

MODULO 3 – DISCIPLINE E TIC
LABORATORIO “I MOTORI DI RICERCA”

Autore Roberto Trincherò

Premessa

L3.1.1. Cosa sono i motori di ricerca?

I motori di ricerca sono grossi archivi interrogabili dall'utente che ci dicono dove è possibile reperire informazioni in rete su argomenti specifici. L'interrogazione viene fatta attraverso delle *parole chiave*, ossia dei termini che identificano l'argomento che ci interessa. Le parole chiave vengono inserite in un'unica casellina e la ricerca viene effettuata in tutto il testo della pagina Web indicizzata.

L3.1.2. Come vengono acquisite le informazioni dai motori di ricerca?

I motori di ricerca indicizzano le informazioni presenti sul Web mediante programmi automatici detti *spider*. Questi programmi vengono attivati a richiesta da chi desidera farsi indicizzare sul motore di ricerca stesso, oppure dai gestori del motore per indicizzare periodicamente determinati domini. Esempi di motori di questo tipo sono Google www.google.it, AllTheWeb www.alltheweb.com e Altavista www.altavista.com

L3.1.3. Le *Web directories*

A differenza dei motori veri e propri, le *Web directories* forniscono elenchi di indirizzi Web selezionati da esperti e ordinati per categoria. Ad esempio, se cerchiamo su una *Web directory* link relativi ai metodi di studio, li troveremo probabilmente all'interno delle categorie Istruzione/Scuola/. Esempi di siti di questo tipo sono Yahoo www.yahoo.it, Arianna www.arianna.it e Virgilio www.virgilio.it.

L3.1.4. Motori di ricerca generalisti e specialistici

I motori di ricerca si dividono poi in motori di ricerca *generalisti*, che indicizzano informazioni relative a tutti gli ambiti della conoscenza, e *specialistici*, che indicizzano informazioni relative ad ambiti specifici (ad esempio la medicina, la biologia, ecc.).

Figura – Alcuni motori di ricerca e Web directories, generalisti e specialistici (in grassetto i più utilizzati)

Motore	Indirizzo http
37.com	www.37.com
AllTheWeb	www.alltheweb.com
Altavista	www.altavista.it
Arianna	arianna.libero.it
Ask Jeeves	www.askjeeves.com
Britannica.com	www.britannica.com
Diabolos	www.diabolos.com
Education World	www.education-world.com
Euroseek	www.euroseek.com
Excite	www.excite.it
Final Search	www.finalsearch.com
FindInfo	www.findinfo.com
Godado	www.godado.it
Google	www.google.it
HotBot	www.hotbot.lycos.it
Il Trovatore	www.iltrovatore.it
LookSmart	www.looksmart.com
Lycos	www.lycos.it
Mamma	www.mamma.com
MSN Web Search	www.msn.it
Nbci	www.nbci.com
Net Search	www.netsearch.org
Nomade	www.nomade.fr
Northern Light	www.northerlight.com
Point Communications	www.pointcom.com
Scirus	www.scirus.com
Search4science	www.search4science.com
SearchUp	www.searchup.com
Seek On	www.seekon.com
ShareLook	www.sharelook.it
ShinySeek	www.shinyseek.it
TuttoNet	www.tuttonet.com
Virgilio	www.virgilio.it
Voilà	www.voila.fr
WebCrawler	www.webcrawler.com
WiseNut	www.wisenut.com
Yahoo.it	www.yahoo.it

L3.1.5. I metamotori

I *metamotori* sono dei motori di ricerca che cercano le parole chiave digitate su più motori di ricerca. Tra i metamotori più famosi segnaliamo Ez-Find www.theriver.com/TheRiver/Explore/ezfind.html, Find-It www.itools.com/find-it/find-it.html, Copernic www.copernic.com.

L3.1.6. I metasiti e i portali

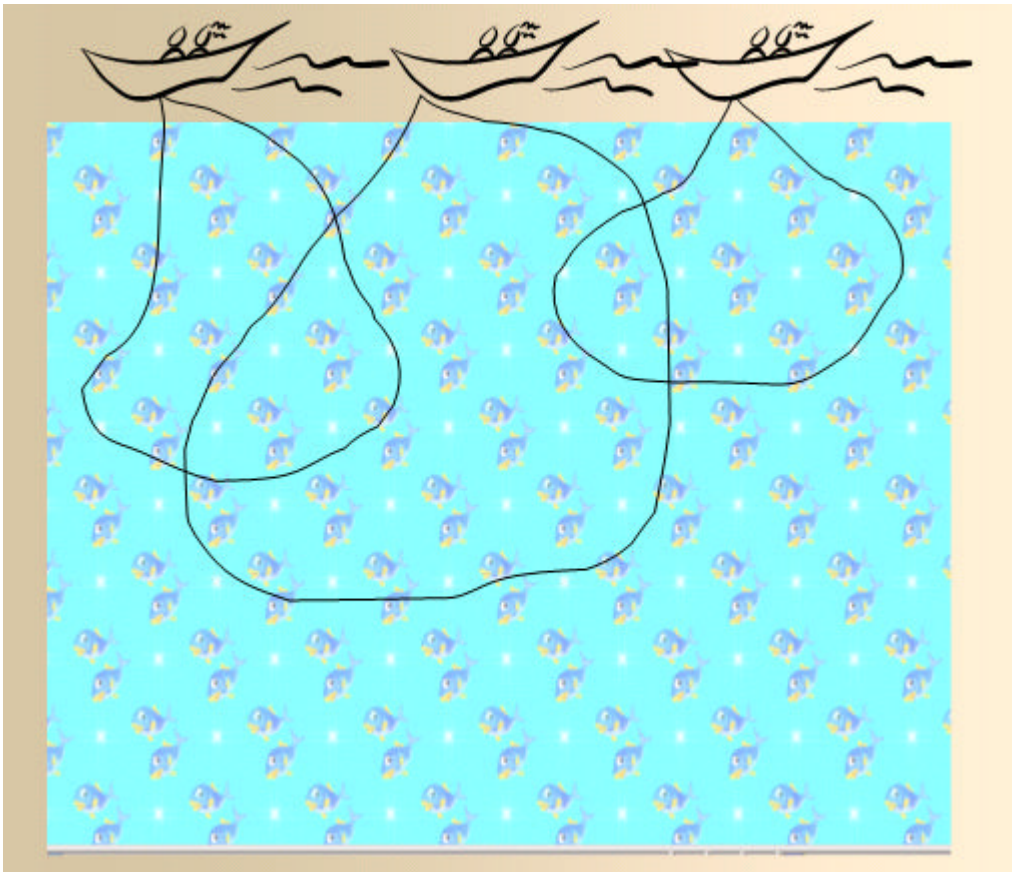
I metasiti sono siti che contengono raccolte ragionate e commentate di altri siti, compilate da esperti del settore. Un esempio di metasito è la sezione Didattica di Formazioneadulti.it www.formazioneadulti.it/links/didattica.htm. I portali sono siti che offrono servizi multipli e che consentono il rilancio verso altri siti in modo ragionato. Un esempio di portale di risorse didattiche è il portale Far – Formazione Aperta in Rete www.far.unito.it.

L3.1.7. Il miglior motore è sempre l'ultimo uscito

Il panorama dei motori di ricerca cambia molto rapidamente. Nuove tecnologie rendono obsoleti motori molto usati fino a poco tempo prima. E' pertanto importante essere costantemente aggiornati sui nuovi motori che compaiono in rete e sulle loro prestazioni. Consigliamo quindi la periodica consultazione dei siti www.motoridiricerca.it www.cerca.com e www.searchenginewatch.com.

L3.1.8. Come funzionano i motori di ricerca?

Ciascun motore segue criteri propri nella scelta delle pagine da indicizzare. Cercare tramite parole chiave su un solo motore non garantisce di trovare tutto ciò che è presente su Web. Motori di ricerca diversi trovano spesso pagine diverse. Una ricerca lanciata con un motore convenzionale corrisponde a pescare con una rete. Ogni motore è una rete diversa. E' buona pratica quindi cercare utilizzando più motori di ricerca.



L3.1.9. Web Profondo (Deep Web)

Web Profondo (*Deep Web*) è un termine per indicare quella parte del Web che i comuni motori di ricerca non possono indicizzare. I motori di ricerca convenzionali indicizzano solo le pagine Web "statiche", ossia quelle che risiedono sui server in formato Html/XML, ma non quelle dinamiche, generate solo su specifica richiesta dell'utente, ad esempio quando si interroga un database (pagine Asp o Php).

Per saperne di più:

<http://library.albany.edu/internet/deep.html>

<http://houtek.com/Deep%20Web.htm>

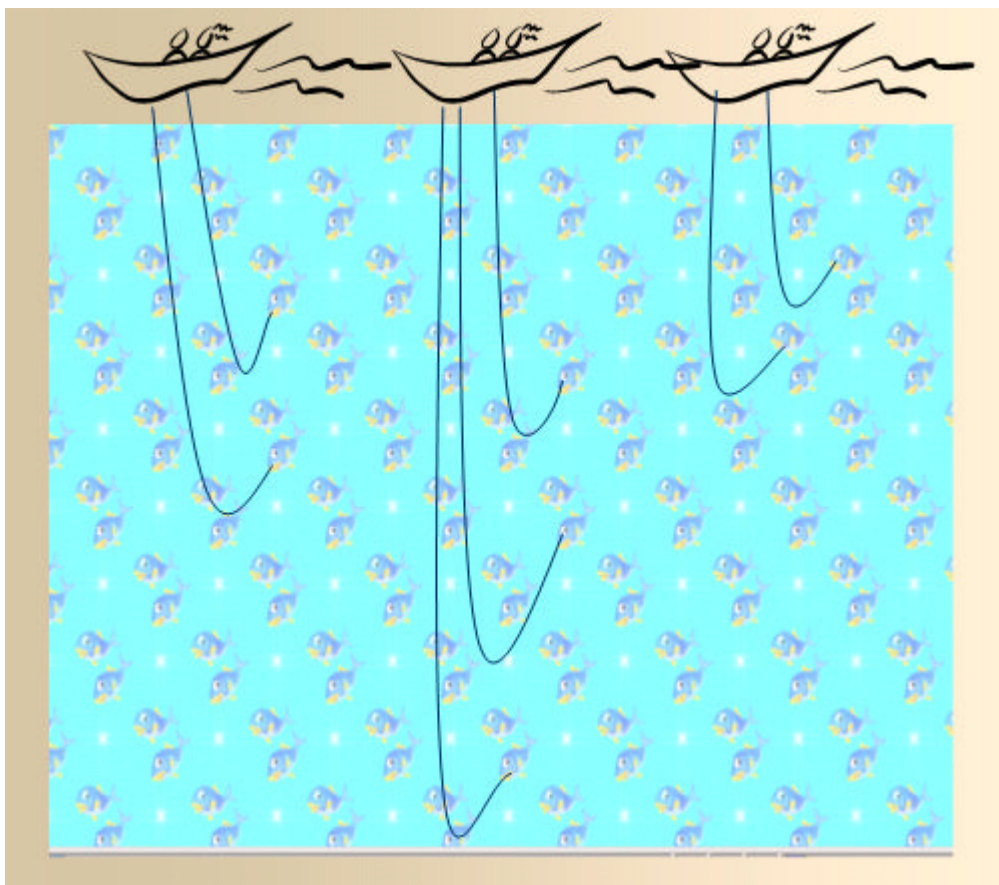
<http://websearch.about.com/cs/invisibleweb/index.htm>

L3.1.10. Cosa si trova nel Web Profondo?

Cercando nel Web Profondo si possono trovare: files non Html/XML (ad esempio files Pdf, Doc, ecc.), databases o archivi consultabili via web (riviste, newsletters, ecc.), siti con accesso protetto tramite login e password, pagine Web create dinamicamente, strumenti interattivi (es. programmi Javascript, applet Java, ecc.), programmi software scaricabili.

L3.1.11. Come funzionano i motori Deep Web?

Cercare nel Web Profondo significa rintracciare l'informazione nel luogo dove è presente. Questo è possibile tramite: motori di ricerca specialistici *Web directories* che raccolgono link a materiali selezionati manualmente, motori dedicati al Deep Web, i quali indicizzano o lanciano interrogazioni su databases anziché indicizzare pagine Web classiche.



L3.1.12. Il programma TeeMot

TeeMot è un programma, sviluppato presso il Dipartimento di Scienze dell'Educazione e della Formazione dell'Università degli studi di Torino, che consente la ricerca su più motori di ricerca di parole chiave inserite dall'utente, parole presenti nel Thesaurus Europeo dell'Educazione www.bdp.it/rubriche/bdpnews/tee.htm oppure combinazioni di esse. Il programma è liberamente fruibile all'indirizzo www.far.unito.it/teemot.

L3.1.13. Strategie di ricerca su Web (1/11)

Se le parole chiave per la ricerca non sono calibrate molto bene il rischio è quello di trovare troppi documenti, molti dei quali poco attinenti con le nostre reali esigenze. I motori di ricerca sono dotati di algoritmi che calcolano automaticamente la pertinenza del contenuto delle pagine trovate alle parole chiave inserite dall'utente. Più parole chiave vengono inserite più il campo di ricerca viene ristretto.

L3.1.14. Strategie di ricerca su Web (2/11)

E' possibile delimitare il campo di ricerca inserendo più parole chiave. L'obiettivo è quello di giungere a ricerche che diano poche decine di risultati, un numero limitato che consenta l'esplorazione delle pagine trovate una per una da parte di chi compie la ricerca. Suggerimenti sulle parole chiave utilizzabili possono venire dall'uso di tesauri.

L3.1.15. Strategie di ricerca su Web (3/11)

Una possibile strategia è quella di partire con un numero elevato di parole chiave, ad esempio *"strategie per ottenere un apprendimento significativo negli allievi della scuola secondaria"*, e poi ridurle se i risultati trovati sono troppo pochi. I risultati della prima ricerca

saranno probabilmente molto pochi ma, se la parole chiave sono state scelte in modo opportuno, saranno molto pertinenti col tema che ci interessa.

L3.1.16. Strategie di ricerca su Web (4/11)

Alcuni motori (es. Google) offrono la possibilità di cercare una frase esattamente come è scritta (ossia con quell'esatta sequenza di caratteri, spazi inclusi), racchiudendo la frase tra virgolette (ossia doppi apici ") quando viene digitata nella casella di ricerca. Altri consentono di attivare questa opzione all'interno di una schermata di "Ricerca avanzata".

L3.1.17. Strategie di ricerca su Web (5/11)

Se i risultati della ricerca sono troppo pochi, nei motori che lo prevedono (come ad esempio Altavista, in modalità Ricerca avanzata), è possibile usare l'operatore NEAR (es. scrivendo *strategie NEAR apprendimento NEAR significativo NEAR "scuola secondaria"*). Tale operatore indica al motore di trovare i documenti che hanno le parole chiave indicate all'interno di un intervallo di 10 parole.

L3.1.18. Strategie di ricerca su Web (6/11)

Se i risultati sono ancora pochi, è possibile nei motori di ricerca che lo prevedono (es. Altavista, in modalità Ricerca avanzata) utilizzare l'operatore AND (es. scrivendo *strategie AND apprendimento AND significativo AND "scuola secondaria"*), oppure scrivendo le parole chiave e selezionando l'opzione "Cerca tutte le parole" (es. Arianna, in modalità ricerca avanzata).

L3.1.19. Strategie di ricerca su Web (7/11)

L'operatore AND consente di trovare i documenti che hanno tutte quelle parole chiave all'interno del documento, in qualsiasi posizione esse si trovino. Nei motori di ricerca che non prevedono opzioni di ricerca avanzata (e sono la maggior parte) l'operatore AND è quello di *default*, ossia quello utilizzato in assenza di altre specificazioni.

L3.1.20. Strategie di ricerca su Web (8/11)

Se non si è trovato ciò che serve occorre chiedersi se si sono usate le migliori parole chiave possibili, e provare a ridurre il numero (ad esempio cercando *apprendimento significativo scuola secondaria*) o provare con parole chiave diverse (es. *studiare in modo efficace nella scuola media superiore*).

L3.1.21. Strategie di ricerca su Web (9/11)

Le parole chiave utilizzate sono legate al tipo di lessico utilizzato nei documenti che si cercano: le parole chiave *apprendimento significativo scuola secondaria* faranno trovare probabilmente documenti redatti da esperti di apprendimento (il termine *apprendimento significativo* non è comune nel linguaggio ordinario) e rivolti ad altri esperti; *studiare in modo efficace nella scuola media superiore*, porterà a trovare con maggior probabilità documenti rivolti agli studenti.

L3.1.22. Strategie di ricerca su Web (10/11)

Se ancora non si è trovato nulla sull'argomento che interessa, si può provare a cambiare motore di ricerca (motori di ricerca differenti trovano spesso cose differenti) o provare a cercare documenti in lingua diversa (es. inglese o francese), scrivendo le parole chiave in quella lingua.

Figura - Una possibile strategia per l'individuazione di parole chiave

L3.1.23. Strategie di ricerca su Web (11/11)

Una strategia diversa e complementare è quella di individuare ed esplorare metasiti, *Web directories* e portali. I metasiti e i portali, oppure di cercare l'informazione sui motori di ricerca specialistici o *Deep Web*. Queste risorse possono essere trovate con i motori di ricerca generalisti.

L3.1.24. Algoritmi di pertinenza

Quasi tutti i motori citati usano algoritmi di pertinenza del contenuto e presentano per prime le pagine trovate in cui le parole chiave siano: a) maggiormente frequenti nel testo; b) presenti nei titoli; c) più vicine tra di loro all'interno del testo; d) contenute in pagine visitate da più persone e più frequentemente (e quindi presumibilmente maggiormente inerenti al tema della ricerca).

L3.1.25. Valutare la qualità dell'informazione trovata

Non tutte le informazioni presenti sul Web sono affidabili o di buona qualità scientifica. E' necessaria una valutazione che può avvenire sulla base dell'autorevolezza delle fonti, dell'autorevolezza dei riferimenti teorici, della scientificità e della chiarezza dell'esposizione, della possibilità di interazione con chi ha creato il sito.

L3.1.26. Autorevolezza delle fonti

E' data dal sito che ospita le pagine (l'informazione trovata su un sito corrispondente ad un dominio universitario è sicuramente più autorevole rispetto a quella trovata su un server che offre pagine Web gratuite) ma anche dal prestigio scientifico degli autori (è utile in molti casi cercare in rete informazioni sugli autori per vedere quali sono le loro pubblicazioni, dove lavorano, ecc.);

L3.1.27. Autorevolezza dei riferimenti teorici

E' data dalla qualità e dalla pertinenza dei riferimenti teorici che gli autori citano all'interno del testo. Se un testo cita, in modo pertinente, autori prestigiosi che si sono occupati di un dato tema, è probabilmente un testo di buona qualità scientifica.

L3.1.28. Scientificità dell'esposizione

E' data dall'utilizzo appropriato e coerente del linguaggio scientifico, dalla controllabilità delle affermazioni espresse, dal mettere a disposizione del lettore tutti gli elementi per comprendere i passaggi logici del testo.

L3.1.29. Chiarezza dell'esposizione

E' data dall'attenzione all'uso di un linguaggio comprensibile dall'utenza a cui il testo si rivolge. Scientificità non vuole dire ermetismo: l'uso di un linguaggio chiaro e sintetico, dove i

“tecnicismi” vengono utilizzati solo quando è indispensabile è un buon indicatore di qualità del documento.

L3.1.30. Possibilità di interagire con chi ha creato il sito

La presenza dell'e-mail dell'autore consente di ottenere da lui dei feedback e maggiori chiarimenti sull'argomento, quindi, unito agli altri requisiti, è un buon indicatore di qualità dell'informazione.

L3.1.31. Tesauri

I tesauri sono raccolte di parole chiave, organizzate in una struttura gerarchica, che facilitano l'individuazione di opportune parole chiave per la ricerca su Internet e su banche dati. Attraverso il tesoro è possibile identificare parole chiave situate nello stesso ambito semantico di un dato termine ed effettuare ricerche molto più efficaci.

Figura – Esempio di tesoro strutturato gerarchicamente (tratto dal Thesaurus Europeo dell'Educazione)

(01) Insegnamento

attività scolastiche

- RT apprendimento (02)
- RT lavoro (33)
- RT metodo di studio (02)
- . attività creative
 - RT attività artistica (15)
- . . laboratorio creativo
- . attività guidate
 - . . prelettura
 - . . prescrittura
- . compiti di casa
- . conversazione
- . corrispondenza scolastica

Attività

• Scheda attività 1. Cercare informazioni su Web

1. Scegliete un tema, legato o meno al vostro ambito di insegnamento, identificate un primo nucleo di parole chiave, scegliete un motore di ricerca e cercate materiali su Web inerenti il tema in oggetto.
2. Componete una raccolta ragionata di materiali (almeno 6 siti), classificandoli in fasce di rilevanza con il tema prescelto: alta, media, bassa.
3. Leggendo i materiali della fascia di rilevanza alta identificate un nuovo nucleo di parole chiave che identifichi meglio il tema in oggetto e ripetete la ricerca.
4. Cosa è cambiato tra la prima e la seconda ricerca? Combinazioni diverse di parole chiave portano a ricerche più efficaci? Discutete l'esperienza con i colleghi attraverso il forum.
5. Provate a chiedere a dei vostri colleghi di ripetere l'esercizio sullo stesso tema utilizzando un loro nucleo di parole chiave. I risultati ottenuti sono gli stessi? La classificazione in fasce di rilevanza operata dai vostri colleghi ha portato agli stessi risultati della vostra? Discutete attraverso il forum della vostra esperienza.

• Scheda attività 2. Confrontare più motori

1. Scegliete un tema, legato o meno al vostro ambito di insegnamento, identificate un nucleo di parole chiave, scegliete un motore di ricerca e cercate materiali su Web inerenti il tema in oggetto.
2. Identificate alcuni tra i siti trovati che vi sembrano più rilevanti e componete un elenco.
3. Avviate una ricerca su un altro motore con le stesse parole chiave utilizzate nella ricerca precedente. Potete utilizzare il software TeeMot www.far.unito.it/teemot digitando le parole chiave nella casellina "Ambito / altri descrittori", scegliendo il motore desiderato e premendo il tasto "Vai". I risultati sono gli stessi della ricerca precedente? Avete trovato qualche sito pertinente che prima non avevate trovato? I siti pertinenti precedentemente trovati compaiono? In quale posizione? Discutete l'esperienza con i colleghi attraverso il forum.

• Scheda attività 3. Valutare la qualità dell'informazione trovata.

1. Scegliete un tema, legato o meno al vostro ambito di insegnamento, identificate un nucleo di parole chiave, scegliete un motore di ricerca e cercate materiali su Web inerenti il tema in oggetto.
2. Una volta identificati almeno 6 siti pertinenti valutatene la qualità in base ai criteri definiti nella premessa e divideteli in tre fasce: qualità alta, media, bassa.
3. Proponete i siti trovati ad un collega e chiedete e lui di valutarli in base ai criteri suddetti. Gli esiti della valutazione sono gli stessi? Se differiscono, in cosa differiscono? Discutete le vostre valutazioni attraverso il forum.

• Scheda attività 4. L'utilità dei tesauri

1. Scegliete un tema legato all'educazione o alla didattica, identificate un nucleo di parole chiave, scegliete un motore di ricerca e cercate materiali su Web inerenti il tema in oggetto.
2. Componete una raccolta ragionata di materiali (almeno 6 siti), classificandoli in fasce di rilevanza con il tema prescelto: alta, media, bassa.
3. Consultate il Thesaurus Europeo dell'Educazione www.bdp.it/rubriche/bdpnews/tee.htm (potete anche utilizzare il software TeeMot www.far.unito.it/teemot) e cercate nel tesoro le parole chiave che avete utilizzato o parole chiave che riguardino il tema in oggetto.

4. La consultazione del tesaurus vi ha suggerito altre parole chiave? L'organizzazione tassonomica del tesaurus vi ha suggerito nuove possibilità di ricerca? Ripetete la ricerca sullo stesso motore con le nuove parole chiave trovate e controllate se la pertinenza dei materiali trovati è aumentata. Discutete attraverso il forum della vostra esperienza.

- **Scheda attività 5. Navigare nel Web Profondo**

1. Scegliete un tema, legato o meno al vostro ambito di insegnamento, identificate un nucleo di parole chiave, scegliete un motore di ricerca e cercate materiali su Web inerenti il tema in oggetto.
2. Componete una raccolta ragionata di materiali (almeno 6 siti), classificandoli in fasce di rilevanza con il tema prescelto: alta, media, bassa.
3. Andate sul sito della BrightPlanet www.lexibot.com, scaricate la versione demo di LexiBot, uno dei migliori motori di ricerca per Web Profondo e installatela sul vostro computer.
4. Cercate materiali servendovi di LexiBot e delle parole chiave precedentemente scelte. La ricerca avviene in modo diverso? Che differenza c'è tra i materiali trovati con il motore di ricerca tradizionale e quelli trovati con LexiBot? Il numero, la pertinenza e la qualità dei materiali trovati è differente? Discutete attraverso il forum della vostra esperienza.