

Maria Teresa Biagetti

# La Scienza dell'informazione

Identità ed esperienze

**FrancoAngeli**

## Informazioni per il lettore

Questo file PDF è una versione gratuita di sole 20 pagine ed è leggibile con **Adobe Acrobat Reader**



La versione completa dell'e-book (a pagamento) è leggibile **con Adobe Digital Editions**.

Per tutte le informazioni sulle condizioni dei nostri e-book (con quali dispositivi leggerli e quali funzioni sono consentite) consulta [cliccando qui](#) le nostre F.A.Q.

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: [www.francoangeli.it](http://www.francoangeli.it) e iscriversi nella home page al servizio “Informatemi” per ricevere via e.mail le segnalazioni delle novità o scrivere, inviando il loro indirizzo, a “FrancoAngeli, viale Monza 106, 20127 Milano”.

Maria Teresa Biagetti

# La Scienza dell'informazione

Identità ed esperienze

FrancoAngeli

Copyright © 2022 by FrancoAngeli s.r.l., Milano, Italy.

*L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito [www.francoangeli.it](http://www.francoangeli.it).*

# Indice

**Prologo** pag. 7

## **Parte prima Il panorama disciplinare**

- 1. Genesi e identità scientifica della Scienza dell'informazione** » 15
  - 1. Le aree di ricerca del settore disciplinare » 36
  - 2. La natura dell'informazione nella Scienza dell'informazione » 41
    - 2.1 La teoria matematica della comunicazione » 51
- 2. Le discipline alla base della Scienza dell'informazione** » 57
  - 1. La Bibliografia: le mappe del sapere e gli indici » 58
  - 2. La Documentazione: i documenti e i fatti » 62
    - 2.1 Il movimento documentalista negli Stati Uniti » 70
  - 3. La Library Science: la fruizione del sapere » 75
    - 3.1 La Special Librarianship » 80

## **Parte seconda Ricerca delle informazioni e organizzazione della conoscenza**

- 1. L'Information Retrieval** » 85
  - 1. Il modello system-oriented » 107
  - 2. Il relevance ranking » 114
  - 3. Il relevance feedback » 116
- 2. I concetti di rilevanza e di pertinenza** » 121
  - 1. Il giudizio di rilevanza » 122
  - 2. La rilevanza cognitiva » 125

<b>3. Il modello bibliometrico</b>	» 131
1. La citazione bibliografica	» 133
2. Eugene Garfield e i citation indexes	» 135
3. I database citazionali: Web of Science, Scopus e Google Scholar	» 140
4. Bibliographic couplig e co-citation	» 148
<b>4. La Library and Information Science</b>	» 151
1. Orientamenti e prospettive della Library and Information Science	» 160
1.1 Il punto di vista cognitivo	» 160
1.2 Il modello socio-cognitivo: la Domain analysis	» 163
1.3 Il punto di vista ermeneutico	» 165
1.4 Il nuovo modello storico-bibliografico	» 166
2. Gli studi sul comportamento dell'utenza nella ricerca	» 172
2.1 Modelli di Information seeking behavior	» 178
<b>5. L'organizzazione della conoscenza registrata – Knowledge Organization</b>	» 183
1. Information Retrieval e OPAC	» 191
1.1 OPAC e relevance ranking	» 193
2. Il modello user-oriented	» 196
<b>6. L'organizzazione della conoscenza in ambiente digitale</b>	» 201
1. Le basi di dati bibliografiche	» 205
2. L'ampliamento delle funzioni per la ricerca	» 207
3. Dai documenti ai dati	» 210
4. Le ontologie come strumenti per l'organizzazione della conoscenza	» 216
<b>Riferimenti bibliografici</b>	» 223
<b>Principali Dipartimenti universitari di Library and Information Science</b>	» 242
<b>Indice analitico</b>	» 243

## Prologo

Il sociologo della conoscenza Richard Whitley, nel corso dei suoi studi dedicati all'analisi degli sviluppi delle discipline nei diversi contesti sociali, ha evidenziato come le diverse aree disciplinari siano caratterizzate dalla presenza più o meno marcata dei rapporti di mutua dipendenza che intercorrono tra gli scienziati. Tutti coloro che partecipano ad un campo disciplinare sono obbligati al mantenimento nel tempo della propria reputazione scientifica, alla conservazione del più ampio consenso e al controllo della credibilità, e tutti quindi sono spinti ad una forte competizione. Le discipline sociali e le Scienze umane, come egli ha chiarito con grande efficacia, mostrano un livello particolarmente basso di mutua dipendenza tra gli scienziati e allo stesso tempo un alto livello di incertezza degli obiettivi scientifici. Whitley definisce infatti queste discipline *fragmented adhocracies*, facendo riferimento al fatto che in queste aree gli studiosi tradizionalmente preferiscono lavorare in solitudine, non amano condividere i risultati delle loro ricerche fino a che non siano concluse, ed evitano in genere le ricerche coordinate attorno ad un obiettivo comune, al punto che si possono determinare molteplici interpretazioni delle identità disciplinari, con differenze in alcuni casi anche sostanziali, fino a mostrare a volte marcate idiosincrasie, che possono ripercuotersi negativamente sugli sviluppi delle discipline stesse.

La Scienza dell'informazione, che fa parte del vasto settore delle Scienze sociali, non costituisce una eccezione al quadro delineato così efficacemente da Whitley. Questo settore disciplinare manifesta un livello di frammentazione epistemologica particolarmente alto e mostra infatti, oggi, la necessità di superare questa eccessiva dispersione in una molteplicità di visioni scientifiche, che in alcuni casi possono essere state determinate da particolari tradizioni di ricerca che possono essersi instaurate anche in risposta a strategie legate a specifiche esigenze, successivamente mutate o cessate del tutto.

La storia e lo sviluppo di questo settore si presentano particolarmente complessi, e per certi aspetti confusi e disordinati, in un intreccio di modelli

ed esperienze che si susseguono durante tutto il Novecento e proseguono nel nuovo millennio, con tentativi di sistemazione teoretica non sempre efficaci e soprattutto non condivisi da tutti gli studiosi che hanno partecipato e partecipano oggi al settore disciplinare.

Questo libro intende fornire un contributo per fare ordine – se è possibile – in questo campo scientifico considerato nella sua interezza, e mettere a fuoco l'identità disciplinare della Scienza dell'informazione definendo le coordinate scientifiche che caratterizzano il settore. Si propone l'obiettivo di delineare i contorni e i confini dei sotto-campi disciplinari, individuare quei nuclei di esperienze che negli anni sono emerse, a volte costituendo da sole delle sotto-discipline, precisare i rapporti con i settori di ricerca vicini e con quelli con cui esso è entrato in contatto, e soprattutto di mettere in luce i diversi paradigmi che nei decenni sono stati di volta in volta seguiti da un numero sostanzioso di studiosi. La ricerca ha fatto emergere un numero davvero rilevante di contributi, su riviste scientifiche e in atti di congressi, ma anche di monografie, che oggi sono dimenticati e del tutto sconosciuti alle nuove generazioni di studiosi e di professionisti.

L'obiettivo, ambizioso, è mettere in evidenza la genesi e gli sviluppi di un settore disciplinare difficile da definire, focalizzarsi sulle diverse esperienze scientifiche, tecnologiche ed applicative riconducibili sotto l'egida della Scienza dell'informazione, che ha assunto sfumature diverse nel corso degli anni, e far emergere i legami profondi, oltre che con la Biblioteconomia e la Documentazione, con la Bibliografia intesa come Scienza degli indici.

Se il legame con la Documentazione, e con le tecniche per il sostegno della ricerca dei contenuti informativi, è stato più volte messo in luce, ad esempio da Buckland, e la connessione con la Biblioteconomia ha dato luogo al concretizzarsi del modello della Library and Information Science, il legame con la Bibliografia è stato, sì, esaminato, ad esempio da Patrick Wilson, ma non ha ricevuto ancora la considerazione che merita la sua dimensione indicale, quel settore della Bibliografia, cioè, che si occupa degli indici, in particolare di quelli semantici, dell'indicizzazione realizzata per facilitare lo sviluppo delle conoscenze dei lettori e dell'organizzazione dei contenuti conoscitivi presenti nelle memorie collettive.

Nel suo libro *Introduzione alla Scienza dell'Informazione*, Alberto Salarelli (2012), uno tra i pochissimi studiosi italiani che hanno offerto un esame della disciplina, definisce la Scienza dell'informazione un campo disciplinare dedicato allo “studio sistematico dell'intero ciclo di vita dell'informazione e delle sue implicazioni con i diversi campi del sapere” (p. 9), e ne individua gli elementi fondativi in primo luogo nel concetto di informazione e poi in quello di documento. Salarelli considera la nascita della Scienza

dell'informazione strettamente connessa con gli sviluppi della tecnologia informatica e riconosce come elemento discriminante l'utilizzo dei computer nella gestione dei documenti e delle pubblicazioni. Lo stesso approccio era stato in sostanza seguito dall'Autore nella voce dedicata a *Scienza dell'informazione*,<sup>1</sup> in cui la nascita del settore disciplinare veniva ricondotta sostanzialmente alla fase di esordio dell'uso dei computer e legata alla teoria della comunicazione di Shannon e alla nascita della Cibernetica.

Salarelli individua nell'approccio quantitativo intrapreso da Shannon e Weaver e nel modello da loro stabilito, e relativo alla possibilità di misurare il processo comunicativo che si determina utilizzando un codice, sulla base delle scelte degli elementi informativi tra un insieme di elementi che abbiano la stessa possibilità di essere selezionati, il tronco da cui scaturiscono le attività e gli sviluppi dell'Information Science intesa principalmente come Information Retrieval. Il secondo approccio che Salarelli individua è quello cognitivo, che riconosce l'importanza del significato, del valore che l'informazione assume per il singolo individuo che la riceve, e sui diversi contesti nei quali la significazione opera, e si sofferma brevemente su alcune teorie, il costruttivismo, il collettivismo socio-cognitivo e il costruzionismo, come particolarmente produttive per gli sviluppi teoretici dell'area.

L'orientamento che ha seguito Salarelli per delineare le caratteristiche della Scienza dell'informazione, incentrato sul concetto di informazione, e che assume come elemento fondativo la teoria della comunicazione di Shannon e Weaver che lo conduce a riconoscere nell'uso del computer il fattore di snodo per il trattamento e l'organizzazione dei contenuti informativi, non è l'unica angolatura attraverso la quale considerare l'intera area disciplinare.

Dall'esame delle diverse anime che compongono il variegato scenario disciplinare, emerge in primo luogo come non sia tanto il concetto di informazione in sé, qualsiasi tipologia di 'informazione', a costituire il *focus* del settore, quanto il trattamento a tutti i livelli dell'informazione bibliografica. Dai sistemi tradizionali di Information Retrieval alle interfacce avanzate di alcune basi di dati con funzionalità innovative per la ricerca e alle interfacce per l'utente dei cataloghi elettronici, tutto sembra ruotare attorno alla categoria bibliografica. Almeno una delle anime della Scienza dell'informazione ha come oggetto di studio l'informazione bibliografica.

Come ha rilevato nel 1999 Marcia Bates, studiosa del settore, e come è in seguito emerso grazie alla ricerca internazionale coordinata da Chaim Zins (2003-2005), la Information Science nel suo complesso si può considerare un settore meta-disciplinare, che si è assunto il compito di trattare e organiz-

<sup>1</sup> *Biblioteconomia. Guida classificata* (2007), pp. 173-178.

zare i contenuti delle altre scienze. Come meta-disciplina la Scienza dell'informazione può dunque essere considerata un corpo unitario, un unico tronco ramificatosi negli anni, che ha fatto emergere settori significativi, dotati di caratteristiche proprie, con riviste scientifiche e congressi dedicati, ma riconducibili ad una unica matrice.

Sulla base degli argomenti presentati in questo libro, è possibile riflettere su quali siano le radici su cui si fonda la Scienza dell'informazione e quali siano state le diverse esperienze che sono state condotte in tempi e contesti diversi.

Il primo orientamento si mette in evidenza seguendo la linea che collega la Documentazione, come è stata concepita alla fine dell'Ottocento e nei primi anni del Novecento da Paul Otlet insieme a Henri La Fontaine, e le tecniche sviluppate a sostegno dello scambio delle informazioni scientifiche avendo come obiettivo lo sviluppo della ricerca. La derivazione dalla Documentazione assegna a questo filone disciplinare connotati che permettono di definirlo un settore orientato all'individuazione e al trattamento delle informazioni fattuali – fatti, dati, statistiche, come affermava Otlet – e non interessato all'indicizzazione dei contenuti dei libri e degli articoli, e a mettere in evidenza gli argomenti, come accade nel settore bibliografico.

Importante sottolineare come questo primo orientamento si colleghi all'attività negli Stati Uniti di Watson Davis e all'istituzione dell'American Documentation Institute (ADI) nel 1937. Davis era interessato soprattutto allo scambio delle informazioni scientifiche tra scienziati e all'impiego dei microfilm. Bisogna inoltre rilevare come dagli anni Cinquanta la Information Science abbia incrementato la sua dimensione documentale dovendo affrontare il trattamento dei rapporti tecnici, dei brevetti, ma anche dei documenti derivanti dall'attività bellica, requisiti ai nemici o inviati dagli alleati, come ha messo bene in evidenza Irene Farkas-Conn nel suo libro (1990), che dovevano essere indicizzati con metodologie nuove che le biblioteche in quel momento non potevano garantire, e di cui si occupò in particolare Mortimer Taube. Si tratta fondamentalmente di due esperienze diverse, quella europea di Otlet e La Fontaine e quella statunitense di Davis e di American Documentation Institute.

Sulle attività di disseminazione delle informazioni scientifiche dette allora di Information Science – in particolare attraverso microfilm e abstract – si tennero sia la prima conferenza sull'informazione scientifica, a Londra nel 1946 curata dalla Royal Society, la seconda nel 1948 sempre a Londra, e quella tenuta a Washington nel 1958 sull'informazione scientifica, focalizzata sui processi di *storage* e *retrieval* delle informazioni. Rientrano in questa anima della Scienza dell'informazione le nuove metodologie di indicizzazione (Mooers, Taube, Luhn) e di ritrovamento di documenti che hanno trovato una concreta rappresentazione nell'ambito dell'Information Retrieval, che per un periodo di

tempo sostanzialmente coincide con l'Information Science *tout court*, comprese anche le metodologie introdotte da Gerard Salton, l'indicizzazione automatica che adotta i 'pesi' per i termini assegnati e il matching tra le richieste degli utenti e i documenti usando la metodologia basata sui vettori e applicando coefficienti di correlazione.

Un fondamentale e fertile settore, Library and Information Science (LIS), emerse alla fine degli anni Sessanta dalla convergenza tra Biblioteconomia, Documentazione e Information Retrieval, come ha suggerito in particolare Boyd Rayward. Significativa, in questo caso, la distanza dalla teoria della comunicazione di Shannon e Weaver e il riconoscimento della centralità degli studi del settore della Semantica. Il campo disciplinare si orienta verso il riconoscimento della preminenza delle attività realizzate dalle biblioteche a favore degli utenti e si concentra sulle analisi dell'applicazione delle funzionalità dell'Information Retrieval agli strumenti catalografici e in particolare a tutte le attività di ricerca e applicative che trovano albergo sotto l'etichetta di Knowledge Organization, vale a dire tutti i procedimenti di indicizzazione e di classificazione bibliografica e tutti i sistemi per l'organizzazione e per l'accesso alle conoscenze registrate sviluppati sia in ambito bibliotecario che documentario: i thesauri, ad esempio, sono espressione del settore documentario ma hanno fatto percepire la loro influenza anche in quello bibliotecario. Negli anni Novanta è stato soprattutto Peter Ingwersen, sulla scia della concezione di informazione manifestata da Belkin, a definire le caratteristiche della Scienza dell'informazione come Library and Information Science: centralità dell'elemento umano nel processo di connessione tra Information Retrieval e i singoli individui nel loro ambiente sociale e concentrazione delle attività di organizzazione delle fonti informative nelle biblioteche e nei centri di informazione o di documentazione.

Attorno a questi nuclei principali ruotano altri ambiti di riflessione più specifici, tra cui gli studi relativi al concetto di informazione, di stampo filosofico, gli studi sull'utenza e le strategie di ricerca, gli studi relativi al trattamento dei documenti e, oggi, dei dati in ambiente digitale, la descrizione delle risorse sul Web e il collegamento dei dati con tecnologie Linked data.

Infine, è particolarmente significativo il filone che si evidenzia seguendo la linea che dalla Bibliografia come disciplina conduce alla Scienza dell'informazione, linea e collegamento meno appariscenti e difficilmente riconosciuti. La Bibliografia presenta una valenza culturale in quanto costituisce il riflesso delle memorie scritte, e una valenza indicale come mappa dei saperi e indice delle connessioni tra le discipline e può servire come strumento per l'organizzazione della conoscenza. Le procedure citazionali sviluppate a partire dall'intuizione di Garfield, nella sostanza, offrono la possibilità di realizzare mappe del sapere, collegamenti basati su una congruenza semantica,

così come le procedure co-citazionali e il *bibliographic coupling*. L'ampliamento delle funzionalità per la ricerca, soprattutto nelle basi di dati, attraverso l'applicazione di thesauri e ontologie, che consentono di compiere ricerche su insiemi di termini semanticamente correlati, sostanzialmente è orientato verso la ricostruzione di mappe di contenuti. Inoltre, il superamento recente del concetto di similarità dei documenti basato sul computo della frequenza dei termini, e l'adozione della misura della similarità basata sulla presenza di citazioni in comune tra articoli, il *bibliographic coupling*, ma anche il potenziamento delle funzionalità della ricerca nelle basi di dati bibliografiche (PsycINFO in particolare) attraverso l'uso degli indici semantici tradizionali, affondano le loro radici nella Bibliografia indicale.

Ma la Bibliografia assume rilievo anche nell'ambito dell'indicizzazione vera e propria, in quanto portatrice di un bagaglio di logiche indicizzatorie sviluppate nei secoli per l'organizzazione della comunicazione scritta. La Bibliografia indicale, come Alfredo Serrai l'ha definita e presentata, può assumere a mio avviso un valore fondativo per la Scienza dell'informazione e per il settore che di essa fa parte, la Knowledge Organization, e raccoglie dietro di sé tutte le attività indicizzatorie e le logiche catalografiche e citazionali che nei secoli sono state significative. Le tecniche di indicizzazione semantica approfondita che derivano dalla tradizione gesneriana, ad esempio, volte a mettere in evidenza e a permettere di reperire facilmente gli argomenti trattati nei testi, e le connessioni tra di essi, attraverso opportuni rinvii, possono costituire un modello per la Scienza dell'informazione lontano dalle tecniche dell'Information Retrieval e orientato invece a soddisfare le necessità del lettore che, attraverso i collegamenti logici fatti emergere dalla trama degli indici, intende fare proprie le risorse contenutistiche dei testi che consulta, permettendo ad esse di entrare in contatto con il proprio sistema cognitivo.

Parte prima  
*Il panorama disciplinare*



# *1. Genesi e identità scientifica della Scienza dell'informazione*

Il panorama scientifico-disciplinare offerto dalla Information Science è stato definito da Shera e Cleveland un'aggrovigliata matassa composta da molti fili che si intrecciano, "a tangled skein of many intertwining threads"<sup>1</sup>, e certamente questa è un'immagine che si adatta bene al campo di cui ci occupiamo. Si tratta di un campo di ricerca scientifica e di pratica professionale certamente vasto e molto articolato, che inizia ad affermarsi come settore autonomo a partire dagli anni Cinquanta del secolo scorso e si sviluppa accogliendo parte delle esperienze condotte nel tempo da altre fondamentali discipline, come la Bibliografia, la Biblioteconomia, la Documentazione, e recependo le nuove esperienze maturate a sostegno dell'informazione scientifica. Seguendo l'opinione dei due autori citati, possiamo affermare che sostanzialmente esso è dedicato allo studio delle modalità per l'accesso intellettuale all'informazione registrata e alla definizione di modelli applicativi per il ritrovamento dei contenuti informativi rilevanti per un determinato obiettivo, per il loro utilizzo e la loro diffusione, attività, queste, che hanno costituito un *continuum* nel corso dei secoli.

Attraverso la denominazione di Information Science sono stati designati, negli anni, campi scientifici e applicativi diversi. Come ha sottolineato Fairthorne,<sup>2</sup> si tratta di un termine 'ombrello', usato per denominare attività differenti, che comprendono sia il processo di acquisizione di nuovi concetti con l'arricchimento delle conoscenze da parte degli esseri umani, sia l'informazione selettiva concepita nell'ambito dell'ingegneria della comunicazione, e in particolare nella teoria della comunicazione di Claude Shannon, come vedremo; ma anche per fare riferimento all'attività di estrazione ed immagazzinamento di 'informazioni' concepite come 'cose' o come 'fatti' di per sé significativi.

<sup>1</sup> Jesse H. Shera, Donald Cleveland (1977), p. 249.

<sup>2</sup> Robert Fairthorne (1975).

La Scienza dell'informazione presenta una natura fondamentale multidisciplinare (o crossdisciplinare).<sup>3</sup> Wersig e Neveling<sup>4</sup> hanno riconosciuto per primi che essa si è sviluppata come scienza in sé coagulando in un unico tronco le caratteristiche di diverse discipline, tra cui, a loro giudizio, la Computer Science, la Library Science, la Linguistica, la Filosofia (e la Tassonomia), la Teoria dell'informazione, ma anche la Matematica e la Cibernetica. Questa molteplicità di prospettive ha fatto sì che il nuovo campo di studio e applicativo si presentasse disomogeneo e frastagliato, e che si manifestassero differenti orientamenti, quello teorico, stretto intorno al concetto di informazione, quello applicativo, concentrato sui mezzi ritenuti utili nel lavoro informativo. Secondo i due autori è emerso tuttavia un orientamento che riconosce la Scienza dell'informazione come il campo del lavoro pratico che possa soccorrere alcuni bisogni e necessità sociali. Il nuovo campo disciplinare si è sviluppato a seguito dell'incremento delle pubblicazioni primarie e secondarie e a seguito dello sviluppo di nuove e più complesse tecnologie della comunicazione. La nuova disciplina emerge quindi per sostenere necessità di tipo sociale, la trasmissione della conoscenza a chi ne ha bisogno, e questa responsabilità nei confronti della società ne costituisce l'elemento fondativo:

Thus a new discipline developed [...] because of a new necessity to study a problem which had completely changed its relevance for society. Nowadays the problem of transmitting knowledge to those who need it is a social responsibility, and this social responsibility seems to be the real background of "information science".<sup>5</sup>

A partire dalla fine degli anni Sessanta del secolo scorso il settore disciplinare ha manifestato chiaramente il suo legame con le attività svolte all'interno degli istituti bibliotecari, e da quegli anni comincia ad essere definito

<sup>3</sup> Per la definizione dei campi interdisciplinare, multidisciplinare, crossdisciplinare e transdisciplinare, ci possiamo avvalere della terminologia adottata dalla European Peer Review Guide (ESF), che a p. 39 presenta le seguenti definizioni: "Multidisciplinary research studies a topic within one driving discipline, with support from complementing disciplines and brings together multiple dimensions. Interdisciplinary research is concerned with the analysis of a topic in different disciplines and according to different approaches. Crossdisciplinarity, instead, means the study of a topic at the intersection of different disciplines. Transdisciplinary research is dealing with topics that are across and beyond all the disciplines, aiming to the unity of knowledge". La Library and Information Science, considerata come campo di studio, presenta le caratteristiche della multidisciplinarietà e della crossdisciplinarietà. Si veda Member Organisation Forum, 2011, disponibile a [http://www.scienceonthenet.eu/files/european\\_peer\\_review\\_guide\\_01.pdf](http://www.scienceonthenet.eu/files/european_peer_review_guide_01.pdf)

<sup>4</sup> Gernot Wersig, Ulrich Neveling (1975), reperibile a: <https://sigir.org/files/museum/pub-13/18.pdf>

<sup>5</sup> *Ivi*, p. 28.

Library and Information Science, denominazione il cui impiego sottolinea la convergenza avvenuta tra il mondo bibliotecario e le tecniche adottate per sostenere la ricerca dei contenuti informativi. Attualmente la maggior parte degli studiosi del settore utilizza la denominazione breve Information Science anche per indicare la Library and Information Science<sup>6</sup>.

L'emergere della Scienza dell'informazione come campo disciplinare unitario viene spesso messo in riferimento con l'attività pionieristica di Vannevar Bush, professore al MIT di Boston tra il 1923 e il 1938, ed esperto di sistemi per la ricerca dell'informazione per il governo degli Stati Uniti durante la seconda guerra mondiale, il quale già negli anni Trenta aveva concepito una soluzione tecnologica non tanto per affrontare il problema dell'esplosione delle pubblicazioni scientifiche, quanto per organizzare i contenuti informativi in un modo confacente alle necessità della ricerca.<sup>7</sup> Convinto che i sistemi di classificazione che si possono adottare per organizzare i contenuti dei libri e degli articoli siano particolarmente rigidi, e non permettano di ritrovare i contenuti desiderati perché distribuiti in classi e sottoclassi diverse, in quanto gli argomenti sono proposti sotto i diversi punti di vista delle classi disciplinari, Bush progetta Memex, una macchina per la gestione personale delle informazioni e dei contenuti che dovrebbe permettere di sovvertire il meccanismo della classificazione in favore di una organizzazione basata sulla possibilità della mente umana di associare e correlare i concetti e i contenuti dei libri.<sup>8</sup> Il progetto Memex prevedeva la realizzazione di una macchina – in sostanza una scrivania da lavoro – dotata di dispositivi elettromeccanici che, nelle intenzioni del suo progettatore, lavorando sugli abstract microfilm degli articoli, avrebbe dovuto permettere di identificare le informazioni e i concetti al fine di poterli poi ritrovare in modi diversi, basandosi sull'idea dell'indicizzazione secondo le associazioni. La macchina, che non fu mai costruita, doveva essere uno strumento per fornire aiuto alla memoria umana, permettere di scrivere note e commenti, e creare *link* tra gli argomenti collegati nei diversi articoli scientifici, in modo da poterli riutilizzare nel modo desiderato. Oggi è considerata l'antecedente teorico della memoria del personal computer e dell'ipertesto.

Dal momento che era stata progettata in un'epoca in cui il computer non era ancora una realtà operativa, Bush ha adottato per la sua macchina i microfilm, soluzione che, come vedremo, fa parte delle innovazioni tecnologiche a corredo della impostazione bibliografico-documentaria di Paul Otlet. I

<sup>6</sup> In un recentissimo articolo, Ávila Araújo (2021) sintetizza i fondamenti della Scienza dell'informazione.

<sup>7</sup> Vannevar Bush (1945).

<sup>8</sup> Per un'analisi approfondita del progetto di Bush si vedano le pp. 99-110 di Paola Castellucci (2009).

microfilm sono a disposizione nella scrivania, selezionabili attraverso pulsanti, leggibili con lettori di microfilm su tre schermi diversi, per poter operare confronti e associazioni; i concetti rilevanti avrebbero potuto essere individuati dal singolo ricercatore e segnati attraverso l'uso di codici sul bordo della pellicola, in modo da poter essere recuperati tramite lettori basati su cellule fotoelettriche. Tramite le marcature sui bordi dei microfilm il ricercatore avrebbe potuto recuperare tutti i documenti in cui è trattato un certo argomento. In sostanza, Bush ha utilizzato un sistema di indicizzazione per concetti e argomenti impiegando i microfilm (quindi direttamente i documenti) invece di sfruttare la mediazione delle schede catalografiche, anticipando di anni le realizzazioni di Calvin Mooers, sul quale ci soffermeremo più avanti, e gli strumenti per l'Information Retrieval.

Non sappiamo quale sarebbe stata, nel caso del Memex, la caratura degli argomenti, quale l'estensione dei concetti, o dei soggetti; si avverte la mancanza di una teoria e di una metodologia per l'indicizzazione semantica che guidi la selezione degli argomenti e sostenga il collegamento tra i contenuti permettendo di mettere in relazione contenuti dello stesso spessore e dello stesso livello di approfondimento nei diversi documenti, evitando di collegare documenti in cui un concetto è solo citato con documenti nei quali invece esso è completamente sviscerato. Questo, invece, è l'elemento che caratterizza la ricerca attraverso gli indici semantici che vengono realizzati in campo bibliografico e catalografico.

In sostanza, l'attività pionieristica di Bush con il Memex, secondo il mio punto di vista, si esplica al livello delle pratiche di indicizzazione, è diretta a realizzare uno strumento innovativo che faciliti il possibile recupero dei documenti nei quali gli argomenti vengono trattati; si pone sullo stesso piano dei tentativi di indicizzazione in cui si utilizzano le schede cartacee con indentature sul bordo, proposte da Mooers per il recupero dei documenti, come vedremo. Anche in quel caso si tratta di tentativi che non contemplano alcuna metodologia di indicizzazione semantica, proprio perché destinati a considerare singoli documenti, brevetti, rapporti tecnici, e non libri che presentano una trama di contenuti complessa. Anche la geniale intuizione di Eugene Garfield, negli anni Sessanta, di utilizzare le citazioni bibliografiche istituite dagli stessi autori degli articoli per indicizzarne i contenuti in modo da evitare la soggettività della indicizzazione da parte dei catalogatori mediatori, fondando il nuovo paradigma bibliometrico, come vedremo, costituisce, in partenza, una riflessione innovativa che fonda le sue radici sempre nell'ambito dell'indicizzazione semantica.

L'importanza del contributo di Bush risiede piuttosto nell'attività di organizzatore del trasferimento e dello scambio delle informazioni scientifiche tra gli scienziati. Durante la seconda guerra mondiale infatti gli scienziati

dell'informazione e gli ingegneri lavorarono per gli obiettivi militari. L'Office of Scientific Research and Development (OSRD), diretto da Vannevar Bush, coordina la ricerca scientifica dei laboratori universitari e assicura lo scambio delle informazioni scientifiche tra gli scienziati. Alla fine della guerra, il presidente Roosevelt decise di riutilizzare l'esperienza dell'OSRD e propose a Bush di riorganizzare una infrastruttura per l'informazione scientifica, riconvertendo a scopi civili le strutture preesistenti. Sarà il presidente Truman a chiudere OSRD e a varare un piano per la comunicazione scientifica, affidato a John R. Steelman, incentrato sulle discipline 'dure'.<sup>9</sup> Un efficace processo di trattamento documentario era stato messo in moto ancora durante la guerra, con una collaborazione tra gli alleati: a Londra, ad esempio, fu istituito l'Air Documents Research Center (ADRC), cui parteciparono la US Navy e il British Air Ministry, che si occupava di raccogliere e riprodurre i documenti sottratti al nemico. Negli Stati Uniti, nel 1945, Vannevar Bush suggerì al presidente di istituire un Publication Board con il compito di trattare e gestire i documenti raccolti durante le attività belliche: fu istituito l'Office of Technical Service (OTS), con il supporto dell'American Documentation Institute e di Watson Davis, personaggio del quale torneremo a parlare.

L'informazione scientifica e i servizi per incrementare l'efficacia della sua disseminazione sono stati al centro dell'attività di ricerca e dello sviluppo di attività applicative durante tutto il Novecento. A Londra nel 1946, a cura della Royal Society, si tenne una prima conferenza dedicata ai servizi per l'informazione scientifica, Royal Society Empire Scientific Conference, durante la quale la Royal Society fu invitata a curare una seconda conferenza interamente dedicata all'esame dei servizi di informazione scientifica. Nel 1948, dal 21 giugno al 2 luglio, la Royal Society ospitò dunque una seconda conferenza, Royal Society Scientific Information Conference,<sup>10</sup> cui parteciparono centotrenta delegati provenienti dal Regno Unito, che fu dedicata proprio allo studio dell'uso dell'informazione scientifica e all'analisi dei servizi di informazione. Ralph Shaw, bibliotecario del US Department of Agriculture, rappresentava il governo degli Stati Uniti alla conferenza del 1948 e pubblicò un resoconto dettagliato dei lavori.<sup>11</sup> La conferenza fu organizzata in quattro sezioni, nelle quali furono indagati i seguenti argomenti, tutti considerati sotto la particolare ottica dell'utilizzazione e del servizio offerto agli scienziati: l'attività di pubblicazione e di distribuzione dei lavori scientifici

<sup>9</sup> Le informazioni sono contenute nel libro di Irene Farkas-Conn (1990), alle pp. 102-109.

<sup>10</sup> The Royal Society Scientific Information Conference (1948).

<sup>11</sup> Ralph R. Shaw (1948). Per un dettagliatissimo resoconto di molti dei paper presentati, si veda la rassegna di J. D. Bernal, David Chadwick, J. E. Holmstrom e H. Munro Fox, i quattro editor delle sessioni della conferenza (1948).

originali, insieme all'analisi della pubblicazione con la tecnica del microfilm; la realizzazione di abstract soprattutto nel caso della letteratura scientifica pubblicata nelle lingue meno diffuse o note; l'attività di indicizzazione da parte delle biblioteche, l'uso della Universal Decimal Classification, l'indicizzazione alfabetica, le possibilità di adozione di strumenti per la selezione meccanica degli argomenti e degli indici, e la realizzazione di servizi bibliotecari, tra cui la fornitura di copie di articoli a distanza; l'ultima sessione riguardò le rassegne e i rapporti annuali. Bisogna rilevare che durante la terza sessione la conferenza affrontò un tema generale di grandissimo rilievo, cioè il problema dell'osservanza del copyright in relazione alle necessità degli studiosi e raccomandò, sotto gli auspici della Royal Society, che in futuro si accordasse la più vasta possibilità agli studiosi di avere copie del materiale necessario per i loro studi.

A Washington nel 1958, a cura della National Science Foundation e di altri istituti scientifici, si tenne la International Conference on Scientific Information.<sup>12</sup> La conferenza, di ampio respiro, alla quale parteciparono circa mille delegati, si ricollegava a quella tenuta dieci anni prima presso la Royal Society, ma si focalizzò invece, come rileva nella sua prolusione Sir Lindor Brown, senior Secretary della Royal Society, sul segmento finale del processo dell'informazione, cioè la sua archiviazione e le possibilità di recupero (*storage and retrieval*). I temi trattati spaziarono quindi in diversi campi: l'analisi delle necessità degli scienziati di qualsiasi area disciplinare di accedere alla letteratura scientifica, l'esame degli strumenti e dei servizi informativi a disposizione degli studiosi, l'analisi dei requisiti dei servizi di abstract, ad esempio, il livello di copertura dei campi disciplinari e la qualità del lavoro di abstract, l'analisi dell'uso delle informazioni reperite da parte degli scienziati, la definizione dei requisiti della recensione dei lavori degli altri scienziati e l'importanza dell'attività di review.

Tornando ora a riflettere sulla genesi del campo di ricerca e di applicazioni di cui ci occupiamo, conoscere la storia della nascita del termine che lo denomina può fornire qualche indicazione per comprendere la genesi di questo complesso campo disciplinare e poi le sue trasformazioni nel corso del Novecento. È infatti unanimemente riconosciuto che l'impiego della denominazione Information Science risalgia al 1958 e all'istituzione da parte di Jason Farradane,<sup>13</sup> nel Regno Unito, dell'Institute of Information Scientists dedicato alla

<sup>12</sup> International Conference on Scientific Information (Washington, D. C. November 16–21 1958). I *Proceedings* (1959), sono disponibili a <https://www.nap.edu/download/10866>.

<sup>13</sup> Jason Farradane, nato Jason Lewkowitsch (1906 – 1989), di origine polacca, laureato in Chimica e poi documentalista, si presenta sulla scena internazionale come studioso durante la Royal Society Scientific Information Conference del 1948. Fece parte del Classification