

APPUNTI SULLE TECNICHE DI RICERCA ONLINE di Luca Morisi

L'utilizzo del Web per il reperimento di dati e informazioni a supporto delle attività accademiche di studio e di ricerca e, in generale, a supporto delle attività intellettuali, può utilmente snodarsi lungo alcune fasi ordinate al fine di risultare il più possibile efficace e pertinente:

1. preparare la ricerca;
2. effettuare la ricerca con logica adeguata;
3. selezionare e valutare i risultati della ricerca;
4. organizzare e archiviare i risultati della ricerca.

L'acronimo "P.E.S.O." può aiutare a memorizzare i quattro punti indicati, che sotto vengono esaminati nel dettaglio.

1. preparare la ricerca

È opportuno pensare e seguire percorsi strutturati preliminarmente, ponendosi obiettivi precisi, al fine di contrastare l'inevitabile carattere "magmatico" del Web, con le continue distrazioni derivanti dalla struttura ipertestuale e dalle suggestioni multimediali: se è senza dubbio un vantaggio navigare liberamente da un riferimento ad un altro, arricchendo così la propria ricerca di spunti e suggerimenti inattesi, appare altrettanto conveniente fissare una traccia preparatoria di ricerca che aiuti a non deviare rispetto agli obiettivi.

In questo senso, al di là della necessaria chiarezza mentale di partenza, appare utile fissare su carta uno schema generale del cammino di ricerca prefigurato: rispetto alle controparti elettroniche, infatti, il foglio aiuta a focalizzare meglio l'attenzione, induce ad una maggiore riflessività, a ponderare e organizzare i pensieri in modo più accurato. Partendo da una buona consapevolezza delle proprie conoscenze si potrà evitare il senso di frustrazione da *overload* informativo che spesso colpisce l'esecutore della ricerca, turbato dalla rapidità e dalla immediatezza con cui si può accedere ad un oceano di dati dalle dimensioni sterminate.

Mantenere la bussola durante la navigazione, facendo costante riferimento all'imprescindibile patrimonio di fonti, conoscenze ed esperienze *offline*, spingerà poi ad una maggiore scrupolosità nella valutazione dell'autorevolezza e dell'attendibilità delle risorse informative trovate sul Web, il quale - allo stato di evoluzione attuale della cultura digitale - rappresenta un *medium* di archiviazione e comunicazione non autosufficiente dal punto di vista contenutistico, benché fornito di eclatanti virtù operative e clamorosi vantaggi pratici rispetto ai *media* precedenti.

2. effettuare la ricerca con logica adeguata

- scelta dello strumento di ricerca:

- motori di ricerca (es. Google, Yahoo!, Microsoft Bing (ex MSN, ex Live): rappresentano la via maestra per ogni ricerca libera sul Web, in grado di coniugare la potenza dei linguaggi di interrogazione informatici (dal semplice uso di parole chiave all'uso di operatori *booleani*, filtri e altre opzioni avanzate) con la possibilità di estendere la ricerca all'intero universo di contenuti online (testi in pagine Web e file, grafica, video ed altri elementi mediatici), organizzato in indici aggiornati in tempo reale da appositi software *spider* che scandagliano incessantemente la rete;
- directory (o indici) (es. Yahoo! Directory [non più attiva in lingua italiana dalla metà del 2010, permane invece in inglese], Google/DMOZ-Open Directory, ecc.): archivi di link Web strutturati attraverso categorie e suddivisioni gerarchiche più o meno complesse, compilati da redazioni umane; per i progetti di directory più autorevoli, l'intervento della mediazione umana dovrebbe garantire risorse di maggiore qualità, pertinenza e autorevolezza, rispetto a quelle liberamente rintracciabili attraverso un motore di ricerca (i cui archivi sono generati da software, pur se sempre più potenti e caratterizzati da tecniche di intelligenza artificiale); a causa della sempre maggiore efficienza dei motori di ricerca, le directory hanno progressivamente perso di popolarità e utilità, soprattutto nella loro forma 'generalista' (mentre diverso è il discorso 'settoriale', soprattutto in ambito molto specialistici, dove l'intervento "umano" può ancora fare la differenza);

- newsgroup (o gruppi di discussione / forum): i gruppi Usenet "storici", newsgroup per antonomasia (nati nei primi anni '80, oggi si contano nell'ordine di svariate migliaia), sono consultabili nella sezione "Gruppi" di Google; vertono sugli argomenti più vari e costituiscono la soluzione ideale nel caso in cui si intenda discutere con persone "reali", scambiando esperienze e commenti - anche su tematiche molto specifiche - difficilmente pubblicati sul Web e, dunque, difficili da trovare attraverso motori di ricerca e directory. I gruppi di discussione funzionano in modalità "asincrona", consentendo ai partecipanti di comunicare anche senza la contemporanea presenza su Internet: il modello è quello di una "bacheca" digitale in cui affiggere messaggi nuovi e rispondere a messaggi già affissi;
- database (o banche dati): benché in linea teorica ogni fonte di dati strutturata (e dunque anche gli indici di motori di ricerca e directory) possa essere considerata un database, con questo termine ci si riferisce qui a particolari archivi online di "dati ordinati" quali dizionari, cataloghi di biblioteche (OPAC), archivi di testi digitali (es. Google Books, Gallica, Europeana) enciclopedie in formato elettronico (es. Wikipedia), a volte nati da corrispondenti istituzioni tradizionali, che da un lato consentono di effettuare ricerche e reperire contenuti digitalizzati (es. per scaricare testi e immagini o ascoltare la pronuncia di un termine straniero), dall'altro di orientarsi meglio nel mondo delle strutture "fisiche" (es. per conoscere la collocazione di un libro nel mondo delle biblioteche reali).

Motori di ricerca, directory e archivi di varia natura possono essere "generalisti", come implicitamente si sono intesi sopra, ma anche "settoriali": ad esempio, motori di ricerca che filtrano automaticamente i propri database, circoscrivendoli a determinati argomenti, oppure directory specialistiche dove vengono incluse risorse Web appartenenti esclusivamente ad un certo ambito.

Esiste, inoltre, una grande quantità di forum di discussione, gestiti attraverso siti Web e paragonabili ai newsgroup "canonici", che stimolano la creazione di "comunità virtuali" dedite a interessi comuni. In questa direzione vanno anche le newsletter, ovvero liste di distribuzione via email che consentono di ricevere periodicamente informazioni aggiornate di varia natura, i cosiddetti blog, sorta di "diari" di riflessioni personali, spesso aperti anche al commento da parte dei visitatori (e in questo senso simili ai forum) e, in generale, tutti quei servizi definiti comunemente come siti o piattaforme di social-networking (Facebook, Twitter, Flickr, YouTube, ecc.). In questo ambito un metodo particolare di comunicazione è quello della chat (o messaggistica immediata), attraverso stanze pubbliche (come il sistema IRC) o canali privati di instant messaging e teleconferenza audio-video (come ICQ, Microsoft Messenger, Yahoo! Messenger, Skype e molti altri) che consentono un dialogo diretto e sincrono tra uno o più utenti: al di là dei molti usi ludici, questi strumenti possono essere funzionali agli scambi intellettuali e, dunque, all'attività di ricerca filosofica, senza barriere geografiche e con costi irrisori.

L'ambito dei contenuti generati dagli utenti viene spesso definito come UGC (User-Generated Content).

- scelta delle parole chiave e formulazione della ricerca (ciò vale soprattutto per i motori):

- scegliere termini pertinenti, almeno in numero di due (ai quali, di norma, i motori di ricerca applicano l'operatore logico "AND", recuperando così solo i documenti che contengono *entrambi* i termini); evitare termini singoli o formulazioni troppo generiche in modo da circoscrivere il più possibile i risultati all'ambito desiderato;
- avvalersi delle funzionalità di ricerca avanzata, in particolare della ricerca per frase esatta (delimitando il testo con doppie virgolette, ""), della ricerca di pagine in una specifica lingua e della ricerca di file (es. .pdf, .doc, ecc.) per trovare documenti più organici e strutturati di una semplice pagina Web;
- usare creatività e immaginazione: per trovare contenuti pertinenti è utile immaginare quali frasi potrebbero essere presenti nel tipo di documento che si cerca, affiancando più parole chiave e/o frasi esatte (es. se si cerca un documento che parli delle differenze tra il pensiero di Platone e Aristotele, potrebbe essere utile cercare: "*tra Platone e Aristotele*" *differenze*; oppure "*tra Aristotele e Platone*"; oppure "*between Plato and Aristotle*" (anziché semplicemente: *Aristotele Platone*);
- scegliere la lingua più "vicina" all'ambito culturale della ricerca da effettuare;
- nella consultazione di directory, è utile effettuare ricerche con parole chiave generiche (es. "filosofia") per trovare subito le categorie più pertinenti dell'indice (es. in DMOZ, "World: Italiano: Società: Filosofia") e procedere così ad una navigazione diretta tra le varie risorse recensite dalla redazione. La stessa tecnica è utile nei motori di ricerca per trovare immediatamente un sito Web riguardante un soggetto definito e preciso (es. un ateneo, una

biblioteca, una casa editrice, un istituto culturale, un'azienda, ecc.) del quale non si conosce l'URL ma che per certo si sa essere presente in rete. In questi casi si parla di query o interrogazioni "navigazionali", mirate cioè a rapido rintracciamento di uno specifico indirizzo.

3. selezionare e valutare i risultati della ricerca

- prima selezione dei risultati e perfezionamento della ricerca:
 - o imparare a riconoscere "a colpo d'occhio" i risultati più appropriati e ad escludere quelli assolutamente non pertinenti, evitando così di dover aprire la pagina Web o il file correlato; a tal scopo soffermarsi in particolare sui titoli delle pagine trovate, sugli URL (indirizzi Web) e sugli "estratti" contestuali che - ad esempio in Google - vengono riportati sotto il titolo delle pagine trovate e mostrano il contesto in cui compaiono le parole chiave ricercate;
 - o utilizzare i risultati della ricerca per perfezionarla, ad esempio rilevando tra i risultati di un motore di ricerca - ma anche nelle descrizioni delle directory o nei testi consultati attraverso newsgroup e altri strumenti -, termini chiave ricorrenti che non si erano presi in considerazione e ripetere la ricerca provando ad affiancarli ai termini originari;
- valutazione dei contenuti di un sito Web o di un file trovato:
 - o acquisire informazioni sulla fonte, sull'autore e sul soggetto - istituzione, ente, azienda, individuo privato - che pubblica il sito su cui i contenuti si trovano, cercando di valutarne autorevolezza e attendibilità (es. risorse di tipo filosofico pubblicate da Dipartimenti di filosofia di atenei italiani potranno essere considerate ragionevolmente attendibili o, perlomeno, pertinenti e degne di considerazione per ulteriori approfondimenti); commenti e suggerimenti trovati in newsgroup e forum possono altresì essere utili per questo tipo di valutazioni; indici innovativi quali il *pagerank* di Google (evidenziato nell'utilissimo strumento *Google Toolbar*) aiutano poi a misurare, se non l'attendibilità, perlomeno la popolarità di un sito Web, in base alla quantità di altre risorse online che ad esso fanno riferimento (tramite link e citazioni);
 - o confrontare tra di loro più fonti online, facendo però attenzione ai rischi di replicazione degli errori e di autoreferenzialità che il Web, grazie all'infinita riproducibilità e flessibilità del digitale, porta con sé;
 - o confrontare le fonti online con fonti *offline* (es. libri e pubblicazioni cartacee); è fuori discussione che, al momento attuale, il mondo editoriale "tradizionale", forte dei propri sistemi e criteri di *gatekeeping* (considerati qui in senso positivo come garanzie di scientificità e pertinenza per il passaggio, ad esempio, dall'idea dell'autore alla pubblicazione), offra garanzie molto superiori di affidabilità e qualità dei contenuti o, perlomeno, ne renda più facile la valutazione, grazie alla chiara riconoscibilità delle fonti, dei contesti e degli attori coinvolti;
 - o porre attenzione ad alcune caratteristiche macroscopiche di un sito Web (aspetto grafico, strutturazione di menu di navigazione e dei contenuti, ma anche completezza, coerenza d'insieme, ecc.); benché a volte le apparenze ingannino, i siti di aspetto più professionale e con una migliore organizzazione tendono ad avere un maggior livello di attendibilità;
 - o valutare la frequenza di aggiornamento dei contenuti, ad esempio per capire se un certo sito Web sia o meno una risorsa "viva" (ciò è utile soprattutto per pubblicazioni che si propongono come luoghi di attività *online* e dibattito aggiornati o punti di riferimento, ad esempio, per raccogliere informazioni sull'attività scientifica e accademica - con news su convegni e pubblicazioni, ecc.)

4. organizzare e archiviare i risultati della ricerca

La facilità di accesso, la natura ipertestuale e multimediale del Web, spingono al disimpegno e alla fluttuazione libera tra i contenuti. Se è utile una preliminare pianificazione dei percorsi della ricerca sul Web, è fondamentale una sua adeguata finalizzazione, attraverso un'accurata opera di archiviazione e organizzazione dei risultati, come del resto è d'uopo anche nell'ambito delle tradizionali ricerche *offline*.

Gli strumenti digitali offrono possibilità di archiviazione e organizzazione dei risultati straordinariamente efficienti, efficaci e flessibili: spazi di memorizzazione virtualmente infiniti; flessibilità e malleabilità dei dati; rapida trasferibilità da un dispositivo all'altro o da un utente all'altro tramite mezzi quali la posta elettronica; indicizzazioni automatiche; estrema velocità e raffinatezza di interrogazione per il recupero delle informazioni archiviate. Un utilizzo accorto di questi potenti strumenti consente la creazione di archivi strutturati e ordinati che, come tali, possono essere agevolmente integrati e modificati nel tempo, divenendo una risorsa di supporto durevole per qualsiasi attività di ricerca.

- consigli operativi generali:

- apprendere la capacità di "orientarsi" nel mondo informatico e digitale, favorendo la pratica e l'uso concreto degli strumenti rispetto a impostazioni nozionistiche (es. lettura di manuali) che, in ogni caso, non garantiscono una conoscenza "stabile" rispetto ad una realtà tecnologica in rapidissima evoluzione; standard e convenzioni diffuse (operative, grafiche, ecc.) consentono poi di presiedere alle attività di gestione dei contenuti su un ampio numero di dispositivi diversi ma resi compatibili dalla comune natura digitale;
- organizzare i siti visitati attraverso la funzione "Aggiungi a preferiti" (o analoghe) presente in tutti i browser, denominando opportunamente i collegamenti e organizzandoli in cartelle e sottocartelle;
- condurre la ricerca avvalendosi della funzione del browser "Apri in altra finestra" (o analoghe), cliccando sui collegamenti Web con il tasto destro del mouse (o con il tasto del mouse + tasto CTRL su sistemi Mac); ciò è utile, ad esempio, per mantenere in sottofondo una finestra di *browsing* contenente un elenco di risorse Web (come i risultati di una ricerca) che verranno via via aperte in nuove finestre (in alternativa, le nuove versioni dei browser più usati (Internet Explorer, Mozilla Firefox, Google Chrome, Safari), offrono un'interfaccia a "linguette" per caricare e tenere aperte contemporaneamente più pagine e passare agevolmente dall'una all'altra);
- utilizzare sistemi di videoscrittura (es. Microsoft Word) parallelamente alla ricerca sul Web: copiare e incollare testi dal browser al *word processor* consente di creare file stabili e ben organizzati, fondendo in un unico archivio testi provenienti da più fonti (che si indicheranno, copiando e incollando l'URL della risorsa online); si ricordino i comandi standard, via tastiera: CTRL+C, per copiare in memoria testi e oggetti selezionati da una certa fonte; CTRL+V, per incollare in una certa destinazione testi e oggetti presenti in memoria (nei sistemi Mac le combinazioni di tasti sono *Mela+C* e *Mela+V*);
- normalizzare i testi provenienti da pagine Web nel programma di videoscrittura, eliminando caratteristiche di formattazione come font, dimensioni, colori dei caratteri, allineamento dei paragrafi, ecc., attraverso la funzione di Word "Incolla speciale → Testo non formattato" (ciò è utile per testi "semplici", privi di formattazione rilevante, collegamenti ipertestuali e altre caratteristiche originarie da preservare, ecc.);
- nel caso di pagine Web di cui si desideri mantenere la formattazione e la struttura ipertestuale, utilizzare la funzione "Salva con nome" (o analoghe); per il *download* (scaricamento) di file da salvare e archiviare (pdf, doc, ecc.) utilizzare la funzione "Salva oggetto con nome" (o analoghi);
- creare e organizzare cartelle e sottocartelle, secondo un'opportuna e ragionata logica di archiviazione (es. creare una gerarchia su tre livelli costituita da cartelle nidificate, con una principale denominata "Filosofia", al cui interno si potrebbe inserire "Filosofia antica"; dentro quest'ultima potrebbero essere create cartelle denominate "Aristotele", "Platone", "Sofisti", ecc.); la profondità di strutturazione e articolazione sarà opportunamente valutata in base alla quantità e alla qualità delle risorse che si intendono archiviare;
- denominare file e cartelle (regole di *naming*) secondo criteri precisi, che garantiscano una rapida identificazione dei contenuti e ne agevolino il reperimento, l'ordinamento e la ricerca (es. usare il termine "heidegger" come prefisso iniziale per tutti i file riguardanti il noto filosofo, aspetto utile ad esempio per ritrovare i file in lunghi elenchi alfabetici di cartelle affollate).

- strumenti hardware e software:
 - o installare la "barra di Google" (*Google Toolbar*), per i sistemi Windows, che, tra le altre cose, consente di evidenziare immediatamente con *highlights* di vari colori le parole chiave cercate all'interno di una pagina Web, evitando così di dover usare la funzione "Modifica -> Trova (in questa pagina)" o analoghe (il browser Safari per Mac offre di default di una funzionalità Google analoga);
 - o utilizzare gli strumenti per le lingue di Google, Yahoo! (*Babelfish*) o altri portali, che consentono di ottenere la traduzione di parole, frasi e intere pagine Web; si tratta perlopiù di traduzioni automatiche assai inaccurate che, però, possono aiutare a capire il significato generale di un testo;
 - o dotare il proprio PC di adeguati sistemi antivirus (con scansione in tempo reale) e *firewall* (per proteggersi da attacchi esterni di virus, *worm*, utenti malintenzionati, ecc.); eseguire aggiornamenti costanti del sistema operativo, in particolare quelli periodicamente rilasciati dai produttori software per colmare eventuali "buchi" nella sicurezza del sistema o contrastare nuove minacce virali;
 - o adottare adeguate strategie di *backup*, salvando ogni pochi minuti i file in fase di redazione ed elaborazione (per tutelarsi contro perdite accidentali a causa di comandi errati, blocchi hardware e software, blackout, ecc.) e, soprattutto, tenendo una copia aggiornata dei file di lavoro importanti (utilizzando, ad esempio, supporti di memorizzazione esterni come CD, DVD, hard disk esterni, "chiavi" USB, dischi virtuali su Web, ecc.);
 - o dotarsi di altre *utility* essenziali quali software per la videoscrittura (es. Microsoft Word), la compressione/decompressione di file (es. Winzip, PkZip, ecc.), la visualizzazione di file nel formato *cross-platform* PDF (Acrobat Reader), la lettura e gestione di newsgroup (es. Outlook Express).

* * *

Motori di ricerca e directory generalisti

- Google (www.google.com, www.google.it), il principale motore di ricerca; [Google Toolbar, strumento di supporto al browser (<http://toolbar.google.com/intl/it/>)];
- Yahoo!, motore di ricerca e directory (<http://dir.yahoo.com/>);
- Bing (ex Live, ex MSN - Microsoft Network), motore di ricerca e directory Microsoft (www.bing.it);
- DMOZ - Open Directory Project (www.dmoz.org);
- Google Gruppi (Usenet) (groups.google.it);

Nota: pur esistendo un gran numero di portali Web, nazionali e internazionali, che offrono funzioni di ricerca generalista (Alice, Virgilio, Tiscali, Arianna, Excite, Lycos, Altavista, AOL, ecc.), la quasi totalità di essi trae i propri risultati dai database di Google, Yahoo! e Microsoft (Bing), le uniche grandi organizzazioni dotate di software *spider* proprietari per l'indicizzazione della Rete (rispettivamente *Googlebot*, *Slurp* e *MSNBot*) rimaste sul mercato dopo una frenetica attività di acquisizioni/fusioni che ha interessato la pluralità di soggetti presenti nei primi anni del Web. I risultati forniti dai vari portali vengono poi, a seconda dei casi, incrociati con quelli di directory proprietarie e/o mescolati con risultati "a pagamento" (secondo logiche *pay-per-click* e altri criteri di *advertising* offerti agli inserzionisti), anch'essi di norma provenienti dai grandi *player* sopra citati (*Google AdWords* è il punto di riferimento per il settore).

Alcune risorse Web utili in ambito filosofico

I link proposti sono a puro titolo esemplificativo e si riferiscono ad alcune tra le risorse italiane e internazionali più note. Gli URL spesso cambiano nel corso del tempo, quindi anche quelli qui riportati non sono da considerarsi "stabili". I contenuti disponibili sulla Rete sono in costante evoluzione e, per rimanere sempre aggiornati, bisogna avvalersi dei motori di ricerca e delle directory, sia generalisti che settoriali.

Nota bene: negli URL (indirizzi Web) riportati è sempre sottinteso il prefisso "http://".	
Directory e archivi di link	
Risorse filosofiche in rete, Dip. Fil. Univ. Genova	www.dif.unige.it/risorse/index.htm
Risorse filosofiche internazionali (a cura di A. Varzi)	www.columbia.edu/~av72/links.html
EpistemeLinks.com	epistemelinks.com/Main/MainPers.aspx
Directory of Open Access Journals	www.doaj.org
Portali	
Filosofico.net	www.filosofico.net/filos.html
Portale della filosofia	www.portalefilosofia.com/
Enciclopedia multimediale delle Scienze filosofiche	www.emsf.rai.it
Filosofia.it	www.filosofia.it
Riflessioni.it	www.riflessioni.it
Enciclopedie online	
Treccani	www.treccani.it
Wikipedia	it.wikipedia.org/wiki/Filosofia - www.wikipedia.it
Stanford Encyclopedia of Philosophy	plato.stanford.edu
Internet Encyclopedia of Philosophy	www.iep.utm.edu
Europeana	www.europeana.eu
Altri siti	
Dipartimenti di Filosofia in Italia e nel mondo	www.dif.unige.it/risorse/2istitu.htm
IISF - Istituto Italiano per gli Studi filosofici	www.iisf.it
Società filosofica italiana	www.sfi.it
Giornale di filosofia	www.giornaledifilosofia.net
Google Scholar	scholar.google.it
Testi in digitale	
Google Libri	books.google.it
Liber Liber - Progetto Manuzio	www.liberliber.it
Links to digital text projects	www.ilt.columbia.edu/publications/digitext.html
Project Gutenberg	promo.net/pg
Projekt Gutenberg - DE	gutenberg.spiegel.de
People with online papers in Philosophy	consc.net/people.html
Alex Catalogue of Electronic Texts	www.infomotions.com/alex
Biblioteche online	
Opac italiani (Sistema Bibliotecario Nazionale)	opac.sbn.it/ - www.internetculturale.it
Catalogo collettivo Università di Verona	opac.univr.it (o accesso da www.univr.it)
MAI MetalOAC Azalai Italiano	www.aib.it/aib/opac/mai2.htm3
CIBIT - Biblioteca italiana telematica	www.bibliotecaitaliana.it
Gallica	gallica.bnf.fr
The British Library	www.bl.uk
The Library of Congress	www.loc.gov