

Jane Devine  
Francine Egger-Sider

### **Going beyond Google. The invisible web in learning and teaching**

London, Facet Publishing, 2009,  
p. 162, ISBN 978-1-85604-658-9

Sulla sommità di una montagna arriva trafelato lo sciatore-ricercatore che chiede finalmente al guru-asceta quale sia il significato della vita. E l'asceta risponde: "Il significato della vita? Cosa dovrei saperne io? Interroga Google!". Si apre all'insegna di questa cinica vignetta il libro-manuale delle due autrici esperte di *information retrieval* ma anche attive sul campo della professione bibliotecaria: Jane Devine è *chief librarian* alla La Guardia Community College Library mentre Francine Egger-Sider copre il ruolo di *coordinator of technical services* allo stesso College. Nonostante la graffiante ironia dell'incipit, l'obiettivo perseguito costantemente per tutto il volume non è "dissacrare Google"; semplicemente ci si prefigge di ridimensionarlo inserendolo in una corretta prospettiva, spiegando in modo semplice e leggero come l'architettura del web sia una realtà complessa, articolata, profonda e organizzata che non può essere indagata in modo qualitativamente utile con un uni-

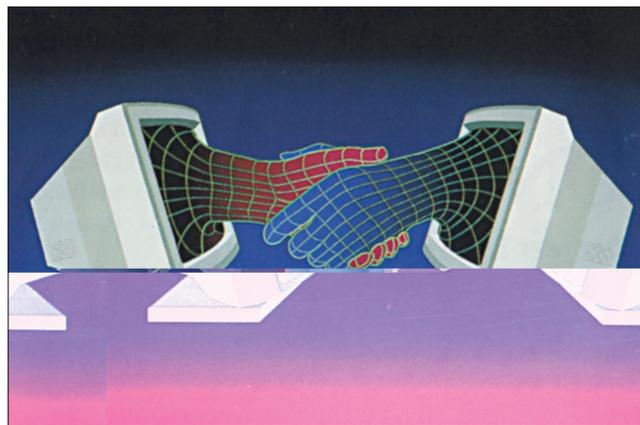
co e generalizzato strumento. *Going Beyond Google* può essere considerato una guida che mira a fornire suggerimenti e direttive di metodo per bibliotecari o docenti interessati a fornire ai giovani studenti e ricercatori le conoscenze e abilità richieste per ricercare informazioni di qualità nel web. Capacità che, come rilevavano Shapiro e Hughes,<sup>1</sup> oggi giorno dovrebbero essere parte del "trivio", vale a dire parte ineludibile delle conoscenze culturali di base per l'intellettuale del nostro millennio. Il problema dell'*information retrieval*, che ritorna sovente nelle pubblicazioni più o meno specialistiche di area americana e inglese, pare più sentito nel mondo anglosassone, perché in questo contesto c'è sicuramente uno sviluppo maggiore delle discipline biblioteconomiche e soprattutto un più avanzato stato di integrazione con la sfera della *computer science* a tutti i livelli disciplinari; tuttavia il problema della povertà e dell'appiattimento delle ricerche scolastiche e universitarie e dello scarso sfruttamento delle risorse online da parte della massa di studenti e ricercatori è grave così nel mondo anglosassone come in Italia.<sup>2</sup> Non è più una novità che lo studente dopo aver fatto un grossolana ricerca in Google e aver

recepito le informazioni per lo più dall'enciclopedia condivisa Wikipedia sull'argomento di interesse e aver creato la propria tesina con un copia e incolla ripreso dai primi link che più gli sembravano seri e opportuni si presenti in biblioteca alla ricerca di qualche testo da mettere in bibliografia.<sup>3</sup>

Google, preso come riferimento di tutta una categoria, perché di fatto è il *general-purpose search engine* più diffuso, è diventato negli ultimi anni il punto di riferimento per scovare ogni tipo di informazione, portando ad amplificare due false credenze rispetto alle risorse presenti nel web: "cercare è semplice" (*searching is easy*) e "tutto ciò che esiste di importante è accessibile gratuitamente" (*everything important is free*). È proprio a questo scopo che il libro è dedicato: smascherare quanto ingannevoli siano queste idee, cercando di fornire strumenti agli addetti ai lavori per spiegare cos'è il web cosiddetto invisibile e insegnare come riuscire a utilizzarlo grazie a strumenti alternativi a Google e specifici del mondo dell'*information retrieval*. Il libro è organizzato in tre parti: la prima è una sezione teorico-esplicativa, la seconda si focalizza sui criteri di ricerca del web nascosto e sui metodi per insegnare a

fare ricerche mirate, e infine la terza parte propone una rassegna di teorie, pratiche e strumenti in sviluppo presso aziende private o promosse da fondazioni o istituzioni pubbliche per cercare di scandagliare sempre più a fondo il web nascosto. Il volume è inoltre corredato di appendici dove ad un'aggiornata seppur sintetica bibliografia segue una serie di materiali (immagini, siti, proposte didattiche) adatti per l'insegnamento dell'invisible web; infine sono riproposti gli standard dell'Association of College and Research Libraries (ACRL) sull'*information literacy*.<sup>4</sup>

La prima parte, *Understanding the division between the visible and invisible web*, è dedicata alla descrizione del web invisibile che si definisce in maniera dinamica proprio in base all'esistenza e al funzionamento dei motori di ricerca generalisti che utilizzano dei software specifici, detti *spiders* o *crawlers* o *robots*, per recuperare le informazioni che non si basano certo sulla qualità dell'informazione ma piuttosto sui sistemi di *ranking*, *linking*, di quantità di citazioni... È proprio in virtù del fatto che certe informazioni non sono rilevate dai motori di ricerca come Google che si è battezzato questo materiale web "invisibile", o "profondo", o "nascosto". Tipicamente le risorse nascoste sono materiali presenti in database o in siti ricchissimi e organizzati con una gran varietà di pagine. Da uno studio di Bergman, commissionato dalla Brightplanet e pubblicato nel 2001 ma aggiornato pochi mesi fa,<sup>5</sup> si deduce che le informazioni nascoste ai motori di ricerca generalisti siano più di 1.000-2.000 volte il materiale del web visibile. Dopo aver chiarito l'archi-



tettura del web e spiegato come mai esistano così tante preziose informazioni nascoste a Google & co., le autrici fanno il punto della situazione sull'uso del web per ricerca. Con l'ausilio di alcuni studi sul comportamento di studenti in varie aree geografiche e per varie fasce generazionali, si dimostra che l'uso pressoché costante e indifferenziato dei motori di ricerca generalisti per fare ricerca è l'abitudine più diffusa tra i "nuovi ricercatori", che producono di fatto tesine e *papers* piuttosto banali e non soddisfacenti. Preso atto di questo malcostume, diventa necessario che figure quali insegnanti e bibliotecari svolgano un ruolo di mediazione e filtro tra studente e informazione, ovvero si preoccupino di formare le nuove leve all'*information literacy*. La seconda parte, *Finding and utilizing the contents of the invisible web*, passa infatti al concreto: i capitoli 3-4 danno veri e propri consigli didattici e di metodo utili a tutti gli operatori che devono insegnare ai giovani come riconoscere, accedere e valutare le informazioni raccolte da una ricerca che comprenda anche l'*invisible web*. I punti di arrivo ideali sono gli standard formulati dall'ACRL e dall'AASL.<sup>6</sup> le autrici raccolgono in due tabelle esplicative gli obiettivi raggiungibili con il percorso metodologico suggerito e i risultati auspicati dalle due associazioni. Il capitolo 5 propone un vero e proprio caso studio: Amanda e la sua ricerca sullo sviluppo internazionale del concetto di "microfinanza". Il capitolo 6 passa in rassegna una ristretta e generale, seppur significativa, porzione di strumenti utili per indagare il web nascosto.<sup>7</sup>

La terza parte, *Narrowing the gap between the visible and invisible web*, propone infine una panoramica di progetti e iniziative messi in atto da grandi aziende come la stessa Google o da istituzioni pubbliche per cercare strumenti, protocolli, strategie (per esempio gli studi sul *semantic web*) che riescano a fornire all'utente del web quello che realmente sta cercando, cioè per fare in modo che l'invisibile diventi sempre più accessibile con un semplice unico strumento.

La conclusione è però che al momento l'intermediazione delle istituzioni per educare all'*information retrieval* o ancor meglio all'*information literacy* resta fondamentale, perché oggi Google non basta. Cercare, cogliere e interpretare il valore di un'informazione resta alla fine una capacità e abilità del singolo individuo.

In generale il libro è sicuramente utile come manuale semplice e fruibile per gli addetti ai lavori, cioè per quelle figure di docenti o bibliotecari chiamati a svolgere didattica sulle tematiche del fare ricerca in internet; tuttavia per chi volesse saperne di più sulla parte teorica della struttura del web e i problemi e le soluzioni avanzate per ovviare all'*hidden web* questo libro può essere solo un punto di partenza.

Linda Spinazzè

Università Ca' Foscari  
lin.lin@libero.it

<sup>1</sup> JEREMY J. SHAPIRO – SHELLEY K. HUGHES, *Information Literacy as a liberal art*, "Educom Review", 31/2, 1996, <<http://net.educause.edu/apps/er/review/reviewArticles/31231.html>>.

<sup>2</sup> FABIO METTIERI, *Il grande inganno del web 2.0*, Roma-Bari, Laterza, 2009, p. VI e 8; si veda anche MARCELLA PERUZZI – MARGARITA PEREZ PULIDO, *Risorse elettro-*

*niche: inchiesta sull'uso*, "Biblioteche oggi", 26 (2008), 9, p. 17-29.

<sup>3</sup> Cfr. FABIO METTIERI, *Il grande inganno del web 2.0*, cit., p. 114.

<sup>4</sup> Una delle migliori definizioni di *information literacy* è quella di Mettieri: "interpretazione del valore dell'informazione e suo uso consapevole", in FABIO METTIERI, *Fonti on-line: dalla valutazione alla contestualizzazione*, "Biblioteche oggi", 26 (2008), 10, p. 16.

<sup>5</sup> M. BERGMAN, *The deep web: surfacing hidden value*. White paper. BrightPlanet, 2001 <<http://brightplanet.com/images/stories/pdf/deepwebwhitepaper.pdf>>; sembra una sorta di aggiornamento il white paper firmato Brightplanet, *Harvest deep data that Google is Missing*, may 2009, <<http://www.brightplanet.com/images/stories/pdf/AboutBrightPlanet.pdf>>; l'aggiornamento di questi dati ha avuto anche una risonanza mediatica, si veda l'articolo di Alex Wright sul "New York Times", 22/02/2009, <[http://www.nytimes.com/2009/02/23/technology/internet/23search.html?\\_r=1&ref=business](http://www.nytimes.com/2009/02/23/technology/internet/23search.html?_r=1&ref=business)>.

<sup>6</sup> Association of College & Research Libraries, <<http://www.ala.org/ala/mgrps/divs/acrl/standards/index.cfm>>; American Association of School Librarian, <<http://www.ala.org/ala/mgrps/divs/aaasl/guidelinesandstandards/learningstandards/standards.cfm>>.

<sup>7</sup> È naturale che in un libro a carattere generale come questo non ci possa essere esaustività a livello disciplinare con elenchi dettagliati di tutte le risorse possibili; è infatti un lavoro specifico di ogni specialista disciplinare formarsi una lista di siti, database, risorse specifiche per il proprio settore.

