

Nuova Periodici  
 PC World  
 Macworld  
 Gamestar  
 Network World  
 Computerworld  
 Yoga Journal  
 Spotlight  
 Fiction-TV  
 Nuov@ Periodici Shop

**COMPUTERWORLD**  
 www.cwi.it online

Visualizza gli articoli per

 cerca

Knowledge Center Bankworld Top 100 Time Machine PMI PLM CIO CSO Scenari & Trend Tecnologie Partner Zone Webcast Whitepaper Blog  
 Applicazioni Business Intelligence Governance Hardware Innovazione Mercato Mobile Networking PA People Servizi IT Sicurezza Software Storage Tlc

SEI QUI: [Computerworld](#) > [Tecnologie](#)

TAG: Internet - Web 2.0 - CNR

Applicazioni mobili Cloud  
 computing ERP  
 Esperienze aziendali  
 Fusioni e  
 acquisizioni Gare e  
 contratti Internet  
 Management Motori di  
 ricerca Netbook  
 Notebook  
 Outsourcing

## Ecco come scegliamo i 'tag' su internet. Il CNR: opportunità rivoluzionarie

Ricercatori italiani e francesi hanno creato un modello applicabile a pubblicità online, lotta allo spam, e catalogazione automatica di enormi moli di informazioni

Mercoledì 8 Luglio 2009

Alcuni ricercatori italiani, insieme a colleghi francesi, hanno messo a punto un modello che spiega il fenomeno del **tagging**, cioè il modo in cui chi naviga su internet, e specialmente nei siti di **social networking**, associa contenuti e parole chiave.

Secondo tale modello, non solo il tagging non è casuale, ma si possono mettere a punto algoritmi matematici al confine tra fisica e statistica in grado di replicarne i risultati, a partire da idee intuitive sul funzionamento della mente umana.

Si tratta di un risultato tutt'altro che astratto. Come spiega una nota del CNR, sapere cosa guida gli utenti internet quando associano contenuti e parole chiave aprirebbe la strada a molte innovazioni nei servizi online. Per esempio sistemi efficaci di **protezione dallo spam** (capaci di trovare le associazioni 'sbagliate' di parole entro email fraudolente), o servizi di **pubblicità online** capaci di proporre all'utente prodotti calibrati con precisione, o sistemi efficaci di **catalogazione** di enormi quantità di dati.

In questo scenario, il modello di cui parliamo oggi, descritto nell'ultimo numero della rivista specializzata **PNAS** (Proceedings of the National Academy of Science) con l'articolo 'Collective dynamics of social annotation', è un primo importante passo.

Gli autori, che sono Vittorio Loreto di **INFN-CNR** e **Università La Sapienza** di Roma, Andrea Baldassarri della Sapienza, Ciro Cattuto della **Fondazione ISI di Torino**, e Alain Barrat e Gregory Schehr del CNRS francese, hanno infatti creato un modello di tagging capace di produrre risultati confrontabili con i dati reali estratti da siti di social networking, nello specifico **del.icio.us** e **Bibsonomy**.

I processi che regolano il tagging, spiega la nota, sono ancora piuttosto oscuri. Quello che intuitivamente si sa è che non è casuale, ma tende a funzionare per catene di "affinità semantica" (come cielo-nuvola, nuvola-bianco, e così via). Questo semplice meccanismo è stato usato dai ricercatori come base per il loro modello, che è un sistema in grado di generare post virtuali che mimano per struttura quelli reali.

Nell'abstract dell'articolo, i cinque ricercatori spiegano che il processo di tagging (definito '**social annotation**') può essere visto come "un'esplorazione collettiva ma non coordinata di uno spazio semantico". Per analizzarlo serve quindi un approccio multidisciplinare, per cui sono stati usati concetti di fisica statistica come le cosiddette "passeggiate aleatorie" (random walks), e di teoria delle reti complesse.

In pratica, centinaia di migliaia di post virtuali sono stati creati e confrontati con altrettanti provenienti da siti reali (**del.icio.us** e **Bibsonomy**), con risultati sorprendenti: i due gruppi condividono una stessa struttura complessiva, e rispondono alle stesse leggi statistiche.

Questa grande coincidenza, sottolinea la nota del CNR, ha sorpreso i ricercatori. Il loro modello, volutamente generico, ha infatti ricostruito in maniera quasi perfetta la struttura dei dati reali, rispecchiandone molte importanti caratteristiche, come la co-occorrenza di tag (cioè quante volte un tag è usato assieme a un altro, si pensi a "mare" e "acqua"), o la legge che regola l'aumento del numero complessivo di tag usati dagli utenti.

Un risultato brillante, conclude la nota, che i ricercatori stanno però già cercando di migliorare, perfezionando ulteriormente il modello con vari correttivi che meglio rispecchiano il comportamento degli utenti reali.

[Computerworld e IBM presentano il Cost Saving Center »](#)

[Home](#) - [Chi siamo](#) - [Abbonamenti](#) - [RSS](#)

Copyright © 2006 - 2009 Nuov@ Periodici Italia S.r.l. - tutti i diritti riservati - P. IVA 06868120152

E' vietato riprodurre i contenuti di questo sito senza l'autorizzazione scritta dell'editore.

I prodotti e i marchi citati sono di proprietà dei rispettivi titolari