



@domani/search

Ormai siamo sottoposti a quantità sempre crescenti di informazioni; la nostra capacità di cercare informazioni e di trovare quelle giuste influenza sempre più la possibilità di muoverci con successo nell'ambiente circostante, creare situazioni favorevoli, accrescere il nostro capitale informativo e quindi la nostra ricchezza.

Il fattore tempo è del resto una discriminante essenziale: l'esigenza continua di muoverci e prendere delle decisioni ci impone di trovare le informazioni nel minor tempo possibile.

In questo quadro, l'abbondanza di informazioni, quando non organizzate e difficili da trovare, può diventare, anziché un valore aggiunto, una pericolosa arma in grado di rallentare e rendere meno efficienti i processi di gestione della conoscenza.

Quali sono gli strumenti a disposizione, oggi, per muoverci in questa complessità e quali le direzioni di sviluppo?

Una strada di ricerca cruciale è quella, oggi, di affiancare il sistema delle descrizioni oggi esistenti sulle pagine web con altri sistemi più sofisticati in grado di tener conto in modo più complesso dei contenuti all'interno dei documenti, delle relazioni tra documenti e anche delle valutazioni date a questi documenti dagli utenti.

Si parla, in questo senso, dello sviluppo di un **web semantico**.

In questa direzione, se lo standard **XML** oggi utilizzato in molte pagine web, si rivolge alla descrizione di **documenti**, il nuovo **RDF** (e le sue evoluzioni, come l'**OWL**), standard emergente promosso dal CERN di Tim Berners Lee, sono maggiormente indicati per rappresentare **dati** e quindi descrivere in modo più preciso e puntuale i contenuti presenti all'interno di un documento.

Ma i passaggi sembrano poter essere ancora molti. Per esempio, si studiano le possibilità per integrare le informazioni relative ai dati in relazione all'esperienza di fruizione.

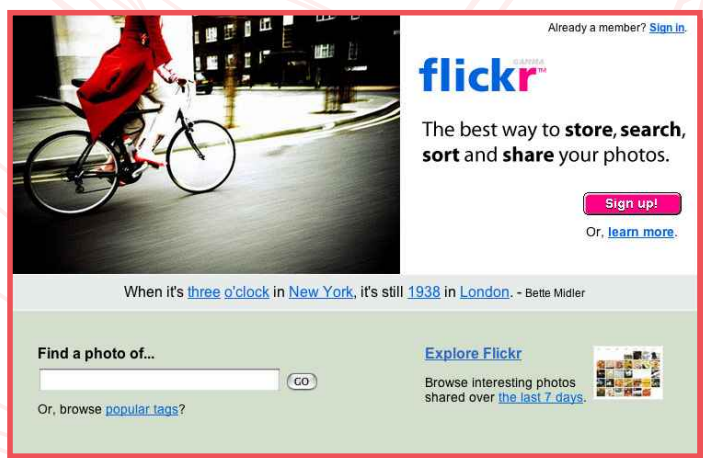
Attraverso **Agenti intelligenti**, sistemi esperti, **sistemi di collaborative filtering**, *knowbots (knowledge robots)* e altri metodi, si sviluppano le capacità dei sistemi di ricerca di imparare dalla ricerca stessa, e dagli input più o meno volontari provenienti da fonti diverse, per perfezionare progressivamente la propria efficacia.



@domani/search

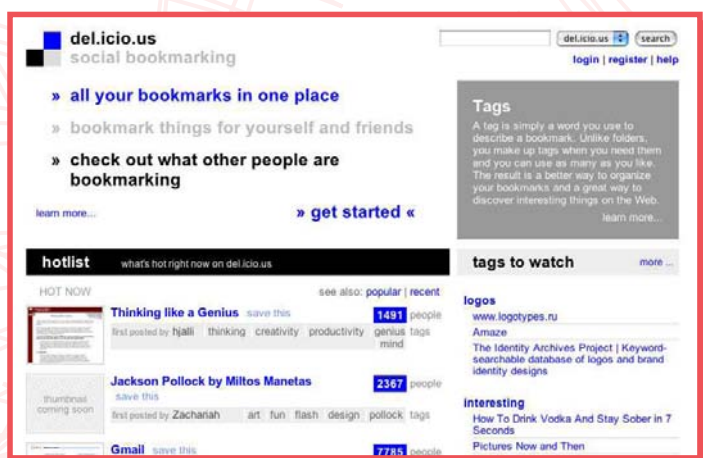
Per esempio, la libreria virtuale Amazon permette, grazie a un sistema di *collaborative filtering*, di ricevere consigli personalizzati di lettura in base ai libri precedentemente acquistati. Se per esempio abbiamo acquistato un libro su Truffault e un altro sul cinema giapponese, il sistema ci segnalerà cosa, utenti che hanno letto i nostri stessi libri, hanno acquistato successivamente, in modo da metterci in contatto con un insieme di suggerimenti provenienti da utenti simili a noi.

Questa idea di sfruttare le conoscenze dei fruitori per arricchire le funzionalità della ricerca è anche portata avanti, con strumenti differenti, dai sistemi cosiddetti di *social tagging*.



Attraverso tali sistemi, presenti, per esempio in siti come Del.icio.us o Flickr, gli utenti possono inserire dei descrittori, definiti in inglese TAG, per organizzare le risorse in rete secondo criteri personali.

L'aggregazione statistica di un TAG intorno a una risorsa (immagine, sito o documento che sia) dà maggior peso a quella descrizione che quindi diventerà più importante come chiave di ricerca all'interno dell'insieme generale.



In generale, si raffinano tutte le modalità di *Relevance ranking* che permettono di organizzare le risorse in base a diversi ordini di importanza: non più in relazione solo ai contenuti di un documento, ma anche, per esempio, al numero di link che puntano a quella risorsa.



@domani/search

Filtri collaborativi online

SITI COMMERCIALI

Tipologia di contenuti

Amazon*	varie
Barnes and Nobles	varie
Loomia	podcasts
Movielink**	film
Netflix***	film

SITI NON COMMERCIALI

Tipologia di contenuti

Alongtail	film
Movielens****	film
MusicStrands	musica
Gnod	varie
StumbleUpon	siti web

*Amazon utilizza un sistema di *collaborative filtering* basato sul prodotto (*item-based*).

**Movielink, il popolare servizio Usa di download di film, utilizza la tecnologia di ChoiceStream, il cui sistema di personalizzazione non si basa sul *collaborative filtering*, bensì su un approccio *content-based*: si chiama *Attributized bayesian choice modeling* (Abcm).

***Netflix mescola al suo sistema di raccomandazioni alcune funzioni più da community: gli utenti possono scrivere recensioni, sbirciare sulle altrui liste di film noleggiati, dare voti ai film e vedere come votano gli altri.

****MovieLens nasce dal GroupLens Research Project dell'Università del Minnesota ed è un tipico caso di *collaborative filtering* incentrato sull'utente (*user-based*). Raccoglie le preferenze degli utenti per i film e in base a queste raggruppa gli stessi utenti; dopodiché, basandosi sui rating espressi da tutti i componenti di un gruppo, tenta di predire i gusti individuali per i singoli film.

Molto interessanti, in questa direzione, sono i cosiddetti **index of popularity**.

Gli *index of popularity*, come Technorati, PopDex o Daypop, infatti, monitorano quotidianamente la rete, in particolare i blog, e attraverso la mappatura dei link contenuti al loro interno forniscono un quadro dinamico dei blog più centrali, in un dato momento, nella rete e, conseguentemente, dei temi e delle discussioni più "calde" in quel momento.

In questo modo, ai criteri automatici *content based* stabiliti da algoritmi informatici, si aggiungono sistemi umani, basati sulla reputazione e sulla fiducia, che permettono di selezionare le informazioni non più solo in base a descrizioni "oggettive" fornite da coloro che le informazioni le hanno prodotte, ma anche in base a criteri "soggettivi", quali il livello di gradimento e popolarità che quella data risorsa ha ottenuto.

Sebbene questi sistemi siano soggetti a un alto livello di discrezionalità, questo non deve essere considerato necessariamente uno svantaggio, visto che è stato dimostrato che le persone prendono la maggior parte delle decisioni, spesso, secondo processi soggettivi e stocastici; non proprio, quindi, in base a criteri di razionalità pura.

@link

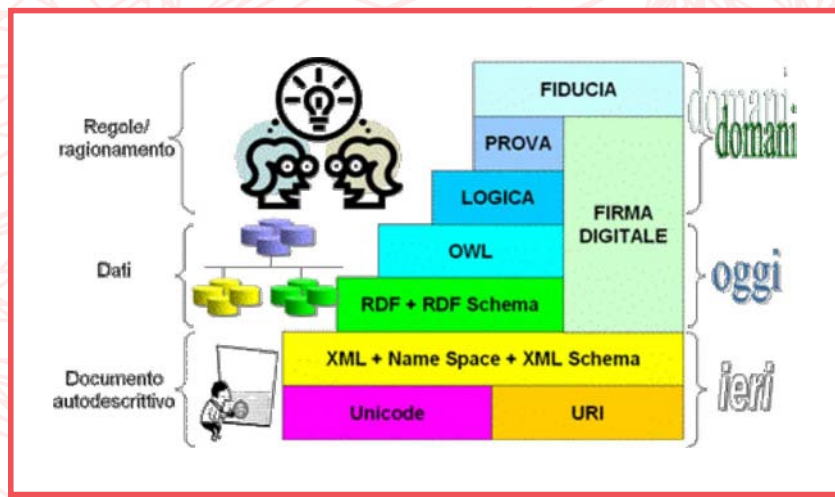
Alongtail www.alongtail.com
 Amazon www.amazon.com
 ChoiceStream www.choicestream.com
 Last.fm www.last.fm
 Loomia www.loomia.com
 Movielens movielens.umn.edu
 MusicStrands www.musicstrands.com
 Netflix www.netflix.com
 Opinity www.opinity.com
 Pandora www.pandora.com
 Real-ity www.real-ity.com

Tratto da www.visionblog.it/

Tratto da www.visionblog.it/



@domani/search



Sistemi “umani” di selezione e valutazione delle informazioni, come la fiducia e la popolarità, se integrati in modo equilibrato con altri sistemi di descrizione delle informazioni, quali le più razionali classificazioni, tassonomie o tipologie, potrebbero quindi rilevarsi uno strumento più “morbido” e sofisticato a nostra disposizione per la quotidiana caccia del nostro millennio: quella alle informazioni.