletra.uoc.edu/ - PLEC: http://www.plec.es/ - Premis literaris: http://www.epdlp.com/framesvpremios.php - Va de llibres: http://www.edu365.cat/vadellibres/ - XTEC juvenil: http://www.xtec.es/recursos/lit_inf/index.htm
Ètica	- Xarxanet: www.xarxanet.org

Esports	-TV3: http://www.elsesports.cat/ - LFP: www.lfp.es -CSD: http://www.csd.gob.es/
---------	---

QUÈ ÉS INTERNET INVISIBLE?

Fins ara hem parlat de la part visible d'Internet o, el que és el mateix, allò que podem trobar directament a través de les diferents eines de cerca d'informació. Però existeix un gran volum d'informació emmagatzemada en bases de dades (que solen ser d'universitats i centres d'investigació) i que no poden trobar els motors de cerca. Aquesta informació no accessible a través de cercadors és el que s'anomena Internet Invisible.

Existeixen serveis que funcionen com a porta d'accés a aquestes bases de dades. Permeten realitzar consultes sobre els termes a cercar i des de la seva pàgina ens dirigeixen a la base de dades que conté informació sobre el tema que hem demanat.

Aquí teniu alguns exemples:

- www.internetinvisible.com (Accés a bases de dades)
- www.completeplanet.com (Accés a bases de dades)
- www.incywincy.com
- infomine.ucr.edu
- www.intute.ac.uk/
- www.freepint.com/gary/direct.htm
- www.archive.org
- library.rider.edu/scholarly/rlackie/Invisible/Inv_Web.html

BUSCADORS ESPECIALITZATS

- Buscadors especialitzats en Cursos i Formació
 - <http://www.educaweb.com/>
 - <http://www.emagister.com/>
 - <http://www.infocurso.com/>
 - <http://www.aprendemas.com/>
 - <http://www.monografias.com/>
 - <http://www.lafacu.com/>
- Universitats:
 - <http://www.universia.net/>
- Salut:
 - <http://www.buscasalud.com/>
 - <http://www.buscamed.com/>
- Biografies:
 - <http://www.buscabiografias.com/>
- Cine:
 - <http://www.buscacine.com/>
- Empreses:
 - <http://www.portalempresas.com/>
 - <http://www.elcorredor.com/>
 - <http://www.guiacom.es/empresas.asp>
 - <http://www.pymesite.com/empresas/>
 - <http://www.guiame.net/>
 - <http://www.paginas-amarillas.es>
 - <http://www.bigyellow.com/>

- Treball i Ocupació:
 - <http://www.trabajos.com/>
 - <http://www.trabajofacil.com/>
 - <http://www.jobpilot.com/>
 - <http://www.laboris.net>
 - <http://www.todotrabajo.com/>
 - <http://www.infojobs.net>
 - <http://www.tecnoempleo.com/>
 - <http://www.oficinaempleo.com>
 - <http://www.monster.es/>
 - <http://www.bolsadetrabajo.com>
 - <http://www.empleo.com.ar/>
 - <http://www.canaltrabajo.com/>
 - <http://www.trabajo.com.mx/>
- Recursos gratis. Programes
 - <http://www.recursosgratis.com/>
 - <http://www.logratis.com/>
- Llistat de Buscadors.
 - <http://buscadores.buscopio.com>
- Esports:
 - <http://www.buscadeporte.com>
- Finances i borsa
 - <http://www.buscafinanzas.com/>
- Legislació
 - <http://www.todalaley.com/>
 - <http://www.jurisweb.com/legislacion/>
 - <http://www.laley.net/>
 - <http://www.vlex.com>
- La legislació de la Unió Europea
 - <http://europa.eu/>
- Administracions públiques
 - <http://www.admiweb.org/>
- Política
 - <http://www.buscadorpolitico.com.ar/>
- Llibres
 - <http://www.amazon.com>
 - <http://www.loslibros.com>
 - <http://www.booksfactory.com/>
- Fotografia
 - <http://www.fotoportal.net/>
- Música en format Mp3
 - <http://mp3.xasa.com/>
 - <http://www.mp3hispania.com/>
 - <http://www.elmusical.com/search.html>
 - <http://www.logratis.com/buscamp3.shtml>

- Literatura
 - <http://www.elaleph.com/buscador/>
 - <http://cervantesvirtual.com/index.shtml>
 - http://www.cervantesvirtual.com/bib_mundo/biblioteca_mundo.shtml
- Turisme
 - <http://www.infocamping.com/>
 - <http://www.buscaturismo.com/>
- Jocs
 - <http://www.solojuegos.com/>
 - <http://www.dondejugar.com/>
- Diccionaris
 - <http://www.diccionarios.com/>
 - <http://www.elmundo.es/diccionarios/>
- De qui és un domini o nom de pàgina
 - <http://cgi.elmundo.es/whois/cgi/whois.html>
- Notícies
 - <http://www.titulares.com/>
 - <http://www.elmundo.es/hemeroteca/>
 - <http://www.noticias.com/>
- Diaris i revistes
 - <http://www.kiosco.net/>
- Cites. Articles científics
 - <http://citeseer.nj.nec.com/cs>
- Informàtica
 - <http://www.canalsw.com/>
 - <http://www.trucostecnicos.com/>
 - <http://www.tucows.com/>
 - <http://www.lawebdelprogramador.com/>
 - <http://www.programacion.com/>
- Articles en revistes
 - <http://ask.elibrary.com/>