

La sfida della crescita esponenziale delle raccolte nelle biblioteche di ricerca: i primi sistemi di *optical electric information retrieval*

di Fabio Venuda

«It is technology that has got the library into present difficulty, and only technology offers any promise of getting it out again. Only a joint attack on library problems by the librarian and the electronics specialist, working hand in hand, is likely to succeed»¹. Nel 1951, Luis N. Ridenour, fisico e *assistant director* del MIT Radiation Lab, un profano – così si presenta nel testo – della biblioteconomia, propone di risolvere con la tecnologia i problemi derivanti dalle preoccupanti proiezioni, elaborate da Fremont Rider nel 1944, sul costante ed esponenziale aumento delle collezioni nelle grandi biblioteche di ricerca americane². Rider aveva infatti calcolato, sulla base della progressione dei dati relativi al patrimonio librario di *college* e università americane dal 1831 (per le più ‘recenti’ dal 1876) al 1938, che le collezioni delle biblioteche di ricerca raddoppiavano ogni 16 anni, sostenendo anzi che le dimensioni delle raccolte di dieci tra le più grandi università fondate a partire dal 1849 erano raddoppiate non ogni 16, ma ogni 9 anni e mezzo³. Oltre a ciò, tra i molti esempi, dati, medie, calcoli e tabelle portati a supporto della sua tesi, Rider sostiene che, all’inizio del XVIII secolo, la Yale Library probabilmente possedeva circa 1.000 volumi e che, se le sue collezioni fossero effettivamente raddoppiate ogni 16 anni, nel 1938 avrebbe dovuto possedere circa 2.600.000 volumi e questo, a fronte dei 2.748.000 che possedeva realmente, rappresentava, secon-

FABIO VENUDA, Università degli studi di Milano, Dipartimento di Studi storici, Milano, e-mail fabio.venuda@unimi.it.

Ultima consultazione dei siti web: 14 luglio 2022.

1 Louis N. Ridenour, *Bibliography in an age of science*. In: Louis N. Ridenour; Ralph R. Shaw; Albert G. Hill, *Bibliography in an age of science*. Urbana: University of Illinois Press, 1951, p. 5-35: p. 6.

2 Fremont Rider, *The scholar and the future of the research library: a problem and its solution*. New York: Hadham Press, 1944, <<https://hdl.handle.net/2027/mdp.39015033591101>>.

3 *Ivi*, p. 4-6.



do Rider, una sorprendente conferma delle sue proiezioni, «an amazing close correspondence with the 'standard' rate of growth»⁴. Nonostante lo studio di Rider sia stato in parte smentito a posteriori dalla realtà e messo in discussione, soprattutto rispetto al metodo utilizzato, da studi condotti nei decenni successivi⁵, è stato per molto tempo considerato innovativo e la decelerazione dei ritmi iniziali di crescita rilevata negli anni Ottanta, non è stata considerata sufficiente per accantonare i risultati da lui ottenuti, ritenuti comunque da Warren Seibert «weightier and near-venerable»⁶.

Ridenour colloca la causa della crescita esponenziale delle raccolte all'interno dello sviluppo tecnologico nei diversi ambiti disciplinari a cui si stava assistendo già dalla fine dell'Ottocento, ricordiamo a questo proposito che nel 1884 e nel 1887 vengono registrati i due brevetti con cui Herman Hollerith protesse la sua invenzione dell'Electric Tabulating System a schede perforate⁷. «We live in a technological society», sostiene Ridenour, e l'aspetto più significativo dello sviluppo tecnologico della società è il cambiamento, profondo e rapido, dovuto principalmente alla preferenza che le persone sempre più hanno per i beni e i servizi che la tecnologia può produrre e portare nelle loro vite⁸. Lo sviluppo della tecnologia ha comportato, d'altro canto, la produzione di quantità enormi di nuova letteratura scientifica che le biblioteche di ricerca devono immagazzinare, indicizzare e gestire, senza però diminuire i propri standard di servizio, soffrendo per questo cambiamento, mantenendo inalterate le tecniche di immagazzinamento e indicizzazione, e beneficiando poco o nulla delle innovazioni tecnologiche sopraggiunte. In questo scenario le biblioteche hanno continuato a funzionare e lavorare secondo gli stessi principi e le stesse tecniche utilizzate negli ultimi cinquant'anni, in una sorta di conservatorismo comprensibile solo alla luce dei milioni di volumi e altro tipo di materiali che devono gestire e mettere a disposizione, a fronte dei quali anche la più piccola e apparentemente insignificante modifica alle procedure di funzionamento potrebbe risultare estremamente costosa in termini di tempo e risorse, tanto da non poter essere sottovalutata⁹.

Normalmente, anche oggi, per informatizzare dei processi è necessario procedere a un'attenta analisi del modo in cui vengono normalmente svolte le procedure, per scomporle poi in un flusso di operazioni che consenta di definire il modo in cui invece 'dovrebbero' essere svolte con il supporto di sistemi informatici.

Al fine di individuare le tecnologie più adatte per affrontare in modo corretto i problemi derivanti dal sovraccarico delle raccolte che affliggevano le biblioteche, Ridenour, in qualità di studioso e tecnologo estraneo al mondo bibliotecario, procede allo stesso modo ed effettua un'analisi delle principali procedure di funzionamento della biblioteca, ossia, nella sua visione, le acquisizioni (*procurement* o *acqui-*

4 *Ivi*, p. 10.

5 Robert E. Molyneux, *What did rider do? An inquiry into the methodology of Fremont Rider's "The scholar and the future of the research library"*, «Libraries & culture», 29 (1994), n. 3, p. 297-325, <<http://www.jstor.org/stable/25542663>>.

6 Warren F. Seibert, *How libraries grow: a brief look backward (and forward)*, «The journal of academic librarianship», 11 (1985), n. 1, p. 19-23: p. 22.

7 Herman Hollerith, *Apparatus for compiling statistics*, U.S. Patent 395,783A, filed 23 Sept. 1884, <<http://bit.ly/Appstat>>; *Id.*, *Art of compiling statistics*, U.S. Patent 395,781A, filed 8 June 1887, <<http://www.webcitation.org/75ZPBn9EP>>.

8 L. N. Ridenour, *Bibliography in an age of science* cit., p. 13.

9 *Ibidem*.

sition), l'immagazzinamento (*storage*), la catalogazione (*indexing*) e il reference (inteso come messa a disposizione in tempi rapidi del materiale necessario alla ricerca)¹⁰.

Riguardo alle acquisizioni, Ridenour sostiene che, come accade oggi, le collezioni delle biblioteche di ricerca devono sostenere gli interessi dei propri studiosi e quindi devono acquisire e mettere a disposizione tutte le risorse nelle quali viene registrata la conoscenza, di cui i libri sono solo una componente. Tuttavia, viste le risorse limitate, le biblioteche devono necessariamente selezionare le risorse che acquistano, col risultato che le biblioteche di ricerca americane, in realtà, possiedono solo una parte di tutti i libri stampati al mondo, con un elevato grado di sovrapposizione fra le collezioni.

A sostegno di questa tesi, Ridenour porta i risultati dell'indagine, condotta da Leroy C. Merritt e pubblicata nel 1942¹¹, sulla produzione mondiale di libri a stampa fino al 1940, finalizzata a stabilire il rapporto tra quanto è stato pubblicato al mondo e quanto posseduto dalle biblioteche americane, con particolare attenzione alle biblioteche di ricerca. Basandosi e aggiornando il precedente studio di Boleslas Iwinski del 1911¹² – condotto a supporto del progetto di Paul Otlet per la compilazione di una bibliografia universale – che aveva quantificato in 10.378.365 libri e 71.248 periodici le pubblicazioni impresse in tutto il mondo dal momento dell'invenzione della tecnica di stampa a caratteri mobili ideata da Gutenberg nel 1455¹³, Merritt calcola che nel 1940 la produzione mondiale di libri a stampa, non considerando quanto prodotto in Cina, poteva verosimilmente ammontare a oltre 15 milioni di volumi (15.377.276). Il calcolo dei volumi stampati al mondo è, nel complesso lavoro di Merritt, funzionale a effettuare un confronto con quanto posseduto dalle biblioteche americane, delle quali nessuna sembrava arrivare a possederne nemmeno la metà, ad esempio la Biblioteca del Congresso ne possedeva, sempre nel 1940, poco più di 6 milioni e, secondo la stima di Merritt, l'insieme delle biblioteche americane arrivava a possedere circa 10 milioni di titoli differenti, rendendo disponibili negli Stati Uniti circa due terzi delle risorse librarie mondiali¹⁴. Effettuando, tuttavia, un confronto a campione tra le schede appartenenti ai cataloghi delle 46 biblioteche appartenenti all'Association of Research Libraries e quelle del Library of Congress (LoC) Union Catalog, Merritt rileva che il livello di sovrapposizione tra i patrimoni delle biblioteche considerate era estremamente alto: solo sette biblioteche possedevano più dell'1% di titoli non presenti nel catalogo della Library of Congress (LoC), la sola Harvard Library avrebbe aggiunto al catalogo ben il 4% di libri differenti, mentre il patrimonio della grande biblioteca della Princeton University si differenziava dal posseduto rappresentato nel catalogo della LoC solamente per lo 0,56%¹⁵.

In sostanza, per poter rispondere nel più breve tempo possibile – tollerato in termini di minuti o al massimo qualche ora – ai bisogni informativi dei propri studio-

10 *Ivi*, p. 19.

11 C. LeRoy Merritt, *Resources of american libraries: a quantitative picture*. In: *Union catalogs in the United States*, edited by Robert Bingham Downs. Chicago: American Library Association, 1942, p. 58-96: p. 77-82.

12 Boleslas Iwinski, *La statistique internationale des imprimés*, «Bulletin de l'Institut International de Bibliographie», 1911, publication n. 109, p. 1-139, <<https://bit.ly/Iwinski>>. Si veda anche Donald W. Krummel, *The heritage of Boleslas Iwinski*, «Library trends», 62 (2013), n. 2, p. 456-464, DOI: 10.1353/lib.2013.0044.

13 D. W. Krummel, *The heritage of Boleslas Iwinski* cit., p. 456.

14 C. L. Merritt, *Resources of american libraries* cit., p. 82. Cfr. anche L. N. Ridenour, *Bibliography in an age of science* cit., p. 18.

15 C. L. Merritt, *Resources of american libraries* cit., p. 68; L. N. Ridenour, *Bibliography in an age of science* cit., p. 21.

si, tutte le biblioteche di ricerca ritenevano necessario avere a disposizione nei propri depositi tutti i libri necessari a sostenere la ricerca, e la necessità di erogare questo servizio costituiva un ulteriore motivo di incremento delle dimensioni delle raccolte creando nel contempo problemi di gestione dello spazio.

Ancora oggi, tra le biblioteche di ricerca italiane, si possono incontrare situazioni simili, con livelli molto alti di sovrapposizione di titoli tra le collezioni e la presenza di una quantità eccessiva di copie degli stessi libri. Infatti, se consideriamo la genesi dei sistemi bibliotecari degli atenei italiani, nati dalla fusione delle biblioteche di cattedra, confluite nelle biblioteche di istituto e successivamente in quelle dei dipartimenti, in qualche caso poi riorganizzate nelle biblioteche di area disciplinare, nelle quali venivano e vengono tuttora acquistati i libri necessari ai diversi gruppi di ricerca, possiamo immaginare come, a fronte di rari costosi interventi di scarto, si siano sommate nel tempo più copie degli stessi libri, andando a ridurre la capacità informativa di queste biblioteche e saturando nel contempo i sempre insufficienti spazi a loro disposizione.

Un esempio in tal senso ci viene offerto da una recente indagine, condotta dal Servizio bibliotecario d'ateneo dell'Università degli studi di Milano al fine di considerare una possibile procedura di scarto: la ricerca ha rilevato nel patrimonio di oltre 1 milione di titoli (1.298.311) di monografie, possedute dalle 17 biblioteche del sistema, una percentuale di circa il 14% (174.469) di sovrapposizioni, ossia di titoli presenti in più di una biblioteca, e a fronte delle 1.730.671 copie totali, la presenza di oltre 432.000 copie multiple a rappresentare circa il 25% del patrimonio complessivo di esemplari¹⁶. Considerando che i sistemi bibliotecari delle università italiane si sono, nel tempo, formati allo stesso modo, è verosimile pensare che le condizioni di duplicazione di titoli e stratificazione di copie, simili a quelle rilevate dall'ateneo milanese, possano essersi riprodotte in altri sistemi bibliotecari e, in generale, fra le biblioteche di ricerca italiane. Una possibile soluzione da proporre oggi potrebbe considerare l'avvio di procedure di scarto delle copie eccedenti, unite a un servizio di prestito intra-sistemico efficiente e a un potenziamento e conseguente maggiore fiducia nel prestito interbibliotecario, che assicurino la circolazione delle risorse e garantiscano in tempi rapidi l'approvvigionamento scientifico di cui i ricercatori hanno bisogno. Inoltre, la definizione e applicazione di una carta delle collezioni che permetta al sistema bibliotecario di progettare uno sviluppo coordinato delle raccolte tra le biblioteche di ricerca che lo costituiscono rappresenterebbe un altro strumento estremamente utile per razionalizzare le acquisizioni future.

Fremont Rider proponeva come soluzione al problema, sia delle acquisizioni sia alla seconda funzione di *storage* individuata poi da Ridenour, il ricorso alle *micro-card*, costituite da schede di dimensioni diverse – tra le quali anche quelle standard (7,5x12,5 cm) delle schede da collocare a catalogo – sulle quali venivano micro-stampati i libri, grazie alla riduzione fotografica, per poter poi essere collocate a scaffale dentro ad apposite buste; se le schede fossero state delle dimensioni standard e dotate di una sintetica descrizione in testa alla scheda, potevano essere inserite nel catalogo della biblioteca, integrando in questo caso i metadati descrittivi e il contenuto della risorsa su un unico supporto, in entrambi i casi risparmiando dello spazio prezioso (Figura 1)¹⁷. Le *micro-card* consentivano quindi un forte risparmio rispetto ai costi di acquisto e azzeravano, secondo Rider, i costi di *storage* (Figura 2)¹⁸, potendo poi leggere il testo per mezzo di un visore che lo ingrandiva sullo schermo.

16 Dati forniti dalla Direzione del Servizio bibliotecario d'ateneo dell'Università degli studi di Milano.

17 Immagine tratta da Martin Jamison, *The microcard: Fremont Rider's precomputer revolution*, «Libraries & culture», 23 (1988), n. 1, p. 1-17: p. 4, <<http://www.jstor.org/stable/25542007>>.

18 Immagine tratta da F. Rider, *The scholar and the future of the research library* cit., p. 118.



Figura 1 – Da sinistra: Readex microprint, *micro-card*, microfiche e microfilm

118 THE SCHOLAR AND THE FUTURE

	500-page book, in printed, bound form	500-page book, in micro-card form (assuming that 2 cards are required)
1. Original purchase cost	\$5.00	\$.30
2. Binding	—	—
3. Minor physical "proc- essing"	.15	—
4. Cataloging	1.02	—
5. Card filing	.05	.06
6. Storage	.44	—
Total cost	\$6.64	\$.36

Figura 2 – Tabella dei costi per un libro rilegato e in microprint

La soluzione al problema della sovrapposizione tra le collezioni e, di conseguenza, del sovraffollamento dei depositi e dello *storage* viene invece affrontata da Ridenour in una prospettiva più tecnologica. Alla domanda su come evitare che le biblioteche statunitensi si sovrappongano nelle loro acquisizioni, risponde nella maniera oggi più logica: «if libraries were connected by a communications network over which complicated research material could be promptly sent from any point to any other, there would be no need whatever for duplication in acquisition»¹⁹.

L'idea di collegare tutte le biblioteche a un *network* in modo che possano scambiarsi e ottenere i documenti che servono ai lettori nel momento in cui servono, senza necessariamente acquistarli preventivamente in tutte le biblioteche, razionalizzando in tal modo le acquisizioni, anticipa di decenni la logica del *just in time*. Una logica produttiva nata nell'ambito manifatturiero, che comincerà a essere considerata come evoluzione dei servizi di biblioteca negli anni Novanta, nel momento in cui la tecnologia ha reso accessibili online in cataloghi e creato le condizioni per poter trasmettere i documenti²⁰; una logica che perseguiamo e a cui continuiamo a tendere anche oggi, entrata nell'uso quotidiano nelle biblioteche di ricerca di tutto il mondo per la condivisione degli articoli scientifici, ma che stenta a realizzarsi per le monografie, rallentata dalle politiche commerciali – legittime o meno non è questa la sede per dibatterlo – degli editori e dalle non secondarie difficoltà, ancora non superate da molti, di leggere i libri su dispositivi digitali.

Razionalizzare lo *storage*, l'*indexing* e il reference, nella visione tecnologica di Ridenour, viene presentato in una prospettiva estremamente semplice, almeno per oggi, quasi disarmante. «In the computing machines of the present day, much use is made of magnetic materials for storing information»²¹ e la soluzione proposta da Ridenour per ridurre i problemi di spazio nei magazzini è proprio quella di digitalizzare i libri: con-

19 L. N. Ridenour, *Bibliography in an age of science* cit., p. 21.

20 Carla Basili; Corrado Pettenati, *La biblioteca virtuale*. Milano: Editrice bibliografica, 1994, p. 113, 116; Elizabeth A. Fuseler, *Providing access to journals: just in time or just in case?*, «College & research libraries news», 55 (1994), n. 3, p. 130-132, 148: p. 130, DOI: 10.5860/crln.55.3.130.

21 L. N. Ridenour, *Bibliography in an age of science* cit., p. 24.

siderando che ogni carattere a stampa è equivalente a 7 *binary digit* (*bit*) e che un libro medio composto da 40.000 parole di cinque lettere 'contiene' 1,4 milioni di *bit* di informazione, le collezioni della Library of Congress avrebbero contenuto circa 100.000 miliardi di *bit*, mentre le biblioteche normalmente immagazzinano libri a stampa raggiungendo a malapena 50 milioni di *bit* per *cubic foot* di spazio a scaffale. Con il grado di sviluppo della tecnologia di *magnetic storage* dell'epoca, sarebbe stato quindi possibile ridurre di dieci volte lo spazio occupato a scaffale dai libri a stampa, offrendo la possibilità di registrare in formato digitale un milione di volumi occupando uno spazio di soli *four cubic feet*²². Questa soluzione sarebbe stata anche complementare alla razionalizzazione delle acquisizioni, permettendo di trasmettere tra le biblioteche collegate al *network* i libri o gli articoli nel momento in cui ne avessero avuto bisogno.

La procedura di 'catalogazione' risulta molto complessa e costosa perché deve essere fatta da persone che analizzano il libro e che devono essere dotate di una considerevole formazione ed esperienza. Centralizzare in un'unica biblioteca le raccolte e le operazioni di catalogazione avrebbe ridotto le operazioni ripetitive che venivano effettuate dalle diverse biblioteche che acquistavano gli stessi libri, ma, pur riducendolo, il costo sarebbe rimasto comunque alto²³. A meno che il lavoro di catalogazione non venisse fatto da una macchina, scansionando il testo per mezzo di una *electronic pencil*²⁴ e facendo riconoscere gli elementi informativi dal sistema informatico, con l'unico intervento umano finalizzato alla configurazione del sistema che avrebbe dovuto effettuare la catalogazione. Questo è quanto viene proposto nel lavoro di Ridenour, sogno angoscioso di generazioni di bibliotecari che intravedevano nei primi sistemi di elaborazione, e anche nei secondi, strumenti che avrebbero potuto agevolarne e potenziarne il lavoro, oppure sostituirli, opzione questa che sappiamo non praticabile.

In realtà, l'interpretazione e la comprensione automatica di elementi bibliografici dal testo in modo da potersi sostituire alla mediazione umana si sono sempre rivelate molto difficili da attuare. Alcuni progetti di conversione retrospettiva hanno rivelato che tale interpretazione risultava troppo dipendente dal formato, dai caratteri e dalla struttura della scheda anche disponendo di dati strutturati in registrazioni catalografiche; la difficoltà aumenta, inevitabilmente, dovendo trattare un testo a stampa. Ad esempio, il progetto europeo Facit (Fast Automatic Conversion with Integrated Tools 1993-1995)²⁵ si proponeva di realizzare un processo per la conversione dei cataloghi cartacei tramite scansione e OCR (*optical character recognition*)

22 Un *cubic foot* corrisponde a circa 0,028 metri cubi, quindi 4 *cubic feet* di supporti magnetici corrisponderebbero a circa 0,11 metri cubi. Vedi *Ivi*, p. 23-24.

23 *Ivi*, p. 25.

24 L'*electronic pencil* è un interessante dispositivo sviluppato in modo sperimentale nella seconda metà degli anni Quaranta dalla RCA (Radio Corporation of America) per permettere il riconoscimento e la lettura di un testo. Costruito in due versioni, una sorta di *scanpen* che, fatta scorrere sulle righe di testo, sillabava le parole lettera per lettera, oppure emetteva dei suoni differenti per ogni lettera. Il riconoscimento del testo ha fatto ipotizzare un utilizzo nell'ambito bibliografico, come inteso da Ridenour, oppure per estrarre testo e creare delle matrici che consentissero di riprodurlo in più copie. *Ivi*, p. 55-56. Si veda anche V. K. Zworykin; L. E. Flory, *An electronic reading aid for the blind*, «Proceeding of the American Philosophical Society», 91 (1947), n. 2, p. 139-142: p. 142, <<http://www.jstor.org/stable/3143599>>.

25 Il progetto Facit venne elaborato e condotto da un consorzio di biblioteche italiane (Biblioteca nazionale centrale di Firenze, Biblioteca nazionale di Napoli) danesi (lo Statensbibliotekstjeneste e la Stat-sbiblioteket di Aarhus) e greche (l'Ethnike Bibhotheke tes Hellados di Atene). Gian Luca Corradi; Clau-

delle schede, al fine di individuare, interpretare e correggere automaticamente il testo scansionato per poi formattare gli elementi informativi e di accesso rilevati in una registrazione elettronica. Il metodo elaborato, che si scontrò con la difformità strutturale e catalografica delle schede, fu utilizzato per la conversione del catalogo Palatino della Biblioteca nazionale centrale di Firenze, incontrando molte difficoltà a essere utilizzato per la conversione di altri cataloghi.

Sul fornire agli studiosi i testi di cui hanno bisogno, servizio indicato come *reference*, Ridenour non ha molto da aggiungere, il servizio non presentava grosse criticità; tuttavia, le soluzioni indicate per le altre procedure avrebbero potuto configurare un servizio migliore. Considerando uno *storage* centralizzato e 'compresso' in forma di *bit*, il documento richiesto poteva essere inviato dalla biblioteca centrale a quella richiedente che lo avrebbe riprodotto e consegnato allo studioso, non fornendo però «the priceless original of the document he seeks, but a cheap, durable, and faithful copy»²⁶, senza dunque attivare il prestito, secondo un sistema che oggi chiameremmo di *print on demand*.

Domare la crescita delle risorse informative: gli *information retrieval system*

All'interno di questo panorama pionieristico e innovativo, oltre alla necessità di controllare la crescita esponenziale delle raccolte nelle biblioteche di ricerca, si pone il problema di fronteggiare ciò che verrà successivamente definito l'*info glut*, ossia l'incapacità di utilizzare le informazioni causata dall'*information overload*: il sovraccarico informativo generato dalla rappresentazione della grande quantità di risorse che le biblioteche erano chiamate ad acquisire e gestire. Su questo tema sono frequenti le grida di allarme lanciate dagli studiosi, il filosofo José Ortega Y Gasset nel suo articolo *Man must tame the book*, pubblicato nel 1936, esprime le sue preoccupazioni per l'accumulo di materiale a stampa in costante aumento e la necessità di controllarlo, sostenendo che «la cultura che aveva liberato l'uomo dalla giungla lo avrebbe presto gettato in una giungla di libri»²⁷, giungla non solo fisica ma anche informativa. Le stesse considerazioni vengono riprese da Mortimer Taube, a capo dello Science and Technology Project della Library of Congress, che nel suo lavoro del 1948²⁸ pone la questione di sperimentare nuovi strumenti per controllare, gestire e, soprattutto, essere in grado di utilizzare la massa di conoscenza registrata che si andava accumulando.

In quegli anni, vengono infatti sviluppati numerosi sistemi per il recupero dell'informazione (*information retrieval system*, IR) per reperire i documenti posseduti da biblioteche o centri di documentazione e gli elementi per poterli recuperare dai depositi, realizzati codificando le informazioni in fori, tacche e posizioni in vari tipi di scheda, gestiti e interrogati con modalità manuali, o supportate da sistemi alimentati dall'elettricità. Si ricordano in questo ambito i primi sistemi di indicizzazione e ricerca post-coordinata manuale costituiti, a partire dal 1932, dalle schede perforate nei bordi, le *edge notched punched card*²⁹, nelle quali i fori rappresentavano

dia Miconi; Giuseppe Vitiello, *Dalla sperimentazione del progetto Facit alla riconversione del Catalogo Palatino della Biblioteca nazionale centrale di Firenze*, «Bollettino AIB», 36 (1996), n. 4, p. 437-444.

26 L. N. Ridenour, *Bibliography in an age of science* cit., p. 27.

27 José Ortega y Gasset, *Man must tame the book*, «Wilson bulletin for librarians», 10 (1936), n. 5, p. 305-307.

28 Mortimer Taube, *New tools for the control and use of reserarch materials*, «Proceedings of the American Philosophical Society», 93 (1949), n. 3, p. 248-252: p. 248, <<https://www.jstor.org/stable/3143474>>.

i descrittori, mentre nel corpo della scheda veniva descritto il documento: l'attivazione dei descrittori in grado di rappresentare gli argomenti trattati nel documento avveniva praticando una tacca (*notch*) che apriva verso il bordo il foro corrispondente al descrittore; in questo modo la selezione operata inserendo un ferro sottile in determinati fori di un insieme di schede, non necessariamente ordinate, consentiva di 'sollevare' e quindi eliminare dalla selezione le schede con il foro 'chiuso', rimaste 'appese' al ferro di selezione, e lasciare nel cassetto le schede con i fori dei descrittori selezionati 'aperti', ottenendo le informazioni sui documenti che trattavano gli argomenti rappresentati da tali descrittori.

Tuttavia, tra i primi sistemi di recupero dell'informazione elaborati a partire dagli anni Quaranta, erano relativamente pochi quelli in grado di integrare in un unico supporto gli elementi descrittivi del documento e il documento stesso, o una sua sintesi, per tentare di limitare la crescita e l'occupazione di spazio delle raccolte e rendere più immediata la fornitura dei documenti. Tra questi, oltre alle *micro-card* ideate da Fremont Rider, vengono progettati in via sperimentale alcuni dispositivi, talvolta entrati in produzione, che possono in linea generale rientrare nelle *aperture card*, categoria trasversale ai sistemi di IR, manuali ed elettronici che utilizzava e integrava la microfilmatura dei documenti nella scheda, e nei sistemi di IR che invece combinavano segnali ottici con quelli elettrici elaborando i dati registrati su una pellicola fotografica.

Le *aperture card* furono inventate nel 1943 da John F. Langan, responsabile della Pictorial Records Division dell'Office of Strategic Services (OSS)³⁰ nel corso della Seconda guerra mondiale, quando il suo ufficio si trovò a dover gestire milioni di fotografie, non standard e non indicizzate, in seguito a un appello che chiedeva ai cittadini di inviare qualsiasi immagine o informazione del nemico e dell'Europa occupata che potesse essere utile³¹. Si trattava di un dispositivo sviluppato – in particolare nell'ambito della ricerca e reperimento delle informazioni – per registrare, ordinare, reperire e visualizzare dei documenti, come ad esempio fotografie, grafici e progetti, ma anche documenti di testo microfilmati. Un foro rettangolare, una sorta di finestrella, ritagliato in una scheda, permetteva di custodire tra due foglietti, oppure in una tasca, di pellicola trasparente, uno o più *frame* di microfilm nei quali era stato riprodotto il documento, integrando, quindi, il documento con la sua descrizione e con gli elementi di accesso costituiti dai descrittori. Nelle prime applicazioni delle *aperture card* il microfilm era stato inserito nelle *edge notched card* realizzate dalla ditta statunitense McBee Co.; tuttavia, l'impiego più diffuso di queste schede vide il microfilm inserito in una scheda perforata che poteva essere elaborata, ordinata e selezionata con uno dei 'moderni' calcolatori sviluppati a seguito dell'invenzione della *electric tabula-*

29 Fabio Venuda, *Alle origini dell' informatizzazione: Herman Hollerith e i sistemi per l' analisi e il reperimento dei dati*, «Il Bibliotecario», 2009, n. 2/3, p. 143-175: p. 155, <<http://hdl.handle.net/2434/73354>>; Frederick G. Kilgour, *Origins of coordinate searching*, «Journal of the American Society for Information Science», 48 (1997), n. 4, p. 340-348: p. 340, DOI: 10.1002/(SICI)1097-4571(199704)48:4<340::AID-AS17>3.0.CO;2-W.

30 L'Office of Strategic Services (OSS) era il servizio di *intelligence* delle forze armate statunitensi nel periodo della Seconda guerra mondiale, fondato l'11 luglio del 1942. Rhodri Jeffreys-Jones, *Hyping the sideshow: Wild Bill Donovan and the OSS*. In: *Id.*, *Cloak and dollar: a history of American secret intelligence*. London: Yale University Press, 2002, p. 131-153: p. 143.

31 Susan A. Cady, *Microfilm technology and information systems*. In: *Proceedings of the 1998 Conference on the history and heritage of science information systems*, edited by Mary Ellen Bowden [et al.]. Medford (NJ): Information Today, 1999, p. 177-186: p. 182, <http://webdoc.sub.gwdg.de/ebook/s/2001/chf/www.chemheritage.org/historicals-services/asis_documents/asisbook.pdf#page=189>.

ting machine di Herman Hollerith, una «electronic data processing card that could be key-punched and machine sorted»³². L'utilizzo di tale strumento di registrazione e reperimento di informazioni e documenti era talmente diffuso che i maggiori centri di intelligence e aziende statunitensi, come la US Navy, l'Air Force, la Central Intelligence Agency, la Western Electric, la Remington Rand e l'IBM, collaborarono per cercare di normalizzare dimensioni, posizione e caratteristiche delle *aperture card*, raggiungendo un accordo sulle specifiche che venne rilasciato nel 1960³³ (Figura 3)³⁴.



Figura 3 – Apertura card

Nell'ambito dei sistemi di recupero dell'informazione in grado di integrare i metadati descrittivi di un documento con il documento stesso, rivestono particolare importanza, almeno dal punto di vista progettuale, alcuni sistemi in grado di avvalersi delle tecnologie disponibili tra il 1930 e il 1950, per sfruttare le proprietà di componenti ottici armonizzate con meccanismi e segnali governati dall'energia elettrica. In realtà, si tratta di una serie di dispositivi basati sugli stessi principi di funzionamento, fotocellule elettriche, microfilm e fotografia ad alta velocità, inventato, reinventato e reingegnerizzato, del quale vennero realizzati alcuni prototipi fino a quando, ripensato completamente da un altro inventore, non entrò in produzione e venne utilizzato con successo a livello commerciale.

Emanuel Goldberg e la *Statistical machine*

Il 29 dicembre 1931 Emanuel Goldberg, in quel periodo a capo della più grande industria di macchine e attrezzature fotografiche del mondo, la Zeiss Ikon, ottiene negli Stati Uniti il brevetto per quello che viene considerato il primo sistema di *document storage and retrieval*: la *Statistical machine*³⁵. Il sistema a cui Goldberg iniziò a lavorare

³² *Ibidem*.

³³ *Ivi*, p. 183.

³⁴ Immagine tratta da Perfect Image. Precision Scan Processing, *Aperture cards format*, 1999, <<http://www.pimage.com/appcards.htm>>.

³⁵ Michael K. Buckland, *Histories, heritages and the past: the case of Emanuel Goldberg*. In: *The history and heritage of scientific and technological information systems: proceedings of the 2002 conference*, edited by W. Boyd Rayward, Mary Ellen Bowden. Medford: Information Today, 2004, p. 39-45: p. 40, <<https://bit.ly/3QcJHy8>>. Emanuel Goldberg, *Statistical machine*, U.S. Patent 1,838,389/1931, and issued December 29, <<https://patents.google.com/patent/US1838389A/en>>.

dal 1927 e per il quale, nello stesso anno, chiese e ottenne il brevetto tedesco³⁶, prevedeva di utilizzare una pellicola fotografica nella quale veniva microfilmato un documento, una citazione con abstract o una scheda bibliografica, occupando la metà di un fotogramma; sull'altra metà veniva, invece, fotografato un codice, costituito da una serie di punti opachi, che rappresentava l'argomento di cui trattava il testo.

Per effettuare una ricerca era necessario utilizzare una scheda di cartoncino nella quale l'argomento era codificato in una serie di fori trasparenti alla luce. Una sorgente luminosa illuminava costantemente la scheda di ricerca, mentre la pellicola veniva fatta scorrere ad alta velocità: nel momento in cui i raggi di luce che uscivano dai fori del codice registrato nella scheda di ricerca coincidevano con i punti opachi del codice registrato nel fotogramma, una fotocellula registrava la completa assenza di luce, lo scorrimento della pellicola cessava e il documento microfilmato veniva proiettato su uno schermo (Figura 4)³⁷. Di questo sistema vennero realizzati due prototipi dalla Zeiss Ikon, ma non venne mai messo in produzione.

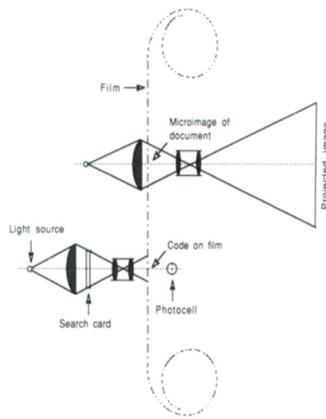


Figura 4 – Schema di funzionamento della *Statistical machine* di Goldberg³⁸

Vannevar Bush e il *Microfilm rapid selector*

Vannevar Bush pubblicò nel 1945 il suo visionario e molto citato articolo *As we may think*³⁹ nel quale, con molta enfasi, descrive un «imaginary personalized, desktop document retrieval system»⁴⁰ che decide di chiamare «to coin one at random»⁴¹ Memex, ricordato per aver ispirato molti studiosi e precorso le modalità di naviga-

³⁶ Michael K. Buckland, *Emanuel Goldberg, electronic document retrieval, and Vannevar Bush's Memex*, «Journal of the American Society for Information Science», 43 (1992), n. 4, p. 284-294: p. 289, 291, DOI: 10.1002/(SICI)1097-4571(199205)43:4%3C284::AID-ASI3%3E3.o.CO;2-o.

³⁷ *Ivi*, p. 287.

³⁸ *Ibidem*.

³⁹ Vannevar Bush, *As we may think*, «The Atlantic monthly», 176 (1945), n. 1, p. 101-108, <<https://bit.ly/3ztjVPW>>.

⁴⁰ M. K. Buckland, *Histories, heritages and the past* cit., p. 40.

⁴¹ V. Bush, *As we may think* cit., p. 106.

zione del web, con la possibilità per il lettore di creare associazioni e collegamenti ipertestuali reciproci tra i documenti. Tuttavia, il saggio che presenta il Memex pubblicato nel 1945, era già in lavorazione dal 1939 e, secondo Michael Buckland, l'immaginazione di Bush fu ispirata dal *Microfilm rapid selector* «a document storage-and-retrieval device» a cui stava lavorando dal 1938, nel periodo della sua collaborazione con il MIT di Boston⁴². Il principio di ricerca e selezione dei documenti microfilmati su pellicola era molto simile a quella della *Statistical machine* di Goldberg: una fonte di luce, una scheda di ricerca con i descrittori codificati in una serie di fori, una pellicola che scorre ad alta velocità, con riprodotti in ciascun fotogramma i documenti e i codici che rappresentavano i descrittori. Il dispositivo di Bush, secondo Buckland, differiva dal progetto di Goldberg per alcune soluzioni tecniche: i codici codificati nei fotogrammi della pellicola non erano opachi, ma trasparenti, e il *match* con i raggi di luce provenienti dalla scheda di ricerca veniva rilevato dalla presenza di luce nelle posizioni determinate dal codice di ricerca utilizzato e rilevate da «a bank of photoelectric cells, one for each possible position [...] and just the right combination of electric outputs had to be detected»⁴³. Inoltre, invece di proiettare l'immagine del documento su uno schermo, grazie all'utilizzo delle nuove lampade stroboscopiche il dispositivo di Bush poteva riprodurre il documento o l'abstract selezionato senza bloccare la pellicola e registrarlo su una nuova pellicola per mezzo di una fotocamera a scatto rapido. In questo modo il risultato della ricerca, potremmo definirlo la bibliografia, impresso nella nuova pellicola, poteva venire sviluppato e visualizzato con un lettore di microfilm, oppure stampato su carta⁴⁴. Come il Memex non venne mai realizzato, allo stesso modo il Microfilm selector di Bush non divenne mai operativo⁴⁵, rimanendo allo stadio di prototipo e Bush non poté brevettare il suo *selector* a causa del brevetto già depositato da Goldberg. Si è molto discusso in letteratura se Bush avesse inventato il *Microfilm rapid selector* oppure avesse avuto modo di conoscere il lavoro di Goldberg. Nel lavoro di Buckland vengono riportate testimonianze sul fatto che Bush non avesse mai fatto riferimento al dispositivo di Goldberg, anche se, secondo Jana Varlejs, Bush ne era al corrente⁴⁶. Inoltre, quando Ralph Shaw incontrò Goldberg nel 1949 e venne a conoscenza del suo brevetto per la *Statistical machine*, gli confessò che non era a conoscenza della sua invenzione e da quel momento cominciò a citarlo nei suoi lavori. Tuttavia, secondo Buckland, il dispositivo di Goldberg era senz'altro conosciuto almeno da due importanti centri di ricerca statunitensi, ossia dall'IBM e dagli Eastman Kodak Research Laboratories, quest'ultimo tra i finanziatori dei progetti di Bush, lasciando quindi la domanda senza una risposta definitiva.

A complicare la situazione si inserisce la figura di Watson Davis che, secondo Alistair Black, descrisse nel 1935, in un suo lavoro per lo Smithsonian Institution Science Service, un *microfilm selector device* «that bore a strong resemblance to the prototype Rapid Selector apparatus developed by the American engineer Vannevar

42 M. K. Buckland, *Histories, heritages and the past* cit., p. 40.

43 M. K. Buckland, *Emanuel Goldberg, electronic document retrieval, and Vannevar Bush's Memex* cit., p. 291.

44 C. D. Gull, *Instrumentation*, «Library trends», 2 (1953), n. 1, p. 103-126: p. 111.

45 M. K. Buckland, *Emanuel Goldberg, electronic document retrieval, and Vannevar Bush's Memex* cit., p. 287.

46 Jana Varlejs, *Ralph Shaw and the rapid selector*. In: *Proceedings of the 1998 Conference on the history and heritage of science information systems* cit., p.148-155: p. 151.

Bush»⁴⁷, apparato di cui Bush costruì un prototipo nel 1940 mentre lavorava al MIT. Per Black non è chiaro se fu Bush a trarre l'idea del Rapid selector da Davis, incontrato nel 1932, o se fu invece Davis a venire influenzato dalle idee di Bush, ma sembra probabile che i due scienziati possano essersi influenzati a vicenda, ancora una volta senza citare l'uno il lavoro dell'altro⁴⁸.

Subito dopo la fine della Seconda guerra mondiale, l'azienda statunitense Engineering Research Associates (ERA) ricevette l'incarico dal Dipartimento del commercio di costruire, con il supporto di Vannevar Bush, un nuovo *microfilm selector* e Ralph R. Shaw, bibliotecario, studioso e direttore della National Agriculture Library, collaborò con entusiasmo alla messa a punto del nuovo *selector* fornendo il materiale da microfilmare e codificare, testandolo anche come supporto alla produzione della *Bibliography of Agriculture*. Il prototipo che venne sviluppato con il contributo di Shaw venne consegnato al Dipartimento dell'agricoltura nel 1949, ma una serie di insuccessi, dovuti principalmente a errori concettuali di indicizzazione e codifica sottovolutati da Shaw, piuttosto che a problemi meccanici o di progettazione, portarono nel 1952 Shaw ad abbandonare il progetto del *rapid selector*⁴⁹ (Figure 5-6)⁵⁰.

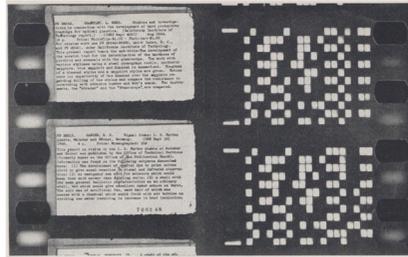


Figura 5 – Rapid selector, ingrandimento del fotogramma con testo e codice di ricerca

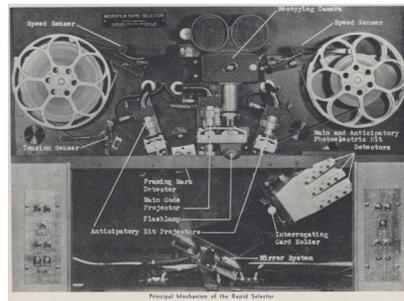


Figura 6 – Rapid selector, struttura interna e meccanismi di funzionamento

⁴⁷ Alistair Black, *Mechanization in libraries and information retrieval: punched cards and microfilm before the widespread adoption of computer technology in libraries*, «Library history», 23 (2007), n. 4, p. 291-299: p. 296, DOI: 10.1179/174581607X254785.

⁴⁸ *Ibidem*.

⁴⁹ J. Varlejs, *Ralph Shaw and the Rapid selector* cit., p. 151.

⁵⁰ Immagini del *rapid selector* tratte da John C. Green, *The rapid selector: an automatic library*, «The military engineer», 41 (1949), n. 283, p. 350-352: p. 351, <<http://www.jstor.org/stable/44603817>>.

Il sistema Filmorex di Jacques Samain

Nel 1940 il dott. Jacques Samain progettò un sistema per il recupero delle informazioni basato sulle schede perforate ideate da Herman Hollerith nel 1889, cercando di superarne le limitazioni date dall'utilizzo di campi fissi per la codifica delle informazioni. Il Samain's electronic selector si proponeva di codificare i descrittori di un determinato documento suddividendo la scheda perforata in 24 sezioni da 40 posizioni per riga modo da poter codificare un vocabolario artificiale di 60 milioni di parole differenti costituite da 6 lettere. Nel progetto, la lettura e la selezione delle schede avrebbero dovuto essere eseguite tramite cellule fotoelettriche, invece di utilizzare la 'spazzola' di lettura, che anche nei modelli successivi della Electric tabulating machine di Hollerith avrebbe dovuto attivare dei contatti elettrici passando attraverso i fori della scheda. In realtà, secondo Ralph Shaw, le fotocellule non vennero usate nei prototipi, ma i dati codificati nei fori della scheda venivano letti dalla usuale spazzola di lettura «just as in normal electrical Hollerith searching», anche se, a differenza di questa, gli impulsi elettrici venivano registrati in una memoria elettronica che provvedeva al confronto con la combinazione di impulsi impostati per la ricerca, raggruppando le schede pertinenti in un raccoglitore, specificando ancora «just as is the card in normal Hollerith searching». Shaw conclude dichiarando che si era trattato di un interessante esperimento che non giunse mai a essere realizzato e che Samain lo abbandonò per dedicarsi alla modifica del *rapid selector*⁵¹.

Mentre lavorava per il CNRS (Centre national de la recherche scientifique) a Parigi, Semain avvia nel 1954 lo sviluppo di un dispositivo di ricerca bibliografica in grado di fornire ai ricercatori le informazioni sulle pubblicazioni scientifiche di cui avevano bisogno. Il CNRS pubblicava già il *Bulletin signalétique*, che ogni anno portava all'attenzione degli studiosi circa 130.000 abstract di lavori condotti nei più diversi ambiti disciplinari e nel quale le pubblicazioni erano riassunte e ordinate secondo un sistema di intestazioni principali. Il dispositivo di ricerca avrebbe dovuto sostituire la realizzazione e la pubblicazione degli indici dei lavori pubblicati nel *Bulletin* a supporto delle attività di ricerca. Semain riprende e perfeziona il meccanismo di selezione del *Rapid selector*, basato sulla codifica di descrittori in sequenze di punti opachi o traslucidi su mezzo fotogramma di un microfilm e nei fori prodotti nella scheda di ricerca, tralasciando la funzione di proiezione e di riproduzione tramite macchina fotografica a scatto rapido.

Il nuovo sistema, infatti, era orientato unicamente a fornire un servizio di *document information retrieval*, ed era costituito, nella descrizione fatta dallo stesso Samain nel 1957, da una scheda di acetato pesante delle dimensioni di 60x35 mm – ricavata da una pellicola fotografica – divisa in due zone, la zona A sulla quale veniva microfilmato il testo del documento, o un suo abstract, e la zona B nella quale erano riprodotti i codici che rappresentavano i descrittori di concetti e caratteristiche della risorsa microfilmata (Figura 7)⁵². I descrittori costituiti da parole o numeri venivano codificati e registrati sotto forma di tratti opachi o trasparenti che potevano essere scansionati dal *selector*.

51 *The state of the library art*, edited by R. Ralph Shaw. New Brunswick (NJ): Rutgers State University, Graduate School of Library Service, 1961, p. 162-163. Cfr. anche Jacques Samain, *La Mécano-mémoire*, «Hommes et techniques» (Dec. 1945), p. 59-61.

52 Jacques Samain, *Documentation by the Filmorex technique*. In: *Information systems in documentation: based on the symposium on systems for information retrieval, held at Western Reserve University, Cleveland Ohio, in April 1957*, edited by Jesse Hauk Shera [et al.]. New York: Interscience Publishers, 1957, p. 478-487: p. 481, <<http://hdl.handle.net/2027/mdp.3901502353996>>.

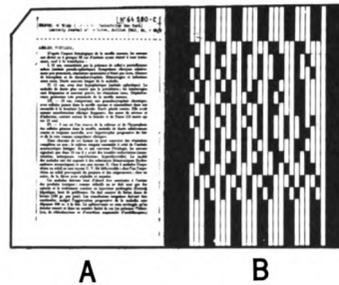


Figura 7 – Scheda del sistema Filmorex

La parte codificata della scheda era costituita da 25 colonne parallele, ciascuna divisa in 6 gruppi di 6 posizioni e ciascuna posizione poteva accogliere 2 punti opachi e 3 trasparenti. La combinazione di questi 5 punti (2 opachi e 3 trasparenti) consentiva di rappresentare una cifra del codice.

La ricerca veniva effettuata inserendo un pacco di schede non ordinato nel *selector*, (Figura 8) nel quale le schede venivano scansionate in sequenza, alla velocità di circa 700 schede al minuto; la scansione permetteva di riconoscere la combinazione di punti opachi e trasparenti che rappresentavano il codice richiesto dalla ricerca e la *microfilm card* selezionata veniva dirottata in un apposito raccoglitore, mentre quelle che non rispondevano ai criteri di ricerca finivano in un altro contenitore. Infine, le schede selezionate potevano venire lette con un normale lettore di microfilm, oppure riprodotte su un'altra pellicola, o stampate su carta fotografica⁵³.

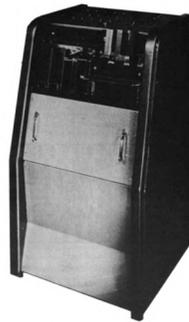


Figura 8 – Selettore del sistema Filmorex

Pur presentando significative similitudini con i dispositivi ideati da Goldberg e da Bush e Shaw, il Filmorex offriva delle caratteristiche di solidità e velocità, risultando molto affidabile. Tra i vantaggi vengono ricordate la robustezza delle schede di acetato che potevano passare nel selettore migliaia di volta senza danneggiarsi; le dimensioni contenute delle schede a fronte della capacità di registrare una quantità molto alta di informazioni; il processo di selezione semplice e quasi del tutto automatico; l'alta velocità di selezione; il costo dell'attrezzatura relativamente basso⁵⁴. Samain aveva di certo risolto i problemi di codifica e di indicizzazione, che avevano

⁵³ *Ivi*, p. 484.

⁵⁴ *The state of the library art cit.*, p. 186.

ostacolato il funzionamento del prototipo di *rapid selector* su cui aveva lavorato Ralph Shaw per la realizzazione della Bibliography of agriculture: il Filmorex entrò, quindi, in produzione e venne commercializzato. Nel 1958, il Dr. Jacques Samain partecipò a un convegno come membro del *discussion panel* nel quale la sua società di appartenenza è indicata nel Departement Filmorex, Société Ceramique et Mecanique, Paris, France⁵⁵. Un'altra conferma sull'utilizzo e sulla, non ancora ben definita, diffusione del Filmorex è fornita da una testimonianza portata al Congresso sulla documentazione scientifico-tecnica, tenuto a Roma nel 1964: in quella sede il rappresentante del Centro de informacion y documentacion del Patronato *Juan de la Cierva* di Madrid descrive l'utilizzo del Filmorex fatto dal suo centro di documentazione per la pubblicazione mensile degli *Indices de revistas cientificas y tecnicas* che registravano la traduzione in lingua spagnola degli articoli apparsi in oltre 1.200 riviste scientifiche nell'ambito della chimica, dell'ingegneria elettrica, elettronica e delle telecomunicazioni e dell'ingegneria meccanica⁵⁶.

Conclusioni

Dei nomi di Emanuel Goldberg, primo inventore della *Statistical machine*, di Watson Davis che la descrisse in un rapporto dello Smithsonian science Sservice, di Ralph Shaw che diede un forte contributo alla realizzazione e al funzionamento del *microfilm rapid selector*, ideato da Bush, e di Jacques Samain che inventò e riuscì a rendere commercializzabile un dispositivo, il *Filmorex*, ispirato agli stessi principi della *Statistical machine* e del *Microfilm rapid selector*, perfezionati, reingegnerizzati e resi perfettamente funzionanti, in letteratura si trovano poche tracce, relativamente poche sono le citazioni dei loro lavori e dei dispositivi ideati, possiamo dire che siano praticamente sconosciuti, o dimenticati. Al contrario, Vannevar Bush, grazie al suo articolo visionario che raccontava di dispositivi mai realizzati, è l'unico che è stato apprezzato e celebrato, come una persona e uno studioso geniale, che ha percorso e aperto la strada a immaginazioni e realizzazioni future; anche se, a tutt'oggi, non sono ancora del tutto state chiarite alcune ombre sulle suggestioni che avrebbe potuto derivare dalla conoscenza del lavoro di chi lo aveva preceduto e che fu, da lui, e non solo da lui, completamente ignorato.

L'assenza di questi studiosi dal panorama delle pubblicazioni scientifiche, ricordati o citati principalmente per sottolinearne la mancanza o rammentarne il ruolo, denota, purtroppo, una scarsa attenzione a una tematica importante che, nell'epoca attuale dominata dal digitale, acquista una rinnovata centralità; l'obiettivo di creare una rete di biblioteche per lo scambio del materiale, l'intento di collaborazione tra biblioteche al fine di servire l'utente in modo più efficiente riducendo i costi, la possibilità di archiviazione e recupero dei documenti sono tutte operazioni prioritarie anche oggi.

Ripercorrere le strade intraprese per arrivare a processi di codifica e indicizzazione dei contenuti, analizzare i metodi di immagazzinamento e gestione dei dati bibliografici, al di là della loro effettiva messa in pratica, può fornirci elementi di riflessione utili in fase di dematerializzazione e digitalizzazione delle risorse informative.

55 International Conference on Scientific Information, *Proceedings of the international conference on scientific information: Washington, DC. Nov. 16-21*. Washington, DC: National Academy of Sciences, 1958, <<https://nap.nationalacademies.org/read/10866/chapter/45>>.

56 José Ramon Perez Alvarez-Ossorio, *Experienced obtained with the Filmorex equipment in an industrial information service*. In *Congresso internazionale sulla documentazione e informazione scientifico-tecnica: Comitato nazionale per la produttività, Roma Palazzo della civiltà del lavoro, 2-11 febbraio 1964*. Roma: s.n., 1964, vol. 1.

ABSTRACT AIB studi, vol. 62 n. 2 (maggio/agosto 2022), p. 287-302. DOI 10.2426/aibstudi-13750
ISSN: 2280-9112, E-ISSN: 2239-6152 - Copyright © 2022 Fabio Venuda

FABIO VENUDA, Università degli studi di Milano, Dipartimento di Studi storici, Milano, e-mail fabio.venuda@unimi.it.

La sfida della crescita esponenziale delle raccolte nelle biblioteche di ricerca: i primi sistemi di *optical electric information retrieval*

Dalla fine dell'Ottocento nelle biblioteche e nei centri di ricerca statunitensi si registra una crescita delle raccolte librerie che, nel 1944, viene calcolata come esponenziale. Crescita che subisce un forte incremento nel secondo dopoguerra, rilevando, nel contempo, un elevato tasso di sovrapposizione tra le raccolte, che riduce il posseduto complessivo a circa due terzi di quanto pubblicato al mondo. Utilizzando la tecnologia dell'epoca, basata essenzialmente sulle procedure di microfilmatura dei documenti, vengono da più parti elaborate proposte per la deduplicazione delle raccolte e la riduzione dello spazio occupato a scaffale, mentre vengono progettati e, in alcuni casi, realizzati dispositivi elettrici, o elettronici, che uniscono l'elemento ottico, i microfilm, con tecniche di *information retrieval* per ottenere il reperimento delle informazioni, integrato con il recupero dei documenti. In realtà si tratta di una serie di dispositivi basati sugli stessi principi di funzionamento, fotocellule elettriche, microfilm e fotografia ad alta velocità, inventati, reinventati e reingegnerizzati, dei quali vennero realizzati solo alcuni prototipi fino a quando, ripensati completamente da un altro inventore, uno di essi non entrò in produzione e venne utilizzato con successo a livello commerciale. Alcune delle figure che gravitarono attorno a questi dispositivi sono rimaste nell'ombra oppure sono state dimenticate: questo lavoro, oltre a documentare un periodo pionieristico molto significativo per la biblioteconomia e per gli sviluppi dei futuri sistemi di recupero delle informazioni, intende rappresentare il ruolo e l'importanza che queste persone hanno avuto.

The challenge of the exponential growth of collections in research libraries: the first optical electric information retrieval systems

Since the end of the nineteenth century in U.S. libraries and research centers there has been a growth in book collections that, in 1944, is calculated as exponential. Growth that undergoes a sharp increase after World War II, noting, at the same time, a high rate of overlap between collections, which reduces the total holdings to about two-thirds of the world's published holdings. Using the technology of the time, based essentially on procedures for microfilming documents, proposals for deduplicating collections and reducing the amount of space occupied on the shelf are being developed from a lot of quarters, while electrical, or electronic, devices are being designed and, at times, realized combining the optical element, microfilms, with information retrieval techniques to achieve information discovery integrated with document retrieval. It was actually a series of devices based on the same operating principles, such as electrical photocells, microfilm and high-speed photography, invented, reinvented and reengineered, from which some prototypes were created, until, completely rethought by another inventor, one went into production and was successfully and commercially used. Some of the figures who gravitated around these devices have either remained in the shadows or have been forgotten: this paper, beyond documenting a pioneering period that was very significant for librarianship and for the developments of future information retrieval systems, aims at representing the importance and the role that these people played.