



Número 5, desembre 2000

Tipologia documental citada en tesis doctorals d'informàtica: bases empíriques per a la gestió equilibrada de col·leccions

[\[versió castellana\]](#)

[Cristóbal Urbano](#)

Facultat de Biblioteconomia i Documentació

Universitat de Barcelona

urbano@bd.ub.es

Resum

Anàlisi de citacions en tesis doctorals d'informàtica de la Universitat Politècnica de Catalunya (UPC), presentades entre 1996–1998, per determinar l'ús d'informació segons la tipologia i el suport documental. Com a punt de partida per al disseny del projecte desenvolupat, s'analitzen els fonaments del mètode emprat en vista d'una revisió de la bibliografia sobre desenvolupament de col·leccions i estudis d'usuaris. Les dades obtingudes entre els doctorands de la UPC són consistents amb les característiques apuntades en la bibliografia sobre els investigadors de l'àrea de ciències de la computació: una elevada obsolescència de la bibliografia professional, el control de la bibliografia mitjançant les citacions en les lectures efectuades, la prioritat per les fonts més pròximes i accessibles, la consulta de cercadors de recursos web, la visita a llocs web d'investigadors i institucions de solvència en el seu terreny, i el contacte amb els col·legues mitjançant la comunicació electrònica, configuren un ús per tipus de documents que fa prevaler els congressos i la literatura grisa en més gran mesura que en altres àrees científicotècniques.

1 Introducció

Amb freqüència, en l'àmbit de les biblioteques universitàries i d'investigació la gestió de col·leccions es troba excessivament focalitzada en el procés de selecció de títols de revistes. La indubtable importància d'aquest tipus de documents en el sistema d'informació de la ciència i la tecnologia, el plus de garantia de qualitat que tradicionalment s'assigna als treballs publicats en revistes, l'elevat cost d'adquisició de gran part de l'edició científica de major impacte, així com la vinculació entre avaluació de la producció científica i visibilitat dels treballs publicats en revistes, poden arribar a ser factors que distorsionin l'equilibri pressupostari ideal entre les diverses classes de documents que cal adquirir en una biblioteca. La pregunta sobre a quins títols de revistes s'ha de subscriure una determinada biblioteca ha estat, sovint, objecte de més atenció que la pregunta prèvia sobre quin percentatge del pressupost s'havia de dedicar a cada categoria documental o, si es prefereix situar el tema més enllà de les adquisicions, quina atenció s'havia de dedicar al control bibliogràfic dels diversos formats i suports documentals. Precisament, el major desenvolupament dels instruments de control bibliogràfic de les revistes científiques ha estat un dels factors que pot haver col·laborat en més gran mesura a un possible sobredimensionament d'aquests documents en els debats professionals sobre desenvolupament de col·leccions. Com a prova d'això, valgui el següent paràgraf en el qual es contraposa l'adquisició de revistes a la de documents secundaris per al control de les revistes:

"The director of a well known university library, when asked why he had not subscribed to the *Science citation index*, answered as follows. 'I think the *SCI* is a fine retrieval tool. But I asked the faculty members — which would you prefer, the *SCI* or 40 new journal subscriptions? We have finite budget and choices must be made.' In reply to this seemingly reasonable approach, I asked, 'Which 40 journals?'" (Garfield, 1968).

Certament, la pregunta de Garfield era, i continua sent, plenament pertinent. De tota manera, davant les grans transformacions que Internet està introduint en els mecanismes de comunicació científica, sembla adequat tornar a reflexionar sobre la ràtio revistes/monografies en els pressupostos d'adquisicions des d'un context més ampli. Això és, a més de planificar la distribució del pressupost d'adquisicions segons la tipologia documental i per títols concrets de sèries, potser cal començar a plantejar-se quins han de ser els recursos econòmics destinats a facilitar el control i l'accés a la informació, independentment de si es tracta de recursos que cal adquirir de forma permanent, o no.

Per tant, conèixer què usen realment els investigadors d'una determinada àrea de coneixement en una universitat pot ser una dada de gran interès en el marc dels processos de redefinició de les polítiques d'adquisicions que exigeix el nou entorn digital. La constatació empírica de l'ús d'informació des d'un enfocament local, juntament amb altres elements més globals sobre valoració de les fonts d'informació, és una condició necessària per dur a terme una gestió de col·leccions centrada en els usuaris.

En aquest sentit, en el nostre treball es vol conèixer l'ús d'informació dels investigadors en informàtica de la Universitat Politècnica de Catalunya amb la finalitat d'obtenir informació empírica de base amb la qual es puguin alimentar els processos relacionats amb la gestió de les col·leccions de la biblioteca. Aquest article forma part d'un estudi més ampli en el qual s'analitzà en profunditat el mètode de l'anàlisi de citacions en publicacions d'usuaris com a mecanisme per a l'estudi d'usos i necessitats d'informació en biblioteques universitàries (Urbano, 2000), i en el qual es va desenvolupar una aplicació pràctica sobre tesis doctorals en informàtica de la UPC. Tanmateix, s'ha considerat pertinent i d'especial valor presentar de manera independent en aquesta contribució a la revista *bid* les dades obtingudes sobre la tipologia documental utilitzada pels usuaris.

2 Fonaments i antecedents

2.1 Anàlisi de citacions en publicacions d'usuaris i gestió de col·leccions

En la gestió de col·leccions són de gran utilitat tant els estudis d'ús d'informació com els de necessitats, i encara que s'han de considerar de naturalesa diferent estan íntimament relacionats en l'àmbit que ens ocupa. Els primers tenen en essència un caràcter retrospectiu, de descripció d'allò que ha succeït; els segons persegueixen anticipar les possibles demandes dels usuaris. No obstant això, són molts els autors que vinculen ús i necessitats d'informació basant-se en el principi que "els usos efectuats en el passat són una bona font d'informació per pronosticar l'ús present o futur" (Baker i Lancaster, 1991, p. 82). Si bé es pot considerar com un principi discutible, el que no ofereix cap dubte és que els usos efectuats serveixen, com a mínim, per detectar si un determinat grup d'usuaris ha evolucionat en els seus hàbits, ja sigui a causa de canvis en l'oferta de recursos i serveis de la biblioteca, ja sigui per canvis en les seves característiques com a investigadors. Així doncs, l'evolució de l'ús és de gran interès en l'estudi de l'evolució de les necessitats.

En aquest sentit, en aquest treball s'ha optat per l'anàlisi de citacions en publicacions d'usuaris com a mètode per determinar-ne l'ús d'informació, i considerant que els mètodes habitualment utilitzats en l'avaluació de l'ús de col·leccions estan excessivament limitats per la condició de l'accés físic, i que difícilment mesuren l'ús que els lectors realitzen amb fonts externes a la col·lecció. En efecte, la majoria de manuals i directius sobre l'avaluació bibliotecària (Lancaster, 1996; Poll i Boekhost, 1998 entre d'altres) es recolzen en una definició d'ús centrada en la col·lecció de la biblioteca i en el recompte dels usos que impliquen un moviment físic dels documents. Aquest enfocament és molt objectiu i, conscientment, es limita a valorar l'ús concret de la col·lecció, però no entra en la investigació de l'ús d'informació que els usuaris locals realitzen i que no necessàriament es manifesta de forma única en l'ús de la col·lecció.

D'altra banda, si es vol realitzar un estudi per mètodes indirectes sobre la distribució de l'ús entre les diverses classes de documents utilitzades per un grup d'usuaris, l'anàlisi de citacions en les seves publicacions és, sens dubte, una de les opcions més eficaces. Malgrat les discussions que es plantegen entorn de si citar un document indica la seva lectura real, s'entén que, com a mètode de mesurar l'ús, el recompte de les citacions rebudes pels diversos tipus de documents és més adequat que el recompte dels documents mobilitzats en una sala de lectura quan els usuaris estudiats són investigadors. En definitiva, es pretén obtenir una imatge de l'ús d'informació que realitza l'usuari i no necessàriament de l'ús realitzat a partir de la col·lecció.

En el cas concret del recompte de referències en les bibliografies d'usuaris de biblioteques, com que es tracta d'un mètode que estudia l'ús d'informació citada, però no necessàriament consultada a la biblioteca des de la qual es realitza l'estudi, podria dir-se que es tracta d'un mètode que parcialment permet estudiar tant l'ús com les necessitats: les fonts citades i no disponibles en la col·lecció de la biblioteca de referència per a l'usuari són un bon mètode per explorar les demandes no satisfetes per la biblioteca. De tota manera, convé recordar que la validesa de l'anàlisi de citacions en els estudis d'ús d'informació per part d'usuaris de biblioteques és objecte d'un llarg debat en el qual es poden distingir tres posicions. Per alguns autors les dades de citació de col·lectius nacionals, o que afecten tota la bibliografia que circula sobre una matèria, són molt més fàcils d'obtenir i normalment igual, o més significatives, que les dels usuaris per establir les seves necessitats d'informació (Broadus, 1985a i 1985b; Lancaster, 1996). No obstant això, les objeccions a la validesa de dades externes enfront de les dels usuaris locals són nombroses (Line, 1977 i 1985; Joswick i Stierman, 1997), i això explica que, per molts autors, l'ús de publicacions dels usuaris en combinació, o no, amb altres fonts de citacions externes sigui imprescindible per a l'aplicació del mètode de l'anàlisi de citacions en estudis locals de necessitats i ús de fonts d'informació bibliogràfiques¹ (entre d'altres: McCain i Bobick, 1981; Buchanan i Herubel, 1994; Jiménez Contreres et al., 1994; Hughes, 1995; Sylvia i Leshner, 1995; Zipp, 1996; Hurd, Blečić i Vishwanathan, 1999).

Finalment, altres autors destaquen que l'anàlisi de citacions, especialment quan es basa en dades generals i no en documents font dels usuaris, manté un nivell baix de correlació amb l'ús real mesurat directament a les biblioteques (Scales, 1976; Satariano, 1978; Line, 1985a) i, per tant, comparativament defensen l'avaluació directa de l'ús a la sala, l'anàlisi de les estadístiques de circulació, o fins i tot, el "judici expert" dels bibliotecaris (Pan, 1978).

Però en definitiva, com diu Broadus (1977, p. 309), hi ha una opinió general que considera que sempre que es treballi amb grans quantitats de citacions,² el seu recompte i anàlisi amb finalitats bibliotecàries té un considerable valor i fiabilitat (Linda C. Smith, 1981, p. 84), especialment si es tenen en consideració dades locals com a complement d'estudis de més gran abast. Així, per exemple, Sylvia i Leshner (1995) reconeixen la temptació d'usar estudis previs i les dades generals de les grans bases de dades de citacions, però alerten de les seves limitacions:

"Citation analysis has been done through nationally collected citation statistics, and proposals have been made to use these analyses for library collection and evaluation. The easiest

approach is to simply look at previous studies and use the findings to govern one's own collection development strategies. There are, however, some very strong arguments for doing a local study. Maurice B. Line argued that every library has a differing clientele, and that clients have different information requirements so that national surveys may have little relevance for individual libraries." (Sylvia i Leshner, 1995, p. 314)

De la seva banda, Zipp (1996), en un treball sobre la validesa de les tesis doctorals com a font per estudiar l'ús d'informació del conjunt dels investigadors d'una universitat, afirma en comparar l'anàlisi de citacions amb els estudis d'ús real per mitjà d'estadístiques de préstec o similars:

"Studies of circulation data and in-library use contribute to our understanding of what parts of a library collection are being used. The link between demand for research materials and use is most easily achieved by direct analysis of the scholarly communication process. Evaluation of the titles used in the process of discovery or problem identification yields valuable information. The titles in which researchers publish and those that they cite in their works also serve to elucidate real and potential demand on library collections." (Zipp, 1996, p. 335)

De tota manera els resultats d'estudis de citacions d'abast local també s'han de llegir amb cautela i tenint en compte els seus punts dèbils. Lancaster (1996) els resumeix alertant fonamentalment sobre tres aspectes: no totes les tipologies documentals són igualment citables, les col·leccions personals poden distorsionar la imatge de l'ús real de la col·lecció, i la més gran accessibilitat de la col·lecció de la biblioteca respecte a publicacions que no hi són localitzades pot generar un biaix de les citacions cap a la col·lecció. Un altre dels punts clau en la discussió metodològica sobre l'anàlisi de citacions en publicacions d'usuaris de biblioteques és la selecció dels documents font per analitzar. En aquest sentit, les tesis doctorals es consideren en nombrosos estudis una font molt significativa per a l'anàlisi de les necessitats dels investigadors en general. Són nombrosos els autors que consideren explícitament que l'anàlisi de les tesis pot ser una font idònia per obtenir dades del consum d'informació del conjunt d'investigadors d'un departament, estudiants de tercer cicle, becaris d'investigació i professorat ajudant inclòs. Així l'estableixen autors com ara McCain i Bobick (1981), Zipp (1996) i Barry (1999), qui bé per mètodes estadístics, o bé per mitjà de sòlides argumentacions, afirmen que el consum dels estudiants de doctorat és molt representatiu de les necessitats i ús d'informació del conjunt de personal docent i investigador en un entorn universitari. En la mateixa línia Walcott (1994) observà que els estudiants de tercer cicle eren els que utilitzaven més intensament les publicacions seriades a la biblioteca de biologia de la State University of New York. D'altra banda, en contrast amb la tendència dels articles de revista a de recollir principalment citacions a altres articles de revista, les tesis contenen una més gran varietat i quantitat de documents citats. Per tant, la seva utilització com a document font per a l'anàlisi de citacions permet una representació millor de la varietat de fonts consultades pels usuaris d'una biblioteca. El fet que una gran fracció de la bibliografia científica no sigui citada en les revistes buidades en els índexs de citacions de l'ISI no significa necessàriament que no sigui llegida, sinó que per les limitacions de nombre de pàgines i de referències que marquen les pautes de publicació de les revistes científiques, no són efectivament citades.

2.2 La ràtio monografies / publicacions en sèrie en la política d'adquisicions

Conèixer la varietat i distribució dels documents citats és molt útil en una biblioteca. El problema de les adquisicions en les biblioteques no es limita a una selecció de títols, sinó que hi ha un problema previ que consisteix a decidir l'equilibri que una col·lecció ha de mantenir entre llibres, revistes, congressos, etc. Des del punt de vista del servei d'informació bibliogràfica i de l'adquisició de documents secundaris, l'interès dels usuaris per uns o altres tipus de documents condicionarà la intensitat en l'ús dels instruments de control bibliogràfic. D'altra banda, una de les formes tradicionals de caracteritzar els usuaris ha estat mitjançant el percentatge d'ús de cada una de les categories documentals. Així, en ciències aplicades i en tecnologia els percentatges de citació de congressos, patents o normes són més grans que en ciència bàsica, en la qual el domini de la revista científica és molt més gran que en la resta d'àrees del coneixement.

L'avaluació de la col·lecció també comporta la determinació de la ràtio ideal entre el pressupost dedicat a revistes i el dedicat a monografies. Aquest és un dels territoris en els qual els resultats de l'anàlisi de citacions amb enfocament local són més satisfactoris. Es tracta d'un problema clàssic amb el qual s'han d'enfrontar les biblioteques universitàries, difícil de resoldre per les implicacions que té en la definició de les partides pressupostàries d'una biblioteca. Amb freqüència es traslladen de manera mecànica percentatges de citació de cada una de les categories documentals observades en estudis generals a percentatges de participació en el pressupost d'adquisicions, sense atendre la demanda del conjunt de la clientela d'una biblioteca, ni els seus hàbits concrets i detallats.

En aquest sentit, el treball de Devin i Kellogg (1990) sobre la ràtio revistes/monografies observada en un total de 70 estudis d'ús de col·leccions resulta molt clarificador: els resultats obtinguts en els diversos estudis revisats eren molt variats en funció de les matèries, les institucions, el període analitzat, o la definició de la tipologia documental³ sobre la base de la qual es realitzava el recompte. Una cosa similar es podria observar analitzant la gran varietat que mostren les diverses institucions espanyoles en relació amb aquest indicador en les estadístiques anuals sobre biblioteques universitàries que publica Rebiun (*Anuari estadístic...*, 2000, p. 67). No obstant això, el que no es pot posar en dubte en el desenvolupament d'una col·lecció de suport a la recerca, concebuda per tant des d'una lectura segregada dels pressupostos, és que ha d'existir una certa correlació entre els percentatges d'ús dels diversos tipus de materials i els corresponents a les partides pressupostàries per a cada categoria documental. Així ho conclouen Devin i Kellogg:

"The authors believe that it is possible to determine guidelines for developing a serial/monograph ratio for research library budgeting based on objective criteria. We propose

that, for a research library, the percentage of serials and monographs purchased should correlate with the characteristics of literature usage in each subject area. [...] In conclusion, we find that there is no one optimum serial/monograph ratio for all research libraries. Rahter, the ratio for each library will vary depending on its own priorities and emphases. For each subject, however, the serial/monograph ratio should be based on the use of the literature by researchers in that subject area as determined by citation studies." (Devin i Kellogg 1990, p. 54)

Per això el millor mètode del qual es disposa és l'anàlisi de les citacions en treballs dels usuaris i el contrast de les dades obtingudes en estudis de citacions d'abast general. Un exemple d'una reassignació pressupostària en funció de l'anàlisi de citacions en tesis d'estudiants d'enginyeria es pot trobar al treball de Kritz (1978), qui observà un nivell d'ús dels llibres superior a l'inicialment previst i actuà incrementant la partida destinada a la seva adquisició en detriment de les revistes.

2.3 Ús d'informació en l'àrea d'informàtica

Amb l'elecció de la Universitat Politècnica de Catalunya i de l'àrea d'informàtica per a la realització del treball es pretén realitzar una contribució a l'estudi de dos àmbits en els quals la bibliografia analitzada mostra un nombre reduït de treballs:⁴ els informàtics com a usuaris de biblioteques i els estudis biomètrics en l'àrea de la informàtica.

D'alguna manera resulta sorprenent el poc interès que els investigadors d'aquesta àrea han despertat entre els bibliotecaris i documentalistes, ja que la bibliografia especialitzada en biblioteconomia i documentació no els ha considerat sinó dins d'estudis genèrics de la ciència i la tecnologia: si bé es poden trobar estudis d'usuaris en els quals els informàtics són estudiats juntament amb altres usuaris, són pocs els treballs que se'ls dediquen de manera monogràfica.

En general, tots els treballs coincideixen a dibuixar un perfil similar dels investigadors de l'àrea: baix nivell d'ús de les bases de dades bibliogràfiques; preferència per les col·leccions personals de documents enfront dels fons d'una biblioteca; tècniques d'interrogació a bases de dades basades majoritàriament en la combinació de paraules clau en llenguatge lliure; control de la bibliografia mitjançant les citacions en les lectures efectuades, el rastreig⁵ de les fonts més pròximes, o el recurs a mitjans electrònics poc formalitzats bibliogràficament com ara la consulta de cercadors de recursos web, la visita a llocs web d'investigadors i institucions de solvència en el seu terreny, i el contacte amb els col·legues mitjançant la comunicació electrònica (correu electrònic, llistes de distribució i grups de notícies).

En definitiva, i prenent com a referència el treball de Cunningham i Connaway (1996), els informàtics són molt selectius en la quantitat d'informació "bibliogràfica" que consumeixen, encara que no en termes de qualitat i pertinència, sinó en termes d'accessibilitat i integració al seu entorn electrònic de treball. Quant a les fonts documentals primàries més utilitzades, destaca l'important paper atribuït a les comunicacions de congressos i els informes distribuïts per la web, no tant perquè es considerin fonts de tota solvència, sinó per la seva gran actualitat enfront dels lents mecanismes de publicació de les revistes científiques.

Un altre dels punts que ha motivat el nostre interès és la doble aplicació que el treball podia tenir en l'àmbit bibliotecari i en el de l'avaluació de la producció científica. Després de la lectura de documents referits al sistema d'avaluació de la UPC, hem observat que els departaments de les àrees d'arquitectura de computadors i de llenguatges i sistemes informàtics atorguen una gran importància als congressos com a forma de publicació rellevant, i altament visible, en un possible sistema d'avaluació de la producció científica de la universitat. Per això aquest departament reclamaren de la Comissió de Publicacions Notables del Vicerectorat d'Investigació que en aquestes àrees es tinguessin en consideració, tant o més que certes revistes, determinats congressos de prestigi. La Comissió és la responsable de la determinació d'una llista de publicacions complementària al *Journal Citation Report* que permeti una avaluació qualitativa i no merament quantitativa de les publicacions generades per a cada investigador (UPC. Vicerectorat de Recerca 1997 i 1999). Per tant, comprovar si realment els congressos jugaven el paper que se'ls assignava des dels departaments semblava un tema de gran interès.

3 Objectius i metodologia

Els objectius concrets que es pretenen són:

1. Obtenir la ràtio corresponent a la proporció entre les citacions a publicacions adquirides per la biblioteca en forma de subscripció i les que reben documents adquirits "com" a monografies. L'obtenció d'aquesta dada s'ha de realitzar en funció del model d'adquisició que actualment porta a terme la biblioteca, i no segons la naturalesa dels diversos tipus de documents.
2. Confirmar les aportacions que els departaments estudiats realitzaren a la Comissió de Publicacions Notables en el sentit de considerar més gran la visibilitat dels congressos en l'àmbit de la informàtica en comparació amb la resta de categories documentals. També es pretén comprovar si l'ús de literatura grisa per part d'aquest tipus d'investigadors és tan elevat com la bibliografia acostuma a descriure.
3. Considerant que el col·lectiu analitzat no presenta problemes significatius d'accés als recursos disponibles a Internet, es pretén quantificar la presència explícita de recursos electrònics i la presència de determinats documents citats en paper, però que en realitat són documents electrònics —com és el cas de bona part de la literatura grisa citada—, o que posseeixen un potencial molt alt de

migració al format electrònic.

Per assolir aquests objectius, enfront de l'alternativa de realitzar un estudi per mètodes directes d'entrevista o enquesta, l'anàlisi bibliomètrica de la bibliografia citada en tesis doctorals ha cridat la nostra atenció perquè es tracta d'un mètode que no interfereix en el comportament dels usuaris mentre es prenen les dades i que permet una disponibilitat relativament senzilla d'aquests: no es requereix la cooperació d'un entrevistat o enquestat, ni la resposta apareix contaminada pel procés d'estudi.

S'ha treballat sobre el total de les tesis del període 1996–1998 presentades al Departament d'Arquitectura de Computadors i al Departament de Llenguatges i Sistemes Informàtics, sense realitzar cap tipus de mostreig. El total de tesis i referències escrutades ha estat de 54 i 6.807 respectivament. Les dades s'han introduït en una base de dades gestionada per un programa relacional a partir de la qual s'han generat resultats de tipus descriptiu i matrius de dades per fer-ne una explotació amb un paquet estadístic.

S'han comptabilitzat les citacions a partir de les llistes de referències que figuren al final de les tesis en forma de bibliografia citada. Es plantejà l'exclusió de les bibliografies que recollissin bibliografia consultada però no citada, però no sobsevà cap cas amb aquestes característiques. També es plantejà l'exclusió d'aquelles tesis que tinguessin la llista de referències segregada i duplicada per capítols sense una llista unificada final de referències citades, però tampoc s'observà cap cas en el conjunt analitzat.

El recompte s'ha realitzat comptabilitzant una vegada cada document citat, entenent per document, en el cas de les sèries, cada un dels articles o les comunicacions que formen part d'una publicació.

Les publicacions s'han considerat en funció de la representació que se'n feia per mitjà de la referència bibliogràfica consignada en la bibliografia. Així, les revistes s'han distingit per anar acompanyades d'una designació de fascicle, o per coincidir amb algun dels títols o abreviatures que figuraven a la base de dades que es tenia en el moment d'iniciar el recompte de citacions.⁸ Per als congressos es comptà, en un inici, amb la llista de congressos notables UPC per alimentar l'aplicació informàtica utilitzada en el recompte, i alhora, en general, la designació cronològica i/o numèrica és ben característica i diferenciada de les revistes. El punt més conflictiu ha estat la consideració que s'havia de donar a determinades sèries monogràfiques que els autors en ocasions citaven com a sèrie, en altres ocasions com a congrés, i també en altres, com a monografia: finalment, en aquelles que ha estat possible, s'han considerat dins de la categoria de monografies.

De tota manera, la falta d'una definició clara i uniforme dels diferents tipus de documents, especialment en el sentit de l'abast donat al terme publicació seriada, és un dels principals problemes a l'hora de contrastar entre els diversos treballs publicats els resultats relatius a percentatges segons la tipologia documental. Així, Walcott (1991) en el seu estudi sobre tesis doctorals de geologia observa un 80% de publicacions seriades enfront d'un 20% de monografies, sense que quedí clar on s'inclouen els congressos citats. D'altra banda, Glänzel i Schoepflin (1999), en un estudi monogràfic sobre la proporció de citacions a sèries en diverses àrees del coneixement en el qual processaren amb ordinador deu milions de referències de les bases de dades de l'ISI corresponents a 1993, consideraren publicació en sèrie totes les referències en les quals figurava un número de volum vàlid i la indicació de primera pàgina de la part component, per tant, sembla que els congressos no entrarien en la seva definició automàtica de publicació seriada que, *de facto*, tenia en compte únicament revistes.

4 Resultats

Els departaments seleccionats han generat al llarg dels tres anys considerats un total de 54 tesis (taula 1). La contribució per part de cada un d'aquests en percentatge sobre el total és bastant equilibrada excepte l'any 1998, en el qual la diferència s'eleva a més de 35 punts, fet que ocasiona que la diferència en el conjunt d'anys s'elevi a 18,5 punts. Al llarg dels tres anys analitzats el nombre de tesis llegides es manté uniforme al Departament d'Arquitectura de Computadors (AC), mentre que al Departament de Llenguatges i Sistemes Informàtics (LSI) es produeix un salt considerable entre el primer any i els dos següents.

Taula 1. Tesis llegides per departament i any de lectura

Departaments	Total	%	1996	%	1997	%	1998	%
AC	22	40,74	7	50,00	9	42,86	6	31,58
LSI	32	59,26	7	50,00	12	57,14	13	68,42
Total	54	100,00	14	100,00	21	100,00	19	100,00

Tabla 2. Total de referències i mitjana de referències per tesis segons departament i any de lectura

Dep.	1996		1997		1998							
	Total Ref.	%	Total Ref.	%	Total Ref.	%						
AC	2485	36,50	113	1.002	51,65	143	1.032	41,20	115	451	19,10	75
LSI	4.322	63,50	135	938	48,35	134	1.473	58,80	123	1.911	80,90	147

Total	6.807	100,00	126	1.940	100,00	139	2.505	100,00	119	2.362	100,00	124
--------------	-------	--------	-----	-------	--------	-----	-------	--------	-----	-------	--------	-----

Tanmateix, si s'atén al nombre de referències totals incloses en les tesis (taula 2) el Departament LSI quasi dobla el Departament AC, amb una diferència percentual de 27 punts. Aquesta diferència s'ha d'imputar tant al major nombre de tesis, com al major nombre de citacions per tesis al Departament LSI, i està en línia amb un major consum d'informació bibliogràfica en l'àrea de llenguatges.

4.1 Tipologia documental

La distribució de les fonts citades (fig. 1) presenta uns resultats que, a primera vista, podrien semblar sorprenents: un 31,9% de les citacions corresponen a comunicacions de congressos, conferències o *workshops*, i aquest tipus de document figura com el més citat per sobre dels articles de revista. En segon lloc, figuren els articles de revista amb un 31,2% i, en tercer, els llibres amb el 13,3%, percentatge que sumat al 6,5% dels capítols de llibre conforma un volum de citacions en aquest tipus de publicació molt elevat pel que és habitual en ciència i tecnologia.

El percentatge de documents citats que la Biblioteca de la UPC adquireix per subscripció és d'un 31,2%. Aquesta xifra s'ajusta amb molta precisió a la distribució actual del pressupost d'adquisicions de la UPC, que, segons les dades que figuren a les estadístiques Rebiu corresponents a l'any 1999 (*Anuari estadístic... 2000*, 67), destina el 64% de la despesa a l'adquisició de monografies, entre les quals s'inclouen les actes de congressos.

Figura 1. Distribució de les citacions segons la tipologia documental

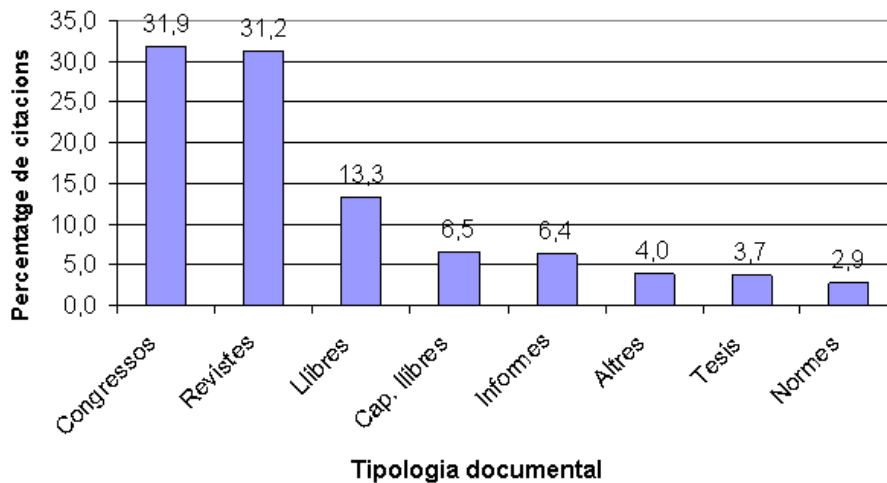


Figura 2. Comparació de l'ús de les diverses categories documentals per part de cada departament una vegada ponderat el pes de cada categoria en el total de citacions de cada departament

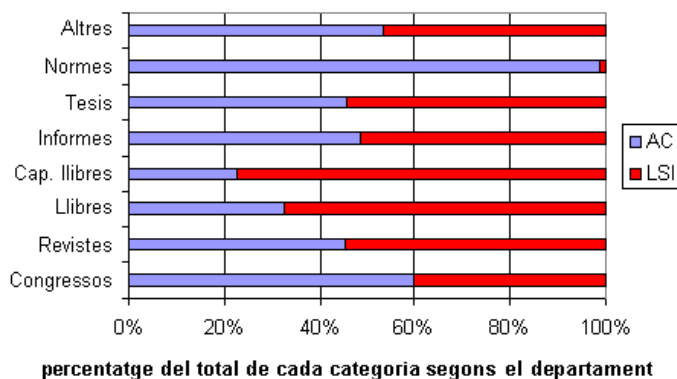


Tabla 3. Tipologia de fonts citades i distribució segons els departaments

Fonts	Total	%	AC	%	LSI	%
Comunicacions a congressos	2.171	31,89	1.000	40,24	1171	27,09

Articles de revista	2.126	31,23	687	27,65	1439	33,29
Llibres	908	13,34	197	7,93	711	16,45
Capítols de llibre	442	6,49	63	2,54	379	8,77
Informes	438	6,43	154	6,20	284	6,57
Tesis	252	3,70	82	3,30	170	3,93
Normes	196	2,88	192	7,73	4	0,09
Inf. comercial i de producte	79	1,16	54	2,17	25	0,58
Preprints i doc. de treball	76	1,12	16	0,64	60	1,39
Manuscrits	53	0,78	8	0,32	45	1,04
Notes tècniques i d'aplicació	21	0,31	12	0,48	9	0,21
Projectes final de carrera	6	0,09	3	0,12	3	0,07
Comunicacions personals	4	0,06	3	0,12	1	0,02
Patents	1	0,01	1	0,04	0	0,00
Sense determinar	34	0,50	13	0,52	21	0,49
Total	6.807	100,00	2.485	100,00	4.322	100,00

La distribució per tipus de document és significativament diferent segons si es tracta d'un departament o d'un altre (fig. 2 i taula 3). La major utilització de fonts d'informació bibliogràfica en l'àrea de llenguatges informàtics concorda amb una utilització de fonts més formalitzades, com ara les revistes científiques: amb un 33,29% els articles de revista són la tipologia documental més citada al Departament LSI, enfront d'un 27,65% en el cas del Departament AC. En canvi el domini dels congressos en les tesis del Departament AC amb un 40,08% de les citacions és molt destacable i concorda amb una àrea de coneixement en la qual l'actualitat de la informació preval sobre qualsevol altra consideració (fig. 3). No s'observen grans oscil·lacions al llarg del temps en el nombre de citacions que rep cada classe de documents (taula 4), excepte en el cas de les normes tècniques, amb una punta fora del normal el 1997, i en el dels preprints el 1998, amb un increment important però dins d'uns límits normals. El cas de les normes s'explica per la distorsió creada per una tesi en la qual aquesta forma documental apareix citada de manera anormalment dominant. Per explicar l'increment dels preprints el 1998 es podria especular amb diversos factors, entre els quals destaca la major i millor distribució que ofereixen a mesura que es consolida Internet, i en particular el web, com a medi habitual de difusió d'aquest tipus de fonts.

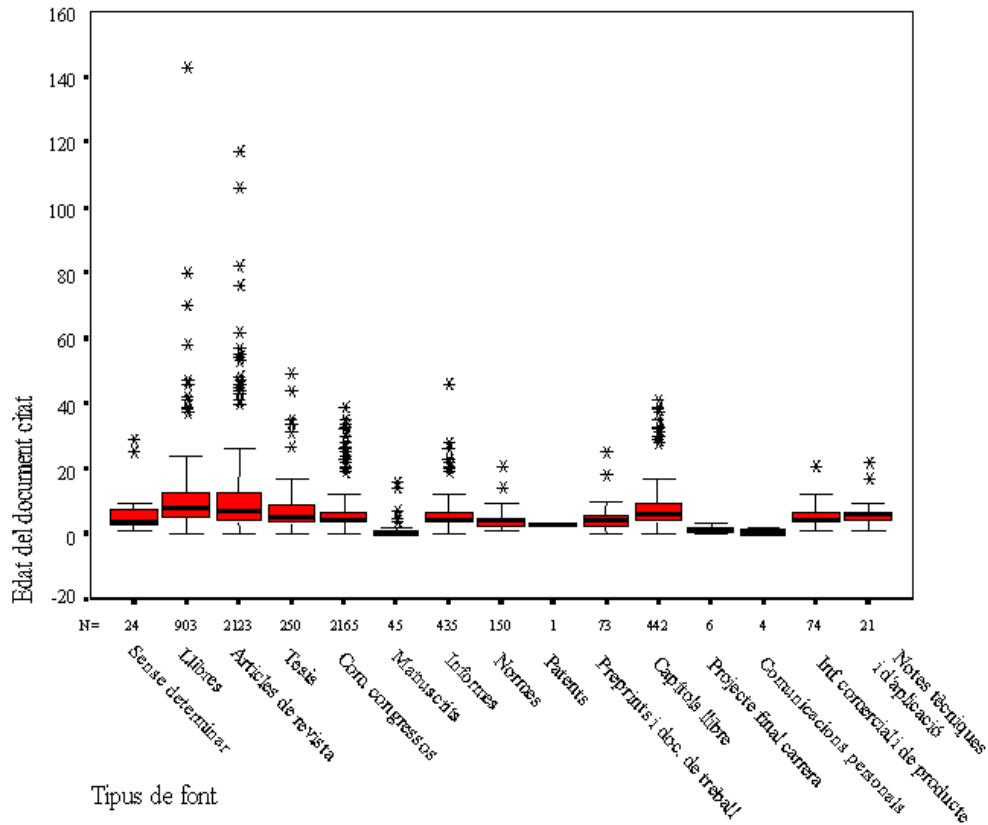
Tabla 4. Tipologia de fonts citades i distribució per anys de lectura de les tesis

Fonts	Total	%	1996	%	1997	%	1998	%
Comunicacions a congressos	2.171	31,89	641	33,04	747	29,82	783	33,15
Articles de revista	2.126	31,23	622	32,06	781	31,18	723	30,61
Llibres	908	13,34	292	15,05	283	11,30	333	14,10
Capítols de llibre	442	6,49	103	5,31	216	8,62	123	5,21
Informes	438	6,43	140	7,22	142	5,67	156	6,60
Tesis	252	3,70	73	3,76	85	3,39	94	3,98
Normes	196	2,88	4	0,21	169	6,75	23	0,97
Inf. comercial i de producte	79	1,16	24	1,24	30	1,20	25	1,06
Preprints i doc. de treball	76	1,12	15	0,77	7	0,28	54	2,29
Manuscrits	53	0,78	15	0,77	24	0,96	14	0,59
Notes tècniques i d'aplicació	21	0,31	4	0,21	5	0,20	12	0,51
Projectes final de carrera	6	0,09	0	0,00	4	0,16	2	0,08
Comunicacions personals	4	0,06	0	0,00	3	0,12	1	0,04
Patents	1	0,01	1	0,05	0	0,00	0	0,00
Sense determinar	34	0,50	6	0,31	9	0,36	19	0,80
Total	6.807	100,00	1.940	100,00	2505	100,00	2.362	100

L'anàlisi detallada de l'actualitat de la bibliografia citada i de l'índex d'envelliment per a cada un dels tipus de documents considerats en l'estudi es pot observar gràficament per mitjà del diagrama de caixes que mostra la figura 3. El diagrama de caixes ofereix una visió resumida molt simple i clara del comportament de la variable edat entre els diversos tipus de documents. Es tracta d'una comparació gràfica que permet

representar múltiples característiques de cada una de les distribucions.

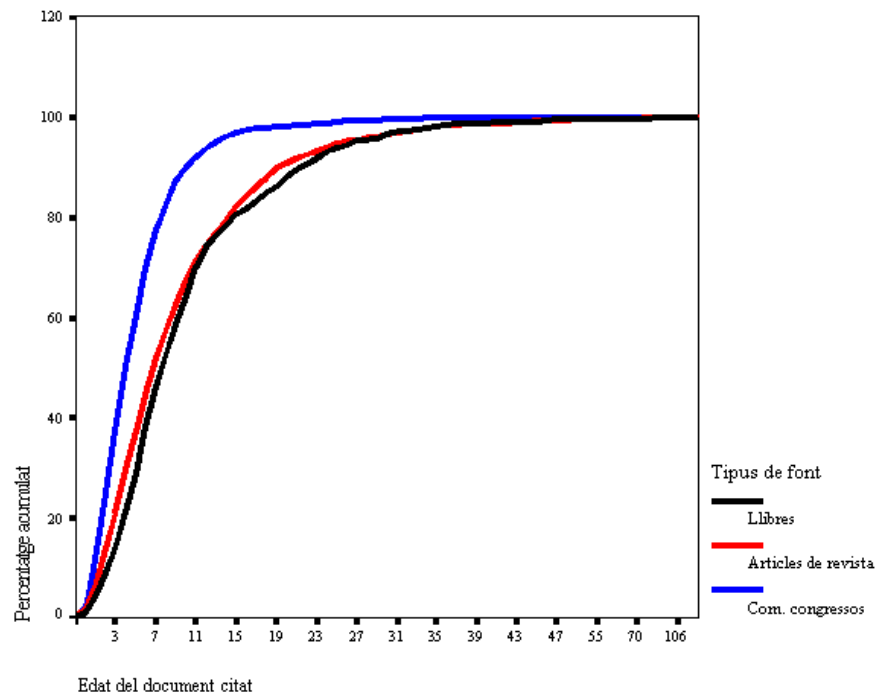
Figura 3. Diagrama de caixa i bigoti de l'edat segons els diversos tipus de documents



Els límits de la caixa vermella estan marcats pel primer i tercer quartil de la distribució, motiu pel qual el 50% dels casos estan compresos dins de la caixa, mentre que la ratlla negra horitzontal dins d'aquesta representa la mitjana. El valor més petit que estigui dins del primer quartil menys 1,5 vegades el rang interquartil marca el límit inferior, representat per la ratlla horitzontal que marca l'inici de la línia vertical que travessa la caixa, mentre que el límit superior el marca el valor més gran que estigui dins del tercer quartil més 1,5 vegades el rang interquartil. Els valors que no arriben o superen aquests límits es mostren en forma d'estrella.

En observar la figura 3, es pot afirmar, a simple vista, que els llibres, els capítols de llibre i les revistes són les publicacions que mostren més resistència a l'envelliment, si bé el nombre de casos extrems i el rang de casos entre els límits és major per a les revistes. El contrast entre l'obsolescència dels congressos i la dels llibres i revistes és més que notable: el percentatge de documents citats que tenen de zero a cinc anys d'edat (índex de Price) per als primers és del 60,23%, amb un valor per a la revistes del 38,81% i per als llibres del 28,05%.

Figura 4. Percentatge acumulat de citacions segons l'edat per als tres tipus de documents més nombrosos



En aquest sentit, es pot establir una relació entre el grau d'actualitat de la bibliografia citada i el major o menor domini dels congressos com a font més citada: el Departament AC manifesta una demanda d'informació lleugerament més actualitzada, amb una mitjana d'edat dels documents citats de 5 anys, per 6 del Departament LSI, i presenta un major percentatge de congressos citats que el Departament LSI.

4.2 Suport documental

S'ha volgut prestar especial atenció a l'observació de l'ús de fonts electròniques enfront del paper, però l'anàlisi de citacions s'ha mostrat insuficient per si mateix: en la majoria dels casos, la referència bibliogràfica del document citat no permet discriminar amb seguretat el suport documental.

Per aquesta raó es comptabilitzen les dades obtingudes de l'anàlisi de citacions considerant com a document electrònic únicament aquell que apareix explícitament descrit com a tal (taula 5), malgrat que se suposa que el nombre real és més gran.

Taula 5. Suports documentals citats i distribució segons departaments

Suport documental	Total	%	AC	%	LSI	%
Paper	6.727	98,82	2.447	98,47	4.280	99,03
Electrònic explícit	47	0,69	26	1,05	21	0,49
- Tangible (cd-rom, etc.)	1	0,01	1	0,04	0	0,00
- Internet FTP	5	0,07	1	0,04	4	0,09
- Internet web	33	0,48	16	0,64	17	0,39
- Internet sense determinar	8	0,12	8	0,32	0	0,00
Sense determinar, o no aplicable	33	0,48	12	0,48	21	0,49
Total	6.807	100,00	2.485	100,00	4.322	100,00

Amb les dades de citacions a la mà, s'observa un domini aclaparador del suport paper o "assimilable", és a dir, de documents que potser estan disponibles en suport electrònic però que responen formalment a la tradició del paper, o que conviuen amb una versió en paper a la qual serveixen de facsimil: revistes en format PDF que tenen doble o triple presentació (paper, CD-ROM, en línia), informes tècnics i d'investigació, i notes tècniques. De tota manera s'ha de destacar que el percentatge de documents comptabilitzats com a documents impresos disminueix, encara que lleument, al llarg del període analitzat (taula 6).

Taula 6. Suports documentals citats i distribució segons anys de lectura

Suport documental	Total	%	1996	%	1997	%	1998	%
Paper	6.727	98,82	1.934	99,69	2.469	98,56	2.324	98,39

Electrònic explícit	47	0,69	2	0,10	26	1,04	19	0,80
- Tangible (cd-rom, etc.)	1	0,01	0	0,00	1	0,04	0	0,00
- Internet FTP	5	0,07	1	0,05	1	0,04	3	0,13
- Internet web	33	0,48	1	0,05	16	0,64	16	0,68
- Internet sense determinar	8	0,12	0	0,00	8	0,32	0	0,00
Sense determinar, o no aplicable	33	0,48	4	0,21	10	0,40	19	0,80
Total	6.807	100,00	1.940	100,00	2505	100,00	2.362	100

Tabla 7. Documents en suport electrònic segons tipologia documental

Tipologia documental	Total citacions	citacions doc. electrònics	% citacions doc. electrònics
Comunicacions a congressos	2.171	4	0,18
Articles de revista	2.126	2	0,09
Libres	908	0	0,00
Capítols de llibre	442	0	0,00
Informes	438	6	1,37
Tesis	252	0	0,00
Normes	196	11	5,61
Inf. comercial i de productw	79	10	12,66
Preprints i doc. De treball	76	9	11,84
Manuscrits	53	1	1,89
Notes tècniques i d'aplicació	21	0	0,00
Projectes final de carrera	6	0	0,00
Comunicacions personals	4	0	0,00
Patents	1	0	0,00
Sense determinar	34	4	11,76
Total	6.807	47	0,69

Més destacable és la vinculació de determinats documents amb l'edició electrònica a Internet: el 5,61% de normes tècniques, l'11,84% dels preprints i el 12,66% de la informació comercial i de producte són documents citats explícitament com a documents d'Internet. Durant el període analitzat (1996-1998) aquestes xifres no fan altra cosa que indicar l'inici d'una tendència que s'ha d'intensificar en els pròxims anys.

5 Discussió

5.1 Tipologia documental

La distribució observada segons la tipologia documental mostra un gran contrast entre les dades d'anàlisi de citacions en tesis doctorals obtingudes en altres estudis i els patrons observats en l'ús de bibliografia científicotècnica en general.⁷ Però si s'analiza des del coneixement de les característiques de la bibliografia informàtica i dels hàbits de consum dels investigadors en aquesta àrea de coneixement, la distribució es mostra bastant ajustada al que és paradigmàtic (Olander, 1992; Cunningham i Connaway, 1996).

En efecte, la primera posició dels congressos entre les tipologies documentals citades sembla concordar amb la importància que els professionals i investigadors de la informàtica atorguen a aquesta font. Així, per exemple, el programa de publicacions notables de la UPC, que té en compte tant revistes com congressos, recull en l'àrea denominada "Tecnologies de la informació i la comunicació i electrònica" un total de 132 títols de congressos, enfront dels 175 que figuren en el conjunt de la llistada congressos notables format per publicacions de cinc àrees diferents: els informàtics han atorgat més importància als congressos que els altres investigadors de la UPC en un procés d'avaluació de la investigació basat en les publicacions.

En part, també es pot explicar el gran nombre de congressos citats sobre la base del fet que es tracta del tipus de document en el qual més han publicat prèviament els autors de les tesis, així com els seus

directors i col·legues d'equip d'investigació.⁸ És evident que sempre hi ha un biaix cap a les fonts en les quals els autors i els seus col·legues més immediats publiquen. Tant per haver publicat amb més freqüència, com pel fet d'una accessibilitat més gran en les biblioteques dels departaments,⁹ els congressos són considerats més accessibles, i això podria explicar també el percentatge superior de citacions d'aquesta tipologia enfront de les revistes.

L'elevat grau d'obsolescència que presenten les citacions als congressos, especialment en comparació amb les revistes (fig. 3), és una altra dada que cal tenir en consideració en el fet que els congressos siguin la tipologia documental més citada: es pot observar que les publicacions que presenten un índex més gran d'obsolescència són precisament els congressos. Aquesta dada és molt important en una àrea de coneixement en la qual els canvis són molt ràpids; tot indica que els congressos, com a tipologia documental, s'adaptin millor a les necessitats de comunicació dels informàtics.

Si s'observa el percentatge acumulat de citacions en funció de l'increment d'edat dels documents citats per a les tres classes de fonts més citades (fig. 4) sorprèn el comportament quasi idèntic de llibres i revistes, i alhora destaca l'actualitat de la bibliografia corresponent als congressos. El fet que en el camp de la informàtica es necessiti una informació molt actualitzada i immediata explica el predomini dels congressos com a forma documental, a causa del lent procés de filtratge, revisió, edició i distribució que suporten les revistes. Més enllà del predomini d'una forma documental, l'elevat nombre de congressos existents en l'àmbit de la informàtica, així com la seva gran especialització, expliquen per si mateixos el paper de la comunicació interpersonal per mitjà de reunions científiques en aquesta àrea. Aquesta immediatesa, juntament amb una accessibilitat superior dels congressos en biblioteques dels departaments o en col·leccions personals, pròpies o de col·legues,¹⁰ explica, en bona mesura, l'elevat nombre de citacions que rep aquesta tipologia documental.

També resulta destacable l'alt percentatge que acumulen les citacions a llibres i capítols de llibre conjuntament, amb un gens insignificant 19,83%. Aquesta xifra s'ha de matisar en l'apartat "Capítols de llibres", ja que s'hi han considerat contribucions individuals en volums de sèries monogràfiques que amb freqüència recullen comunicacions seleccionades de reunions i congressos. Tot i així, les xifres d'aquestes dues tipologies resulten bastant elevades pel que és habitual en la bibliografia tecnològica i es podrien explicar tenint en consideració l'enfocament més teòric i de fonamentació que els departaments de la UPC volen donar al doctorat i la tesi, enfront de l'orientació més aplicada dels projectes de final de carrera. Aquesta afirmació es reforça pel baix volum de citacions que reben les patents i els projectes de final de carrera.

Quant a l'estudi del volum de literatura grisa citada, el primer obstacle que es presenta és que no hi ha una definició totalment clara del que s'entén per literatura grisa, ni tan sols es disposa d'un consens en castellà respecte al terme que la designa. A nivell operatiu, en aquest treball en el qual s'analitzen citacions —sense tenir el document a la mà— s'ha optat per considerar dins de l'apartat "literatura grisa" fonts d'informació que mostren un control bibliogràfic dèbil: informes tècnics i d'investigació, tesis, informació comercial i de producte, preprints i documents de treball, manuscrits, notes tècniques i d'aplicació, projectes de final de carrera, i comunicacions personals.

Realitzades aquestes consideracions, el volum de literatura grisa es pot considerar més que notable: representa un 13,65% de les citacions, amb una xifra de 929 citacions sobre un total de 6.807. Aquesta xifra podria ser més gran si es tinguessin en compte dins d'aquest apartat algunes contribucions a congressos, que habitualment no assoleixen una publicació i distribució editorial "normal" i que es queden formalment a nivell de *papers*. Aquest alt volum de citacions a publicacions de literatura grisa concorda amb les característiques de l'ús d'informació observades entre els tecnòlegs en general (Terrada, 1994, p. 94). El volum de consum d'aquest tipus de fonts es pot dir que és directament proporcional a la distància que té una determinada àrea d'investigació respecte a allò que podríem denominar *enfocaments teòrics*, o del que es coneix com a *investigació bàsica*: com més aplicada és una disciplina científica, més nombrosa és la presència de la literatura grisa en la seva bibliografia.

En aquest sentit, una lectura atenta de la figura 2 permet matisar el menor nombre de referències per tesi en el Departament AC amb la distribució d'aquestes segons la tipologia documental: el consum d'informació és menor, però respecte al total de citacions rebudes per congressos, informes, tesis, normes i altres, la quota de participació del Departament AC és més gran, tant en valor absolut en alguns casos com ara les normes, com en valor relatiu una vegada descomptat el menor nombre de referències generades per les tesis del departament. Per tant, no sols es dona un nombre menor de referències per tesi en el Departament AC, sinó que el consum d'informació es basa en més gran mesura en fonts de la denominada *literatura grisa*. En estudis locals d'usuaris com aquest, és de gran interès la comparació dels resultats obtinguts amb les dades disponibles per al mateix col·lectiu en àmbits nacionals o internacionals.

En aquest sentit, les ciències de la computació tenen un instrument de control bibliogràfic de primer ordre que pot ser útil per analitzar si la distribució de citacions observades segons la tipologia documental concorden amb el que succeeix en el conjunt de la bibliografia informàtica. Es tracta de la bibliografia de bibliografies d'informàtica titulada *The collection of computer science bibliographies* (Achilles, 1999), a la qual es pot accedir lliurement per Internet. Al marge de funcionar com una bibliografia de bibliografies clàssica, amb una potència de recuperació d'informació més que notable, també permet obtenir en temps real les estadístiques de totes i cadascuna de les bibliografies indexades, amb dades sobre la tipologia documental i les dates de publicació. Igualment ofereix les estadístiques agregades de tot el conjunt, de tal manera que es pot obtenir una imatge força interessant a mig camí entre el que es publica sobre informàtica i el que es publicita per mitjà de bibliografies selectives i bibliografies temàtiques. Les dades representen una anàlisi d'un milió de referències bibliogràfiques, algunes de les quals, per descomptat, estan duplicades.

En comparar les dades observades en aquest estudi amb aquelles que figuren a la base de dades

d'Achilles (1999) sobre perfils temàtics similars als dels departaments estudiats, s'observa una gran similitud amb les tipologies documentals citades en les tesis. D'altra banda, el resum general del conjunt de la bibliografia de bibliografies d'Achilles ofereix una dada reveladora de l'estructura de la comunicació científica en aquestes àrees de coneixement: sobre un total d'1.038.616 notícies,¹¹ els articles de revistes representaven un 46,70%, les comunicacions a congressos un 29,18%, i els informes tècnics un 10,37%.

Aquestes dades, obtingudes de bibliografies de propòsit informatiu i no de citacions en treballs, informen sobre la producció editorial i no sobre l'ús, i alhora fan prevaler el valor permanent de la informació. Per això, en les dades agregades de conjunt, els articles estan sobrerrepresentats. De tota manera, allò que és rellevant és que els congressos representen quasi un 30% de la bibliografia circulant, percentatge no gaire habitual en altres disciplines científiques.

A grans trets, la tipologia de fonts citades en les tesis i la seva distribució percentual és consistent amb el que recull la bibliografia que estudia per mètodes directes l'ús d'informació dels informàtics (Takayama, 1986; Olander, 1992; Cunningham i Connaway, 1996), en què es confirma la imatge obtinguda en aquest estudi segons la qual una part molt significativa de la comunicació científica es realitza per la via de congressos i de literatura grisa. S'ha de destacar que, en conjunt, les revistes tan sols representen un 31,2% del total de citacions: es tracta d'una dada molt significativa, ja que, en general, s'assumeixen per a la ciència i la tecnologia percentatges d'ús més alts per a aquesta tipologia documental, percentatges que amb freqüència s'apliquen de manera molt mecànica en el desenvolupament de col·leccions.

En efecte, els resultats obtinguts són realment rellevants en l'avaluació de polítiques d'adquisició. Si bé els congressos es podrien considerar dins del capítol de publicacions seriades, la veritat és que, en general, la seva adquisició no es realitza de manera consolidada com a publicació en sèrie, amb una imputació de l'adquisició en la majoria de les ocasions a la partida pressupostària de monografies. Tant per aquesta raó, com pel significatiu percentatge de llibres i capítols de llibre citats, les dades que s'han obtingut en aquest apartat poden ser motiu de reflexió en el terreny de la ràtio entre monografies i revistes del pressupost d'adquisicions.

5.2 Suport documental

La xifra de documents electrònics accessibles en línia per Internet, dins del seu reduït nivell, és ben significativa si es compara amb estudis anteriors sobre l'ús de recursos d'Internet fet per investigadors en ciència i tecnologia basats en l'anàlisi de citacions (Urbano, 1997) i en els quals es constata un nombre encara més baix de citacions explícites a documents electrònics. En aquest cas, s'observa un percentatge del 0,66% sobre el total de citacions, amb una tendència a l'increment al llarg del temps (taula 6) dels recursos web que podria relacionar-se amb el paper creixent de l'edició no tradicional vinculada a Internet. Comparat amb l'1,9% de citacions que reben els recursos electrònics en l'estudi d'Harter i Kim (1996b), realitzat a partir d'articles publicats en revistes electròniques, el 0,66% en les tesis de la UPC és certament rellevant.

Com pot observar-se a la taula 7, la presència de documents en suport electrònic no és uniforme, ja que és molt més notable en l'àmbit de la literatura grisa. S'ha de tenir en consideració que la revisió i l'arbitratge com a condició per publicar en ciència i tecnologia està molt més consolidat i és més exigent que en moltes disciplines de les humanitats i de les ciències socials. Aquest fet provoca que s'atorgui molt més valor a fonts reconegudes, assentades tradicionalment en models de publicació pre-Internet. Així, com a revista electrònica d'accés en línia i de distribució gratuïta tan sols s'ha identificat *NHSE review* amb dues citacions. Cap de les revistes electròniques de matèries afins als usuaris estudiats i amb un nombre de citacions més gran en l'estudi d'Harter i Kim (1996b) figura en la llista de les citades en les tesis de la UPC.¹²

No és sorprenent que en el moment en què les condicions d'accés i de capacitació en l'ús d'Internet dels investigadors en ciències socials i humanitats s'igualen als de ciències "dures" i tecnologia, el nivell de recursos d'Internet citats sigui més gran en les àrees de coneixement amb mecanismes de comunicació científica menys estrictes. Així, un estudi de les citacions en publicacions de distribució normalitzada, com ara els llibres i les revistes, mostraria que es poden trobar nombroses cites a documents electrònics en treballs que estudien Internet en relació amb aspectes socials i d'aplicació, però en menor mesura en treballs d'investigació tecnològica reconeguda i de primera línia, ja que la Xarxa és una bona font d'informació sobre si mateixa en termes de normalització i difusió de la seva tecnologia, així com per l'estudi del seu impacte social. Això explica que els estudis sobre l'ús d'Internet en termes d'anàlisi sociològica o, com en el cas dels bibliotecaris documentalistes, sobre la radical transformació d'una pràctica professional a causa de la Xarxa, són els que mostren un nombre de citacions més gran a recursos d'Internet (Masip, 1997; Estivill i Urbano, 1998).

En qualsevol cas, a partir de les característiques de la col·lecció, de les observacions del personal bibliotecari encarregat de les publicacions periòdiques, de les subscripcions electròniques que a nivell personal o institucional mantenen determinats investigadors sense estar consolidades en el catàleg de la biblioteca, i de l'anàlisi dels servidors web dels departaments estudiats, es podria arribar a concloure el que aquí es formula únicament com a suposició: que un bon nombre de les referències bibliogràfiques, malgrat que no es pugui establir amb certitud per la descripció, correspon a consultes de documents en suports diferents al paper (microformes, CD-ROM, recursos en línia comercials i recursos en línia gratuïts).

En aquest terreny, l'anàlisi de citacions s'ha mostrat insuficient i serien necessàries altres tècniques per avaluar l'ús de recursos electrònics com, per exemple, l'anàlisi de connexions a servidors locals o remots vinculats al subministrament de documents primaris. Aquest tema és especialment interessant, ja que hi ha dos factors recents que poden modificar el comportament futur en l'ús de la col·lecció. D'una banda, des de l'any 1999 tant el Servei de Biblioteques de la UPC com el Consorci de Biblioteques Universitàries de

Catalunya ofereixen accés a revistes electròniques des d'ordinadors amb IP de la universitat. D'altra banda, la Biblioteca Rector Gabriel Ferraté, central del campus Nord, disposa de la base de dades *IEL (IEE/IEEE Electronic Library)*¹³ des de l'any 1997 en CD-ROM amb accés local, fet que ha comportat una progressiva cancel·lació de subscripcions a revistes en paper de l'IEEL, cancel·lacions que es van accentuar clarament el 1999.

Si les dades sobre citacions dels doctorands de la UPC es rectificuessin tan sols en el que es refereix a bibliografia d'informes i *preprints* posteriors a 1990,¹⁴ tot considerant-los recursos electrònics accessibles per Internet, es comptarien un total de 404 cites, xifra que representaria quasi un 6% sobre el total de les 6.807 comptabilitzades en l'estudi. Les xifres sobre ús de documents disponibles via Internet que s'obtidrien realitzant una projecció com l'enunciada concorden millor amb les estadístiques que dona Achilles (1999) a *The collection of computer science bibliographies*. En total, sobre el milió de notícies analitzades, 92.543 contenen l'element URL en la referència bibliogràfica com a camp de descripció documental, raó per la qual no és aventurat afirmar que actualment un 9% de la bibliografia circulant sobre informàtica a nivell mundial pertany a l'àmbit de l'edició no comercial en xarxa.

El gran volum de literatura grisa citada ha de ser tingut en consideració tant des del servei d'informació bibliogràfica, com des de la política d'adquisicions i de catalogació de recursos electrònics. El debat sobre com s'ha d'actuar en relació amb els recursos electrònics disponibles a Internet és especialment rellevant en el terreny d'aquest grup de documents: s'han d'identificar els llocs web que es dediquen a recollir aquest tipus de documents; s'ha de tenir coneixement dels llocs web més productius en la publicació d'informes i preprints, i especialment els més utilitzats entre els usuaris de la biblioteca; s'ha d'analitzar el grau de cobertura amb el qual són tinguts en compte pels sistemes clàssics de control bibliogràfic; i, per acabar, s'ha de decidir si la biblioteca ha d'oferir-hi algun mecanisme d'accés, i si aquest mecanisme passa pel catàleg, o per repertoris independents de recursos web.¹⁵

En qualsevol cas, les dades obtingudes demostren que en l'àmbit de les ciències de la computació el paper que té la comunicació científica entre col·legues, enfront d'aquella que canalitzen les grans editorials i les revistes científiques, és realment significatiu. D'altra banda, en l'àrea de la informàtica tant els instruments de control bibliogràfic presents a la Xarxa, com les col·leccions digitals de relleu —ja siguin de lliure accés o de pagament— tenen un nivell de cobertura i de qualitat molt remarcable.¹⁶

Com a exemples destacats de recursos gratuïts es poden citar: *ResearchIndex* (Lawrence, Giles, Bollacker, 1999a i 1999b), índex electrònic de citacions de publicacions a Internet; *The collection of computer science bibliographies* (Achilles, 1999), bibliografia de bibliografies d'informàtica en format electrònic; i *Networked computer science technical reference library* (NCSTRL, 1999), repertori d'informes tècnics i d'investigació. Per part de l'edició comercial, els dos editors més importants en el camp de la informàtica —confirmats en vista del que s'ha observat en aquest estudi en funció del nombre de citacions que reben les seves publicacions— disposen de seus webs destinats a distribuir totes les publicacions: l'ACM (Association of Computing Machinery) amb l'*ACM digital library* (ACM, 1999) i l'IEEE (Institute of Electrical and Electronics Engineers) amb la *IEEE/IEE electronic library online (IEL online...)*, 1999).

En definitiva, les possibilitats de treballar únicament amb recursos en línia són reals i, per tant, aquest fet haurà de ser tingut en consideració per les biblioteques que vulguin afegir valor a aquests documents, únic mètode perquè siguin valorades com a mecanismes eficients d'accés a la informació en el nou escenari d'edició a la Xarxa que es dibuixa tant per als canals no comercials com per als comercials.

6 Conclusions

La tipologia documental més citada és la que correspon a les comunicacions de congressos, cosa que confirma el paper clau atorgat per la Comissió de Publicacions Notables a aquesta categoria documental en l'àrea de la informàtica a requeriment dels dos departaments estudiats. Aquest domini en conjunt dels congressos, juntament amb la presència d'un important volum de literatura grisa, així com de llibres i capítols de llibres, ha de ser tinguda molt en consideració per no sobrevalorar en la política d'adquisicions sectorial de la informàtica la importància de la revista científica.

La bibliografia citada mostra una elevada obsolescència. De tota manera, en realitzar-se l'anàlisi de les citacions a partir de tesis doctorals com a font, s'han recollit mencions a documents d'una certa antiguitat, ja que una de les característiques de la investigació portada a terme pels doctorands acostuma a ser la revisió de la bibliografia sobre un tema considerant-ne els fonaments i els primers desenvolupaments.

El mètode utilitzat no s'ha demostrat vàlid en la determinació de l'ús de documents electrònics, ja que, en moltes ocasions, les referències no s'han realitzat consignant la naturalesa del suport. No obstant això, el volum de citacions explícites a documents electrònics accessibles per Internet és ben destacable tenint en consideració el període cobert (1996-1998) i prenent com a referència el que és habitual en publicacions científiques d'un cert nivell de l'àmbit de les ciències "dures" i tecnològiques.

7 Referències

Achilles, Alf-Christian (1999). *The collection of computer science bibliographies* [en línia]. [Alemanya]: l'autor, 1995-1999. <<http://iinwww.ira.uka.de/bibliography>> [Consulta: 11 nov. 2000].

ACM [Association for Computing Machinery] (2000). *ACM digital library* [en línia]. New York: ACM. <<http://www.acm.org/dl>> [Consulta: 13 nov. 2000].

- Anuario estadístico de las bibliotecas universitarias y científicas españolas 1999* [en línia]. (2000). Rebiun. Madrid: Universidad Autónoma de Madrid: CRUE (Conferencia de Rectores de las Universidades Españolas). 78 p. <<http://www.uma.es/rebiun/informeEstadistico99.pdf>> [Consulta: 12 nov. 2000].
- Baker, Sharon L.; Lancaster, F.W. (1991). *The measurement and evaluation of library services*. 2nd ed. Arlington, Virginia: Information Resources Press. XVIII, 411 p.
- Bandyopadhyay, A.K. (1996). "Citation analysis of doctoral dissertations in mathematics using dBase 3 plus", *Annals of library science and documentation*, vol. 43, no. 3 (1996), p. 81-107.
- Barry, Christine A. (1999). "Las habilidades de información en un mundo electrónico: la formación investigadora de los estudiantes de doctorado", *Anales de documentación*, nº 2 (1999), p. 237-258. [Traducció de l'article publicat el 1997: "Information skills for an electronic world: training doctoral research students", *Journal of information science*, vol. 23, no. 3 (1997), p. 225-238.]
- Bowman, Michael. (1991). "Format citation patterns and their implications for collection development in research libraries", *Collection building*, vol.11, no. 1 (1991), p. 2-8.
- Broadus, Robert N. (1977). "The applications of citation analyses to library collection building", *Advances in librarianship*, no.7 (1977), p. 299-335.
- Broadus, Robert N. (1985a). "A proposed method for eliminating titles from periodical subscription lists", *College & research libraries*, vol. 46, no. 1 (1985), p. 30-35.
- Broadus, Robert N. (1985b). "On citations, uses, and informed guesswork: a response to Line", *College & research libraries*, vol. 46, no. 1 (1985), p. 38-39.
- Buchanan, Anne L.; Herubel, Jean Pierre V.M. (1994). "Profiling PhD dissertation bibliographies: serials collection development in political science", *Behavioral and social sciences librarian*, vol.13, no. 1 (1994), p. 1-10.
- Cunningham, S.J.; Connaway, L.S. (1996). "Information searching preferences and practices of computer science researchers". En: *Sixth Australian Conference on Computer-Human Interaction*. [Australia]: IEEE. p. 294-299.
- Deshpande, M.; Rajyalakshmi, D. (1997). "Citation study of dissertations in library and information science", *Annals of library science and documentation*, vol. 44, no. 2 (1997), p. 41-53.
- Devin, Robin B.; Kellogg, Martha (1990). "The serial/monograph ratio in research libraries: budgeting in light of citation studies", *College & research libraries*, vol. 51, no. 1 (1990), p. 46-54.
- Estivill, Assumpció; Urbano, Cristóbal (1998). "Cómo citar recursos electrónicos: del uso a la norma y de la norma al uso: análisis de prácticas de citación en revistas españolas de biblioteconomía". En: *VI Jornadas Españolas de Documentación*. Valencia: Fesabid, p. 245-256.
- Garfield, E. (1968). "Decision making in libraries acquisitions: indexes or journals?", *Current contents*, (April 9th 1968).
- Glänzel, Wolfgang; Schoepflin, Urs (1999). "A bibliometric study of reference literature in the sciences and social sciences", *Information processing & management*, 35 (1999), p. 31-44.
- Harter, Stephen P.; Kim, Hak Joon (1996). *Electronic journals and scholarly communication* [en línia]: a citation and reference study. A paper delivered at the Midyear Meeting of the American Society for Information Science, San Diego, CA, May 20-22, 1996. <<http://php.indiana.edu/~harter/harter-asis96midtxt.html>> [Consulta: 11 nov. 2000].
- Hughes, Janet (1995). "Use of faculty publication lists and ISI citation data to identify a core list of journals with local importance", *Library acquisitions: practice and theory*, vol.19, no. 4 (1995), p. 403-413.
- Hurd, Julie M.; Blecic, Deborah D.; Vishwanatham, Rama (1999). Information use by molecular biologist: implications for library collections and services, *College & research libraries*, vol. 60, no. 1 (1999), p. 31-43.
- IEL online* [en línia]: *IEEE/IEE electronic library online*. (1999). Information Handling Services. <<http://iel.ihs.com>> [Consulta: 11 nov. 2000].
- Jiménez Contreras, E.; et al. (1994). "Determinación de las colecciones básicas de publicaciones periódicas en hemerotecas científicas". En: *IV Jornadas españolas de documentación automatizada*. Oviedo: Universidad de Oviedo, p. 216-221.
- Joswick, Kathleen E.; Stierman, Jeanne K. (1997). "The core list mirage: a comparison of the journals frequently consulted by faculty and students", *College & research libraries*, vol. 58, no. 1 (1997), p. 48-55.
- Kriz, Harry M. (1978). "Subscriptions vs. books in a constant dollar budget", *College & research libraries*, vol. 39, no. 2 (1978), p. 105-109.

- Lancaster, F.W (1996). *Evaluación de la biblioteca*. Traducción por Ramón Abad y Belén Altuna. Madrid: ANABAD, 1996. 374 p. [Traducción de: *If you want to evaluate your library...* 2nd ed. (Londres: The Library Association, 1993)].
- Lawrence, Steve; Giles, C. Lee; Bollacker, Kurt (1999a). "Digital libraries and autonomous citation indexing", *IEEE computer*, vol.32, no. 6 (1999), p. 67-71.
- Lawrence, Steve; Giles, C. Lee; Bollacker, Kurt (1999b). *CiteSeer* [en línia]: *autonomous citation indexing*. NEC Research Institute. <<http://www.neci.nec.com/~lawrence/citeseer.html>> [Consulta: 11 nov. 2000].
- Line, Maurice B. (1977). "On the irrelevance of citation analyses to practical librarianship". En: *EURIM II: an European conference on the application of research in information services in libraries*. London: Aslib, p. 51-53.
- Line, Maurice B. (1985). "Use of citation data for periodicals control in libraries: a response to Broadus", *College & research libraries*, vol. 46, no. 1 (1985), p. 36-37.
- Masip, Pere (1997). "L'ús d'informació digital en ciències socials". En: *6es Jornades Catalanes de Documentació*. Barcelona: Socadi: Col·legi Oficial de Bibliotecaris-Documentalistes, p. 439-446.
- McCain, Katherine W.; Bobick, James E. (1981). "Patterns of journal use in a departamental library: a citation analysis", *Journal of the American Society for Information Science*, vol. 32, no. 4 (1981), p. 257-267.
- Mubeen, M. A. (1996). "Citation analysis of doctoral dissertations in chemistry", *Annals of library science and documentation*, vol. 43, no. 2 (1996), p. 48-58.
- NCSTRL. (1999). *Networked Computer Science Technical Reference Library* [en línia]. NCSTRL. <<http://www.ncstrl.org>> [Consulta: 11 nov. 2000]
- Nederhof, A. J.; Van Raan, A.J. (1987). "Citation theory and the Ortega Hypothesis", *Scientometrics*, vol.12, no. 5-6 (1987), p. 325-328.
- Nisonger, Thomas E. (1998). *Management of serials in libraries*. Englewood, Colorado: Libraries Unlimited. "XXVIII", 433 p.
- Olander, Anna Birgitta (1992). *Personal information management in computer science research*. Ph. D. thesis University of Toronto. 227 p.
- Pan, E. (1978). "Journal citation as a predictor of journal usage in libraries", *Collection management*, vol. 2, no. 1 (1978), p. 29-38.
- Poll, Roswitha; Boekhorst, Peter (1998). *Medición de la calidad: directrices internacionales para la medición del rendimiento en bibliotecas universitarias*. Madrid: ANABAD, 1998. 209 p.
- Satariano, W. A. (1978). "Journal use in sociology: citation analysis versus readership patterns", *Library quarterly*, vol. 48, no. 3 (1978), p. 293-300.
- Scales, P.A. (1976). "Citation analyses as indicators of the use of serials: a comparison of ranked title lists produced by citation counting and from use data", *Journal of documentation*, vol. 32, no. 1 (1976), p. 17-25.
- Smith, Linda C. (1981). "Citation analysis", *Library trends*, vol. 30, no. 1 (1981), p. 83-106.
- Sylvia, Margaret; Leshner, Marcella (1995). "What journals do psychology graduate students need?: a citation analysis of thesis references", *College & research libraries*, vol. 56, no. 4 (1995), p. 313-318.
- Takayama, Masaya (1986). "Analysis of technological information transfer among Japanese computer scientists at a research front", *Information services & use*, vol. 6, no. 1 (1986), p. 9-25.
- Terrada, María Luz (1994). "Documentos, citas y referencias en la comunicación científica". En: Pérez-Álvarez Ossorio, J.R; Goedegebuure (ed.). *New worlds in information and documentation* Amsterdam: Elsevier, p. 89-97.
- UPC. Vicerrectorat de Recerca (1997). *Publicacions notables* [en línia]. <<http://inopia.upc.es:80/sia/cat/butlleti-upc/documents/76-1997.html>> [Consulta: 11 nov. 2000].
- UPC. Vicerrectorat de Recerca (1999). "Sistema d'indicadors de l'activitat de recerca de la UPC". Document accessible en línia en l'aplicació *Fenix* de producció científica de la UPC.
- Urbano, Cristóbal (1997). "Information seeking behaviour of doctoral students in Barcelona since the Internet explosion: groundwork for a project of research". En: *5th International Symposium BOBCATSSS 1997, Budapest*. Amsterdam: Hogeschool van Amsterdam.
- Urbano, Cristóbal (2000). *El análisis de citas en publicaciones de usuarios de bibliotecas universitarias [en línia]: estudio de las tesis doctorales en informática de la Universidad Politécnica de Cataluña, 1996-1998*. Director: Ernest Abadal i Falgueras. 529 p. Tesis doctoral, Universitat de Barcelona, 2000.

<<http://www.tdcat.cesca.es/TDCat-0706100-091717>> [Consulta: 11 nov. 2000].

Walcott, Rosalind (1991). "Characteristics of citations in geoscience doctoral dissertations accepted at United States academic institutions 1981-1985", *Science and technology libraries*, vol. 12, no. 2 (1991), p. 5-16.

Walcott, Rosalind (1994). "Local citation studies: a shortcut to local knowledge", *Science and technology libraries*, vol. 14, no. 3 (1994), p. 1-14.

Zipp, Louise S. (1996). "Thesis and dissertation citations as indicators of faculty research use of university library journal collections", *Library resources and technical services*, vol. 40, no. 4 (1996), p. 335-342.

Notes

- ¹ Es fa referència a l'ús per part del lector, i no necessàriament de la col·lecció: l'anàlisi de citacions com a mètode d'avaluació de l'ús de la col·lecció està més discutit en comparació amb altres sistemes d'avaluació (Baker i Lancaster, 1991).
- ² Així es pot diluir el biaix i determinades distorsions. En aquest sentit Nederhof i Van Raan (1987) afirmen que si es consideren grans quantitats de documents font de manera agregada, les influències omeses i les citades, però no reals, tendeixen a diluir-se i contrarestar-se.
- ³ Nisonger (1998, p. 145) subratlla que els diversos estudis dedicats a establir la ràtio publicacions seriades / monografies no parteixen d'una definició uniforme del que és una publicació seriada.
- ⁴ La informàtica com a disciplina bibliomètricament considerada i els informàtics com a usuaris de biblioteques ha rebut poca atenció tant a Espanya com en l'àmbit internacional (Cunningham i Connaway, 1996, Urbano, 2000).
- ⁵ Es tradueix *to browse* per *rastrear*, si bé en els diferents contextos en els quals pot utilitzar-se el terme anglès se'n podria ajustar la traducció a un terme més específic: *explorar*, *tafanejar*, *fullejar*, *cercar aleatòriament*, etc.
- ⁶ Abans de començar el recompte de citacions es va construir una base de dades relacional i es van importar dades del JCR de l'ISI, com també la llista dels congressos notables recollits en la documentació del *Programa de publicacions notables UPC*.
- ⁷ Com a exemple del domini de les citacions en revistes en altres estudis es poden mencionar els següents percentatges de revistes citades: 80,01% en matemàtiques (Bandyopadhyay, 1996), 91,3% en biologia molecular (Hurd, Blečić i Vishwanathan, 1999), 68,74% en biblioteconomia i documentació (Deshpande i Rajyalakshmi, 1997), 73% en química (Mubeen, 1996), o 62,2% en enginyeria electrònica (Glänzel i Schoepflin, 1999). Tots ells comptabilitzen de manera separada congressos i revistes fent servir aquesta terminologia, diferència que no queda tan clara quan s'utilitza el concepte genèric de publicació seriada que inclou ambdues categories.
- ⁸ Malgrat que la dada de les autocitacions no s'ha comptabilitzat, sí que s'ha observat aquesta pràctica vers els treballs publicats en congressos.
- ⁹ Una primera anàlisi exploratòria realitzada amb una petita mostra permet observar una presència elevada d'aquest tipus de document en les biblioteques departamentals de la UPC.
- ¹⁰ El volum d'autocitacions vers treballs propis o d'altres membres del departament, o grup d'investigació, es concentra fonamentalment en informes de recerca i en comunicacions de congressos, fet que fa suposar com a certa l'afirmació en relació amb les col·leccions personals.
- ¹¹ En l'estudi realitzat per Harter i Kim (1996b) es van identificar en el directori de revistes electròniques de l'Association of Research Libraries un total de 131 revistes científiques electròniques que comptaven amb sistema *peer review*, de les quals 28 eren de ciències, 34 de ciències socials, 31 d'humanitats i 38 es classificaven en una categoria miscel·lània. Educació i literatura amb 13 i 12, respectivament, eren les àrees concretes amb més títols. L'esmentat directori en les dades de l'estudi recollia principalment revistes electròniques de nova creació i lliure accés.
- ¹² La *JEL* en CD-ROM és una base de dades mensual en text complet i imatge que recull les publicacions (revistes, congressos i normes) de l'IEEE (Institute of Electrical and Electronics Engineers) i de l'IEE (Institution of Electrical Engineers). Inclou tant les publicacions en curs com les retrospectives des de 1988. La Biblioteca Rector Gabriel Ferraté de la UPC disposa d'una col·lecció que cobreix des de 1991 fins l'actualitat. Actualment també es comercialitza en versió web (*JEL online...*, 1999).
- ¹³ Tot analitzant les seues web de diversos departaments universitaris de l'àrea, s'ha observat que el 1990 podria ser una data convencional com a frontera a partir de la qual bona part dels informes de recerca i dels preprints comencen a ser disponibles en format electrònic.
- ¹⁴ Les biblioteques de la UPC sembla que han optat, de moment, per la creació d'una base de dades independent del catàleg per al control dels recursos d'Internet (Beumala et al., 1998).

¹⁵ Els números d'abril de 1995 i de 1998 de la revista *Communications of the ACM*, i el número de febrer de 1999 d'*IEEE computer* són monogràfics sobre biblioteques digitals, especialment en l'àrea d'informàtica. Són moltes les pàgines web de biblioteques, o pàgines web personals, dedicades a recollir selectivament aquest tipus de recurs; una selecció d'aquestes es pot trobar a Achilles (1999).

Facultat de Biblioteconomia i Documentació
Universitat de Barcelona
Barcelona, desembre de 2000
<http://www.ub.edu/biblio> •  [Comentaris](#)

 [Citació recomanada](#) • [Metadades](#)
[UB](#) • [Facultat](#) • [BiD](#)