

L'efficacia del recupero

Teoria di una metodologia sperimentale per misurare il valore del risultato della ricerca dal punto di vista dell'utente

di Clare Beghtol

Haaas e Kraft hanno criticato gran parte delle ricerche nel campo della scienza dell'informazione in quanto non portano "alcuna prova utile di nessuna ipotesi teorica"¹ ed hanno raccomandato di impiegare più sovente progetti sperimentali che forniscano cognizioni atte a facilitare la soluzione di problemi pratici. Questo articolo suggerisce un progetto e una metodologia di ricerca sperimentale adattabili a qualsiasi ambiente di recupero di informazioni. La metodologia si serve di misure che permettono un esame rigoroso dell'efficacia del recupero, sicché i risultati degli studi che impiegano la metodologia contribuiscono agli sforzi pratici per migliorare un progetto di sistema di recupero delle informazioni.

Si conviene che l'utilità di un sistema informativo dal punto di vista dell'utente potrebbe essere una misura della sua efficacia di recupero. L'idea è stata variamente rubricata, come l'"utilità",² la "relazione con le circostanze",³ il "giudizio sulla

destinazione",⁴ il "vantaggio possibile"⁵ o il "valore"⁶ dell'informazione recuperata. Sicché la percezione dell'utilità di un sistema da parte degli utenti potrebbe costituire un elemento nelle prove sull'efficacia del recupero. Comunque si è scritto di meno sul modo di misurare in maniera significativa l'aiuto che si riceve individualmente da un sistema. Questo articolo presenta una struttura teorica e una metodologia sperimentale pratica di uso generale per valutare l'utilità dal punto di vista dell'utente. La metodologia è sviluppata con un certo dettaglio mediante l'esperimento dell'efficacia del recupero di un soggetto richiesto come esempio prolungato.

Cooper ha così delineato lo scopo delle metodologie per la valutazione: "Una misura ideale dell'efficacia del recupero deve misurare in qualche modo il valore definitivo del sistema di recupero nel confronto dei suoi utenti. Purtroppo non si può assicurare nessun procedimento pratico per determinare con precisione e finalità assolute questo valore".⁷

CLARE BEGHTOL, *Retrieval Effectiveness: Theory for an Experimental Methodology Measuring User-perceived Value of Search Outcome*, "Libri", 39, 1 (March 1989), p. 18-35. Traduzione di Carlo Revelli. Presentare la traduzione di un articolo che risale a quattro anni fa può destare qualche perplessità: le stesse che abbiamo avuto noi e che abbiamo superato, considerando che l'orientamento dell'informazione dal punto di vista del lettore è uno dei temi maggiormente presenti nell'attività bibliotecaria di oggi. In questo articolo si valuta quanto il rapporto tradizionale tra richiamo e precisione possa risultare modificato dal rapporto tra la quantità dei documenti ritenuti utili in seguito a una ricerca catalografica e quella dei documenti che, in effetti, sono risultati utili al ricercatore.

¹ D.F. HAAS-D. KRAFT, *Experimental and Quasi-experimental Designs for Research in Information Science*, "Information Processing and Management", 20 (1-2): 229-237, 1984, p. 223.

² W.S. COOPER, *A Definition of Relevance for Information Retrieval*, "Information Storage and Retrieval", 7: 19-37, 1971.

³ P. WILSON, "Situational Relevance", "Information Storage and Retrieval", 9: 457-471, 1973.

⁴ T. SARACEVIC, *Relevance: a Review of and a Framework for the Thinking on the Notion in Information Science*, "Journal of the American Society for Information Science", 26 (5) Nov.-Dec.: 321-343, 1975.

⁵ M.J. BUCKLAND, *Relatedness, Relevance and Responsiveness in Retrieval Systems*, "Information Processing and Management", 19 (3): 237-241, 1983.

⁶ W.B. ROUSE-S.H. ROUSE, *Human Information Seeking and Design of Information Systems*, "Information Processing and Management", 20 (1-2): 129-138, 1984.

⁷ W.S. COOPER, *On Selecting a Measure of Retrieval Effectiveness. Part I: The "Subjective" Philosophy of Evaluation; Part II: Implementation of the Philosophy*, "Journal of the American Society for Information Science", 24 (2) Mar.-Apr.: 87-100; 24 (6) Nov.-Dec.: 413-424, 1973, Part I, p. 88.

Per Cooper una misura del "valore definitivo" consiste in una valutazione dell'"utilità del sistema". Il ricercatore intervisterebbe gli utenti durante le loro ricerche per trarne un'analisi dei costi-benefici individuali per ogni documento ritrovato nel senso del valore venale ad esso attribuibile. Si potrebbero accettare risposte positive, nulle o negative. La risposta per ogni documento misurerebbe l'"utilità del documento"; il cumulo delle utilità dei documenti per quell'utente misurerebbe l'"utilità della ricerca", ed il cumulo delle utilità delle ricerche per un certo numero di utenti potrebbe quindi servire a misurare l'"utilità del sistema", ossia il valore definitivo per gli utenti dei documenti trovati da quel sistema ed espresso in dollari.

Cooper chiamò "semplice" questa metodologia ed ammise che è probabilmente inapplicabile. Ciononostante egli sostenne che "in linea di principio non ci può essere valutazione migliore [...] dei benefici di un sistema di recupero che una valutazione soggettiva da parte dei suoi utenti" e che la sua metodologia poteva servire da "modello concettuale con il quale paragonare altre misure di efficacia più pratiche".⁸ Una delle sue misure più pratiche consisteva nel valutare i sistemi in base al numero medio dei riferimenti incontrati da un utente durante la ricerca che risultassero ancora validi al termine del lavoro.⁹ Come nella metodologia "semplice", un utente valuta immediatamente il valore dei riferimenti trovati. Pare comunque che le valutazioni immediate diano luogo a giudizi inadeguati sul valore di un riferimento. La valutazione immediata dei riferimenti è invece una misura del valore *potenziale* del documento indicato dal riferimento, intesa soggettivamente. L'utente potrebbe scoprire che un giudizio immediato era errato qualora il valore di un documento risulti superiore o inferiore rispetto a quanto previsto. Se tali giudizi errati si verificassero con frequenza, l'utilità complessiva del sistema potrebbe risultare notevolmente diversa.

Il problema potrebbe essere superato misurando l'utilità dopo l'impiego dei documenti da parte di chi ha svolto la ricerca. Si può definire l'"utilità definitiva" come il giudizio finale su un documento da parte dell'utente dopo che esso è servito allo scopo per il quale l'utente l'aveva cercato. Si presume qui che per selezione di riferimenti promettenti si intenda che il ricercatore abbia ritenuto che il materiale rappresentato dai riferimenti sia potenzialmente valido. Alla fine tra questi riferimenti potenzialmente validi l'utente potrebbe ritenere che una parte del materiale abbia mantenuto le promesse e sia risultato definitivamente utile. Se si recuperasse più materiale definitivamente utile da un sistema piuttosto che da un altro, si potrebbe concludere che quel sistema è risultato alla fine più utile. Inoltre, se si potesse escogitare il modo di scoprire il giudizio finale di un utente senza domandare quali giudizi si fossero dati in precedenza, si otterrebbe una misura relativamente oggettiva e non condi-

zionata della valutazione soggettiva dell'utente riguardo all'utilità definitiva dei materiali recuperati dal sistema.

L'applicazione di queste misure dipende dall'accesso ai risultati del progetto per il quale l'utente ha condotto la ricerca. Martyn e Lancaster hanno così discusso i risultati: "La sequenza delle immissioni, delle emissioni e dei risultati è complessa in misura crescente. Le immissioni in un servizio di informazioni sono la cosa più facile da misurare. Possono essere espresse quantitativamente [...]. Le emissioni sono più difficili da misurare perché si deve considerare la qualità altrettanto della quantità [...]. I risultati di un servizio di informazioni sono l'aspetto più difficile da studiare, in quanto la nozione dei risultati riporta a quella degli impatti, degli effetti o dei benefici. È più difficile valutare le emissioni che quantificare le immissioni. Probabilmente tutti i tipi di servizi di informazione hanno dati affidabili sulle immissioni, ma pochi hanno dati significativi sulla qualità delle emissioni e probabilmente non esistono dati sui risultati".¹⁰

Una difficoltà nella raccolta dei dati sui risultati consiste nel fatto che sovente si ignora lo scopo per cui viene consultato un sistema di recupero. Inoltre l'impiego finale, se esiste, che si fa dei risultati della ricerca sovente non è disponibile e può non prestarsi ad essere studiato. E poi a quanto pare non ci sono misure certe sulla "bontà del risultato".

Cooper ha postulato un caso ideale in cui "si chiede a un ricercatore di valutare un sistema di recupero di riferimenti o di documenti... [e] utenti reali si rivolgano in gran numero [al sistema] con la speranza di soddisfare bisogni reali di informazione, e che questo sia continuato abbastanza a lungo da rendere omogenea la tipologia degli utenti. È opportuno supporre che sia omogeneo anche il carattere della raccolta dei documenti su cui opera il sistema".¹¹

Si possono porre in tal modo le condizioni ottimali per condurre uno studio sull'efficacia del recupero:

1) in un ambiente informativo che operi su una raccolta omogenea; 2) con un gruppo omogeneo di utenti; 3) che abbiano bisogni reali di informazione; 4) per i quali sia possibile specificare: lo scopo della ricerca e il tempo approssimativo in cui si conoscerà il risultato finale; 5) per lo studio dei quali si possa utilizzare alla fine un risultato chiaro.

Si potrebbero trovare queste condizioni in una biblioteca accademica, in quanto essa possiede una raccolta omogenea di documenti (Condizione 1). I docenti e gli studenti costituiscono un gruppo di utenti con bisogni reali di informazioni (Condizioni 2 e 3). Questi utenti cercano materiale per i loro contributi accademici (Condizione 4). Il contributo stesso, con i suoi riferimenti bibliografici, può essere considerato ragionevolmente un risultato chiaro della ricerca (Condizione 5).

Martyn e Lancaster¹² hanno notato che lo studio dei sistemi stabiliti tende a limitarsi alla teoria, perché i progetti sperimentali interferiscono con l'uso della biblioteca da par-

⁸ W.S. COOPER, *On Selecting... Part I*, cit., p. 100.

⁹ W.S. COOPER, *On Selecting... Part II*, cit., p. 417-418.

¹⁰ J. MARTYN-F.W. LANCASTER, *Investigative Methods in Library Science: an Introduction*, Arlington, Information Resources Press, 1981, p. 199.

¹¹ W.S. COOPER, *On Selecting... Part I*, cit., p. 88.

¹² J. MARTYN-F.W. LANCASTER, *Investigative Methods*, cit., p. 197.

te dei frequentatori. Sembra comunque possibile studiare i ricercatori accademici senza interferenze indebite. L'impegno intellettuale per scrivere un contributo accademico può essere considerato come un processo continuo di raccolta e valutazione di informazioni. Dal lavoro di Cooper sulle metodologie di valutazione viene un sostegno teorico per interpretare il giovamento come utilità definitiva dei materiali per l'utente. Comunque la connotazione dell'utilità definitiva differisce da quella di Cooper; si presume qui che l'utilità definitiva rimanga sconosciuta finché il risultato della ricerca, in questo caso un contributo accademico con i suoi riferimenti bibliografici, sia divenuto disponibile per essere studiato. In queste condizioni la ricerca sperimentale può accertare quale sistema abbia prodotto più materiale utile in via definitiva per le persone impegnate in un progetto noto iniziato in un certo tempo e per il quale ci si aspetti un risultato finale. La metodologia potrebbe confrontare le prestazioni di qualsiasi forma di recupero (ad es. cataloghi o indici automatici, manuali o stampati, oppure basi dati bibliografiche) cercata con qualsiasi sistema (ad es. con vocabolario controllato o no, classificato, con riferimenti, a testo completo o ibrido). Naturalmente le ipotesi per uno studio particolare dipenderebbero dalla struttura teorica di quella ricerca. Comunque per esemplificare la metodologia generale e per delinearne più specificamente le variabili si può ipotizzare una prova con due cataloghi manuali a registrazioni multiple, uno con intestazioni per soggetto e l'altro sistematico. La discussione di questo esempio esige il trattamento specifico dei problemi propri dell'esempio. Per qualunque tipo o forma di sistema si può impiegare la metodologia generale, ma nell'adattare ad ogni prova il modello del progetto si devono applicare i problemi propri di ciascuna situazione sperimentale.

SVILUPPO DELLE VARIABILI E LORO MISURA

Si è ritenuto a lungo che i due scopi del catalogo di una biblioteca siano "rendere possibile il reperimento di un libro del quale [...] si conosca il soggetto" e "mostrare che cosa la biblioteca possiede [...] su un soggetto determinato".¹³ Di solito non viene messo in discussione il valore della ricerca per soggetto, ma le opinioni sui mezzi migliori per ottenerla risultano divise.¹⁴ Di recente gli studi si sono concentrati sull'accesso verbale ma si è anche sperimentato l'accesso sistematico,¹⁵ e le dispute sul valore della visualizzazione ver-

bale e sistematica in linea riaprono il dibattito sui meriti di base dei due accessi.

Ranganathan definì il catalogo sistematico come uno "nel quale alcune registrazioni sono numeriche ed alcune sono verbali" e che perciò contiene "due parti — la Parte Classificata e la Parte Alfabetica". La parte sistematica "aiuta un lettore ad avere una veduta panoramica delle registrazioni di tutti i documenti nel campo particolare del suo interesse, ordinate secondo i loro soggetti specifici in una sequenza utile".¹⁶ La seconda parte contiene registrazioni verbali in ordine alfabetico. Quelle che si riferiscono ai soggetti rinviano gli utenti al punto appropriato per iniziare la ricerca nella serie classificata. La ricerca comporta uno spostamento gerarchico avanti e indietro fino ad incontrare il soggetto esatto.

Ritenere utile la presentazione sistematica dei soggetti affini costituisce un fondamento logico essenziale quanto convenzionale per molte teorie e sistemi di classificazione bibliografica, e questa veduta concorda con le scoperte della ricerca sui processi conoscitivi.^{17, 18} Comunque la convinzione che i dati sistematici migliorino i sistemi di recupero delle informazioni non è sostenuta dalla ricerca in biblioteconomia e nella scienza delle informazioni. Ad esempio, Keen trovò che di solito i sistemi classificati non davano risultati migliori di quelli non classificati. Dal che concluse che sarebbe stata una perdita di tempo sviluppare il recupero classificato. Comunque egli sospettava anche che queste prove presupponessero "un'opinione errata sugli obiettivi dei sistemi di recupero dei documenti", perché non si era tenuto conto a sufficienza delle svariate esigenze dei diversi utenti.¹⁹

Ranganathan argomentava che il nostro obiettivo è di aiutare le persone a trovare quello che cercano e che il catalogo sistematico è ben adatto ad ottenere questo scopo. Se combiniamo il sospetto di Keen che le prove sull'efficacia abbiano messo a fuoco gli obiettivi sbagliati, la posizione di Ranganathan sugli obiettivi corretti per i sistemi di recupero e il lavoro di Cooper sulle metodologie di valutazione, possiamo verificare l'efficacia relativa di recupero dei sistemi per soggetto sulla base di quale risulti alla fine più utile alle singole persone.

Progetto e variabili della ricerca

Si potrebbe raggiungere questo scopo con un esperimento

¹³ C.A. CUTTER, *Rules for a Dictionary Catalog*, 4^a ed., Washington, US Government Printing Office, rist.: London, The Library Association, 1962, p. 12.

¹⁴ J.E.L. FARRADANE, *Subject Headings versus Classified Catalogue (in a Library)*, "Cataloguing Australia", Jan.-June: 5-7, 1976.

¹⁵ K. MARKEY-A. DEMEYER, *Dewey Decimal Classification Online Project: Evaluation of a Library Schedule and Index Integrated into the Subject Searching Capabilities of an Online Catalog*, Final Report to the Council on Library Resources, OCLC/OPR/RR-86/1, Dublin, Ohio, OCLC Online Computer Library Center, Office of Research, 1986.

¹⁶ S.R. RANGANATHAN, *Classified Catalogue Code*, 5^a ed., Bombay, Asia Publishing House, 1964, p. 167.

¹⁷ R.W. SCHVANENVELDT, D.E. MEYER, C.A. BECKER, *Lexical Ambiguity, Semantic Context and Visual Recognition*, "Journal of Experimental Psychology: Human Perception and Performance", 2 (2): 243-256, 1976.

¹⁸ D.E. KIERAS, *The Role of Referents and Sentence Topics in the Construction of Passage Macrostructure*, "Discourse Processes", 4: 1-15, 1981.

¹⁹ E.M. KEEN, *Prospects for Classification Suggested by Evaluation Tests*, in *Classification in the 1970s: a Second Look*, a cura di A. Maltby, London, Clive Bingley, p. 155-156.

sul campo che mettesse a confronto i due sistemi manuali che costituiscono la variabile indipendente dello studio. L'efficacia del recupero è considerata una funzione del numero di ricerche positive condotte a termine. Una ricerca positiva comprende tre distinti elementi successivi. Il ricercatore deve: 1) trovare un riferimento promettente a un documento durante una prima ricerca per soggetto; 2) leggere il documento; 3) citare il documento nel contributo accademico per il quale era stata condotta la ricerca.

Non sembra che queste condizioni rigorose per le ricerche positive siano state proposte in precedenza. Queste tre condizioni per una ricerca positiva costituiscono le tre variabili dipendenti del progetto di studio. Le ipotesi nulle stabilirebbero che nessuno dei sistemi di informazione allo studio darebbe risultati migliori di un altro. Lo studio prevede un progetto per un gruppo di controllo successivo alla prova²⁰ nel quale i partecipanti vengono assegnati a caso ai due trattamenti. La campionatura casuale assicura che le caratteristiche personali dei partecipanti allo studio siano distribuite equamente. La Tabella 1 prende in considerazione il progetto successivo alla prova per lo studio che abbiamo ipotizzato. Verranno delineati in seguito i fondamenti teorici per stabilire la variabile indipendente e le misure dell'efficacia del recupero (variabili dipendenti).

La variabile indipendente

In questo esempio la variabile indipendente consiste in due cataloghi per soggetto, ciascuno dei quali fornisce l'accesso per soggetto a un gruppo di documenti disponibile nella biblioteca o nel servizio di informazioni allo studio. Gli stessi documenti compaiono in entrambi i cataloghi, la descrizione catalografica è identica in entrambi e in entrambi si trovano punti di accesso concettualmente equivalenti per ciascun

documento. I cataloghi si possono così confrontare in tutto tranne che in una cosa: l'ordinamento di base, alfabetico o classificato, dei punti di accesso ai documenti.

In generale, ci sono due generi di variabili indipendenti.²¹ Quando i trattamenti differiscono nel genere si ha una variabile qualitativa. Ad esempio, se si prova l'efficacia dell'aspirina e della codeina la variabile è qualitativa, perché si prova se le "qualità" di una sono migliori di quelle dell'altra. Si ha una variabile indipendente quantitativa se i trattamenti non differiscono nel genere, ma nel grado. Ad esempio, se si provano le varie grandezze di dosatura dell'aspirina, la variabile è quantitativa perché si provano quantità diverse del medesimo trattamento.

Nel nostro caso si considera che la variabile indipendente sia quantitativa. Gli ordinamenti alfabetico e sistematico non sono reciprocamente esclusivi. In un soggettoario o in un thesaurò è caratteristica la presenza di elementi propri della classificazione nelle intestazioni rovesciate, nelle intestazioni suddivise e nella struttura sindetica. Del pari, sovente i sistemi di classificazione forniscono un indice ed elenchi alfabetici delle sottoclassi, quando si ritiene che un ordinamento sistematico ulteriore non sia preferibile. Per tale ragione si può considerare che i cataloghi per soggetto differiscano quantitativamente nel grado di ordinamento sistematico. Questo carattere quantitativo significa che possiamo considerare essenziale solo il fondamento logico dell'ordinamento per soggetto; non è necessario considerare gli elementi sistematici del catalogo alfabetico o gli elementi alfabetici del catalogo sistematico come fattori inquinanti. Mildred riconobbe in questo aspetto dei sistemi di accesso per soggetto la caratteristica di contenere "solamente una variabile importante: se la base dell'ordinamento della serie è costituita da concetti (sistematica) o da simboli arbitrari (alfabetica)".²² Sicché le attività di ricerca nei due gruppi differiscono nel grado, ma non nel genere. ➤

Tab. 1 - Progetto e variabili.

Gruppo considerato	Condizioni di trattamento	Riferimenti promettenti recuperati	Documenti letti	Documenti citati
R1	AC	o1	o2	o3
R2	CC	o4	o5	o6
R: Gruppo scelto a caso		AC: Catalogo alfabetico		
o: Osservazione		CC: Catalogo sistematico		

²⁰ D.T. CAMPBELL-J. C. STANLEY, *Experimental and Quasi-experimental Designs for Research*, Chicago, Rand McNally, 1963, p. 25-27.

²¹ R.E. KIRK, *Experimental Design: Procedures for Behavioral Sciences*, Belmont, Calif., Brooks/Cole, 1968, p. 6-7.

²² J.L. MILSTEAD, *Subject Access Systems: Alternatives in Design*, Orlando, Fla, Academic Press, 1984, p. 45.

Le tre variabili dipendenti

“Riferimenti promettenti recuperati”. La prima variabile dipendente è il numero di “riferimenti promettenti recuperati” da un ricercatore. I dati si raccolgono chiedendo a ogni partecipante di registrare l'informazione bibliografica per ogni riferimento trovato durante una prima ricerca che sembri indicare una possibile pertinenza del documento.

“Rilevanza” e “pertinenza” sono termini onnipresenti: Saracevic ne riporta un sommario delle definizioni e delle connotazioni.²³ Una norma consiste nel distinguere quei concetti sulla base di chi giudica le dimensioni di un confronto tra un documento e una richiesta o tra un documento e un bisogno. Ad esempio, Lancaster impiegò i termini “rilevanza in riferimento a una relazione tra un documento e una richiesta, basata sulla decisione soggettiva di una o più persone, e pertinenza in riferimento a una relazione tra un documento e un bisogno di informazione, ed in questo caso la decisione era presa esclusivamente dalla persona che aveva il bisogno di informazione”.²⁴

In questo caso noi adottiamo la decisione di Lancaster, sicché ci preoccupiamo della “pertinenza” attribuita soggettivamente da un partecipante a un documento rispetto ai propri bisogni di informazione, non della “rilevanza” rispetto alle richieste che un altro potrebbe vedere nel documento.

Comunque un partecipante a questo studio svolgerebbe le ricerche in un catalogo contenente solo sostituti dei documenti. Secondo la distinzione di Lancaster, sarebbe impensabile sostenere che un partecipante ha dato un “giudizio di pertinenza” (ossia un giudizio sul fatto che un certo documento abbia soddisfatto o meno il suo bisogno di informazione) fin tanto che il partecipante non abbia esaminato effettivamente il documento. Belzer, impiegando il termine “rilevanza” per il concetto da noi chiamato “pertinenza”, notò che “se un utente ricevesse qualcosa meno del documento completo — ad es. una citazione o un riassunto — non potrebbe essere del tutto sicuro che il documento intero sia o meno rilevante. Egli dovrebbe quindi fare previsioni e si potrebbe a volte verificare il caso che la previsione risulti errata”.²⁵ Nell'esempio presentato qui i partecipanti fanno previsioni riguardo alla pertinenza dei documenti sulla base di una loro rappresentazione che li sostituisce. Non sarebbe giustificato chiamare “giudizi di pertinenza” la selezione di sostituti, in quanto una ricerca catalografica permette solo di fare “previsioni sulla pertinenza”; per via di principio si presume impossibile che un utente formuli un “giudizio sulla pertinenza” nei confronti di un documento non visto né tanto meno esaminato.

Per questa ragione i riferimenti selezionati da un partecipante sono chiamati “riferimenti promettenti recuperati”. Questo termine riconosce sia la distinzione cronologica tra una previsione di pertinenza e un giudizio di pertinenza che il fatto che gli utenti scelgono i documenti da esaminare sulla base di elementi che la descrizione fa ritenere promettenti per il loro scopo. In questo senso la previsione sulla pertinenza fatta dagli utenti di un sistema sostituito è un atto che implica che la previsione si rivelerà esatta. Essa comunque non implica che l'utente non possa, in seguito all'esame, respingere come inutile il documento reale. Così, le previsioni di pertinenza positive non impegnano l'utente a un futuro giudizio di pertinenza positivo sul documento stesso. I “riferimenti promettenti recuperati” si possono definire come i riferimenti che un utente registra di aver recuperato, sui quali egli intende fare una previsione di pertinenza positiva, condizionata da un esame diretto del documento reale.

“Documenti letti”. La seconda condizione per una ricerca positiva è che un documento rispetto al quale sia stato recuperato un riferimento promettente venga in seguito letto. Belzer, usando “rilevanza” per quello che noi chiamiamo “pertinenza”, notò che “se un utente legge tutto un documento pubblicato, sa se esso è rilevante o no per lui [...]. Qui la rilevanza assume una connotazione funzionale. [...] A lettura ultimata del documento l'utente saprebbe esattamente se esso è rilevante o no per lui”.²⁶

Così, la seconda variabile dipendente sarebbe determinata dal numero dei documenti letti da un partecipante tra quelli per i quali erano stati recuperati riferimenti promettenti durante la prima ricerca al catalogo.

Pare che l'espressione “documenti letti” sia stata impiegata solo una volta. In uno studio del Lucis (London University Central Information Service), Vickery e Batten intervistarono alcuni utenti “alcuni mesi” dopo il completamento delle loro ricerche. Essi domandarono agli utenti “quale proporzione dei nuovi riferimenti rilevanti che avevano trovato [...] avessero letto”.²⁷ Un partecipante a questo studio, Lantz, rispose di aver letto tutti i “riferimenti trovati”, ma non tutti i documenti ai quali essi facevano riferimento. Lantz suggerì che “riferimenti letti” fosse cambiato in “documenti letti”. Sugerì inoltre che il rapporto tra i “riferimenti rilevanti recuperati” (termine che egli sostituì al posto di “riferimenti rilevanti trovati” e che noi chiamiamo “riferimenti promettenti recuperati”) e i “documenti letti” costituirebbe un'utile, doppia valutazione dell'efficacia del recupero perché entrambi sono “ragionevolmente indipendenti dal sistema”.²⁸

Nell'esperimento qui delineato, i dati per i “documenti letti” verrebbero raccolti da un questionario compilato dai

²³ T. SARACEVIC, *Relevance*, cit.

²⁴ F.W. LANCASTER, *Information Retrieval Systems: Characteristics, Testing and Evaluation*, New York, John Wiley, 1979, p. 63.

²⁵ J. BELZER, *Information Theory as a Measure of Information Content*, “Journal of the American Society for Information Science”, 24 (4) July-August: 300-304, 1973, p. 301.

²⁶ J. BELZER, *Information Theory*, cit., p. 301.

²⁷ A. VICKERY-A. BATTEN, *Large-scale Evaluation Study of On-line and Batch Computer Information Services*, University of London: Library Resource Co-ordinating Committee, 1978, p. 43.

²⁸ B.E. LANTZ, *The Relationship between Documents Read and Relevant References Retrieved as Effectiveness Measures for Information Retrieval Systems*, “Journal of Documentation”, 37 (3) Sept.: 134-145, 1981, p. 135.

partecipanti per ciascuno dei loro "documenti promettenti recuperati". Il questionario sarebbe distribuito un po' di tempo dopo la ricerca, in modo da dar tempo al partecipante di leggere un documento se, a un primo esame, lo avesse ritenuto pertinente. Domande essenziali per raccogliere i dati sarebbero "Avete esaminato questo documento?"; "Avete letto completamente questo documento?"; "Avete letto in parte questo documento?". Si chiederebbe ai partecipanti se hanno esaminato ciascun documento cosicché, se la risposta fosse negativa, esso non conterebbe come "documento esaminato ma non letto (e perciò non pertinente)" e potrebbe essere eliminato dall'analisi statistica. Secondo Belzer si presume che un utente decida di leggere un documento dopo averlo esaminato: se, comunque, per qualche ragione un utente non esamina il documento, egli non ne dà un giudizio di pertinenza e non può dire se la previsione di pertinenza si sia o no avverata. Si domanderebbe ai partecipanti se hanno letto in tutto o in parte il documento, perché un partecipante potrebbe giudicare pertinente una parte di un documento (ad es. un capitolo di una monografia, un articolo in una raccolta). In quel caso, probabilmente il partecipante avrebbe letto solo la parte pertinente. Poiché occorre sapere che cosa ha letto il partecipante per quantificare la terza variabile dipendente, la distinzione tra documenti completi e parziali aumenta la precisione dei dati del questionario. Quindi "documenti letti" si può definire come il numero dei documenti completi o parziali che il partecipante ha riferito di aver letto tra i suoi riferimenti promettenti.

"Documenti citati". La terza condizione per una ricerca positiva è che un documento letto in seguito a un recupero di informazione dev'essere citato nel contributo accademico del partecipante. I dati per la terza variabile dipendente verrebbero raccolti dalle note a piè di pagina o dalle bibliografie dei contributi dei partecipanti.

Cronin avvertì che i dati delle citazioni "sono stati utilizzati con fantasia per molteplici usi",²⁹ ma non pare che l'analisi delle citazioni sia stata applicata all'efficacia del recupero. L'analisi, delle citazioni si è attirata una varietà di critiche;³⁰ una delle più importanti è la mancanza di una teoria che spieghi perché gli autori citano le opere che in effetti citano.³¹ Sembra perciò che nessuna posizione teorica o ricerca empirica giustifichi l'impiego delle citazioni in un documento accademico come l'ultima di una serie di condizioni successive per misurare l'efficacia del recupero. Si conviene ugualmente, comunque, che una citazione è almeno in un certo senso un "indicatore di uso".³² Cronin ha osservato che citare il lavoro di altri "non è del tutto casuale o bizzarro"³³ e che "l'evidenza non ci permette di concludere che la pratica è caratterizzata dal caso e dall'incoerenza".³⁴

Questa supposizione — che il comportamento della citazione sia significativo — sta alla base di tutta la letteratura sull'analisi delle citazioni. Si può allora postulare che in un contributo accademico la citazione di un documento per il quale sia stato recuperato un riferimento promettente e che successivamente sia stato letto sia uno dei risultati valutabili nella preparazione del contributo. Questo suggerimento non tenta di motivare una citazione, ma soltanto di offrire una nuova applicazione del presupposto che il comportamento di chi fa una citazione sia razionale e non "bizzarro". Questa applicazione è il suggerimento che si può ragionevolmente considerare una citazione come una riconferma della correttezza della previsione e del giudizio sulla pertinenza, da parte del partecipante, registrati nei dati per le prime due variabili dipendenti. Si potrebbero calcolare le citazioni nei contributi dei partecipanti confrontando i dati relativi ai riferimenti promettenti recuperati e ai documenti letti con le note a piè di pagina o con la bibliografia di ciascun contributo. Sicché si può definire un "documento citato" come un documento per il quale era stato registrato un riferimento promettente, per il quale era stata riferita una lettura totale o parziale, ed al quale è comparsa una citazione nel contributo del partecipante.

ALTRE CONSIDERAZIONI

Presupposti e limitazioni

Presupposti non provati da questa metodologia sono:

- 1) che per le persone la misurazione dell'utilità definitiva dei riferimenti recuperati è un elemento valido per valutare l'efficacia del recupero per soggetto;
- 2) che la combinazione sequenziale delle tre variabili dipendenti è una misura ragionevole dell'utilità definitiva del sistema per gli utenti;
- 3) che un contributo accademico con i suoi riferimenti bibliografici può essere considerato ragionevolmente una conseguenza riconoscibile della prima ricerca di documentazione per soggetto.

McGrath descrisse i punti forti e deboli delle varie strategie di ricerca per illustrare quello che egli chiamava il "dilemma a tre corni" nella scelta di un progetto di ricerca: "Ogni attestazione della ricerca coinvolge delle persone (qui, A per Attori) che fanno qualcosa (qui, B per Behavior) in un ambiente riferibile a tempo, luogo, oggetto (qui, C per Contesto). È *sempre desiderabile (ceteris paribus) massimizzare*: (A) *l'ampiezza del campione* rispetto alle persone, (B) *la precisione* nel controllo e nella misura delle variabili rispetto ai comportamenti dell'oggetto studiato, e C *la realtà* effettiva rispetto ai partecipanti ▶

²⁹ B. CRONIN, *The Need for a Theory of Citing*, "Journal of Documentation", 37 (1) March: 16-24, 1981, p. 16.

³⁰ L.C. SMITH, *Citation Analysis*, "Library Trends", 30 (1): 83-106, 1981.

³¹ B. CRONIN, *The Need for a Theory of Citing*, *ivi*.

³² J. MARTYN-F.W. LANCASTER, *Investigative Methods in Library Science*, cit., p. 88.

³³ B. CRONIN, *The Citation Process: the Role and Significance of Citations in Scientific Communication*, London, Taylor Graham, 1984, p. 83.

³⁴ *Ivi*, p. 84.

del contesto entro il quale si osservano quei comportamenti”³⁵. Massimizzare un dato elemento comporta automaticamente l'impossibilità di massimizzare gli altri due. Ad esempio, gli esperimenti sul campo presentano la realtà effettiva rispetto ai partecipanti, riescono a raggiungere un controllo abbastanza preciso sulle variabili, ma hanno solo una limitata ampiezza del campione al di là degli utenti particolari. “*In linea di principio non c'è alcun modo di massimizzare tutti e tre gli elementi (in conflitto) nell'ambito della strategia della ricerca*”³⁶. Così, si deve scegliere una strategia i cui vantaggi e gli inevitabili svantaggi che la caratterizzano offriranno l'ultimo modo sia pur condizionato per rispondere alla domanda sulla ricerca. Nello stesso tempo qualunque scelta non permetterà inevitabilmente di massimizzare gli altri elementi. Queste limitazioni si applicano alla metodologia qui delineata. Per verificarle, occorrerebbe una ricerca con vari gruppi di persone e con vari sistemi di informazioni.

La variabile indipendente

I cataloghi sperimentali nel nostro esempio hanno due scopi. In primo luogo, devono agire da variabili indipendenti controllate in maniera adeguata. Kerlinger³⁷ avvertì che un inconveniente degli esperimenti sul campo può essere dato dall'incapacità di manipolare questa variabile. In secondo luogo, i cataloghi devono essere “normali” cataloghi di biblioteca. Il risultato di fissare condizioni ottimali per accertare l'utilità definitiva comporta che gli utenti cerchino un sistema “reale” per uno scopo “reale” in un modo “reale”. Per queste ragioni occorre spiegare i principi che stabiliscono le forme fisica e intellettuale dei cataloghi. In teoria i cataloghi dovrebbero differenziarsi solo nell'elemento sostanziale, cioè la base dell'ordinamento delle registrazioni per soggetto. Da questa teoria conseguono due esiti interdipendenti: l'uguaglianza fisica dei cataloghi e la somiglianza intellettuale dell'accesso per soggetto da essi fornito.

Occorrono due cataloghi per i documenti in esame. Si può ottenere una descrizione catalografica identica dall'inventario topografico della biblioteca in esame con metodi automatizzati o manuali (ad es. per fotocopia), ma la difficoltà principale di provvedere due cataloghi per gli stessi documenti consiste nel controllare che la precisione e la profondità dell'analisi per soggetto siano sostanzialmente simili. Si deve mantenere più costante possibile l'analisi concettuale dei documenti, in modo che l'unico fattore di differenza tra i cataloghi sia il loro ordinamento di base. Certamente le conseguenze sollevate da questo punto non sono state esplorate nei precedenti progetti di ricerca sull'efficacia del recupero. Per ottenere un'analisi per soggetto sostanzialmente simile i punti di accesso per ciascun documento dovrebbero esprimere il medesimo contenuto concettuale in ciascun sistema. Il principio di mantenere i cataloghi concettualmente costanti nasce dalle esigenze che governano la ricerca sperimenta-

le. È necessario che si diversifichi solamente la caratteristica del criterio di accesso (ossia, nel nostro caso, l'ordinamento di base dei cataloghi). Se un catalogo si basasse su analisi più complete dell'altro, si potrebbe presumere che esso fornisca un accesso per soggetto migliore, indipendentemente dall'organizzazione dei punti di accesso. Ad esempio, a un livello elementare, se si assegnano a un documento due intestazioni di soggetto nel catalogo alfabetico ma quattro notazioni in quello sistematico, si deve presumere che un utente abbia una probabilità doppia di trovare il documento nel catalogo sistematico. Non pare ragionevole presumere che queste ineguaglianze vengano eliminate dalle variazioni casuali nel numero delle intestazioni assegnate per una grande quantità di documenti. Per eliminare questa disuguaglianza potenziale ogni documento dovrebbe avere lo stesso numero di intestazioni verbali e di notazioni. Questa procedura assicura che un utente che potrebbe trovare promettente un certo riferimento abbia la medesima probabilità di trovarlo in entrambi i cataloghi.

È complesso stabilire per ciascun documento un'effettiva somiglianza concettuale dell'accesso per soggetto. Supposto che i punti di accesso siano concettualmente simili quanto più possibile, quali dovrebbero essere? Da un punto di vista il metodo ideale per fornire una somiglianza concettuale potrebbe sembrare quello di creare due cataloghi del tutto nuovi. Chi conduce l'esperimento analizzerebbe ciascun documento per esprimere quell'analisi impiegando i sistemi in prova. Ciascun sistema fornirebbe lo stesso numero di punti di accesso per ciascun documento e si presupporrebbe che in questi punti di accesso sia presente la medesima analisi concettuale. Comunque da questa soluzione sorgono due inconvenienti. In primo luogo, è dubbia la supposizione che si avrebbe la medesima analisi concettuale. Senza volerlo chi conduce l'esperimento, in quanto preferisce una forma di catalogo rispetto all'altra o è più familiare con un sistema che con l'altro, potrebbe sfruttare meglio le possibilità analitiche di quel sistema. Questa situazione costituirebbe una forma inaccettabile di condizionamento da parte di chi conduce l'esperimento. In secondo luogo, è normale che in una biblioteca avvengano periodicamente delle sostituzioni nei cataloghi. Le nostre condizioni ideali richiedono che studiamo un sistema informativo stabilizzato in un ambiente normale di ricerca di informazioni. Così, i cataloghi nuovi aumentano il condizionamento indesiderato da parte di chi conduce l'esperimento e diminuiscono l'influenza desiderata dell'ambiente reale.

Sembra preferibile un metodo alternativo di produrre i cataloghi. Questo metodo consiste nell'accettare qualsiasi analisi per soggetto sia già stata fatta. Quest'analisi esiste nelle intestazioni per soggetto e nelle segnature già assegnate dai vari catalogatori. Chi conduce l'esperimento accetterebbe le intestazioni per soggetto e la segnatura di un documento già esistenti. I concetti presenti in ciascun sistema verrebbero allora tradotti nell'edizione corrente dell'altro sistema con le

³⁵ J.E. McGRATH, *Dilemmatics, the Study of Research Choices and Dilemmas*, in *Judgement Calls in Research*, a cura di J.E. McGrath, J. Martin e R.A. Kulka, Beverly Hills, Calif., Sage Publication, 1982, p. 74.

³⁶ J.E. McGRATH, *Dilemmatics*, cit., p. 76.

³⁷ F.N. KERLINGER, *Foundations of Behavioral Research*, 2^a ed., New York, Holt Rinehart and Winston, 1973.

sue integrazioni e i cambiamenti ufficiali al tempo in cui si allestivano i cataloghi. Questa procedura assicura automaticamente che in ciascun catalogo compaia lo stesso numero di punti di accesso impiegati per ciascun documento e che la dimensione complessiva dei cataloghi sia la stessa. Inoltre, l'analisi concettuale non dipenderebbe dalla conoscenza dei sistemi o del campo concettuale del documento da parte di chi conduce l'esperimento, sicché una fonte del condizionamento da parte di chi conduce l'esperimento risulterebbe attenuata. Senza contare che le incongruenze nell'approfondimento o gli errori nella precedente analisi per soggetto verrebbero conservati in ambedue i cataloghi e manterrebbero la corrispondenza desiderata con l'ambiente della ricerca di informazioni. Se si rianalizzassero i documenti senza alcun riferimento all'analisi per soggetto precedente, non si potrebbe ritenere che entrambi i cataloghi rivelino somiglianza concettuale e presentino effettivamente condizioni reali anche nel caso in cui il numero delle notazioni eguagliasse il numero delle intestazioni per soggetto per un documento.

Così, la creazione di cataloghi sulla base delle analisi per soggetto verbali e sistematiche già esistenti risponde alle obiezioni sulla creazione di cataloghi nuovi. Comunque questa soluzione esige la capacità di assegnare intestazioni per soggetto e notazioni con una corrispondenza concettuale di uno a uno, impiegando come base la precedente analisi per soggetto di un documento. Questa esigenza solleva un certo numero di questioni. L'autore intraprese una ricerca nel 1986 dimostrando che si poteva ottenere un'effettiva somiglianza concettuale tra la *Classificazione decimale Dewey* (DDC19) e le *Library of Congress Subject Headings* (LCSH 10), a patto di applicare certe modificazioni o estensioni nell'impiego della DDC e delle LCSH, secondo determinate raccomandazioni specifiche fissate allo scopo. Si potrebbero sviluppare raccomandazioni simili per "tradurre" i concetti tra altri sistemi di analisi per soggetto. Lo scopo di tali raccomandazioni è di mantenere le caratteristiche fondamentali dell'accesso per soggetto sistematico e verbale in modo che i cataloghi possano comportarsi da variabile indipendente quantitativa sufficientemente controllata. Questo arricchimento dei soggetti nelle intestazioni catalografiche, sebbene assomigli poco agli studi sui sistemi di accesso per soggetto nei quali la segnatura determina da sola l'accesso sistematico,³⁸ si rivela necessario affinché la probabilità di reperire il documento in un catalogo non sia maggiore della probabilità di reperirlo nell'altro.

In un esempio che confronta un catalogo sistematico ed uno verbale si deve cercare una soluzione diversa. In un'intestazione per soggetto la ricerca catalografica e i vocabolari delle voci sono fusi in un'unica sequenza alfabetica. La "struttura sindetica può essere considerata come una specie di indice all'indice, che guida l'utente dal termine con il quale il sistema è stato interrogato fino alla posizione di quel concetto nel sistema".³⁹ La struttura sindetica per

un catalogo sperimentale con intestazioni di soggetto potrebbe essere generata da quella impiegata nella biblioteca o nel servizio di informazioni allo studio. Comunque in un catalogo sistematico i due vocabolari sono separati. Il vocabolario per la ricerca appare nella serie ordinata concettualmente, ma quello delle voci consiste in un indice alfabetico esterno.

La ricerca in un catalogo sistematico è descritta di solito come un processo in due fasi, perché il ricercatore passa da un indice verbale esterno a una notazione nella sequenza classificata. Per contro, la ricerca in un catalogo con intestazioni di soggetto è descritta come un processo a fase unica, perché l'utente rimane entro la stessa sequenza e cerca la stessa forma di espressione. Comunque questa distinzione sembra valida solo in parte, perché le strutture sindetiche costituiscono un indice interno sia ai termini usati che a quelli non usati. Anche se un ricercatore può individuare un termine appropriato al primo colpo, ed in tal caso si può dire che ha evitato l'indice interno, la presenza di rinvii e di richiami può suggerirgli altri termini utili. Inoltre l'utente può scoprire, dopo aver cercato quella che pareva fosse una "buona" intestazione, che le suddivisioni di quell'intestazione sarebbero state più precise. Per queste ragioni sostenere che una sequenza sistematica richieda due fasi e che una alfabetica ne richieda solo una sembra un eccesso di semplificazione. Pare più ragionevole sostenere che l'accesso ai due cataloghi è essenzialmente simile e che ciascun vocabolario è appropriato al tipo di sistema per il quale è stato sviluppato. Così possiamo ritenere che le diversità che si notano nelle modalità di accesso alle registrazioni da parte degli utenti siano irrilevanti per questa metodologia. Qualunque sia l'indice fornito alla sequenza sistematica (ad es., indice a catena, stringhe PRECIS), per lo scopo di uno studio che impieghi questa metodologia esso può considerarsi analogo alla struttura sindetica di un catalogo con intestazioni per soggetto. Milstead⁴⁰ ha basato su premesse consimili il confronto delle caratteristiche degli ordinamenti sistematico ed alfabetico.

VALORE DELLA METODOLOGIA

Era scopo di questo articolo delineare la teoria ed i principi di una metodologia sperimentale conveniente a perfezionare la ricerca sull'efficacia del recupero e di metterla in grado di valutare l'utilità dei riferimenti recuperati alla luce del risultato finale della ricerca di un utente. I principi che abbiamo esaminato forniscono misure pratiche che si possono confrontare positivamente con la metodologia "semplice" di Cooper.

In questo esempio prolungato la discussione si è concentrata sulla ricerca con accesso per soggetto. Alcuni studiosi sostengono che la presentazione sistematica può aiutare ➤

³⁸ K. MARKEY, *Subject-searching Experiences and Needs of Online Catalog Users: Implications for Library Classification*, "Library Resources & Technical Services", 29 (1) Jan./Mar.: 34-51, 1985.

³⁹ J.L. MILSTEAD, *Subject Access System*, cit., p. 61.

⁴⁰ *Ivi*.

gli utenti⁴¹ ma altri sono più cauti.⁴² Studi recenti sui cataloghi in linea hanno accertato che sovente gli utenti fanno più ricerche per soggetto di quanto non richiedano documenti già conosciuti e che si sono riscontrati risultati simili con una varietà di metodologie come inchieste, questionari, analisi del giornale delle transazioni, interviste a gruppi determinati e individuali, attività di ricerca e studi di casi.⁴³ Con eccezioni parziali per le analisi del giornale delle transazioni e per le attività della ricerca, i ricercatori non hanno sviluppato metodi obiettivi per valutare l'utilità finale per gli utenti dei sistemi di informazione. Inoltre i lavori e l'ambiente della ricerca di informazioni previsti dai progetti precedenti erano per lo più simulati. Se non sono coinvolti attivamente i bisogni e gli interessi dei singoli, non si possono trasferire i risultati a un ambiente di non ricerca. Pejtersen e Austin hanno osservato che nelle prove di simulazione i ricercatori non avevano "alcun sentimento personale sulle esigenze della ricerca che avevano avuto come compito e che perciò sceglievano gli elementi di ricerca 'più facili', senza considerare gli altri concetti".⁴⁴ Essi conclusero che si poteva rispondere alle domande essenziali sul loro sistema solo con un'"ampia investigazione delle ricerche reali".⁴⁵

È utile avere un metodo per valutare il vantaggio definitivo dei sistemi con accesso per soggetto che sia relativamente non soggettivo e non simulato, in una macchina oppure in un ambiente manuale. Si sono sviluppati qui due accorgimenti per la misurazione, allo scopo di aumentare il repertorio delle metodologie sperimentali. Il primo, "documenti letti", come primo controllo sulla previsione di pertinenza fatta da un utente nello scegliere i riferimenti, non è stato esplorato a fondo. In uno studio sul recupero della narrativa, Pejtersen e Austin hanno riconosciuto l'esigenza di questo genere di misurazione: "ci occorre stabilire che cosa pensano gli utenti dei romanzi recuperati dal sistema dopo che li hanno *letti* effettivamente. Il divario tra l'opinione finale su un'opera da parte di un lettore e le sue aspettative basate sulla sua descrizione nell'indice si deve considerare una misura critica del comportamento di qualsiasi sistema di indicizzazione".⁴⁶

Un altro studio⁴⁷ supponeva che se un utente recupera un riferimento a un documento che ha già letto, allora ha ragione di respingere il riferimento come inadatto sulla base di un'antecedente conoscenza del documento. Questa sup-

posizione pare intuitivamente corretta. Markey e Demeyer si sono astenuti dall'impiegare le misure tradizionali del richiamo e della precisione, perché "se i ricercatori avessero fondato le loro valutazioni di rilevanza sul testo completo di un libro, queste avrebbero potuto essere molto diverse".⁴⁸

Non sembra che il secondo accorgimento, "documenti citati" come misura finale della valutazione di pertinenza dei documenti da parte di un utente, sia stato suggerito in precedenza. Il suo valore potenziale è congettuale, ma pare fondato su supposizioni ragionevoli ed accettate sul meccanismo delle citazioni. "Documenti citati" può essere una misura significativa dell'utilità definitiva di un sistema di recupero per gli utenti se il risultato della ricerca è un contributo accademico che comprenda citazioni bibliografiche e che sia disponibile per lo studio.

Come hanno avvertito Martyn e Lancaster,⁴⁹ i dati sui risultati sono rari e possono mancare del tutto, sicché pare in teoria a buon diritto importante una metodologia che vada oltre il recupero iniziale di un riferimento e cerchi di valutare i risultati in una situazione reale. McGrath, Martin e Kulka hanno fatto notare: "È solo quando metodi differenti ci forniscono informazioni convergenti sullo stesso problema che possiamo parlare di accumulo della conoscenza. Ma un simile *accumulo di sostanza convergente da strumenti divergenti*, — quando e se riusciamo ad ottenerlo — è di gran lunga più vigoroso e di più vasta utilità dei risultati che si possano ottenere da qualsiasi studio".⁵⁰ Così, se metodologie differenti possono ottenere risultati simili a quelli di altri studi, aumenta il grado di applicabilità di tutti gli studi con i risultati relativi. I criteri qui suggeriti per una ricerca positiva sono più rigorosi di quelli proposti nei progetti e nei metodi oggettivi non simulati presentati in precedenza per misurare la valutazione definitiva da parte degli utenti del successo della ricerca. Se si fissano questi criteri e se ne applicano le misure si forniranno solide basi teoriche per la futura ricerca sull'efficacia del recupero con l'impiego di un vero progetto sperimentale. Poiché il progetto sperimentale sviluppato qui è indipendente dal sistema, si potrebbe usare la metodologia per confrontare l'efficacia del recupero in qualsiasi numero di tipi e di formati di sistemi. Il cumulo di risultati da più studi potrebbe contribuire notevolmente alla progettazione di sistemi di recupero dell'informazione. ■

⁴¹ E. SVENONIUS, *Use of Classification in Online Retrieval*, "Library Resources & Technical Services", 27 (1) Jan./Mar.: 76-80, 1983.

⁴² J.S. HILL, *Online Classification Number Access: Some Practical Considerations*, "Journal of Academic Librarianship", 10 (1): 17-22, 1984.

⁴³ K. MARKEY, *Subject Searching in Library Catalogues: Before and After the Introduction of Online Catalogs*, Dublin, Ohio, OCLC, 1984.

⁴⁴ A.M. PEJTERSEN-J. AUSTIN, *Fiction Retrieval and Evaluation of a Search System Based on Users' Value Criteria. Parts 1 and 2*, "Journal of Documentation", 39 (4) Dec.: 230-246 e 40 (1) Mar.: 25-35, 1983-84, Part 2, p. 32.

⁴⁵ A.M. PEJTERSEN-J. AUSTIN, *Fiction Retrieval, Part 2*, cit., p. 34.

⁴⁶ *Ivi*.

⁴⁷ D.R. MOREHEAD, A.M. PEJTERSEN, W.B. ROUSE, *The Value of Information and Computer-aided Information Seeking: Problem Formulation and Application to Fiction Retrieval*, "Information Processing and Management", 20 (5-6): 583-601, 1984, p. 598.

⁴⁸ K. MARKEY-A. DEMEYER, *Dewey Decimal Classification Online Project*, cit., p. xxxi.

⁴⁹ J. MARTYN-F.W. LANCASTER, *Investigative Methods in Library Science*, cit.

⁵⁰ J.E. MCGRATH, J. MARTIN, R.A. KULKA, *Some Quasi-rules for Making Judgement Calls in Research*, in *Judgement Calls in Research*, a cura di J.E. McGrath, J. Martin e R.A. Kulka, Beverly Hills, Calif., Sage Publication, 1982, p. 104.