

UN MOTORE DI RICERCA “FAI DA TE”

Simone Torsani

Dottorando in Lingue culture e TIC,

Università degli Studi di Genova

s.torsani@gmail.com, + 39-10-209-5512

Summary: The use of the Internet as a tool for teaching and research has led many, since the beginning of the www, to try to define methods to ascertain the real quality of the information retrieved. These methods are the basis for any web resources evaluation but, being rather general, cannot guarantee good results as regards directories focused on specific subjects. The application introduced in this paper makes it possible to create a personalised web directory in which the administrator can define his/her own categories and evaluation sheet.

Key words: web directory, evaluation, web resources, search engine

Riassunto: L'uso della Rete come strumento didattico e di ricerca ha portato molti a definire metodi per determinare la qualità delle informazioni reperite. Questi metodi sono la base per ogni valutazione delle risorse in rete ma, essendo generici, non possono garantire buoni risultati per quanto riguarda raccolte dedicate ad argomenti molto specifici. L'applicazione presentata permette di creare raccolte indicizzate personalizzate, nelle quali l'amministratore può definire le proprie categorie e una propria scheda di valutazione.

Parole chiave: web directory, valutazione, risorse in rete, motore di ricerca

UN MOTORE DI RICERCA “FAI DA TE”

1. Introduzione

La ricerca di materiale in rete è una delle grandi sfide poste dalla diffusione del www. Il problema più ovvio è quello della quantità e della qualità delle risorse disponibili; di qui la necessità di reperire velocemente il materiale corrispondente all'oggetto della propria ricerca¹: a prescindere dai diversi validi sistemi per reperire le informazioni, rimane il problema di raccogliere, ordinarle e condividerle con altri utenti. In campo didattico questo è ancora più vero, dal momento che molte risorse in rete potrebbero risultare davvero preziose proprio per integrare, in ambiti e con metodologie differenti, per integrare l'attività didattica².

Il mezzo più efficace per reperire informazioni pare essere quello delle raccolte indicizzate³, il cui utilizzo dovrebbe precedere logicamente la più nota ricerca per parole chiave; ha infatti il vantaggio di offrire una lista già compilata (seppure parziale e non aggiornata) delle risorse in questione, mentre una ricerca per parole chiave pone, a fronte di un numero di risultati nettamente maggiore, una serie di problemi, come l'effettiva corrispondenza tra la stringa ricercata (ing. *query*) e le risorse ottenute.

Qui sorge però il problema della valutazione: se è sicuro che i siti recensiti sono pertinenti non è altrettanto sicuro che siano attendibili e soprattutto sino a che punto. Per quanto, infatti, la valutazione delle risorse sia stata, fin dagli inizi della diffusione del www, uno dei temi portanti della discussione sulla Rete, le raccolte offrono poche indicazioni su come vengono valutati i siti e sugli eventuali criteri comuni o specifici ad ogni ambito.

Non è facile rispondere in modo universale e nello stesso tempo mirato a simili esigenze, che possono variare notevolmente da contesto a contesto. Ho cercato comunque di farlo sviluppando un modello il più possibile flessibile

¹ Ackermann e Hartman (1998).

² Per una bibliografia su multimedialità e scuola v. la pagina http://puntoeduft.indire.it/160mila/moduli/9/a_9_bibliografia.htm

³ Per un elenco di *directory* di siti v. http://it.dir.yahoo.com/Informatica_e_Internet/Internet/World_Wide_Web/Ricerca_sul_Web/Directory_di_siti

e modificabile, all'interno del quale un utente possa da una parte definire l'ambito e le categorie e dall'altra i criteri sui quali valutare le risorse stesse. L'applicazione permette di creare una raccolta di siti personale sul modello delle grandi raccolte indicizzate presenti in rete, come Yahoo!. Si tratta, naturalmente, di raccolte di dimensioni ridotte e molto concentrate su argomenti molto specifici, nonché ricche di descrizioni e di valutazioni dei diversi aspetti. A differenza delle grandi raccolte, però, queste, meno estese, permettono una maggior flessibilità nella definizione dei criteri.

Prima di illustrare l'applicazione e le sue caratteristiche, è opportuno riassumere le problematiche fondamentali che ne stanno alla base: quelle dell'indicizzazione, della valutazione e della validazione delle risorse.

2. L'indicizzazione, la valutazione e la validazione delle risorse

In una raccolta di risorse, in modo particolare quelle specializzate, la valutazione è strettamente collegata all'indicizzazione. La valutazione di una risorsa è una catalogazione, nel senso che discrimina tra le informazioni valide da quelle non valide (e tutta la scala di valori compresa tra i due estremi), nello stesso tempo, definire l'ambito all'interno del quale una risorsa deve essere collocata determina il metro sul quale questa va valutata. Insieme a un ampio sforzo volto a definire i criteri per una corretta valutazione delle risorse, infatti, gli anni del www hanno visto lo sviluppo di tecniche e metodologie sempre più avanzate per una corretta indicizzazione dei documenti.

2.1 L'indicizzazione delle risorse

L'indicizzazione⁴ delle risorse in rete si è sviluppata in tre diversi momenti: in principio queste venivano catalogate in raccolte dedicate a uno specifico argomento; in un secondo momento furono introdotti la ricerca per parole chiave e i metadati; infine sono utilizzate metodologie per automatizzare sempre di più le operazioni di indicizzazione. Il motivo è semplice: l'aumento esponenziale delle risorse (pagine) in rete⁵. In un

⁴ Sulla catalogazione delle risorse v. Gorman, 1995

⁵ Per una serie di grafici che mostrano lo sviluppo quantitativo del www v. il sito *Hobbes' Internet*

primo tempo, infatti, il numero delle pagine era tale da permettere un'indicizzazione e una valutazione manuale, tale per cui era possibile che persone visitassero e indicizzassero singole pagine o siti. In seguito, quando il numero delle pagine aumentò, furono introdotti mezzi per la ricerca e l'indicizzazione automatica, come i motori di ricerca per parole chiave e i metadati, dati inseriti nelle pagine che potevano essere letti e interpretati dalle macchine⁶. Gli sviluppi odierni seguono, come prevedibile, queste tracce. Da un parte i motori di ricerca per parole chiave come Google fanno uso di tecniche complesse di estrazione dei dati⁷; dall'altra, più interessanti per il discorso presente, sono i cosiddetti portali verticali (detti anche *vortal*⁸) o raccolte verticali (*vector*) che indicizzano risorse di uno specifico campo, rendendo la navigazione più approfondita. Più evoluta la ricerca sui metadati, come il progetto Dublin Core⁹, che ha sviluppato tecniche per descrivere le caratteristiche fondamentali di una risorsa, per esempio sul contenuto. Lo sviluppo più interessante è, in ogni caso, quello legato al web semantico¹⁰ e all'uso di standard XML per descrivere i dati presenti in una risorsa.

Fin qui gli sforzi delle grandi organizzazioni e dei professionisti della rete; ma la questione tocca tutti gli utenti del www che hanno ugualmente percepito il problema della ricerca e della catalogazione delle risorse (pagine).

Timeline v8.1 all'indirizzo

<http://www.zakon.org/robert/internet/timeline/>

⁶ Per un'introduzione ai metadati v. il sito *Metadata for the masses* all'indirizzo

<http://www.ariadne.ac.uk/issue5/metadata-masses/>

⁷ Una delle caratteristiche più note (e più discusse) è il sistema del *link popularity*, per cui più colleganti puntano a una pagina, maggiore sarà il *ranking* di questa pagina. Questo sistema è stato contestato da alcuni perché le pagine più puntate sono naturalmente quelle di grandi organizzazioni, pagine che, per altro, avendo un *ranking* elevato vengono visitate prima di altre: si sarebbe così creato una sorta di ingorgo antidemocratico. Per un approfondimento v. il sito *PageRank: Google's Original Sin* all'indirizzo <http://www.google-watch.org/pagerank.html>.

⁸ Ovvero *vertical portal*, opposto a portale orizzontale o generico, come Yahoo!, che indicizza siti per diverse categorie.

⁹ V. il sito <http://dublincore.org>, per una descrizione in italiano sul funzionamento di Dublin Core v. il sito *Meta-dati* all'indirizzo

<http://www.uniroma1.it/documentation/metadati.html>

¹⁰ V. il sito del gruppo Gruppo Web semantico

all'indirizzo <http://www-dimat.unipv.it/biblio/isko/doc/gruppowebsemantico.htm>

L'introduzione del www, infatti, è stata accolta dal grande pubblico senza una reale preparazione sulle sue potenzialità, per cui la rete è vista come un'enorme biblioteca o deposito di materiali: questo ha favorito un utilizzo che ricalca quello delle biblioteche tradizionali, nel quale, per esempio, la sitografia sostituisce la bibliografia e il testo in rete riproduce nella struttura il testo a stampa e si definisce ipertesto solo per il fatto di essere in Rete. Col che si fa esattamente quello che non si dovrebbe fare e si utilizza male una risorsa che è nata per gestire grandi quantità di informazioni come singole unità connesse e indipendenti¹¹.

Un primo sistema di catalogazione delle risorse, ad uso esclusivamente personale, è la funzione dei preferiti (ing. *favourites*) o segnalibri¹² (ing. *bookmark*) nelle applicazioni per la navigazione (ing. *browser*). Tali funzioni permettono di raccogliere gli indirizzi dei siti di interesse, dividendoli per argomenti e sottoargomenti; la struttura ad albero delle cartelle, che riprende quella dei sistemi operativi, permette agli utenti di creare una loro versione personale della classificazione per argomenti che riprende le caratteristiche proprie della loro attività. Questo diventa ancora più evidente se si trasforma l'unità di misura dal libro/sito all'informazione/indirizzo: in questo caso la catalogazione personale non è più una riproduzione in piccolo della catalogazione per argomenti tipica della biblioteca, ma una raccolta organizzata di informazioni specifiche raccolte intorno a nuclei molto approfonditi.

Un secondo esempio di catalogazione è dato dalla pagina dei collegamenti dei vari siti nella quale gli autori del sito raccolgono gli indirizzi di pagine che abbiano una qualche attinenza con l'argomento del sito in questione¹³: questa

¹¹ È importante tenere presente la relazione tra ipertesto e www, dei quali il secondo è prodotto del primo che prevede una struttura di collegamenti tra diversi blocchi di informazione. Per una bibliografia sull'ipertesto v. la pagina all'indirizzo <http://www.eastgate.com/Bibliography.html>; in particolare v. LANDOW, GEORGE P. (a cura di), *Hyper/Text/Theory*. Baltimora, Johns Hopkins Press, 1994.

¹² Il secondo termine rende maggiormente la concezione ipertestuale dell'informazione.

¹³ Di grande interesse per i docenti sono le pagine personali (anche amatoriali v. sotto) di altri docenti nelle quali si possono trovare idee, spunti o materiali a volte molto originali, stesso discorso per la sezione dei collegamenti che puntano a pagine sicuramente interessanti e originali alle quali si sarebbe magari giunti difficilmente.

opzione pare limitata a coloro che hanno una qualche competenza informatica e che sanno come creare e gestire un sito (in realtà è possibile fare un sito con strumenti rudimentali come un programma di videoscrittura e metterlo in rete in pochi minuti). Attraverso le caratteristiche del linguaggio HTML è possibile creare una raccolta in linea ordinata di siti rilevanti per un dato argomento (non sfuggirà l'uso didattico di quest'attività): data l'estrema mobilità dell'informazione in Rete queste pagine (come tutte le pagine) devono essere aggiornate costantemente.

Come accennato in precedenza, il modello classico per la gestione in rete degli indirizzi sono i portali, siti che guidano l'utente alla ricerca dell'informazione o sito che gli interessa e che si dividono, sommariamente ma in maniera imprecisa, in motori di ricerca per parole chiave (ing. *keyword*) e raccolte indicizzate di siti. Queste ultime, il cui modello classico è Yahoo!, permettono di navigare tra categorie e sottocategorie fino a trovare quella che corrisponde all'argomento ricercato. Inoltre, a differenza della ricerca per parole chiave, il sistema garantisce che il risultato trovato corrisponda effettivamente (perché catalogato da persone reali) a quello che si sta cercando o, meglio, a quello che il nome della categoria indica. L'esempio classico è quello del corso di lingua: ricercando "corso di lingua" in un motore si trovano milioni di pagine di corsi all'estero, di manuali ecc., in una raccolta indicizzata, al contrario, si può scegliere tra un numero ragionevole di siti corrispondenti alla categoria. Come già evidenziato, tale strumento è soprattutto valido per quanto riguarda la catalogazione, dal momento che garantisce l'effettiva corrispondenza di un sito (ma di un intero sito e non di una singola informazione o pagina, v. sotto) con l'argomento della categoria cui appartiene; dal punto di vista della valutazione, al contrario, il sistema si presenta meno efficace: se anche i siti proposti vengono valutati, i criteri di questa valutazione non sono esplicitati, sono necessariamente generici e probabilmente viene valutata la qualità generale del sito e la sua attinenza all'ambito specifico. Va infatti tenuto conto del fatto che la valutazione di una risorsa specifica a un argomento ha bisogno di criteri particolari, oltre a quelli comunemente considerati come fondamentali, come l'attendibilità della fonte o l'obiettività e l'accuratezza dei contenuti.

2.2 La valutazione delle risorse

Abbiamo accennato ai diversi sforzi di quanti hanno tentato di definire criteri di valutazione per le risorse. Uno dei primi e più noti articoli è *The Good, The Bad & The Ugly: or, Why It's a Good Idea to Evaluate Web Sources*¹⁴ del 1996, che è possibile prendere come riferimento ed esempio. Nell'articolo sono riassunti alcuni punti fondamentali che ricorrono nella bibliografia specializzata: autorità (ing. *authority*), accuratezza (ing. *accuracy*), obiettività (ing. *objectivity*), attualità (ing. *currency*) e copertura (ing. *coverage*). L'autorità fa riferimento a una questione fondamentale, ovvero a chi ha preparato la risorsa; dal momento che chiunque può caricare una pagina in rete (a differenza della stampa) è importante notare se è indicato l'autore della pagina e se è una persona qualificata. L'accuratezza mira a stabilire se l'informazione è attendibile o meno, ovvero se la presenza di errori riveli poca accuratezza e, di conseguenza, pregiudichi il contenuto della pagina; come per il punto precedente bisognerebbe rilevare se esiste un editore che controlla la validità delle informazioni della pagina. Valutare l'obiettività della pagina serve, naturalmente, a determinare se contenga pregiudizi o particolari inclinazioni; l'attualità di una pagina indica, in un contesto di elevata obsolescenza dell'informazione (con la Rete è possibile essere, per esempio, maggiormente aggiornati sulle bibliografie specialistiche), se le informazioni siano o meno aggiornate. L'ultimo parametro valuta l'informazione dal punto di vista del contenuto stesso: come, per esempio, l'argomento è trattato. Un'altra analisi è offerta da *Thinking Critically about World Wide Web Resources*, di Esther Grassian¹⁵, dove, oltre agli elementi citati nell'articolo precedente, si accenna anche alla struttura della pagina, come *Does the document follow good graphic design principles? Do the graphics and art serve a function or are they decorative? Do the icons clearly represent what is intended?*

È interessante sottolineare come questi contributi siano articolati in una serie di domande, spesso accompagnate da una spiegazione, in modo da essere molto vicini al modello del questionario e della griglia di valutazione.

Come si nota, inoltre, molti degli elementi da valutare tendano a stabilire se una risorsa possa effettivamente essere considerata tale ed

¹⁴ <http://lib.nmsu.edu/instruction/evalcrit.html#top>

¹⁵ Grassian Esther, *Thinking Critically about World Wide Web Resources*, UCLA College Library <http://www.library.ucla.edu/libraries/college/help/critic/index.htm> (2000)

eventualmente fruibile: poco spazio è dedicato all'effettiva valutazione del contenuto; o meglio, i termini sono molto generici, se paragonati per esempio ai parametri per valutare l'attualità o l'attendibilità della pagina. Eppure il contenuto dovrebbe essere l'elemento più importante da valutare. Questo per il motivo accennato in precedenza: ogni ambito specifico ha bisogno di criteri specifici per la valutazione del contenuto.

2.3 R3

Il progenitore diretto dell'applicazione è il "repertorio ragionato di risorse" R3, progetto nato sul modello delle raccolte indicizzate e specificamente dedicato ai siti francofoni. Di questo strumento esistono due differenti versioni: la prima era un sito "statico", mentre la seconda uno "dinamico"¹⁶. Nella prima versione gli utenti compilavano una scheda di valutazione preparata dal gruppo di ricerca che avrebbe poi corretto e messa in rete, aggiornando la pagina della categoria pertinente il sito in questione: le diverse pagine (per esempio la pagina relativa alla letteratura francese del XVII secolo) erano realizzate ed aggiornate una ad una e collegate al sito principale. Questo sistema era, naturalmente, molto macchinoso (poiché tutte le operazioni non automatizzate sono a rischio di errore) e lento, poiché prevedeva diverse fasi: la comunicazione all'utente della pagina dove si trovava la scheda, la compilazione della scheda, la consegna, la valutazione, l'aggiornamento della pagina e il suo caricamento nel sito. A molti di questi problemi si ovviò con la realizzazione del sistema del nuovo R3 che, sfruttando le tecnologie dinamiche, automatizzava e velocizzava di molto l'operazione, con l'ovvia conseguenza di aumentare di molto il numero delle pagine indicizzate.

Questo sistema permette, attraverso un pannello di controllo, di generare automaticamente le categorie e le pagine che contengono i siti in questione, di compilare, valutare e pubblicare la scheda direttamente in rete, riducendo i tempi delle operazioni necessarie affinché il sito sia effettivamente indicizzato. Dal punto di vista della

¹⁶ Statico e dinamico si riferiscono alle tecnologie utilizzate per la realizzazione di documenti in rete. Le pagine statiche sono come pagine di giornale: non cambiano mai se non quando sono modificate e ricaricate dall'autore, le pagine dinamiche sfruttano determinate tecnologie che ne modificano la struttura e i contenuti a seconda dei dati immessi dall'utente (es, il motore di ricerca genera la pagina dei risultati a seconda delle parole chiave immesse, non esiste certo una pagina per ogni possibile stringa di ricerca).

valutazione R3 risponde pienamente ai requisiti messi in luce prima, come si nota anche dai parametri definiti per la valutazione del sito:

- Navigabilité (max 5) [*clarté de la structure (max 3), navigabilité interne (max 2)*]
- Interface et aspect graphique (max 3)
- Interactivité (max 17) [*dialogue avec l'utilisateur (max 4), quantité des liens (max 4), qualité des liens (max 5), services offerts (max 4)*]
- Contenus (max 22) [*mise à jour (max 2), qualité et exhaustivité des contenus (max 5), richesse des contenus (max 5), organisation des contenus (max 5), offre de matériel - par exemple matériel à télécharger (max 5)*]
- Dénomination (max 3) [*originalité du nom/facilité de mémorisation (max 2), cohérence et facilité d'emémorisation de l'adresse (max 1)*]

Poiché R3 è una raccolta piuttosto generale (sui siti francofoni) non è possibile ipotizzare criteri di ricerca troppo specifici; se fosse però suddivisa in diverse raccolte, ognuna dedicata a un argomento particolare, sarebbe pensabile definire campi di valutazione molto più ristretti e quindi valutare con molta più precisione. Nell'ambito degli esercizi di lingua, per esempio, sarebbe auspicabile verificare il giudizio una volta concluso un determinato compito (se aggiunge commenti o si limita a dare un punteggio) oppure se e quanto sono presenti aiuti di vario tipo a chi fa l'esercizio.

2.4 R3 e la validazione

R3 si basa su un terzo elemento, la validazione. Il sistema, infatti, prevede che la scheda di un sito sia proposta da un utente e, in seguito, avvallata da un'autorità scientifica del settore, garantendo un filtro adeguato alla valutazione generale e, in particolare, dei contenuti. Dal momento, infatti, che i criteri di valutazione del contenuto variano a seconda del campo, l'autorità del settore dovrebbe essere garante della reale correttezza del giudizio.

Come detto in precedenza, infatti, non è chiaro su quali criteri le raccolte indicizzate valutino le risorse¹⁷; né, nel caso si tratti di risorse scientifiche, a quale titolo. I modelli di valutazione citati in precedenza danno importanza,

¹⁷ La directory *dmoz*, illustra brevemente i criteri che utilizza per la valutazione dei siti. Questi criteri non si allontanano da quelli generali esposti in precedenza (v. il sito <http://dmoz.com/guidelines/include.html>).

per esempio, all'obiettività come elemento imprescindibile in un sito: in ambiti potenzialmente ideologizzabili è importante stabilire se una determinata risorsa sia o meno obiettiva; se ne può avere un esempio nella sezione di *Yahoo! Italia* dedicata al fascismo¹⁸, che compare nella categoria "storia"; su sette siti indicizzati due non possono essere raggiunti, i restanti sono chiaramente di parte (quattro di ispirazione fascista e uno di ispirazione antifascista): non che una risorsa di parte non possa essere valida o interessante, ma dovrebbe per lo meno essere indicato il fatto che non si tratta di materiale obiettivo e, in ogni caso, non vi è alcuna garanzia sul valore del contenuto. Questo è un caso estremo, ma illustra in maniera abbastanza evidente la necessità di una garanzia istituzionale sul contenuto.

3. La risorsa

3.1 Caratteristiche principali di FPI2

FPI2 (Flexible Personal Indexing Interface) è l'erede diretto dei due progetti R3 e ne estende le potenzialità al vasto pubblico della Rete e a quanti intendono utilizzare le raccolte di siti come strumento per la didattica o sono interessati ad avere un "motore di ricerca fai da te" (naturalmente il termine motore di ricerca ha un valore più che altro di *marketing*, trattandosi principalmente di raccolte indicizzate di siti, sia pure in presenza di un motore di ricerca interno). Abbiamo visto che uno degli sviluppi dell'indicizzazione dei materiali in rete è costituito dai cosiddetti "portali verticali": FPI2 permette agli utenti di creare un proprio portale verticale dedicato all'indicizzazione e alla valutazione di risorse specifiche.

L'applicazione consente, al pari di tanti altri siti che offrono servizi, di aprire uno spazio personale e personalizzabile in rete nel quale gestire la propria raccolta di siti, sul modello delle directory verticali menzionate in precedenza. Chi apre l'account è anche amministratore del proprio spazio con specifiche prerogative e specifiche responsabilità: crea le categorie della raccolta, valida i siti proposti dai visitatori, tiene aggiornato lo spazio inserendo notizie.

Questa applicazione combina i differenti aspetti citati in precedenza: la catalogazione, la valutazione e la validazione, elementi interconnessi e interdipendenti.

¹⁸ Sul revisionismo storico in rete v. Borrowman (1999)

La definizione di categorie e sottocategorie ha lo scopo di descrivere gli elementi che compongono l'ambito in questione. Attraverso un menu ad albero è possibile modificare l'albero delle categorie, aggiungendo o eliminando categorie e sottocategorie: questo finché non si ottiene una suddivisione confacente alle proprie necessità. Questa disposizione visiva, molto simile a quella ottenibile attraverso la funzione segnalibri del browser (v. sopra), è molto importante dal momento che presenta una disposizione gerarchica e realistica delle categorie e delle sottocategorie, rendendo evidenti eventuali mancanze o difformità nella distribuzione e nella descrizione del campo in questione.

L'altro aspetto fondamentale è quello della scheda di valutazione che definisce i criteri attraverso i quali una risorsa viene valutata.

È possibile aggiungere un numero indefinito di campi di valutazione alla scheda per la catalogazione della risorsa; questi campi, che possono essere spostati di posizione tra loro, sono di diverso tipo: binario, numerico seriale, numerico arbitrario e testuale, tutti, però con valori numerici. Questi quattro tipi dovrebbero essere sufficienti per definire i diversi parametri per valutare un risorsa (anche se basterebbe solo l'ultimo):

1. **Campo binario:** permette di definire una condizione che può essere soddisfatta o meno, dove uno dei due ha un valore numerico arbitrario e l'altro vale zero. Se, per esempio, il fatto che un sito sia multilingue è considerato un fatto importante, si può decidere di assegnare a questo campo un valore consistente in funzione della valutazione finale.
2. **Campo numerico seriale:** permette di definire una serie di valori numerici in sequenza, per esempio, da 1 a 10.
3. **Campo numerico arbitrario:** permette di definire una serie di valori arbitrari tra cui scegliere. Per esempio, 1, 3, 5 ecc.
4. **Campo testuale:** permette di accoppiare valori numerici a stringhe di testo che compariranno invece del nome. Per esempio, insufficiente (4), sufficiente (6), buono (8) ecc.

Quando viene effettuata la valutazione del sito, questi valori sono utilizzati nei calcoli per definire un graduatoria delle risorse nella categoria e ordinare i siti di conseguenza.

L'amministratore, infine, deve validare le schede proposte modificando i valori immessi dall'utente. Il sistema prevede un modello elastico che permette di definire con precisione l'insieme degli elementi fondamentali della raccolta. Permette di definire l'ambito e le categorie che la compongono e i parametri più adatti a valutare le risorse.

Questo, naturalmente, non risolve il problema della validità del contenuto, perché l'unica garanzia, per ora, rimane l'autorità di chi valida una scheda: si forniscono però gli strumenti per una catalogazione e una valutazione più precise.

3.2 Caratteristiche secondarie di FPI2

Oltre alla raccolta di siti il sistema prevede un semplice sistema di gestione delle pagine, a partire dalla pagina principale; questa funzione di gestione della pagina principale permette all'amministratore di inserire notizie, che i visitatori possono, a discrezione dell'amministratore stesso, commentare (l'amministratore può anche cancellare i commenti dalla pagina di amministrazione). Lo scopo di questi elementi esula dall'ambito strettamente didattico e scientifico e introduce questioni propriamente collegate al *marketing*, che non è un fattore da sottovalutare. Gli esperti del settore ripetono, nelle riviste specializzate¹⁹, che un buon sito non deve solo attirare visitatori per una volta ma invogliarli a ritornare e il sistema più sicuro per ottenere questo effetto è quello di aggiornare continuamente il sito stesso: in questo caso l'amministratore dovrà inserire di volta in volta informazioni, presumibilmente collegate all'argomento della sua raccolta di siti. L'interazione con l'utente è l'altro aspetto da tenere presente, in questo senso il sistema permette ai visitatori di lasciare dei commenti sulle notizie riportate. Il sottosistema è in tutto e per tutto un weblog.

3.3 FPI2 e gli altri portali

Esistono molti sistemi per la creazione di portali personalizzati (il più noto è PHPNuke) che permettono di installare, creare e gestire con numerose funzioni un proprio portale: si tratta di applicazioni piuttosto generali, non orientate a un argomento/funzione in particolare, ma il più universali (e per questo modificabili da parte dell'utente). FPI2, al contrario è fortemente

orientata, espressamente elaborata intorno alla raccolta di siti, la parte più elaborata e complessa: le altre parti fanno da supporto e servono per aumentare la vivacità del sito e arricchirlo di funzioni. A differenza di PHPNuke il sistema predilige la semplicità. Per quanto questo vada a scapito di una "editabilità" del sistema, ha il vantaggio di lasciare poco spazio al dubbio e alla difficoltà e di non appesantire o distogliere dall'ambito principale dell'applicazione. Per chi possiede maggiori competenze o un sito cui vuole aggiungere questa funzionalità è disponibile per una versione completa e scaricabile da integrare nel proprio sito, priva però di tutte le funzioni di supporto.

Per chi fosse intenzionato a scaricare il prodotto può essere interessante una panoramica sulle sue caratteristiche tecniche. La tecnologia utilizzata è quella classica PHP/MySQL. Nella versione 5 di PHP sono state ampliate le caratteristiche per lo sviluppo orientato agli oggetti; sebbene queste caratteristiche restino comunque distanti dagli standard di linguaggi come Java, favoriscono grandemente lo sviluppo di applicazioni robuste e complesse. Nei miei lavori ho sviluppato una serie di classi per la creazione e la gestione sia degli elementi più comuni di questo genere di applicazioni come i *form* o le interazioni con *database SQL*, sia di quegli elementi caratteristici delle diverse applicazioni.

Dal punto di vista della fruizione, infine, bisogna ricordare che molta letteratura scientifica si è occupata della flessibilità del contenuto delle raccolte indicizzate (Chen, Magoulas e Dimakopoulos, 2005), auspicando la possibilità per l'utente di modificare in maniera anche consistente le modalità di navigazione: in questo senso l'applicazione permette di scegliere, tra gli altri, il numero dei risultati per pagina e come visualizzare le diverse categorie di risorse.

4. Studio di un caso: esercizi di lingua in rete

L'utilizzo ottimale di questa risorsa avviene in ambiti piuttosto ristretti perché, come detto in precedenza, questo permette di definire campi più specifici che permettono una valutazione più efficiente e precisa: ne è un ottimo esempio una raccolta indicizzata di esercizi di lingua. La scelta di un ambito così specifico mira a illustrare la necessità di definire categorie e criteri di valutazione particolari.

¹⁹ v. Holzschlag, 2004 pp. 330-354, tra le riviste v., tra le altre, *Internet magazine* n. 110 novembre 2003 pp.22-31

È importante definire però quali risorse vadano indicizzate. La risposta più ovvia sembrerebbe quella di indicizzare siti, ma questo, a uno sguardo più attento, va contro il carattere della rete e dell'ipertesto e, più in particolare, della ricerca che deve essere il più veloce e precisa possibile. Se anche non è ipotizzabile indicizzare singole informazioni (o nodi, nella terminologia dell'ipertesto), è certamente più vantaggioso indicizzare singole pagine relative alla categoria. Esistono in rete molti siti che contengono molti esercizi: per quanto indicizzare uno di questi siti potrebbe essere utile, anche in un ambito piuttosto ristretto (come quello degli esercizi di lingua), sarebbe molto più utile indicizzare i singoli esercizi nelle rispettive categorie.

Il primo passo consiste dunque nella definizione della gerarchia delle categorie. La definizione di categorie e sottocategorie è molto importante dal momento che permette di definire una mappa (una delle possibili) attraverso cui è descrivibile un argomento; non è necessario avere le idee troppo chiare, dal momento che è possibile modificare anche in seguito la propria raccolta (e spostare eventualmente gli indirizzi da una categoria all'altra quando se ne elimina una). Nel caso specifico si effettua per prima cosa la definizione di categorie principali corrispondenti, ad esempio, agli argomenti principali dei manuali come morfologia, sintassi ecc., quindi sottocategorie per ognuna di queste, come esercizi sui tempi e, ancora, esercizi sui singoli tempi. Queste categorie definiscono i livelli e le competenze che l'esercizio dovrebbe aiutare a sviluppare.

La seconda fase consiste nel definire i campi di valutazione della scheda, a seconda dei tipi di campi disponibili. Per esempio può essere utile stabilire se (e quanto) gli esercizi sono corredati da risorse come glossari/dizionari, suggerimenti ecc., se sono graduati, efficaci, se lavorano sulle eccezioni, se spiegano bene in cosa consiste l'esercizio.

Una possibile scheda di valutazione potrebbe essere:

- l'esercizio è chiaro: sì (1), no (0)
- numero di elementi presenti (0-20)
- l'esercizio fornisce un giudizio: nessuno (0), limitato (2), adeguato (4), ampio(6)
- l'esercizio fornisce suggerimenti o strumenti che aiutano l'utente: nessuno (0), limitati (2), adeguati (4), ampi (6)
- l'esercizio tiene conto delle eccezioni della regola o dell'argomento trattato: no (0), poco (1), molto(2)

- l'esercizio copre l'argomento in questione: no (0), poco (3), molto(6)
- l'esercizio è: troppo corto (1), troppo lungo (2), adeguato (3)

Come si nota tutti questi fattori sono propri della valutazione degli esercizi ed esulano dai criteri comunemente definiti per valutare un sito, come l'attendibilità della fonte. Questo è il vero punto di forza dell'applicazione: la sua adattabilità a contesti differenti. Inoltre, la possibilità di connettere valutazioni verbali a valutazioni numeriche permette agli utenti di definire una caratteristica piuttosto che assegnare un numero. Va notata, infine, la possibilità di stabilire valori numerici diversi a seconda dell'importanza del campo: in questo caso elementi come la copertura dell'argomento o il giudizio finale sono più importanti della chiarezza o della lunghezza dell'esercizio. Si tratta, naturalmente, di una valutazione personale ma allo stesso tempo molto precisa e pertinente all'argomento: una raccolta generica difficilmente renderebbe conto di argomenti così importanti nella valutazione di un esercizio di lingua.

Abbiamo citato, all'inizio del contributo, il caso di pagine personali di docenti e notato come queste si sottraggano parzialmente dalle regole solitamente stabilite per la valutazione di una pagina, come la cura tecnica. Esistono oggi numerose applicazioni che per la realizzazione di test ed esercizi esportabili nel formato HTML; non è un caso, infatti, che la maggior parte degli esercizi di lingua presenti in rete siano realizzati con *Hot Potatoes* e inseriti in pagine personali di docenti di lingua. Si tratta di un aspetto molto interessante, dal momento che difficilmente nei grandi motori di ricerca pagine del genere potrebbero trovare spazio, mentre potrebbero essere molto utili per l'attività didattica come materiali già pronti. Inoltre, dal punto di vista dell'educazione alla valutazione²⁰, richiedere di valutare degli esercizi permetterebbe di fornire metodologie che rendono gli utenti più consapevoli dei meccanismi che governano il rapporto tra esercizio e apprendimento delle competenze (in altre parole a capire se un esercizio può essere utile), ma anche, e soprattutto, in ambiti di autoapprendimento o di aggiornamento delle competenze.

²⁰ V. per esempio Cottrell, 2001

Bibliografia

- ACKERMANN, E. C., HARTMAN K., *Searching and Researching on the Internet and the World Wide Web*, Wilsonville, OR: Franklin Beedle, 1998.
- ALEXANDER J. E., TATE M. A., *Web Wisdom: How to Evaluate and Create Information Quality on the Web*, Lawrence Erlbaum Associates, Inc., 1999.
- ARNOLD J., ANDERSON E. ., *Dangling by a slender thread: the lessons and implications of teaching the world wide web to freshmen*. Journal of Academic Librarianship, January 1998, pp. 43-52.
- AUER N., *Bibliography on Evaluating Internet Resources*, Emergency Librarian 25(5) May-June 1998, pp. 23-24.
- BARRON A., KAREN I., *An Internet Research Model in: Call of the North, NECC '96*. Proceedings of the Annual National Educational Computing Conference (17th, Minneapolis, Minnesota, June 11-13, 1996) pp 26-29. ERIC DOCUMENT: ED398880.
- BLANDY S., GRISWOLD P., O'BRIEN L., *As the Cursor Blinks: Electronic Scholarship and Undergraduates in the Library*. Library Trends 44(2) Fall 1995, pp. 279-305.
- BODI SONIA, *Scholarship or Propaganda: How Can Librarians Help Undergraduates Tell the Difference?*, Journal of Academic Librarianship January 1995, pp. 21-25.
- BRANDT D. S, *Evaluating Information on the Internet*. (Techman's Tech Page) Computers in Libraries 16(5) May 1996, pp. 44-47. (see internet site above)
- BRANDT D. S, *Relevancy and Searching the Internet*, Computers in Libraries 16(8) Sept. 1996, pp. 35-38.
- BRANDT D. S, *Constructivism: Teaching for Understanding of the Internet*, Communications of the ACM 40(10) October 1997, pp. 112-117.
- BORROWMAN S., *Critical Surfing: Holocaust Denial and Credibility on the Web*, College Teaching 47.2 (Spring 1999), pp. 44-47.
- BURBULES N., *Paradoxes of the Web: the Ethical Dimensions of Credibility*, Library Trends 49 (3) Winter 2001, pp. 441-453.
- CHEN S.Y, MAGOULAS G.D. DIMAKOPOULOS D., *A Flexible Interface for Web Directories to Accomodate Different Cognitive Stiles*, in *Journal of the American Society for Information Science and Technology*, pp. 70-83, 2005
- CONNELL, H. T., TIPPLE J. E., *Testing the Accuracy of Information on the World Wide Web Using the Alta Vista Search Engine*, Reference & User Services Quarterly. Summer 1999, pp. 360-368.
- COLLINS B., *Beyond Cruising: Reviewing*, Library Journal February 15, 1996, pp. 122-124.
- COOKE A. NEAL-SCHUMAN, *Authoritative Guide to Evaluation Information on the Internet*. Neal-Schuman. 1999.
- Jacobson T. (ed.), *Critical Thinking and the Web: teaching users to evaluate Internet Resources*, Active Learning Series #4. Pittsburgh: Library Instruction Publications, 2000.
- COTTRELL J. R., *Teaching Students to Evaluate Web Sources More Critically: Implications from a Faculty Workshop*. College & Research Libraries News 62(2) February 2001, pg 141-143, 186.
- CUNNINGHAM S., *Teaching Students to Critically Evaluate the Quality of Internet Research Resources*, SIGCSE Bulletin June 1997, pgs.31-34, 38.
- DORAN K., *The Internet: Helping Library Patrons Understand What the Internet Is Not (Yet)*, Computers in Libraries June 1995, pp. 22-26.
- DROBNICKI J.A., ASARO, R., *Historical fabrications on the Internet: recognition, evaluation, and use in bibliographic instruction*, Reference Librarian, v.74, 2001, pp. 121-64.
- FITZGERALD M., *Critical Thinking: Tools for Internet Information Evaluation*, in *Information Rich but Knowledge Poor? Emerging Issues for Schools and Libraries Worldwide. Research and Professional Papers Presented at the Annual Conference of the International Association of School Librarianship Held in Conjunction with the Association for Teacher-Librarianship in Canada* (26th, Vancouver, British Columbia, Canada, July 6-11, 1997), 1997.
- FITZGERALD M. A., *Misinformation on the Internet: Applying Evaluation Skills to Online Information*, Emergency Librarian 24(3) January-February 1997, pp. 9-14.
- FITZGERALD M. A., GALLOWAY C., *Relevance Judging, Evaluation, and Decision Making in Virtual Libraries: A Descriptive Study*, in *Journal of the American Society for Information Science & Technology*. 52(2) Oct 2001, pp. 989-1010.
- FOX L. M., *Some On-line (and Off-line) Resources for Evaluating Information on the World Wide Web*, Colorado Libraries Summer 1996, pp. 46-47.
- FRITC, J. W., CROMWELL R. L., *Evaluating Internet Resources: Identity, Affiliation, and Cognitive Authority in a Networked World*, in *Journal of the American Society for Information Science and Technology*, vol. 52, no. 6 (Apr 2001), pp 499-507.
- GARDNER S. A., BENHAM H, H., NEWELL B. M., *Oh, what a tangled web we've woven! Helping students evaluate sources* English Journal, vol. 89, no. 1 (Sept 1999), pp 39-44.
- GARMAN N., *When Online Is Not Enough*, Online May/June 1995, pp. 6-7.
- GARNES D., MILLS C. V., *Finding and Evaluating Health Sources on the Internet: An Overview*, The Reference Librarian no. 74 (2001) pp. 177-86. [02 Aug 2002]
- GORMAN M., *The Corruption of Cataloging*, Library Journal September 15, 1995, pp. 32-34.
- GOULD C., *Searching Smart on the World Wide Web: Tools and Techniques for Getting Quality Results*, Internet Workshop Series ; No. 8. Berkeley, CA: Library Solutions Press, 1998.
- HAHN S. E., *Internet: Let the User Beware*, Reference Services Review 25 (2) 1997, pp. 7-13.
- HERNON P., *Disinformation and Misinformation through the Internet: Findings of an Exploratory Study*, Government Information Quarterly 12 (2) 1995, pp. 133-139.

- HERRING S. D., *Faculty Acceptance of the World Wide Web for Student Research*, College & Research Libraries 62(3) May 2001, pp. 251-258.
- HOLZSCHLAG M. E., *250 HTML and Web Design Secrets*, Wiley Indianapolis, 2004
- JACOBSON T., COHEN L., *Teaching Students to Evaluate Internet Sites*, The Teaching Professor 11(7), August/September 1997, page 4.
- JANES J. W., AND ROSENFELD L. B., *Networked Information Retrieval and Organization: Issues and Questions*, Journal of the American Society for Information Science 47(9) September 1997, pp. 711-715.
- JUREK R. J., *Don't Be Fooled Again*, Internet World 8(4) April 1997, pp. 47-50.
- KAPOUN J., *Teaching Undergrads WEB Evaluation*, College and Research Libraries News 59(7) July/August 1998, pp. 522-523.
- KIRKWOOD H. P. JR., *Beyond Evaluation: A model for cooperative evaluation of Internet resources*, Online, 22(4) July/August 1998, pp. 66-72.
- KONRAD L., STEMPER J., *Same Game, Different Name: Demystifying Internet Instruction*, Research Strategies, vol.14 no.1, Winter 1996, pp. 4-21.
- LEDERER N., *New form(at): using the Web to teach research and critical thinking skills*, Reference Services Review, vol.28, no.2, 2000, pages 130-153.
- MCKENZIE J., *Making WEB Meaning*, Educational Leadership 54(3) November 1996, pp. 30-32.
- MAKULOWICH J., *Quality Control on the Net*, Database February 1996, pp. 93-94.
- MARTORANA J. AND DOYLE C., *Computers On, Critical Thinking Off: Challenges of Teaching in the Electronic Environment*, Research Strategies 14 (3) Summer 1996, pp. 184-191.
- MCBRIDE K. B., DICKSTEIN R., *The Web Demands Critical Reading by Students*, Chronicle of Higher Education 44(28) March 20, 1998, pg B6.
- MILLER W., *Troubling Myths about On-Line Information*, Chronicle of Higher Education 43(47) August 1, 1997, pg. A44.
- MINKEL W., *Burden of Spoof: Teach Students to Evaluate the Internet with Bogus Web Sites*, School Library Journal 46(10) October 2000, pg. 49.
- MORRISON J. L., STEIN L., *Assuring integrity of information utility in cyber-learning formats*, Reference Services Review, vol. 27, no.4, 1999, pg 317-326.
- NOWICKI S., *Information Literacy and Critical Thinking in the Electronic Environment*, Journal of Instruction Delivery Systems 13 (1) Winter 1999, pp. 25-28.
- OBERMAN C., *Avoiding the Cereal Syndrome, or Critical Thinking in the Electronic Environment*, Library Trends 39 (3) Winter 1991, pp. 189-202.
- OBERMAN C., *Unmasking Technology: A Prelude to Teaching*, Research Strategies 13 (1) Winter 1995, pp. 34-39.
- PACK T., *Use it or Lose it: Jakob Nielsen Champions Content Usability*, EContent 24(4) June 2001, pp. 44-46.
- PAGE M., KESSELMAN M., *Teaching the Internet: Challenges and Opportunities*, Research Strategies 12 (3) Summer 1994, pp. 157-167.
- PAGELL R. A., *Quality and the Internet: An Open Letter*, Online July/August 1995, pp. 7-9.
- PASK J. SNOW C. E., *Undergraduate Instruction and the Internet*, Library Trends 44(2) Fall 1995, pp. 306-317.
- PRATT G. F., FLANNERY P., PERKINS C., *Guidelines for Internet Resource Selection*, College and Research Libraries News March 1996, pg 135.
- RADER H., REINHART B., THOMPSON G., *Evaluating Information: A Basic Checklist*. Chicago: ALA, 1990.
- RETTIG J., *Beyond "Cool": Analog Models for Reviewing Digital Resources*, Online 20 (5) Sept/Oct 1996, pp. 52-64.
- ROSENFELD L. B., *Guides, Clearinghouses, and Value-Added Repackaging: Some Thoughts on How Librarians Can Improve the Internet*, Reference Services Review Winter 1994, pp. 11-16.
- ROTHENBERG D., *How the Web Destroys the Quality of Students' Research Papers*, Chronicle of Higher Education August 15, 1997 pg A44.
- SABOL L., *The Value of Student Evaluation of a Web Site*, Research Strategies 16(1) 1998, pp. 79-84.
- SAFFORD B. RIPP, *The Problem with the Internet: It is NOT the Information Highway*, School Library Media Activities Monthly 13(3) November 1996, pp. 42-43.
- SAFFORD B. RIPP ET AL., *What Do We Teach about the World Wide Web?*, School Library Media Activities Monthly 13(4) December 1996, pp. 44-46.
- SANTA V., EDMUND F., *The Internet as a Reference and Research Tool: A Model for Educators*, The Reference Librarian, no. 41-42 (1994), pp. 225-236.
- SCHOLZ-CRANE A., *Evaluating the Future: A Preliminary Study of the Process of How Undergraduate Students Evaluate Web Sources*, RSR; Reference Services Review 26(3/4) 1998, pgs. 53-60.
- SCHROCK K., *It Must Be True. I Found It On The Internet*, Technology Connection 3(5) September 1996, pp. 12-14.
- SKOV A., *Separating the wheat from the chaff: Internet Quality*, Database August/September 1998, pp. 38-40.
- SMITH A. G., *Testing the Surf: Criteria for Evaluating Internet Information Resources*, The Public-Access Computer Systems Review 8(3) 1997
- STOLL C., *Silicon Snake Oil: Second Thoughts on the Information Highway*. New York: Doubleday, 1995.
- SWEETLAND J. H., *Reviewing the World Wide Web--Theory versus Reality: Three Sets of Web Site Evaluation Criteria Compared with Web Reviews from Choice magazine*, Library Trends 48(4) Spring 2000, pp. 748-68.
- TATE M., ALEXANDER J., *Teaching Critical Evaluation Skills for World Wide Web Resources*, Computers in Libraries 16(10) Nov-Dec 1996, pp. 49-54.
- VALENZA J. KASMAN, *Evaluating Web Resources: Garbage In, Garbage Out!*, www.classroom.com Newsletter February 2002, pp. 4-7.
- WEILER A. (2001). *Two-year college freshmen and the Internet: do they really "know all that stuff?"*, Libraries and the Academy, v. 1, p. 161-67.