

Alcune analisi comportamentali degli utenti dell'Emeroteca Virtuale

MARCO SCARNÒ

Sintesi

Gli utenti di una Emeroteca Virtuale tendono a divenire più esperti nel suo utilizzo. Ricerche e navigazioni più efficienti, ossia risultati (articoli) raggiunti con un minor numero di passi. Osservazioni interessanti, che derivano da un processo di analisi delle tracce lasciate dagli utenti stessi nei server che ospitano il sistema in questione. Un processo, tuttavia, che passa per l'analisi di grandi quantità di dati che debbono essere trasformati in apposite fonti informative.

Introduzione

L'Emeroteca è una Istituzione che offre, con diverse modalità, l'accesso ai servizi di catalogazione e consultazione di giornali, riviste e periodici. Essa fornisce, quindi, inequivocabili spunti di ricerca alla collettività, la quale può così anche mantenersi aggiornata sui progressi nei rispettivi campi d'interesse.

Negli ultimi anni, e quella del CIBER ne è un esempio, si sono sviluppate diverse *Emeroteche Virtuali*, il cui accesso è permesso tramite internet. Esse presentano notevoli vantaggi, tra i quali si possono considerare quelli di natura gestionale (un sistema informatico grande quanto un piccolo armadio può contenere milioni di articoli) e quella legati alla loro accessibilità (basta un computer e una connessione al *World Wide Web* per consultarla).

Introdurre gli svantaggi non è compito facile, anche perché andrebbero considerati i molteplici punti di vista sia degli utenti sia dei bibliotecari; ognuno di questi individui si è trovato, senza dubbio, invischiato nel cambiamento indotto dal *progresso tecnologico* e, forse, le loro posizioni "critiche" potrebbero essere riferite più a questo che alla modifica del loro spazio di ricerca e lavoro.

Certo è che non si può non osservare come, anche in questo caso, il progresso stesso si associ ad una *deumanizzazione* dei rapporti - non è possibile chiedere dei suggerimenti di lettura a una macchina -; probabilmente, però, le generazioni future (o semplicemente quelle più *giovani*) vedranno tali sistemi come l'unica e inevitabile realtà.

Tale considerazione, ad esempio, può essere estesa considerando l'enorme cambiamento avvenuto con la nascita d'Internet e, quindi, con il modo mediante il quale si forma parte della Conoscenza.

Già undici anni fa nel Corriere della Sera era l'articolo "Enciclopedia, morte dei commessi viaggiatori" [1] che raccontava della trasformazione della celebre *Britannica* da cartacea a piccolo CD-ROM; pur se il riferimento era la non necessità dei venditori a porta a porta, si coglieva il cambiamento nelle abitudini di chi ha fatto, trenta anni fa, le ricerche scolastiche sfogliando libri, rispetto a quanti, già da diversi anni, usano strumenti informatici.

In tutto questo risulta senza dubbio interessante valutare come si siano modificate le abitudini degli utenti per *arrivare* all'individuazione del contenuto (prima cartaceo ora elettronico) utile per il proprio lavoro (o svago).

Si può considerare che l'elemento *suggerito da qualcuno* sia rimasto valido anche nell'era virtuale; infatti, tramite un amico, un docente o un altro utente si può arrivare a scaricare il documento comunque d'interesse.

Un'altra modalità invariata nel tempo è quella relativa allo *sfogliare* i contenuti per *tenersi aggiornati*; se prima ci si recava nella biblioteca ora appositi *indici* virtuali consentono di vedere il contenuto di vecchie e nuove riviste.

Quanto è, in realtà, estremamente cambiato in questi anni, ossia con l'avvento del *virtuale*, è relativo alla *ricerca di contenuti* che si avvale di *motori* specifici messi a disposizione all'interno delle pagine internet.

Dietro tali motori, che celano algoritmi anche molto complicati, è un enorme interesse, sia commerciale sia scientifico, in quanto i loro risultati sono in grado di *guidare* le scelte degli utenti verso questo o quel contenuto.

Non è un caso che l'usuale, al momento, Google, sia costantemente insidiato da molti concorrenti. Dietro ciò, oltre il mero interesse economico derivante dai ricavi pubblicitari, è il desiderio degli utenti, spesso aziende, che i propri contenuti appaiano tra i primi risultati delle ricerche.

Anche nell'emeroteca virtuale è definito un *motore* atto a proporre articoli che riguardano maggiormente l'oggetto dell'interesse specificato dagli utenti; si tratta, in breve, di un *interprete* delle parole chiave scritte in appositi campi; queste sono poi ricercate tra quelle che caratterizzano ogni articolo e, laddove significativamente presenti, contribuiscono a portare l'articolo identificato tra quelli risultanti.

La *virtualizzazione* dell'Emeroteca rende possibile l'analisi dei possibili comportamenti dei suoi utenti; ciò mediante lo studio delle tracce da questi lasciate all'interno dei sistemi elettronici che ospitano l'Emeroteca stessa.

Le tracce degli utenti

L'azione compiuta da un utente dell'Emeroteca Virtuale, al pari di quella effettuata su di un qualunque sito Internet, è in questo memorizzata in una riga di un apposito file detto *web log*.

Ovviamente, quel che risulta non è il nome o, comunque, nessun dato sulla persona che effettua la connessione; ciò che viene rilevato è l'indirizzo del *computer* dal quale questi si è collegato. Molto spesso, inoltre, tale indirizzo non si riferisce neppure al singolo pc, bensì a quello che genericamente identifica la rete di appartenenza.

Nel dettaglio del contenuto della traccia lasciata, essa contiene nelle righe l'indicazione su ogni *file* che l'utente ha scaricato. Poiché, generalmente, la singola pagina Internet visualizzata si compone di una molteplicità di elementi, si osserva che, a ogni azione effettuata da un individuo, corrispondono diverse righe nel *file* di log.

Qualora l'utente abbia, poi, scelto di effettuare una ricerca su un determinato campo, la riga del *web log* conterrà l'associazione tra il nome del campo e il valore a questo attribuito; tale riga è scritta con una sintassi particolare, come riportato nella figura 1 (dove si è cercato il nome *Marco* tra gli autori).



```
/cgi-bin/search.pl?Max=10&Start=1&SearchTemplate=search_fora.stiplet&search_field=autore&fields=Author
```

Fig. 1 – Esempio di riga corrispondente all'invio della ricerca della parola *marco* nel campo *autore*

L'analisi dei *web log* costituisce un elemento molto importante nel processo di descrizione dei comportamenti di ricerca; il loro studio ha inizio verso la fine del secolo scorso (si veda, ad esempio, Jansen [2] e Spink [3]). Tuttavia il loro uso all'interno di ricerche comportamentali in sistemi di Emeroteche Virtuali è molto meno diffuso (si veda Wolfram [4]).

Nell'Emeroteca Virtuale i campi che consentono di specificare delle chiavi di ricerca sono diversi (si osservi che nella pagina principale di *Google* ve ne è uno solo); a volte il loro numero consente agli utenti di poter effettuare delle ricerche più *specifiche* all'interno dei contenuti dell'Emeroteca stessa; ciò, ad esempio, è il caso del servizio offerto dal CIBER.

Modalità di ricerca nell'Emeroteca Virtuale

Il collegamento all'Emeroteca Virtuale del CIBER consente di accedere a diverse tipologie di pagine di ricerca, ognuna caratterizzata da potenzialità distinte.

Vi è quella che può essere definita come *simple search* (ricerca semplice), riportata nella

figura 2.

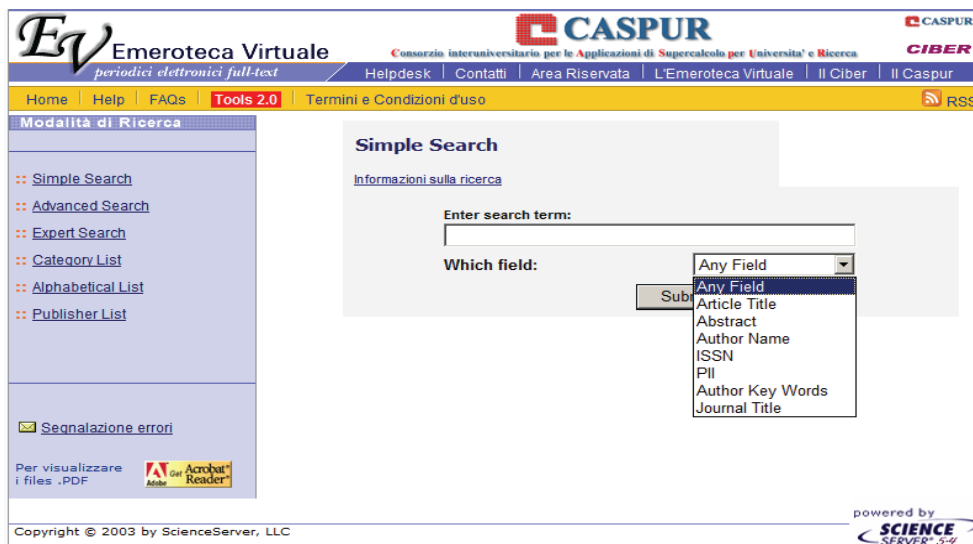


Fig. 2 – Forma attuale della pagina relativa a una *simple search*

Questa consente la specificazione di una o più parole che andranno a essere cercate in uno dei campi selezionati.

Vi è, poi, la modalità *advanced* (avanzata) riportata nella figura 3.

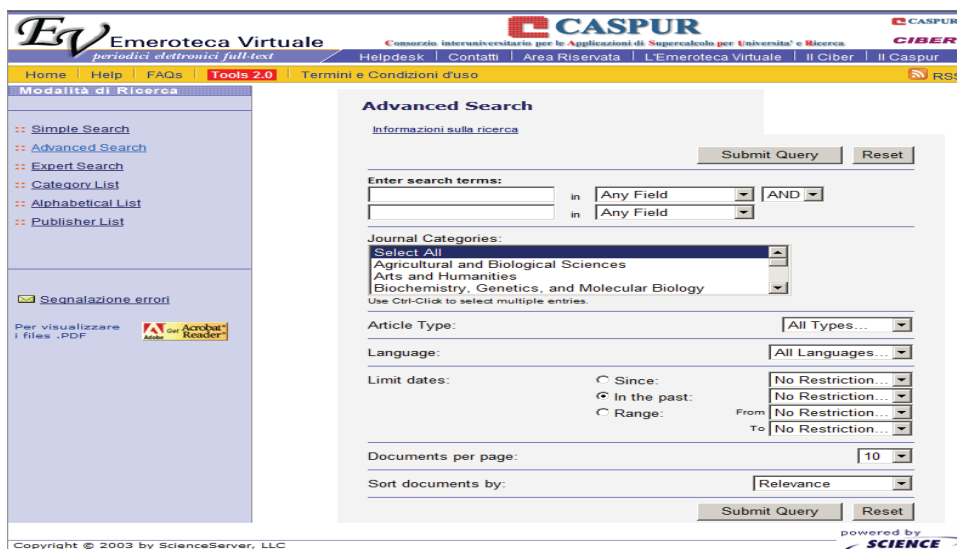


Fig. 3 – Forma attuale della pagina relativa a una *advanced search*

In questa, i termini da ricercare sono associati alle categorie di riviste da considerare, e sono specificati singolarmente.

Infine, c'è la modalità *expert* (esperta) riportata nella figura 4.

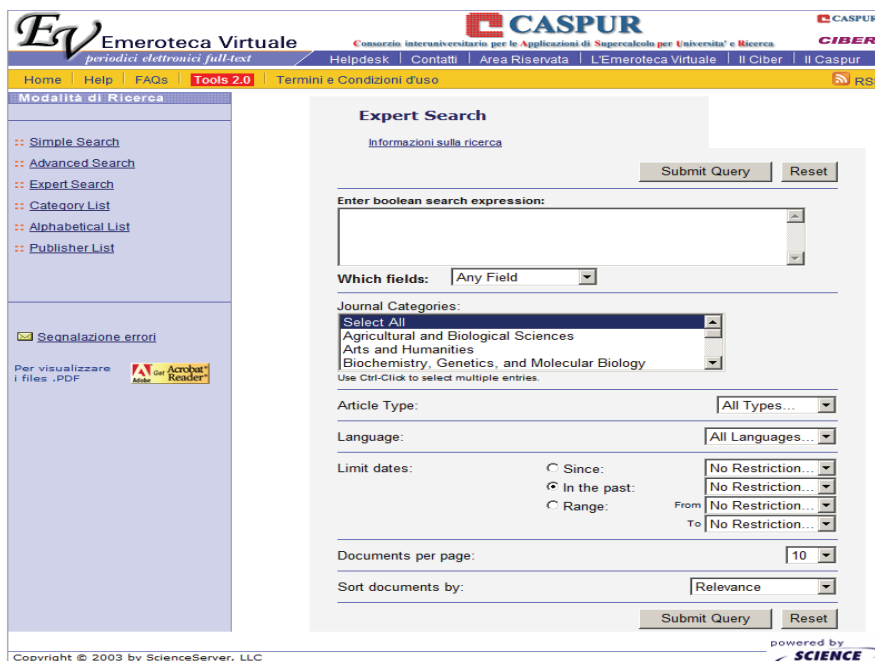


Fig. 4 – Forma attuale della pagina relativa a una *expert search*

Quest'ultima è decisamente molto specifica e prevede la possibilità di scrivere complesse espressioni *boolean* (ossia una serie di condizioni logiche) atte a definire meglio gli ambiti della ricerca stessa.

L'importanza delle possibilità offerte da tali tipologie di pagine risiede nella loro capacità di fornire all'utente degli strumenti atti a *navigare* l'elevata quantità di articoli presenti nell'Emeroteca onde individuare rapidamente quelli che maggiormente possono soddisfare le loro esigenze di studio o svago.

Tuttavia, l'operazione di ricerca non è l'unica modalità di navigazione e, come introdotto precedentemente, anche lo *sfogliare* le riviste è contemplato, in una modalità che in termine tecnico viene definita con *browsing*.

Il *browsing* dei contenuti

Con il termine *browsing* si può intendere la navigazione dei contenuti di un sito tramite la selezione successiva di una serie di collegamenti atti, in ognuno di essi, a indirizzare l'utente verso la sua destinazione finale.

A titolo esemplificativo, si può considerare quanto riportato nella figura 5, ossia la pagina presente nell'Emeroteca Virtuale del CIBER.

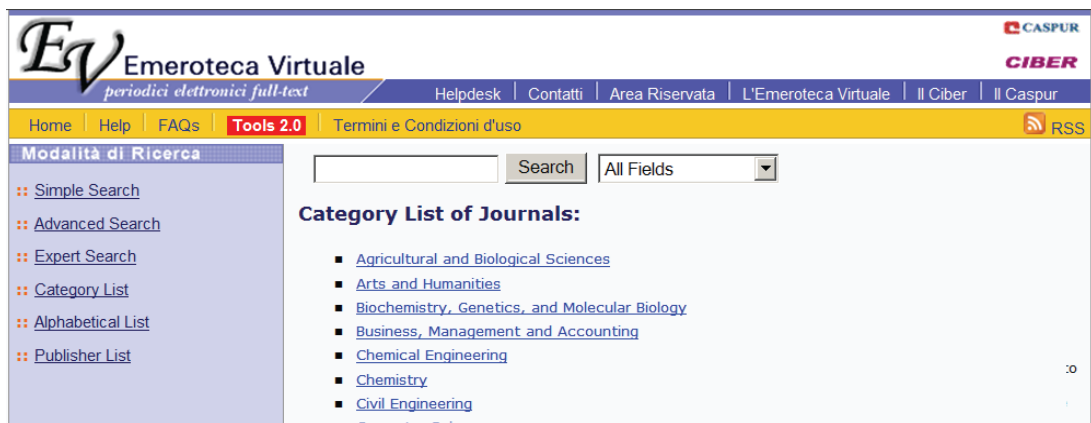


Fig. 5 – Forma attuale della pagina relativa al *browsing*

Come si vede dalla figura 5, la navigazione ha luogo tramite la scelta di una categoria alla quale appartiene la rivista d'interesse; poi si potrà scegliere il nome della rivista e, in seguito, l'anno e il volume desiderato.

In realtà, è possibile anche raggiungere tale risultato considerando liste alfabetiche delle riviste o liste per editore.

La contrapposizione fra ricerca e navigazione dei contenuti

Vi sono diversi autori che hanno osservato una contrapposizione tra l'azione di ricerca e quella del *browsing*, a partire dall'articolo di Evans [5]. Questi ha svolto uno studio su come la disponibilità *online* degli articoli possa influenzare le citazioni effettuate; la sua base era composta da circa 34 milioni di articoli. Le sue conclusioni evidenziano una correlazione negativa tra tali quantità, ossia egli deriva che tante più copie di un giornale sono disponibili *online*, quanto meno quei giornali sono citati; da questo anche la strana nota che giornali liberamente scaricabili sono citati ancora meno. L'autore attribuisce tale fenomeno al fatto che gli utenti effettuano più ricerche che *browsing*.

Anche Tenopir [6] ha osservato che il numero attuale di articoli trovati mediante il *browsing* non è aumentato quanto quello derivante dalle operazioni di ricerca. È doveroso, tuttavia, osservare che i suoi risultati derivano da *interviste dirette*, ossia da informazioni collezionate tramite questionari.

La strategia di analisi delle tracce degli utenti

L'analisi delle tracce degli utenti atta ad individuare i *movimenti* degli utenti dell'Emeroteca Virtuale si basa, come detto, sulle informazioni contenute all'interno dei *web log*.

In realtà, questi costituiscono solo il punto di partenza delle tabelle-dati di tipo statistico che sono in seguito utilizzate; in particolare, ciò di cui si ha bisogno sono delle fonti informative specifiche nelle quali ogni riga sia riferita a un individuo (utente) ed ogni colonna ad una caratteristica su questo rilevata (l'azione che esso ha compiuto).

Ciò vuol dire che il *web log* deve essere *interpretato* da un apposito algoritmo e successivamente trasformato; inoltre la presenza nella sua struttura dei campi sui quali è possibile effettuare delle ricerche e dei valori per essi scelti rende necessaria una notevole flessibilità dell'algoritmo stesso (ciò perché, ad esempio, potrebbero essere state modificate le pagine nel corso degli anni, inserendo o togliendo chiavi di ricerca).

Un possibile processo di analisi dei *web log* potrebbe essere composto dalle seguenti fasi:

- caricamento del file in una tabella statistica;
- eliminazione dei *record* non significativi (tra questi si ricorda anche la necessità di non considerare azioni replicate dovute al *doppio click*⁶⁴);
- identificazione dell'azione compiuta; con riferimento ad una Emeroteca Virtuale si può considerare: *ricerca, browsing, visualizzazione di un articolo*;
- identificazione della sessione di lavoro dell'utente.

Al termine di tali fasi, la tabella statistica ottenuta sarà pronta per essere analizzata; tuttavia proprio l'ultima fase, l'identificazione della sessione, costituisce un interessante elemento che merita uno specifico approfondimento.

⁶⁴ A questo proposito, è doveroso osservare che la pressione o la selezione di qualunque collegamento (mediante un semplice *click* del mouse) in una pagina internet è causa della memorizzazione sul *server* dell'azione effettuata. Spesso accade che l'utente, sia per abitudine sia per sollecitare un risultato che tarda ad arrivare, selezioni più volte tale collegamento. Ciò si traduce nella presenza di più righe nel *web log* che competono alla medesima richiesta e che, quindi, possono fornire dei risultati distorti.

L'identificazione della sessione di lavoro

Secondo Swanson [7], la ricerca di un contenuto può essere interpretata come un processo a *correzione dell'errore* in cui una richiesta è assimilabile ad una ipotesi sulle caratteristiche che il documento desiderato dovrebbe avere; le risposte del *server* hanno lo scopo di migliorare le ipotesi iniziali per ulteriori tentativi; in questo modo gli utenti gradualmente raffinano le loro richieste e i loro obiettivi.

Spink [8] denomina tale processo con il termine *successive search phenomenon* (fenomeno delle ricerche successive), e lo definisce come quel processo di ricerche ripetute nel tempo in relazione a uno specifico, possibilmente evolutivo, problema conoscitivo.

Esso deve, quindi, essere soddisfatto nel corso delle varie interazioni degli utenti con un motore di ricerca; ovviamente, tale esigenza potrebbe, proprio dopo diversi tentativi, giungere a compimento. A questo punto le successive azioni dovrebbero essere riferite all'insorgere di un nuovo problema conoscitivo.

La questione di non facile soluzione è che non è possibile identificare, dall'esame delle *tracce*, a quale obiettivo appartengano le azioni effettuate; in particolare, ci potrebbero essere ricerche dell'*articolo desiderato* che hanno, come durata, un'intera giornata o, al contrario, pochi minuti, per poi nuovamente iniziare con riferimento ad un nuovo argomento da approfondire.

Tale questione è legata al concetto di *sessione di ricerca*, che può essere risolta considerando le seguenti possibilità:

l'utente ha un'unica esigenza informativa, relativa ad un determinato problema, che viene identificata con tutte le ricerche effettuate in un singolo giorno;
vi sono più esigenze di ricerca nella giornata, ognuna delle quali è identificata considerando che una prolungata inattività dell'utente ne definisce di nuove.

In letteratura non c'è un consenso generalizzato sul concetto di *sessione* applicato ad un motore di ricerca (si veda Gayo Avello [9]); tuttavia l'approccio prevalente è quello che si riferisce a Silverstein [10], il quale afferma che una sessione è la serie di richieste effettuate da un singolo utente in un breve intervallo di tempo.

Sfortunatamente, per individuare l'esatta durata del *breve intervallo di tempo* non esistono dei metodi condivisi; tra quelli più impiegati vi sono quelli che si basano sullo studio dei tempi che intercorrono tra richieste successive e sulle caratteristiche della loro distribuzione.

Spink e Jansen [11], ad esempio, osservano che la durata più comune di una sessione di ricerca nel World Wide Web è non superiore a 15 minuti, con una percentuale significativa inferiore ai 5 minuti. Goker [12] suggerisce l'identificazione dell'intervallo tra 11 e 15 minuti.

L'analisi delle tracce provenienti dall'Emeroteca Virtuale del CIBER

Le tracce presenti nei *server* messi a disposizione dal Consorzio CASPUR per ospitare l'Emeroteca Virtuale sono la base di diversi studi effettuati nel corso di questi anni per descrivere il comportamento dei suoi utenti (Scarnò [13]).

Si può innanzitutto osservare che nel periodo compreso tra il primo Gennaio 2006 e il 31 Gennaio 2009 tali tracce ammontano a diversi milioni di righe.

In questo contesto, tuttavia, sono stati considerati solo quei *record*, presenti nei *web log*, che si riferiscono a utenti che hanno acceduto mediante la pagina di autenticazione; ciò consente una esatta identificazione delle loro azioni che non vengono, così, confuse con quelle che possono essere compiute da diversi utenti identificati da un medesimo indirizzo del proprio *computer*. Tale ipotesi consente di ridurre le tracce utilizzate a circa il 10% di quelle iniziali, ossia a poco più di un milione e duecento mila righe riguardanti il dettaglio delle azioni compiute.

La figura 6 riporta l'andamento del numero di utenti autenticati per singolo mese, nel

periodo compreso tra il primo Gennaio 2006 e il 31 Gennaio 2009.



Fig. 6 – Numero degli utenti autenticati nell’Emeroteca Virtuale nel periodo 1/1/2006-31/1/2009

In prima istanza, si può certamente osservare il crescere del loro numero nel corso del periodo; ciò può essere imputato sia a una maggiore diffusione dell’Emeroteca Virtuale negli Enti convenzionati sia a una maggiore presa di conoscenza, da parte degli utenti, dei vantaggi offerti dall’autenticazione nel sistema (possibilità di poter connettersi anche da *computer* pubblici e di poter personalizzare le proprie pagine o salvare le proprie ricerche).

Dopo le citate operazioni di caricamento e pulizia dei dati a disposizione, si può pervenire alla verifica della distribuzione dei tempi che intercorrono tra due azioni successive (in questo contesto sono state considerate sia le ricerche sia le operazioni di *browsing* come elemento discriminante per riconoscere le esigenze conoscitive dell’utente). La figura 7 riporta tale curva in riferimento al periodo esaminato.

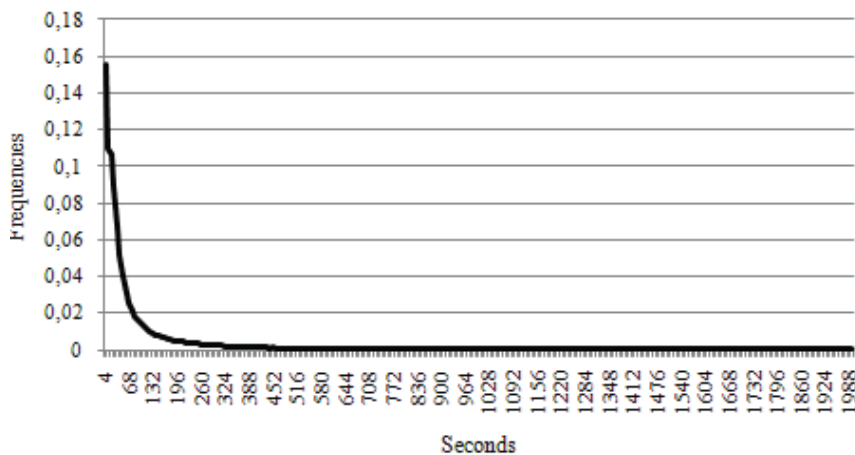


Fig. 7 – distribuzione dei tempi intercorrenti tra due azioni successive (in secondi)

Sebbene vi sia un graduale declino nella distribuzione dei tempi tra azioni, si osserva che non vi è un preciso valore *soglia*; è, comunque, interessante osservare che nell’83% delle volte una richiesta segue la successiva in meno di 3 minuti; nel 93% dei casi l’intervallo tra due richieste avviene in meno di 11 minuti.

Certo è, inoltre, che sarebbe arbitrario considerare un valore soglia comune a tutti gli utenti: si rischierebbe di penalizzare quelli più *veloci* (ossia quelli che si muovono più rapidamente all’interno dell’Emeroteca Virtuale) o quelli più *lenti*.

Una possibile strategia, qui impiegata, fa riferimento alla possibilità di valutare *soglie* personalizzate per utente; in particolare, si è proceduto:

- calcolando il tempo medio tra richieste successive per utente in corrispondenza di ogni giorno in cui questo era presente;
- dalla distribuzione dei vari tempi medi per giorno si è scelto il valore minimo.

Tale valore, unico per ogni utente, è il valore personalizzato della soglia. La media di questo per tutti gli utenti corrisponde a circa 330 secondi (5 minuti e 30 secondi), in linea con i risultati presenti in letteratura.

I comportamenti degli utenti nell'Emeroteca Virtuale del CIBER

Dopo aver identificato le varie sessioni di ricerca, è possibile determinare il loro andamento nel periodo considerato, assieme al numero di articoli *full text* visti (figura 8).

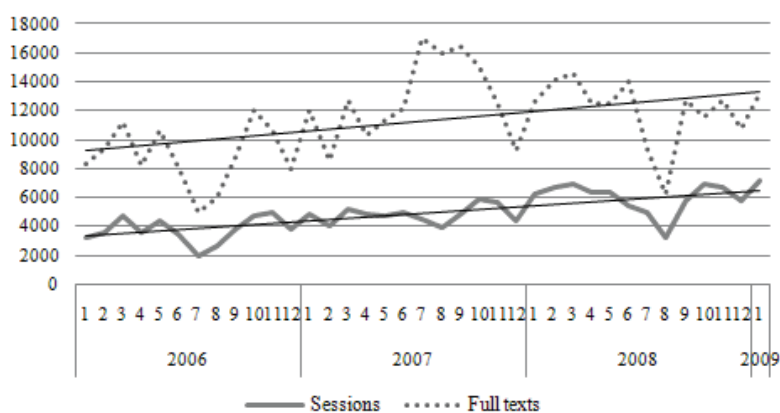


Fig. 8 – sessioni e numero di articoli *full text* visti dagli utenti nel periodo 1/1/2006-31/1/2009 (nella figura sono riportate anche le rette di tendenza)

Si osserva una media, nel periodo, di circa cinque mila sessioni, alle quali corrispondono poco più di undici mila articoli visualizzati (ad ogni sessione corrisponde, quindi, una media di 2,3 articoli visti).

Per quel che riguarda il *primo passo* effettuato dagli utenti considerati dopo l'inizio della loro sessione, si può considerare la seguente distribuzione:

- il 41% delle volte questo corrisponde a una *ricerca semplice*;
- il 22% a una *ricerca avanzata*;
- l'1% a una *ricerca esperta*;
- il restante 35% coincide con l'inizio di una navigazione di tipo *browsing*.

Considerando, invece, il *successive search phenomenon* si può considerare la tabella 1, che riporta le probabilità di transizione tra un passo e quello immediatamente successivo.

	Ricerca semplice	Ricerca avanzata	Ricerca esperta	Browsing	Totale
Ricerca semplice	0,91	0,04	0	0,05	221804
Ricerca avanzata	0,07	0,83	0,01	0,09	150886
Ricerca esperta	0,07	0,09	0,79	0,05	7533
Browsing	0,06	0,07	0	0,87	206190

Tabella 1 - probabilità di transizione tra passi successivi calcolati in relazione agli utenti dell'Emeroteca Virtuale nel periodo 1/1/2006-31/1/2009

La tabella 1 evidenzia che vi è una netta tendenza a *ripetere* il passo precedente; ciò può essere inteso come un probabile *raffinamento* della ricerca onde ottenere l'articolo che più soddisfa le esigenze dell'utente.

Ulteriori considerazioni possono derivare dall'osservare l'andamento, nei mesi considerati, delle tipologie del *primo passo*, così come riportato nella figura 9 (in essa si è utilizzato il rapporto tra questi e il numero delle sessioni, onde evitare l'effetto di distorsione derivante dal variare del numero di queste).

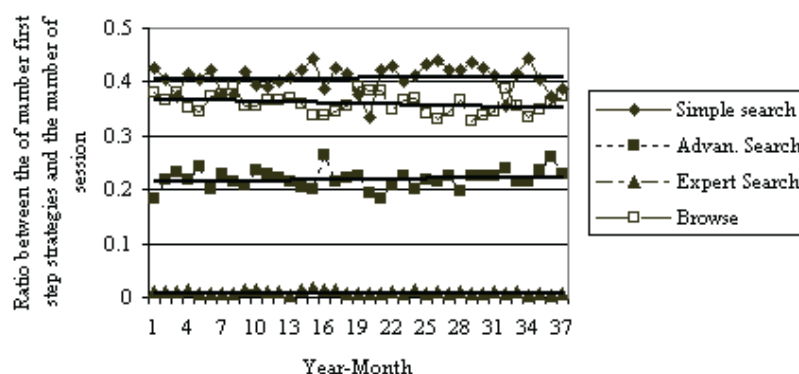


Fig. 9 – Andamento mensile delle tipologie del *primo passo* nell'Emeroteca Virtuale (rapportato al numero di sessioni); nella figura sono indicate anche le linee di tendenza. Con 1 si è indicato il primo mese del periodo considerato (Gennaio 2006), con due il secondo, ecc.

La figura 9 sembra confermare i risultati ottenuti dagli autori citati, i quali riferiscono di un incremento delle ricerche a sfavore del *browsing*.

Tuttavia, qualora si consideri il numero totale di ricerche (indipendentemente dalla loro tipologia e dal fatto che possano essere effettuate al primo o nei successivi passi) e quello relativo al *browsing*, si hanno i risultati riportati nella figura 10.

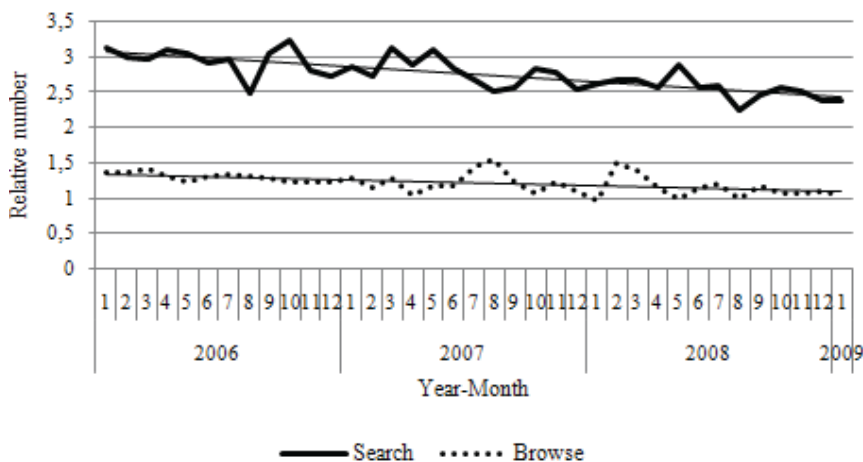


Fig. 10 – Andamento mensile delle ricerche (indipendentemente dalla loro tipologia) e del *browsing*, con riferimento al loro totale tra tutti passi compiuti dagli utenti nel periodo considerato.

Tale figura, in realtà, evidenzia un risultato discordante con il precedente (sia le ricerche sia il *browsing* calano nel tempo, con un decremento maggiore da parte delle ricerche).

Il commento di tale fenomeno risulta difficoltoso data la possibilità di poter fare solo delle supposizioni sui motivi che ne sono causa; in particolare, si potrebbe considerare che gli utenti considerati divengono più efficienti nell'uso dell'Emeroteca Virtuale, riducendo così i passi di ricerca compiuti. Si potrebbe anche ipotizzare che questi tendono anche ad apprezzare la possibilità di *tenersi aggiornati*, ossia di arrivare più rapidamente a navigare i nuovi contenuti delle riviste ritenute d'interesse per le proprie esigenze.

Inoltre, è utile osservare che la possibilità di poter *personalizzare* il proprio ambiente una volta che ci si è autenticati nel sistema consente, ad esempio, di poter *salvare* le ricerche già effettuate. Tale azione, tipica di molti portali Internet, tende a facilitare il *lavoro* degli utenti stessi.

Conclusioni e futuri sviluppi

Il processo di analisi dei comportamenti degli utenti dell'Emeroteca Virtuale del CIBER evidenzia una serie di risultati compatibili con i comportamenti generici di un utente del *World Wide Web*.

In particolare, la durata delle loro sessioni e la tendenza a compiere strategie di raffinamento successivo delle loro ricerche confermano quanto presente in letteratura.

Quel che è interessante è l'osservazione che gli utenti diventano sempre più *efficienti*, ossia diminuiscono il numero d'interazioni con il sistema e giungono prima al termine della loro sessione (si può, quindi, affermare che trovino prima l'oggetto del loro interesse).

Certo è che una questione fondamentale riguarda anche l'enorme quantità d'informazione offerta dal sistema (sono milioni gli articoli presenti nella sola Emeroteca Virtuale); a questa andrebbe anche aggiunta anche quella già presente in Internet.

Da un punto di vista descrittivo sarebbe opportuno approfondire i risultati presentati, anche arricchendoli con contemporanee indagini di tipo qualitativo (questionari), onde verificare quali siano realmente le modificazioni nelle abitudini comportamentali e se l'utente percepisca il sistema come *efficace* nell'individuare l'oggetto d'interesse o se, a volte, subisca la sovrabbondanza d'informazione.

Da un punto di vista del sistema offerto, invece, si potrebbero forse modificare gli strumenti di ricerca e il modo mediante il quale i loro risultati sono forniti. Una possibilità in tal senso potrebbe essere rappresentata dal fornire agli utenti la possibilità di *lasciare un commento* a un determinato articolo; in tal modo le operazioni di ricerca potrebbero far apparire sia la lista degli articoli più *vicini* alle parole ricercate sia, per ognuno di essi, gli eventuali commenti lasciati da altri.

Ciò consentirebbe una facilitazione alla comprensione della reale utilità dell'oggetto che si è scelto di visualizzare (a volte assieme alla lettura di un articolo si potrebbero vedere i commenti onde verificare se altri lo hanno ritenuto utile, magari superato, magari innovativo, ecc.).

Una tale possibilità, inoltre, potrebbe essere vista come il non lasciare che le idee scritte rimangano ferme; infatti i commenti potrebbero risultare anche utili per continuare uno studio, magari con aggiunte, modifiche, domande e risposte successive che vedono anche lo stesso autore impegnato in tale processo.

Con riferimento, inoltre, a quanto esposto nell'introduzione, si potrebbe così avere una maggiore *umanizzazione* del contenuto dell'Emeroteca Virtuale, in quanto in essa gli utenti avrebbero la possibilità di scambiarsi informazioni e suggerimenti.

Bibliografia

1. Corriere della Sera, 7 Febbraio 1998, pagina 34
<archiviostorico.corriere.it/1998/febbraio/07/Enciclopedia_morte_dei_commessi_viaggiatori_co_0_9802078575.shtml>.
2. Jansen, B. J. (1998) "Real Life information Retrieval: a study of user queries on the web", *ACM SIGIR Forum*, Volume 32, 5-17 <jimjansen.tripod.com/academic/pubs/sigirforum98/forum98.html>.
3. Spink, A. et al. (2001), "Searching the Web: the public and their queries", *Journal of the American Society for Information Science and Technology*, Vol. 52, 226-243
<jimjansen.tripod.com/academic/pubs/jasist2001/jasist2001.pdf>.
4. Wolfram, D.; Xie, H. (2002). "Traditional IR for web users: a context for general audience digital libraries", *Information Processing and Management*, Vol. 38, 627-648 <periodici.caspur.it/cgi-bin/sciserv.pl?collection=journals&journal=03064573&issue=v38i0005&article=627_tifwuacfgadl&form=pdf&file=file.pdf>
5. Evans, J. A. (2008). "Electronic Publication and the Narrowing of Science and Scholarship". *Science* 321, no. 5887, 395-399.
6. Tenopir, C.; King, D. W.; Edwards, S.; Wu, L. (2009). "Electronic Journals and Changes in Scholarly Article Seeking and Reading Patterns", *Aslib Proceedings: New Information Perspectives*, 61 (1), 5-32
<www.dlib.org/dlib/november08/tenopir/11tenopir.html>.
7. Swanson, D.R. (1977). "Information retrieval as a trial-and-error process", *Library Quarterly*, Vol.: 47, 128-148.
8. Spink, A. et al. (1998). "Modelling users' successive searches in digital environments", *D-Lib Magazine*
<www.dlib.org/dlib/april98/04spink.html>.
9. Gayo-Avello, D. (2009). "A survey on session detection methods in query logs and a proposal for future evaluation", *Information Sciences*, Vol. 179, issue 12, 1822-1843 <periodici.caspur.it/cgi-bin/sciserv.pl?collection=journals&journal=00200255&issue=v179i0012&article=1822_asosdmaapffe&form=fulltext>.
10. Silverstein C., et al. (1999). "Analysis of a very large Web search engine query log", *ACM SIGIR Forum*, Vol.: 33, 6-12 <www.sigir.org/forum/F99/Silverstein.pdf>.
11. Spink, A.; Jansen, B. J. (2004). *Web search: Public searching of the Web*, Dordrecht: Kluwer eds.
12. Goker, A., & He, D. (2002). "Analysing Web search logs to determine session boundaries for user-oriented learning", *Proceedings of the international conference on adaptive hypermedia and adaptive Web-based systems*, London: Springer-Verlag, 319-322.
13. Scarnò M. (2009) User's behaviour inside a digital library, in atti di QQML2009: Qualitative and Quantitative Methods in Libraries, International Conference
<www.isast.org/proceedingsQQML2009/PAPERS_PDF/Scarno-User_behaviour_inside_a_digital_library_PAPER-QQML2009.pdf>.



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 2.5 License](https://creativecommons.org/licenses/by-nc-sa/2.5/).