

Tendenze e novità degli strumenti per la ricerca

Search Engine Watch
<http://searchenginewatch.com>

Online Inc.
<http://www.onlineinc.com/newspage.html>

Information Today
<http://www.infotoday.com/>

Excite
<http://www.excite.com>

Virgilio
<http://www.virgilio.it>

Amazon
<http://www.amazon.com>

Magellan
<http://www.mckinley.com/>

Mining Company
<http://www.miningco.com/>

Argus Clearinghouse
<http://www.clearinghouse.net/>

Northern Light
<http://www.northernlight.com>

Ask Jeeves
<http://www.askjeeves.com/>

Gli strumenti per la ricerca in Internet, motori di ricerca, cataloghi, metamotori ecc., sono sottoposti a continui processi di revisione e di implementazione. Le variazioni sono dovute ai progressi delle tecnologie, alle necessità del mercato, agli accordi commerciali, alla concorrenza e a moltissimi altri fattori. La mancanza di stabilità del web è un fatto acquisito da tutti gli utenti di Internet, e tale instabilità è inevitabile anche in questi strumenti che, sebbene ancora imperfetti, costituiscono ad oggi la sola possibilità disponibile di identificare almeno una parte dell'informazione presente in rete per poterne fruire.

La lettura di riviste specializzate e la consultazione di siti che si occupano in specifico di moto-

ri di ricerca è indispensabile per mantenere quell'aggiornamento che è necessario per un uso efficace di tali mezzi.

Trascurando di citare i numerosissimi siti che contengono guide ai motori di ricerca, si segnalano innanzitutto alcune fonti, tra le tante possibili, utilizzabili per seguire la novità nel settore degli strumenti di ricerca.

Uno dei siti più noti è Search Engine Watch (<http://searchenginewatch.com>) il cui autore, Danny Sullivan, tiene sotto costante monitoraggio le novità dei principali strumenti di ricerca, organizza la materia in capitoli (rassegne, guide, tecnologia, ecc.) e pubblica un utilissimo bollettino mensile, distribuito via e-mail, che informa delle principali novità. Sebbene rivolto principalmente ai webmaster, tale sito è di indispensabile consultazione anche per i documentalisti perché fornisce una panoramica sempre aggiornata e puntuale ed inoltre suggerisce mensilmente una buona bibliografia.

Chi desidera un aggiornamento quotidiano può consultare il notiziario dell'editore Online Inc., (<http://www.onlineinc.com/newspage.html>), le news sono organizzate in vari capitoli ed uno è dedicato ai motori di ricerca. Online Inc. pubblica inoltre le riviste "Online" e "Database", (link dalla homepage dell'editore <http://www.onlineinc.com>) da anni tra le migliori e più note riviste per l'aggiornamento dei ricercatori delle banche dati online e che vanno sicuramente annoverate tra le fonti indispensabili per una informazione di qualità su nuovi prodotti, tendenze, confronti e suggerimenti sulle tecniche di ricerca anche nel web. Ottimo per gli scopi predetti

anche il sito di Information Today (<http://www.infoday.com/>) che riporta tra l'altro il link alla rivista "Searcher: the magazine for database professionals" (<http://www.infoday.com/searcher/default.htm>) ricca di informazioni molto aggiornate anche sui motori di ricerca.

Tentiamo ora di riassumere e di indicare qualcuna delle principali tendenze dei mezzi di ricerca per il web, con la convinzione che è sempre di fondamentale importanza sapere in che direzione stanno muovendosi mercato e tecnologie al fine di non trovarsi disorientati o peggio sorpassati in una materia che costituisce uno degli assi portanti della nostra professionalità.

La tendenza generale più evidente è senz'altro la trasformazione di gran parte dei motori di ricerca nei cosiddetti servizi "portal", orientati cioè a fornire un accesso non solo alla ricerca sul web, ma a diventare il punto di partenza per una gamma di servizi in rete più completa possibile. Così accanto alle finestre che servono per eseguire la ricerca vera e propria inserendo i termini da ricercare, appaiono in numero sempre maggiore altre possibili scelte che introducono servizi che spaziano dalla posta elettronica gratuita, ai servizi di informazione su condizioni del tempo, quotazioni di borsa, news, shopping, ecc. dove il limite sembra essere solo quello della fantasia dei produttori. Molte delle homepage dei servizi di ricerca sono diventate sempre più personalizzabili per facilitare l'accesso dell'utente ai servizi in rete che gli interessano maggiormente, e solo a quelli. Un esempio per tutti è Excite (<http://www.excite.com>) che consente di avere una pagina con notizie e aggiornamenti filtrati a seconda dei propri interessi specifici. Ogni volta che l'utente si collega ad Excite per eseguire una ricerca, ➤



DEFINISCI DOVE CERCARE:

[Eurosearch](#) | [Con più termini](#) | [In una nazione](#) | [All'interno di un sito](#)



Unreal, battaglie 3D

I segreti di uno dei più interessanti action games degli ultimi mesi. • IN VIRGILIO GIOCHI

- [7 giorni di siti da non perdere](#)
- [Il fascino delle grotte](#)

canali



Ambiente

[Parchi, Animali, Associazioni](#)

Arte

[Musei virtuali, Storia dell'arte](#)

Business

[Commercio elettronico, Finanza](#)

Cinema

[News, Attori, Festival](#)

Computer

[Software, Periferiche, Notizie](#)

Donne

[Bellezza, Lavoro, Cultura](#)

Education

[Corsi online, Scuole, Università](#)

Fantascienza

[Libri, Giochi, Film](#)

Giochi

[Simulatori, Recensioni, Soluzioni](#)

Hobby & Casa

[Motori, Ricette, Audio-Video](#)

Home banking

[Servizi, Demo, Tecnologie](#)

Informazione

[Attualità, Quotidiani, Radio e TV](#)

Lavoro

[Concorsi, Stage, Cerco/Offro](#)

Libri

[Acquisti online, Biblioteche, Autori](#)

Meteo

[Il tempo in Italia e nel mondo](#)

Musica

[Generi, Acquisti, MIDI](#)

Ragazzi

[Fumetti, Eventi, Giocattoli](#)

Salute

[Cura del corpo, Farmaci e cure](#)

Scienze

[Curiosità, Discussioni, Scoperte](#)

Shopping

[Aziende e acquisti online](#)

Società

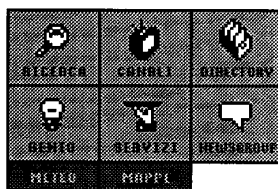
[Politica, Movimenti, Solidarietà](#)

Sport

[Atleti, Federazioni, Attrezzature](#)

Viaggi

[Destinazioni, Guide, Trasporti](#)



[Navigatori Solitari Unitevi!](#)



Smau 98

Preparati all'evento hi-tech dell'anno.

• IN VIRGILIO COMPUTER

PARTI DA VIRGILIO

- Fai di Virgilio [la tua Home Page](#)
- Metti Virgilio tra i tuoi [siti preferiti](#)

VIRGILIO NEWS

Tutto il web in una mail
Abbonati, è gratis!

Inserisci il tuo indirizzo e-mail:

Scopri [Virgilio News!](#)

Consulta [l'archivio online](#).

METEO

Previsioni del tempo delle [principali città italiane](#).

te con Altavista e Hotbot/Lycos fanno sì che sia possibile affiancare alla ricerca tradizionale in modo semplice ed immediato la ricerca nei cataloghi delle librerie, ottenendo un elenco di titoli di libri correlati all'argomento ricercato. Si tratta certamente di una operazione strettamente commerciale, che tuttavia non infastidisce l'utente e con un valore di utilità pratica e culturale rilevanti.

Su questa linea Lycos offre anche la possibilità di ripetere immediatamente la ricerca impostata in modo generale su particolari aree e/o servizi web per individuare oltre che libri, notizie, immagini, rassegne.

Analoga utilità, ma ai fini del reperimento di siti aziendali, ha il servizio di Altavista "Real Name" che facilita appunto l'individuazione degli URL di società registrate. È stato creato un database che contiene tutte le possibili variazioni della denominazione di una azienda; eseguendo la ricerca nel tradizionale form di Altavista, il link alla homepage della società viene fornito in testa alla lista dei risultati della ricerca. Il vantaggio per l'utente è di guadagnare tempo arrivando direttamente alla homepage aziendale senza perdersi tra le pagine, a volte molto numerose, che contengono notizie indirette su di essa; il vantaggio per Altavista è la tassa di registrazione che le aziende pagano per farsi trovare dagli utenti.

Un'altra tendenza molto evidente è la minore differenziazione tra i motori di ricerca propriamente detti e i cataloghi o *directories*. Questa tendenza si spiega con la maggiore facilità dell'uso dei cataloghi rispetto all'uso dei motori. Per gli utenti non esperti è certamente più semplice fare del browsing tra materiale organizzato con un ordine gerarchico, che non impostare una strategia di ricerca. Inoltre il materiale dei cataloghi, sebbene

se ha personalizzato la pagina attraverso una operazione molto semplice e abbastanza rapida, trova sul proprio schermo oltre alla finestra di ricerca, una serie di informazioni aggiornate quotidianamente e selezionate in base alle preferenze da lui fornite. In materia di personalizzazione dobbiamo aspettarci grosse novità dalla tecnologia degli "intelligent agents": Microsoft ha recentemente acquistato la tecnologia Firefly che riguarda appunto sistemi per filtrare e personalizzare la ricerca. Anche l'interesse dei grandi

media tradizionali per le acquisizioni o le alleanze con i produttori di servizi di ricerca testimonia la tendenza a offrire agli utenti servizi più articolati e completi di accesso al web: ultimo esempio in ordine di tempo l'acquisizione di una quota rilevante di Infoseek da parte della Disney. Alcuni servizi nati per la ricerca sono già stati trasformati in veri e propri media. Un esempio per tutti può essere Virgilio (<http://www.virgilio.it>): nato come catalogo di risorse italiane e tuttora il maggiore indice per soggetto italiano (circa 42.000

siti indicizzati), appare oggi come un servizio di accesso al web di tipo quasi televisivo, e relega anche graficamente l'accesso al catalogo in secondo piano. Le alleanze di ordine squisitamente commerciale fanno ben comprendere come i servizi di ricerca stiano cercando di fornire all'utente servizi e risposte più completi con una fusione di interessi commerciali e di servizi per l'utenza a volte veramente interessanti. Gli accordi tra le grandi librerie Amazon (<http://www.amazon.com>) e Barnes & Noble rispettivamente

quantitativamente inferiore rispetto ai motori, ha generalmente subito una maggiore selezione perché c'è stato l'intervento di una redazione (anche se spesso tali redazioni hanno una visione un po' troppo informatica della catalogazione!). Altavista, ad esempio, ha recentemente introdotto sia una directory per consentire il browsing che dei canali (*zones*) di siti selezionati sugli argomenti di maggiore richiamo (salute, turismo, ecc.)

Nel settore dei cataloghi acquistano popolarità servizi molto centrati sulla qualità dei siti indicizzati come Magellan (<http://www.mckinley.com/>), Mining Company (<http://www.miningco.com/>) e Argus Clearinghouse (<http://www.clearinghouse.net/>), all'insegna di contenuti piccoli, ma molto selezionati da redazioni composte da veri esperti delle varie discipline.

Le guide per argomento di Lycos si avvalgono anche della cooperazione degli utenti per valutare la qualità dei siti selezionati: l'utente che visita un sito trovato nelle guide Lycos ha la possibilità di votare facendo sapere al sistema come giudica il sito. I giudizi degli utenti concorrono a stabilire il *rating* del sito; anche se non sempre la qualità coincide con la popolarità, queste forme di "collaborative filtering" stanno prendendo piede e vanno seguite con attenzione.

Per quanto riguarda l'utilizzo di motori di ricerca veri e propri la tendenza è quella di facilitare all'utente anche non esperto la formulazione della sintassi della domanda senza che ciò vada a scapito della completezza della strategia.

Innanzitutto, fortunatamente, si va diffondendo una sintassi di ricerca comune: quasi tutti i principali motori di ricerca accettano ormai gli operatori booleani semplificati + per AND e - per NOT, e le virgolette ("...") per condurre la ri-

cerca di una frase esatta. Altri servizi, come Hot Bot che prevedono la creazione della strategia di ricerca attraverso i menu a bandiera, hanno potenziato e migliorato la possibilità di scelta delle opzioni ed offrono una ricerca praticamente sovrapponibile a quella più "libera" di sistemi come Altavista tradizionalmente preferiti dai professionisti della ricerca dell'informazione per la versatilità nella costruzione delle strategie.

Quasi tutti i servizi hanno aggiunto poi funzioni che forniscono suggerimenti su come migliorare i risultati ottenuti; i risultati di queste funzioni sono ancora molto spesso deludenti, ma è interessante esaminarle perché indubbiamente indicative del cammino che si intende percorrere nell'assistenza all'utente. Una delle forme migliori di assistenza resta per ora fornita da quei motori come Excite e Lycos che offrono la possibilità di ricercare siti "simili" a quelli trovati e ritenuti più pertinenti (funzione "more like this"), mentre la possibilità di escludere e aggiungere alla strategia di ricerca termini suggeriti (ad es. la funzione refine di Altavista) è per il momento quasi sempre di scarsa utilità.

I produttori inoltre stanno lavorando attivamente per migliorare i criteri con cui viene stabilita la rilevanza dei documenti reperiti. L'argomento è estremamente tecnico e potrebbe dar luogo ad infinite discussioni sui criteri per la determinazione della pertinenza e della rilevanza di un documento nonché sui modi con cui i meccanismi di determinazione possono essere aggirati o addomesticati a favore delle proprie pagine web. Una delle tendenze dei prototipi sperimentali per determinare il *ranking*, sembra essere comunque il coinvolgimento del conteggio dei link alle pagine reperite e/o la memorizzazione delle

pagine maggiormente consultate dagli utenti che svolgono una ricerca su un determinato soggetto.

Per essere più vicini agli utenti i grandi produttori degli strumenti di ricerca hanno sviluppato recentemente servizi mirati agli interessi regionali. Aumentano le interfacce di interrogazione con le istruzioni nelle varie lingue, che non sempre però offrono la stessa gamma di servizi e di funzioni degli originali in lingua inglese. Quasi tutti i motori consentono la ricerca per dominio, mentre per il momento solo Altavista e Northern Light (<http://www.northernlight.com/>), che possiedono dei software in grado di riconoscere la lingua in cui è scritto un documento, consentono di ricercare i siti in una certa lingua. Altavista inoltre offre un servizio, per la verità ancora molto rozzo, di traduzione delle pagine reperite.

Ricordiamo inoltre le sezioni regionali di Yahoo, che sebbene ancora meno interessanti dei cataloghi locali "storici" come Virgilio avranno degli sviluppi sicuramente da studiare.

Per avere una panoramica sempre aggiornata delle possibilità offerte dai servizi di ricerca più recenti è necessario anche seguire e testare i nuovi strumenti che normalmente sono quelli che applicano le tecnologie più innovative. Una delle maggiori novità concettuali dell'ultimo anno è il motore di ricerca Northern Light che combina la ricerca sul web a quella su una raccolta selezionata di articoli e notizie chiamata "Special Collection" posseduta dal produttore (licenza IAC) e il cui contenuto è ricercabile gratuitamente, ma consultabile a pagamento (prezzi da 1 a 4 \$ per articolo, con la garanzia soddisfatti o rimborsati — titoli e abstract gratuiti). Inoltre Northern Light fornisce i risultati raggruppati e riordinati in cartelle (*custom folders*) che ne

rendono molto più immediata la consultazione e la lettura. I risultati vengono raggruppati per soggetto, per fonte, per tipo, per lingua e, ovviamente, per appartenenza alla Special Collection. Aprendo ogni cartelletta si hanno delle ulteriori suddivisioni in sotto-cartelle. Si tratta insomma di un esempio di catalogazione automatizzata dei risultati delle ricerche che consente all'utente di avere una visione generale e rapida degli argomenti e del tipo di documenti trovati. Dietro questa catalogazione automatizzata c'è comunque un grosso lavoro preparatorio "umano" per la determinazione di più di 200.000 soggetti che servono come base per la parte automatizzata.

Anche dietro le quinte di un altro servizio, il metamatore Ask Jeeves (<http://www.askjeeves.com/>), c'è una combinazione di intelligenza artificiale e di lavoro preparatorio di esperti ricercatori. Ask Jeeves accetta le domande in linguaggio naturale (con particolare successo se la richiesta consiste in una domanda del tipo who, when, what, where) ed elabora, come prima risposta, una serie di domande guida che aiutano l'utente a focalizzare l'argomento da ricercare.

A conclusione di questa rassegna forzatamente incompleta, vista la quantità e la rapidità dei cambiamenti, una brevissima nota sui riflessi di questa situazione sul nostro lavoro quotidiano. Lo sforzo necessario per seguire, almeno nelle sue linee generali, lo sviluppo degli strumenti richiede senza dubbio impegno di tempo e applicazione, ma è anche una attività molto stimolante, spesso divertente e soprattutto irrinunciabile per continuare a fornire ai nostri utenti un supporto qualitativamente eccellente.

[Ultima consultazione: 10.8.1998]

Mariateresa Pesenti