

L'utilizzo della rete per la ricerca di informazioni affidabili

Giovanni Bonaiuti^a, Giuliano Vivanet^b

^a *Università di Cagliari*, giovanni.bonaiuti@unica.it

^b *Università di Cagliari*, giuliano.vivanet@unica.it

Abstract

Realizzare ricerche affidabili in educazione è oggi un'esigenza per molti: dagli insegnanti e formatori, agli studenti, a chi deve prendere decisioni politiche su questioni finanziarie e organizzative. L'enorme quantità di informazioni presenti in rete può portare a nascondere i problemi connessi all'impostazione di una buona ricerca. Il contributo intende presentare strumenti e metodi capaci di estendere le modalità di ricerca usuali facendo riferimento ad alcune pratiche maturate nell'ambito dell'*evidence based education*.

Parole chiave: ricerca in educazione, archivi in internet, banche dati sull'educazione, educazione basata su prove di efficacia.

Abstract

Achieving reliable research in education is now a requirement for many people: teachers and trainers, students, decision and policy makers. The huge amount of information in Internet can lead to hide the problems related to the setting of a good research. The paper will present tools and methods extending the usual search methods by referring to some practices gained in evidence based education.

Keywords: research in education, Internet archives, databases on education, evidence based education.

Introduzione

La rete offre oggi un incredibile patrimonio di risorse eterogenee: notizie, articoli, commenti, video, dati numerici, immagini, libri digitali, registrazioni audio e molte altre cose. La quantità di informazioni cresce esponenzialmente a un ritmo sorprendente, e questo grazie a fattori quali la sempre più capillare diffusione di internet, la progressiva digitalizzazione di tutte le risorse conoscitive, la crescente disponibilità di risorse prodotte dagli utenti secondo lo spirito del web 2.0. La rete rappresenta, dopo l'invenzione della stampa, la più grande rivoluzione nelle modalità di diffusione del sapere della storia dell'umanità. Se l'invenzione di Gutenberg portò in 50 anni dai circa 30 mila manoscritti disponibili alla fine del Quattrocento ai 10 milioni di libri diffusi in tutta Europa, in un tempo molto minore la rete ha raggiunto i 295 exabytes, una quantità di dati pari a circa 13 strati di libri disposti su un territorio come quello degli Stati Uniti (Hilbert, López, 2011).

Una tal estensione di informazioni rappresenta una preziosissima fonte di conoscenza, ma anche un significativo problema. La possibilità di "ricercare" qualcosa di appropriato, in questa enorme quantità di dati, è più problematico di quello che strumenti di ricerca come Google fanno percepire. Il compito dei motori di ricerca, infatti, è quello di nascondere la complessità, fornendo nel minore tempo possibile risposte plausibili ai quesiti posti. La realtà è che questo prezioso patrimonio di risorse è disponibile solo in minima parte; questo per due motivi: il primo è che i motori di ricerca indicizzano e rendono disponibili solo parte delle risorse presenti in rete (Gulli e Signorini, 2005, stimano che solo il 76% dei contenuti siano indicizzati da Google); il secondo è che fare ricerche efficaci, capaci di individuare esattamente quello che si voleva trovare, non è affatto semplice. Non a caso l'*information literacy*, ovvero la capacità di accedere alle risorse attraverso le tecnologie dell'informazione e della comunicazione, è ritenuta oggi una delle "competenze chiave" da diversi organismi internazionali¹. Tali competenze risultano fondamentali per sfruttare al meglio le risorse disponibili in un ambiente come la rete dove chiunque può pubblicare ciò che crede e dove non è sempre facile disporre di parametri precisi per giudicare quanto trovato dal momento che, spesso, mancano meccanismi di controllo della veridicità e fondatezza dei contenuti.

Il presente lavoro, senza ignorare la complessità del dibattito connesso alla *digital competence*, caratterizzato dall'intreccio di riflessioni di ordine epistemologico, etico, cognitivo e pedagogico, prova a suggerire alcune risposte operative al problema del "fare ricerca efficace" in rete. Gli autori partono dalla premessa che anche persone con sufficienti "competenze tecniche", come uno studente universitario, un insegnante o un giovane ricercatore, possano incontrare numerose difficoltà. La facilità con cui si ottengono risposte dai motori di ricerca rischia, infatti, di essere una potente trappola capace di nascondere dietro alla pertinenza e alla verosimiglianza, risposte parziali se non, addirittura, scorrette o fuorvianti.

¹¹ Si pensi alla raccomandazione del Parlamento Europeo e del Consiglio dell'Unione Europea del 18 dicembre 2006, relativa a competenze chiave per l'apprendimento permanente (2006/962/CE), così come all'Information Literacy Competency Standards for Higher Education, approvati dal Association of College and Research Libraries (ACRL) nel 2000 e sostenuti, tra gli altri, dalla American Association for Higher Education e dal Council of Independent Colleges (ACRL, 2000).

Cercare nella rete

Per fare ricerche in rete si è soliti utilizzare motori di ricerca generalisti come Google, Bing o Yahoo. Oltre il 90% delle persone, secondo una ricerca del Pew Internet & American Life Project, usa questi sistemi per accedere alla rete e l'85% degli utenti sostiene di trovare le informazioni che stavano cercando (Purcell et al., 2012). I motori di ricerca, del resto, restituiscono sempre qualche risultato ed è piuttosto improbabile rimanere del tutto delusi. Ogni qualvolta il processo di ricerca si conclude in modo soddisfacente per l'utente, si crea un'esperienza positiva che si traduce in una fiducia ancora maggiore (Purcell et al., 2012). Solo comparando risultati ottenuti con ricerche diverse è del resto possibile accorgersi della differenza tra una ricerca efficace e una solo apparentemente esauriente.

Per capire come impostare una buona ricerca, così pure per valutare l'attendibilità dei risultati ottenuti è necessario conoscere come funzionino questi strumenti. I motori di ricerca suddividono le loro attività in tre diverse fasi: (i) la scansione (*crawling*) e raccolta delle informazioni dai diversi siti presenti in rete; (ii) la costruzione di indici di rilevanza delle informazioni raccolte e (iii) la risposta alle diverse ricerche fatte dagli utenti. La scansione avviene grazie a programmi software detti anche "spider" (i "ragini" della rete) incaricati di scansionare i file presenti sui diversi server per estrarne le informazioni. Ogni pagina visitata è scomposta in elementi di base, principalmente le parole che la compongono, e associata a indici di "rilevanza".

Tali indici consentiranno, a fronte delle successive ricerche effettuate, di decidere tra i milioni di pagine che contengono i termini ricercati quali siano quelle plausibilmente più importanti (rilevanti) per l'utente. Questa decisione viene presa sulla base di complessi algoritmi, diversi da motore a motore, ma che, in generale, prendono in considerazione elementi quali: il contenuto delle pagine, la loro ricchezza informativa, la quantità e la posizione dei termini nella pagina, la presenza di sinonimi, il numero di altri siti che linkano al sito, il numero di aggiornamenti che il sito ha, il numero di visite che riceve, la presenza di meta descrittori (informazioni nascoste che descrivono la pagina).

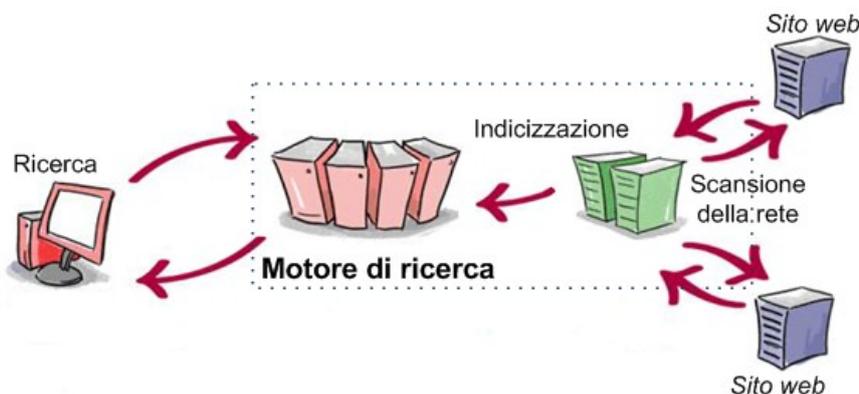


Figura 1. Funzionamento di un generico "motore di ricerca".

Sapere come funziona un motore di ricerca è indispensabile per impostare le ricerche nel modo più corretto possibile considerando anche le criticità sottese. Ognuna delle tre fasi, scansione della rete, indicizzazione e risposta alle domande di ricerca, nasconde, infatti, insidie e criticità. La fase della scansione presenta il problema della difficoltà per gli "spider" di raggiungere tutti i contenuti disponibili in rete. Bergman (2001) parlava di un "web profondo", caratterizzato da pagine dinamiche, da database e risorse private o

collocate su computer connessi attraverso i meccanismi del “peer to peer”, da contenuti protetti da copyright inaccessibili ai motori di ricerca. Tali risorse, invisibili ai motori, sono invece accessibili attraverso strumenti specializzati (si pensi a eMule o bitTorrent per la ricerca di video o file musicali). Questo fatto suggerisce l’impiego, specie nelle ricerche finalizzate all’individuazione di risposte specialistiche, di fare affidamento solo in parte sui motori di ricerca generici privilegiando, piuttosto, banche dati e portali specializzati. La seconda fase, quella dell’indicizzazione, nasconde l’insidia derivante dal fatto che gli algoritmi sviluppati per validare le fonti e contestualizzare le informazioni operano sulla base di ranking automatici istruiti per sovrappesare le risorse più popolari a discapito di quelle di nicchia, anche se scientificamente più accreditate. A meno di non usare combinazioni appropriate di termini specialistici, ricerche semplici (fatte con un solo termine) preferiscono risorse provenienti da siti a carattere divulgativo e di inquadramento generale. Non è un caso se le prime voci di ogni ricerca sono solitamente occupate da siti come Wikipedia o altri dizionari ed enciclopedie online. L’ultima fase, quella del supporto dell’utente alla ricerca nasconde, infine, l’insidia della “riduzione” della percezione della complessità del compito. I sistemi sono progettati per fornire rapidamente risposte convincenti anche a domande generiche o male impostate. Sempre più spesso, infatti, oltre ai collaudati meccanismi di attribuzione di un “peso” ai termini utilizzati rispetto alle risorse presenti nei database del motore di ricerca (ranking) – che posizionano in alto della lista risorse “plausibili” – stanno prendendo campo meccanismi di personalizzazione delle risposte basati sulla cronologia di navigazione dello specifico utente. Grazie ai cookies (informazioni memorizzate sul computer dell’utente) e ad altri sistemi, il motore di ricerca può posizionare, in cima alla lista, oltre a risposte standard, riferimenti a risorse coerenti con gli interessi dell’utilizzatore come quelle provenienti da siti precedentemente consultati. Nella Figura 2 vengono mostrate le differenze della stessa ricerca della parola “educazione” effettuata con Google su due diversi computer. Il numero dei risultati individuati è identico, ma l’ordinamento è differente e risente degli “interessi” di ricerca precedentemente manifestati dall’utente. L’insieme di queste strategie aumentano, negli utilizzatori, la percezione di affidabilità ed efficienza del motore di ricerca e, conseguentemente, fanno diminuire il loro impegno nella formulazione delle domande.

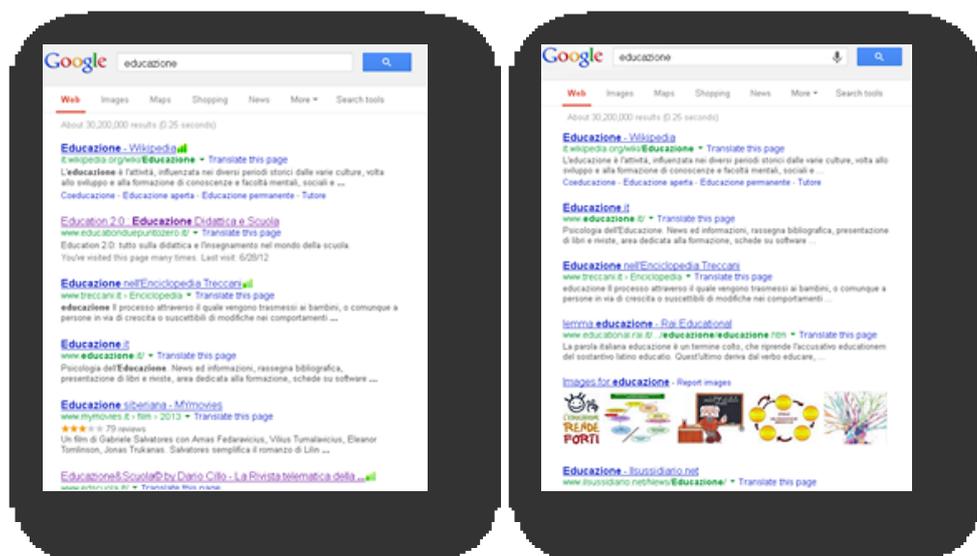


Figura 2. La stessa ricerca effettuata con il termine “educazione” su due diversi computer: il primo (a sinistra) con il sistema dei cookies attivo, il secondo (a destra) no.

Per rispondere a queste insidie è necessario adottare strategie adeguate. In particolare è necessario selezionare e combinare accuratamente i termini da utilizzare nelle ricerche e non affidarsi a un solo strumento facendo verifiche incrociate con l'impiego di strumenti diversi. L'individuazione delle parole chiave è un passaggio fondamentale perché da esso dipende la bontà stessa dell'esito della ricerca. Scegliere di utilizzare una sola parola chiave, magari generica, ambigua o non del tutto capace di rispondere al problema di ricerca porta solitamente a risultati parziali e imprecisi. Una buona ricerca richiede l'impiego di un insieme di più termini specifici digitati con alcuni accorgimenti come, ad esempio, l'uso delle virgolette alte (“ ”) che circoscrivono i risultati di ricerca alle sole risorse che contengono esattamente l'espressione tra essi racchiusa. Ad esempio, se nel campo di input di Google vengono inserite le seguenti parole chiave *evidence*, *based* ed *education*, esso fornirà come risultati tutte le risorse in cui tali termini sono contemporaneamente presenti (ricerca per termini). Se, invece, si inserisce la frase “evidence based education” racchiusa tra apici, Google limiterà i risultati alle sole risorse in cui non solo tutti e tre i termini sono presenti, ma sono presenti in questa esatta sequenza (ricerca per frase).



Figura 3. Uso delle virgolette in una ricerca con Google.

L'uso degli apici non è il solo mezzo a disposizione per affinare la ricerca. In quasi tutti i motori è possibile usare gli operatori logici booleani AND, OR, NOT (e altri comandi) per concatenare, anche attraverso l'uso di parentesi, più parole. Più precisamente l'inserimento di AND tra due termini indica al motore che si desidera ottenere risorse in cui entrambe le parole sono presenti (esso è di norma dato di default, pertanto non è necessario specificarlo). L'uso di OR tra due parole indica che si desidera ottenere risorse in cui compaiano i due termini oppure anche uno solo dei due (si tratta di un operatore che dunque espande la ricerca rispetto all'AND). L'uso di NOT prima di uno dei termini indica, invece, che si desidera limitare la ricerca a risorse in cui non compare la parola chiave da esso preceduta.

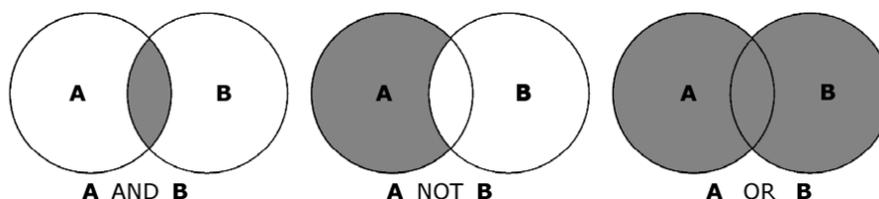


Figura 4. Aree coperte nelle ricerche con gli operatori logici.

Ogni motore di ricerca ha proprie regole sintattiche. In generale gli operatori logici possono essere usati congiuntamente utilizzando delle parentesi per comporre le richieste, come nell'esempio seguente: *autismo AND ("evidence based" OR "research based")*. Inoltre, è possibile usare il simbolo dell'asterisco "*" in sostituzione di qualsiasi carattere, ad esempio se si digita "matematic*" si otterranno tra i risultati le risorse contenenti sia "matematica" sia "matematico" sia "matematici" (una funzione che dunque può essere utile quando si vogliono includere nei risultati maschili e femminili, singolari e plurali², ma anche quando non si è sicuri dell'esatta ortografia del termine, ad esempio nel caso di parole straniere).

AND / +	per ricerche che contengano entrambi i termini
OR	per cercare almeno uno due termini digitati
NOT / -	per escludere termini e parole dalla ricerca
~	per estendere la ricerca a sinonimi (es. ~food facts riporta i risultati anche di "nutrition facts")
""	per ricercare esattamente la frase contenuta tra le virgolette
*	per cercare qualsiasi termine che inizi con la stringa data (es. "disl*" = dislessia, dislessico...)
site:	per cercare termini su un singolo sito (es. "site:uniss.it" all'interno dell'Università di Sassari)
Related:	per trovare siti simili al dominio cercato
Filetype:	digitato dopo un termine restringe a contenuti in un solo tipo di file (es. "EBE filetype: pdf")
Inurl:	per cercare i termini contenuti in un indirizzo web
Intitle:	per cercare il primo termine nel titolo di un articolo, e gli altri termini nel testo
Allintitle:	per cercare tutti i termini contenuti esclusivamente nel titolo di un articolo

Tabella 1. Principali operatori impiegabili con Google.

Ogni strumento di ricerca presenta uno specifico set di comandi che includono l'impiego di operatori logici anche se, con sempre maggiore frequenza, tali opzioni di ricerca sono esercitabili mediante interfacce grafiche³ che permettono di fare ricerche selettive senza ricorrere alla conoscenza e all'uso di questi termini. In ogni caso è importante essere consapevoli del fatto che, data la crescente massa di informazioni disponibili, il modo migliore per fare ricerche è quello di impiegare combinazioni di più parole chiave. Ma quanti e quali termini è meglio utilizzare? Qual è il giusto equilibrio tra generalità e specificità di una parola chiave? Meglio usare delle frasi o dei singoli vocaboli? È preferibile utilizzare parole italiane o inglesi? Non vi sono purtroppo delle regole assolute, valide per qualsiasi tema o problema di ricerca. La risposta a queste domande dipende dal problema di ricerca e trovare le parole chiave più efficaci (quelle che ci consentono di ottenere un sufficiente numero di risultati pertinenti) è una questione di sensibilità che si acquisisce col tempo (Trincherò, 2002). Sicuramente è essenziale conoscere la terminologia specifica dello specifico tema di ricerca. I termini più efficaci potranno essere suggeriti dall'esperienza del ricercatore, ma qualora dovessero mancare conoscenze o esperienze pregresse sullo specifico tema è sicuramente utile una lettura preliminare di testi di base o perfino di documenti non specialistici al fine di acquisire il lessico di base. Si pensi a uno studente che desideri compiere ricerche sulle strategie di sostegno a favore di studenti con disturbi specifici dell'apprendimento, senza avere una conoscenza preliminare del tema. Egli potrebbe dapprima leggere la pagina di Wikipedia sui disturbi dell'apprendimento o quella dedicata allo stesso tema nell'enciclopedia

² Si tenga presente che generalmente gli algoritmi dei motori di ricerca restituiscono varianti del termine cercato, quali appunto varianti di genere e numero, di default, senza la necessità di utilizzare l'asterisco.

³ Si veda ad esempio la pagina di ricerca avanzata di Google http://www.google.it/advanced_search?hl=it.

Treccani online. Attraverso tali letture è facile comprendere che sono parole o espressioni chiave efficaci termini quali: dislessia, disortografia, disgrafia, discalculia come l'acronimo DSA. Dalla consultazione di una pagina di un'enciclopedia online si può avere la fortuna di accedere direttamente ai PDF di uno o più saggi online presenti in rete. La loro lettura permette di definire ulteriormente il quadro terminologico anche perché, la gran parte degli articoli pubblicati sulle riviste scientifiche riportano, nella prima pagina, proprio le parole chiave in grado di descrivere il tema trattato.

Un'ulteriore fonte, per la ricerca di parole chiave, è rappresentata dai thesauri specificamente sviluppati nel dominio delle scienze dell'educazione. Un thesaurus, secondo la definizione fornita dall'International Organization for Standardization (ISO, 1993), è un "vocabolario di un linguaggio d'indicizzazione controllato, organizzato in maniera formale, in modo cioè da rendere esplicite le relazioni a priori fra i concetti". Si tratta, in altre parole, di un vocabolario controllato e strutturato di termini propri di un certo dominio, tra cui sussistono relazioni semantiche e gerarchiche (Spinelli, 2005). In ambito educativo, ne sono stati sviluppati diversi, si veda in questo proposito la tabella seguente.

Catalogo	Indirizzo internet	Contenuto
TESE	eacea.ec.europa.eu/education/eurydice/tese_en.php	Thesaurus europeo dei sistemi educativi
ETB Thesaurus	lreforschools.eun.org/web/guest/lre-thesaurus	Learning Resource Exchange Thesaurus curato dal Vocabulary Bank for Education (VBE) e disponibile in dodici lingue
Thesaurus of ERIC	www.eric.ed.gov/ERICWebPortal/thesaurus/thesaurus.jsp	Thesaurus curato dall'Education Resources Information Center

Tabella 2. Principali dizionari controllati in ambito educativo.

Una volta selezionato un primo insieme di parole chiave (che ovviamente potranno essere riformulate e affinate in ogni fase del processo di ricerca), si presenta l'esigenza di individuare lo strumento di ricerca più adeguato: motore di ricerca, banche dati, OPAC, e così via. Non si tratta di strumenti da intendersi come alternativi; al contrario sarebbe preferibile utilizzarli in maniera complementare, affinché ciascuno di essi possa risultare utile a fornire riferimenti che completano quelli dell'altro.

Gli strumenti oltre "Google"

Google, come abbiamo anticipato, non sempre rappresenta lo strumento migliore per recuperare informazioni specialistiche come quelle provenienti dalla letteratura scientifica. Il sapere scientifico, come noto, è veicolato prevalentemente da pubblicazioni specializzate ed è per molti aspetti più facilmente raggiungibile attraverso strumenti che indicizzano selettivamente questo tipo di risorse. I motori di ricerca generici possono non riuscire a estrarre tutte le informazioni presenti negli archivi ad accesso riservato e, comunque, a causa degli algoritmi di indicizzazione pensati per valorizzare le risorse più popolari, possono considerate meno rilevanti, e quindi presentare in posizioni arretrate, le informazioni di valore scientifico.

Operare ricerche selettive partendo dall'interrogazione di cataloghi e banche dati rappresenta un modo per aggirare il problema circoscrivendo l'ambito d'azione. Questi strumenti raccolgono informazioni provenienti da fonti specializzate quali libri, articoli di

rivista, report, atti di convegni, tesi, relazioni e altre tipologie e, indipendentemente dalla loro natura (risorse digitali o fisiche), ne descrivono il contenuto attraverso campi (descrittori) come autore, titolo, anno, abstract, ecc. e, laddove disponibile, forniscono l'accesso alla lettura integrale delle risorse indicizzate. Gli archivi e le banche dati sono molteplici: ci sono quelle promosse da organismi governativi o emanazione di enti di ricerca, quelle private – tipicamente legate a gruppi editoriali -, quelle settoriali, ad esempio promosse da enti ed associazioni di categoria attivi nei diversi settori scientifici (medicina, ingegneria, scienze umane). Alcune banche dati sono ad accesso aperto, altre richiedono il pagamento di un canone o possono essere usate gratuitamente solo all'interno di biblioteche o reti universitarie. Le opere librarie, invece, sono tipicamente gestite dalle singole biblioteche o dai servizi di catalogazione centralizzati (in Italia SBN, Sistema Bibliotecario Nazionale gestito da ICCU Istituto Centrale per il Catalogo Unico).

L'interrogazione dei cataloghi avviene attraverso OPAC, dall'acronimo On-line Public Access Catalog, strumenti che permettono la ricerca sui diversi campi e l'impiego delle parole chiave mediante operatori booleani e filtri in grado di limitare la ricerca a determinati anni, al tipo di pubblicazione, ecc.. Banche dati e cataloghi consentono di ottenere informazioni di sintesi (es. le informazioni bibliografiche o il sunto dell'opera) e, sempre più spesso, di accedere al full-text ossia alla visualizzazione di quelle risorse disponibili in formato elettronico. L'accesso ai documenti integrali, solitamente PDF, può richiedere il pagamento del costo della versione elettronica per quelle risorse che non sono gratuite. Molte biblioteche universitarie avendo sottoscritto abbonamenti a pacchetto possono offrire ai loro utenti la possibilità di accedere gratuitamente anche a contenuti protetti purché le ricerche e la lettura avvengano all'interno dei computer interni alla loro rete.

La ricerca scientifica può essere svolta – oltre che attraverso le banche dati vere e proprie, anche attraverso servizi ibridi come gli aggregatori di risorse, ovvero sistemi che si preoccupano di indicizzare e aggregare risorse provenienti da fonti diverse, ma anche attraverso strumenti come quelli messi a punto da editori, da distributori e perfino librerie online. Le tabelle che seguono offrono una panoramica di sintesi delle diverse tipologie di strumenti impiegabili per la ricerca nell'ambito delle scienze dell'educazione.

Strumento	Indirizzo internet	Contenuto
AMAZON	www.amazon.it www.amazon.com	Il sito internet di Amazon, come quello dei venditori online IBS, BOL, ecc., consente di localizzare libri, leggerne la recensione, verificarne opere correlate e, attraverso l'opzione "click to look inside", sfogliarne alcune parti (sito americano).
AZALAI	www.aib.it/aib/opac/mai.htm	MetaOpac Italiano capace di interrogare simultaneamente i cataloghi di più biblioteche italiane.
Google Libri	books.google.it	Motore di ricerca di Google specializzato nell'indicizzazione di libri. Oltre alla ricerca bibliografica consente, in molti casi, la lettura di parti dei testi.
SBN	opac.sbn.it	Catalogo italiano di opere librarie
Library of Congress	www.loc.gov/homepage	Catalogo americano di opere librarie

Tabella 3. Siti per la ricerca online di libri.

Strumento	Indirizzo internet	Contenuto
DOAJ	www.doaj.org	Motore di ricerca per l'accesso alle pubblicazioni scientifiche ad "accesso aperto", ovvero redatte nello spirito della libera condivisione dei contenuti.
ED/ITLIB	www.editlib.org	Archivio di risorse digitali nel settore dell'educazione e dell'ICT.
ERIC - Education Resources Information Center	www.eric.ed.gov	Principale banca dati sulla ricerca educativa internazionale in lingua inglese (curata da Dipartimento educazione USA)
Google Scholar	scholar.google.it	Motore di ricerca di Google specializzato nell'indicizzazione di saggi e articoli scientifici. Consente di individuare informazioni bibliografiche e, quando disponibili, i documenti elettronici.
JSTOR	www.jstor.org	Banca dati bibliografica contenente spogli e articoli a testo intero di periodici a partire dalle edizioni del secolo scorso.
PsycLine	www.psychline.org	Sistema per localizzare contributi di ricerca nel campo della psicologia e delle scienze sociali.
EMERALD	www.emeraldinsight.com	Cataloghi commerciali che permettono, attraverso l'aggregazione di più fonti, l'accesso ai contenuti di migliaia di pubblicazioni.
Ingenta Connect	www.ingentaconnect.com	
Scirus	www.scirus.com	
WorldCat	www.worldcat.org	

Tabella 4. Siti per la ricerca online di articoli su periodici.

Strumento	Indirizzo internet	Contenuto
SAGE	online.sagepub.com	Portali di accesso alle riviste scientifiche dei principali editori. Ogni editore fornisce strumenti molto sofisticati di accesso ai contenuti. Alcuni editori (vedi in particolare Scencedirect della Elsevier) indicizzano anche riviste diverse dalle proprie.
Taylor & Francis	www.tandfonline.com	
Elsevier	www.sciencedirect.com	
Springer	link.springer.com	
John Wiley & Sons	onlinelibrary.wiley.com	

Tabella 5. Siti per la ricerca online sui cataloghi dei principali editori di periodici scientifici

Strumento	Indirizzo internet	Contenuto
Mendeley	www.mendeley.com	I "reference manager software" sono strumenti utilizzabili online (o offline) che aiutano i ricercatori ad organizzare le risorse individuate (bibliografie, file, link). Grazie alla condivisione in rete delle raccolte personali, rappresentano un ulteriore modo per fare ricerca
EndNote	endnote.com	
Papers	www.papersapp.com/papers	
RefWorks	www.refworks.com	
Zotero	www.zotero.org	

Tabella 6. Strumenti per la catalogazione personale delle risorse bibliografiche (*reference manager software*). Rappresentano un nuovo modo per accedere e condividere le conoscenze scientifiche.

L'uso di strumenti di ricerca specializzati consente di approfondire ed estendere lo studio grazie alla disponibilità di innovativi strumenti di navigazione e di esplorazione

dell'ampio reticolo bibliografico (esempio citazioni di citazioni), delle opere correlate ovvero che trattano lo stesso tema, delle risorse consultate ed apprezzate da altri ricercatori che avevano esigenze di ricerca simili. Il “supporto sociale” alla ricerca, fornito da molti strumenti, consente – nello spirito del web 2.0 – di accedere rapidamente all'individuazione di ampi repertori bibliografici e di entrare in contatto con ricercatori e studiosi che si stanno occupando delle stesse problematiche (si vedano, in particolare, gli strumenti), rappresenta infatti uno degli ausili più interessanti forniti da molti strumenti per la catalogazione personale delle risorse bibliografiche (tabella 6). I motori di ricerca interni a strumenti come Mendeley e Zotero, basandosi su citazioni inserite dagli utenti, non si limitano ad applicare algoritmi statici, ma organizzando le risorse per aree di interesse, per “tag” e reti di significato, consentono di coadiuvare i processi di costruzione sociale delle conoscenze.

Google non è però estranea ai problemi della ricerca scientifica. Non a caso due dei principali prodotti utilizzabili in questo senso sono proprio dell'azienda produttrice dell'omonimo motore di ricerca: Google Scholar e Google Book. Google Scholar è un motore di ricerca specializzato nella indicizzazione e recupero di risorse scientifiche. L'utilizzo di Google Scholar è del tutto simile a quello di Google: nella ricerca semplice è sufficiente inserire le parole chiave relative alla propria domanda di ricerca nel campo di input. Inoltre, è possibile in esso sfruttare i diversi operatori e prefissi precedentemente citati. Oltre a ciò, Google Scholar offre una serie di funzionalità avanzate che possono essere molto utili per le ricerche bibliografiche. Si possono, ad esempio, compiere delle ricerche per autore, per data e per tipo di pubblicazione. La differenza rispetto al più noto motore di ricerca è che sono indicizzati solo documenti di ricerca (saggi, articoli, atti di convegno ecc.) ignorando tutto il resto (prodotti, servizi, discussioni libere, esperienze varie). I risultati sono quindi molto meno e, soprattutto, più focalizzati. Come mostrato nella Figura 5, la stessa ricerca effettuata con le parole chiave “clickers and classroom” produce in Google Scholar solo 5.970 risultati contro i 1.380.000 di Google. Per ogni opera individuata sono inoltre mostrate informazioni aggiuntive, come il numero di citazioni che il lavoro ha totalizzato (la sua rilevanza), la possibilità di accedere ad opere ad esso connesse, l'estrazione della citazione da inserire in bibliografia e l'accesso al file (se disponibile).



Figura 5. Interfaccia di presentazione dei risultati di Google Scholar.

Impostare la ricerca

La ricerca, in educazione come in qualsiasi altro settore scientifico, non è però solo un problema di strumenti. Paradossalmente è proprio dalla disponibilità di strumenti performanti che può sorgere l'idea che fare ricerca sia semplice. Il fatto che da ogni domanda emergano decine di risposte interessanti e pertinenti porta a dimenticare una delle massime più note dell'information technology: “*garbage in, garbage out*”, ovvero se ad un elaboratore si forniscono “informazioni spazzatura” questi, a sua volta, non può che elaborare e restituire un risultato insensato. In altre parole la domanda di ricerca deve essere chiara, non generica o confusa, ed essere formulata adeguatamente. Inquadrare con precisione il problema di ricerca porta a individuare i termini della questione e, conseguentemente, i termini, o parole chiave, attraverso i quali fare ricerca. Eisenberg et al. (2000), parlando di strategie di ricerca, identificano i seguenti passaggi: (i) definizione del compito; (ii) messa a punto di strategie di ricerca; (iii) localizzazione e accesso alle informazioni; (iv) sintesi dei contenuti in un prodotto finito; (v) valutazione del prodotto finale e del processo. Passaggi che, in rete, possono diventare le seguenti fasi di un processo che può assumere carattere iterativo:

1. definire il problema di ricerca;
2. scegliere le parole chiave da utilizzare;
3. effettuare la ricerca di risorse pertinenti il problema definito;
4. valutare e selezionare le risorse trovate.

Si può parlare di processo iterativo perché, tipicamente, una buona ricerca è sempre caratterizzata da un progressivo affinamento. Lo stadio finale, quello dalla valutazione e selezione delle risposte, apre nuovi orizzonti, può estendere o ridefinire il campo d'azione, aiuta il ricercatore a definire meglio la domanda. I problemi maggiori li hanno, naturalmente, gli utenti inesperti che si avvicinano per la prima volta alla questione da indagare. Studenti universitari, giovani ricercatori o chiunque si avvicini a indagare un problema inedito, può non avere un'idea precisa di quali siano i temi portanti, gli autori più accreditati, gli elementi principali del dibattito in corso. Per iniziare una qualunque ricerca è essenziale partire da una conoscenza almeno di base del tema cui il problema fa riferimento. Saranno necessarie varie letture, all'inizio anche generiche o a carattere divulgativo, prima di riuscire a comprendere quali siano gli aspetti principali e quelli che meritino di essere ulteriormente indagati. L'individuazione stessa delle parole chiave realmente efficaci per la ricerca di informazioni in rete, che è un passaggio fondamentale per l'affidabilità dei risultati di ricerca, richiede un attento lavoro di analisi. Come abbiamo visto l'impiego di una sola parola chiave o di termini generici, ambigui e non pertinenti col problema di ricerca, conduce a risultati non ottimali: una eccessiva quantità di risultati con un basso tasso di precisione, il rischio di ottenere pochi o nessun risultato o, ancora, l'ottenere risultati falsati dalla ambiguità del vocabolario (ad esempio, in presenza di casi di sinonimia o di polisemia) e così via.

Alcuni suggerimenti su come impostare una ricerca efficace vengono dal filone di ricerca dell'*evidence-based education*. Con l'aumentare dell'interesse, a livello internazionale, per la ricerca di risposte oggettive alle questioni conoscitive più rilevanti nei vari ambiti, dall'agricoltura alla medicina ed anche all'educazione, sono state istituite, solitamente a livello governativo, centri specializzati nello svolgimento di ricerche capaci di indirizzare sia le politiche di finanziamento, sia la pratica degli operatori. In altre parole ci sono organismi che operano per effettuare ricerche sulle tematiche di maggiore interesse al fine di dare risposte, tipicamente di sintesi a quanti – insegnanti, educatori, dirigenti – si trovano a dover operare nella pratica e non sempre hanno il tempo o la voglia di dedicarsi

a questo tipo di attività. Centri come il “What Works Clearinghouse” negli Stati Uniti o l’EPPI-Centre (Evidence for Policy and Practice Information and Co-ordinating Centre) nel Regno Unito rispondono a questa esigenza. I ricercatori che operano all’interno di queste organizzazioni, a partire da problemi conoscitivi, producono delle analisi della letteratura esistente nella forma di systematic reviews, scoping reviews, and meta-analyses. Ognuno di questi approcci porta questi operatori a seguire dei protocolli standardizzati attraverso i quali svolgere le ricerche finalizzate a selezionare le risposte affidabili e autorevoli relativamente alle diverse domande. La prima fase di una ricerca richiede sempre di concentrarsi sul formulare correttamente la domanda relativa all’esigenza conoscitiva (Richardson et al, 1995; Rosenberg & Donald 1995, Sackett & Rosenberg, 1995) in maniera da poter intervenire con azioni verosimilmente “migliori di altre” all’interno di un determinato dominio applicativo. In ambito sanitario si è dato grande attenzione nella formulazione di domande meritevoli di attenzione (Booth 2000; Barrie & Ward, 1997; Ellis et al, 2000; Ely et al, 1999; Flemming, 1998; Gorman & Helfand, 1997; Swinglehurst & Pierce, 2000) cercando di comprendere quale sia la struttura di una domanda “mirata” all’ottenimento di una risposta possibile (Richardson, 1998; Geddes, 1999) distinguendo nella domanda gli aspetti “in primo piano” da quelli di “sfondo” (Richardson & Wilson, 1997) perché, come anticipato, una formulazione efficace conduce ad una ricerca delle “evidenze” desiderate (Eldredge, 2000b, Snowball, 1997; Villanueva et al, 2001) mentre, al contrario, domande confuse portano solo a risposte confuse (Oxman e Guyatt, 1988). L’EPPI Centre, uno degli organismi più accreditati nello sviluppo di ricerche sistematiche attorno a problemi complessi del mondo dell’educazione ha elaborato dei protocolli di lavoro capaci di accompagnare il lavoro dei ricercatori. La fase iniziale del lavoro è finalizzata all’identificazione della domanda di ricerca e delle parole chiave capaci di intercettarla, degli strumenti da usare (motori di ricerca, database, archivi), delle tipologie di documenti che saranno presi in considerazione (ad esempio quelli redatti in inglese, quelli scritti in un certo intervallo di tempo, quelli che riguardano una certa popolazione o che adottano un certo tipo di metodologia), dei criteri di inclusione ed esclusione dei lavori che verranno selezionati (EPPI, 2007). Ognuna di queste attività è accuratamente formalizzata prima dell’avvio della ricerca perché il tempo e lo sforzo speso in questa fase porta a risparmiare tempo e fatica in seguito.

Nello spirito di dare un supporto concreto ai propri studenti, alcune università forniscono strumenti – spesso nella forma di griglie di lavoro – che, ispirandosi ai principi delle ricerche svolte dai professionisti, ambiscono a guidare il processo. Quella riportata nella Figura 5 è un esempio di utilizzo di una di queste schede messa a punto dalla canadese Thompson Rivers University. Come si può notare i passaggi sono, in forma sintetizzata, quelli previsti da organismi come l’EPPI-Centre: formulazione di una chiara domanda di ricerca, individuazione delle parole chiave, estensione delle parole chiave attraverso l’identificazione di sinonimi, la combinazione delle parole chiave (e dei sinonimi) in stringhe di ricerca mediante operatori logici, l’individuazione dei database su cui effettuare le ricerche, la definizione dei criteri di revisione e rivisitazione dei risultati che saranno poi ottenuti.

THOMPSON RIVERS UNIVERSITY TRU Library's Developing Effective Search Strategies

Steps	Search Strategy					
Write a clear topic statement:	<i>Discuss the prevalence and use of marijuana by Canadian teenagers</i>					
Select the keywords from your topic statement. <i>Usually 2 or 3 keywords give better results.</i>	Keyword A	<i>marijuana</i>				
	Keyword B	<i>teenagers</i>				
	Keyword C	<i>Canadian</i>				
Based on the keywords you identified above, select synonyms, related terms, and alternate forms for each of these keywords.	Keyword A Synonyms	<i>teens, adolescents, adolescence, youth, young people</i>				
	Keyword B Synonyms	<i>cannabis, pot, weed, reefer, dope, doobie, sticky icky...</i>				
	Keyword C Synonyms	<i>Canada, Canadians, British Columbia, Alberta, etc...</i>				
Search Strategy 1: Formulate a search strategy using Boolean operators ('AND' / 'OR') to connect your keywords. • Combine synonyms with OR . • Combine different ideas/concepts with AND .	Keyword A Synonyms	<i>teens</i>	OR	<i>teenagers</i>	OR	<i>adolescents</i>
	AND					
	Keyword B Synonyms	<i>marijuana</i>	OR	<i>cannabis</i>	OR	<i>pot</i>
	AND					
Search Strategy 2: Formulate a search strategy using Boolean operators ('AND' / 'OR') along with your keywords. • Combine synonyms with OR . • Combine different ideas/concepts with AND .	Keyword C Synonyms	<i>Canada</i>	OR	<i>Canadians</i>	OR	<i>Alberta</i>
	AND					
	Keyword A Synonyms	<i>adolescents</i>	OR	<i>young people</i>	OR	<i>youth</i>
	AND					
Select appropriate search tools (i.e. library catalogue, an article database, etc.).	Keyword B Synonyms	<i>marijuana</i>	OR	<i>pot</i>	OR	<i>weed</i>
	AND					
	Keyword C Synonyms	<i>British Columbia</i>	OR	<i>Canada</i>	OR	<i>Ontario</i>
Select appropriate search tools (i.e. library catalogue, an article database, etc.).	<i>TRU library catalogue</i>		<i>article database: CBCA Current Events</i>			
	<i>article database: Academic Search Complete</i>		<i>Google??</i>			
Search	<i>Take your search strategies from above and insert them in the databases you selected.</i>					
Review results and revise search.	<i>Not enough results? Revise your search terms and/or remove one of your keywords to broaden your results.</i> <i>Too many results? Use limits (scholarly journals, or by publishing year, etc...) and/or add another keyword to narrow your results.</i>					

This is what **search strategy #1** will look like in an article database, such as Academic Search Complete.

Searching: **Academic Search Complete** | [Choose Databases >](#)

 in

in

in [Add Row](#)

Figura 6. Esempio di utilizzo di una scheda di supporto alla ricerca (http://www.tru.ca/library/pdf/developing_effective_search_strategies.pdf).

Prima di concludere questo contributo è opportuno sottolineare come uno degli aspetti a prima vista meno evidenti, ma su cui maggiormente insiste il dibattito di chi si occupa di definizione dei criteri per l'individuazione di valide risposte alle domande conoscitive è relativo ai criteri di selezione delle risorse individuate. La questione riguarda aspetti formali, come la verifica della "fonte" (ad esempio il fatto che il contributo sia stato pubblicato su una rivista qualificata e sia passato al vaglio di un processo di referaggio), ad aspetti connessi a problemi epistemologici (come il fatto che il lavoro espliciti chiaramente le assunzioni teoriche e valoriali da cui muove e definisca chiaramente la domanda di ricerca), ad aspetti metodologici quali la numerosità del campione su cui si basano le indagini, gli strumenti impiegati (un questionario sulla percezione non è la

stessa cosa di una misurazione dei risultati di una performance), il luogo e il periodo di svolgimento, le procedure impiegate per raccogliere i dati e molte altre cose. Il dibattito in atto, a livello internazionale, sulle “evidenze” in educazione muove proprio dalla constatazione che troppo spesso, in questo settore, si fanno affermazioni – ad esempio sull’efficacia di una tecnologia o di un metodo di insegnamento – senza basarli su dati, ma solo su argomentazioni logiche o, se si espongono dei dati, questi non sono magari adeguati o sufficienti a dimostrare quanto si vorrebbe affermare.

Conclusioni

Riuscire a realizzare ricerche affidabili attingendo all’enorme patrimonio di risorse disponibili in rete è oggi un compito al contempo facile e problematico. La facilità con cui è possibile ottenere risposte ai più disparati problemi di ricerca può condurre a sottovalutare la complessità del compito. In questo contributo abbiamo cercato di indicare alcune possibilità che la rete offre per estendere e migliorare le modalità di ricerca.

Riconoscimenti

Il presente contributo è stato concepito collegialmente. In ogni caso i paragrafi Introduzione, Gli strumenti oltre “Google”, Conclusioni sono attribuibili a Giovanni Bonaiuti, mentre l’abstract e i paragrafi Cercare nella rete e Impostare la ricerca a Giuliano Vivanet. Il contributo di Giuliano Vivanet alla presente pubblicazione è stato prodotto durante l’attività di ricerca finanziata con le risorse del P.O.R. SARDEGNA F.S.E. 2007-2013 - Obiettivo competitività regionale e occupazione, Asse IV Capitale umano, Linea di Attività 1.3.1 “Avviso di chiamata per il finanziamento di Assegni di Ricerca”.

Bibliografia

- ACRL (2000). *Information Literacy Competency Standards for Higher Education*. <http://www.ala.org/acrl/standards/informationliteracycompetency> (ver 29.07.13).
- Ary D., Jacobs L. C., Sorensen C. (2009). *Introduction to research in education* (8th ed.). Belmont, CA: Wadsworth.
- Bergman, M. K. (2001). White Paper: The Deep Web: Surfacing Hidden Value. *The Journal of Electronic Publishing*, 7(1). doi:10.3998/3336451.0007.104.
- Booth A. (2004). Formulating answerable questions. In Booth, A., Brice, A. (Eds) *Evidence Based Practice for Information Professionals: A handbook*. London: Facet Publishing, pp. 61-70.
- EPPI (2007). EPPI-Centre Methods for Conducting Systematic Reviews. London. <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=hQBu8y4uVwI=&tabid=88&mid=6162>
- Gough D., Oliver S., Thomas J. (2012). *An Introduction to Systematic Reviews*. London: Sage.

- Gulli, A., & Signorini, A. (2005). The indexable web is more than 11.5 billion pages. *Proceeding of WWW 2005*, May 10–14, 2005, Chiba, Japan. (pp. 5–6). Chiba, Japan: ACM. <http://dl.acm.org/citation.cfm?id=1062789> (ver. 29.07.13).
- Hilbert, M., & López, P. (2011). The World's Technological Capacity. *Science*, 332 (April), pp. 60–65.
- Purcell, K., Brenner, J., Rainie, L. (2012). Search Engine Use 2012. *Pew Internet & American Life project*. <http://www.pewinternet.org/Reports/2012/Search-Engine-Use-2012.aspx> (ver. 29.07.13).
- Richardson, W.S., Wilson, M.C., Nishikawa, J., Hayward, R. S. (1995). The well-built clinical question: a key to evidence based decisions. *ACP Journal Club*, 123 (3), A12-A13.
- Spinelli, S. (2005). *Introduzione ai thesauri*. <http://www.dfl.univr.it/documenti/OccorrenzaIns/matdid/matdid852909.pdf> (ver. 29.07.13).
- Trincherò, R. (2002). *Manuale di ricerca educativa*. Milano: Franco Angeli.