

World Wide Web 2.0 vs World Wide Web 3.0

Il fenomeno della triplice W,¹ nato per mano del matematico Tim Berners Lee nel 1991, destinato a rivoluzionare la fruizione delle risorse dell'informazione, conosce la sua seconda generazione, nota come Web 2.0 e si proietta in un imminente futuro verso il terzo lustro.

Web è, secondo le definizioni più comuni registrate in letteratura, una grande piattaforma per la pubblicazione di documenti e per l'erogazione di servizi di vario genere. Rappresenta, cioè, per l'utente un *mare magnum* dove, citando un classico delle metafore, navigare alla maniera degli esploratori.²

Web 1.0, e la sua evoluzione, rimane ancora e soprattutto un archivio di pagine HTML, formato che ne ha garantito l'affermazione per semplicità d'utilizzo a scapito dell'accuratezza del significato e della qualità e pertinenza delle informazioni reperite. La rintracciabilità delle informazioni ricercate sottolinea la necessità di approfondire il tema delle ontologie, dei significati oggettivi che creano l'ingegneria della conoscenza, ponendo al centro della ricerca stessa l'interrogazione. I concetti alla base della ricerca devono essere messi in relazione tra loro per diventare specchio di una categorizzazione della conoscenza.

Il web si trasforma da piattaforma per la consultazione documentale (archivio) a piattaforma multimediale collaborativa, con tempi di risposta brevi e aggiornamen-

ti immediati. La sua principale peculiarità è la capacità di selezionare le informazioni e di convogliarle all'attenzione dell'utente quando ne ha bisogno in tempi rapidissimi.

Il web diventa, cioè, capace di filtrare, dando al ricercatore-utente, a fronte di un termine da ricercare, la possibilità di recuperare un testo, un video, un'immagine o una canzone (*filtering collaborativo*).³

Nel presente lavoro verrà presa in esame una delle applicazioni più note in ambito editoriale del filtering collaborativo, l'RSS. È usato da molte riviste scientifiche online per disseminare articoli o aggiornamenti su temi precisi. RSS è un acronimo definito nel 2004 da Hammond del gruppo Nature,⁴ l'Idra del mondo web, sempre per usare una metafora classica: sette sono, infatti, le differenti versioni e svariati i suoi impatti nel sociale.

Per tornare alla spiegazione dell'acronimo, abbiamo tre opzioni: Really Simple Syndication, Rich Site Summary e RDF (che sta per Resource Description Framework) Site Summary. La definizione dipende da chi avete o state interrogando e a quale versione di RSS fate riferimento.⁵ La parola chiave di cui si occupa il presente lavoro è il termine *syndication*; dietro a questo concetto si nasconde l'attività di un'agenzia stampa virtuale che produce il flusso di notizie, le organizza secondo una struttura comune: titolo, corpo della notizia, data e ora di

emissione, autore e categoria tematica. Per *syndication* si intende, quindi, la possibilità offerta da un fornitore di contenuti informativi di inviare e di ricevere automaticamente, di ripubblicare o riutilizzare le informazioni prodotte.⁶ Perché tutto ciò avvenga, bisogna che le informazioni siano opportunamente marcate, marcatura finalizzata alla struttura logica e ai contenuti dell'informazione. Tecnicamente l'uso di un RSS sottende la creazione di un *web-feed* o più semplicemente di un feed. Il feed è un documento che contiene gli elementi della storia dell'informazione ricercata.

Nella figura 1 viene riportato un esempio di feed, creato da una ricerca lanciata dal database della National Library of Medicine – PubMed.⁷ Leggiamo tra le righe le informazioni contenute nel feed:

```
<?xml version="1.0" encoding="UTF-8" ?>
```

La prima riga indica che si tratta di un file XML nella sua versione 1.0:

```
– <rss version="2.0">
```

La seconda riga riporta la radice del contenuto del documento e la sua versione:

```
– <channel>
```

```
<title>PubMed: cancer</title>
<link>http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=pubmed&term=cancer</link>
```

In questa sequenza vengono descritti: il canale creato per il feed, il titolo del feed e del suo canale, l'URL (indirizzo) della pagina web a cui il feed è associato.

```
<description>NCBI: db=PubMed; Term=cancer</description>
```

È la descrizione narrativa del feed, *alias* ho ricercato il lemma "cancer" nel database di PubMed, ho creato un feed di cui fornisco l'indirizzo.

RSS è un formato standard

de facto,⁸ dunque, basato sul linguaggio XML a sua volta estensione del linguaggio HTML per gli ipertesti, per inviare o meglio notificare aggiornamenti di contenuti direttamente sul proprio desktop o su dispositivi mobili come telefoni cellulari e palmari.⁹

RSS è il nostro assistente virtuale che si collega ai siti web di nostro interesse, raccoglie i contenuti dei documenti rilevanti e attraverso un file detto "feed" all'interno di un canale rende leggibile e disponibile l'informazione o il suo aggiornamento.¹⁰ La sua peculiarità sta nella capacità d'analisi automatica delle risorse testuali e medialiali, nella mappatura, cioè, delle ontologie – rappresentazioni concettuali standard – rendendo così la ricerca efficace per rilevanza.

Per poter leggere i contenuti nel formato RSS era un tempo necessario installare software *ad hoc* chiamati Reader; oggi questa possibilità è integrata nelle nuove versioni dei browser tra cui il diffuso Internet Explorer 7, oppure direttamente dal sito di Google¹¹ nelle funzioni Reader o Bloglines (quest'ultimo fornisce gli aggiornamenti dei blog). Concettualmente il feed RSS si pone nell'insieme di tecnologie definite *push*, e di fatto è il completamento degli strumenti di *e-alert* presenti nella posta elettronica, evitando però lo *spamming* fastidioso e indesiderato. Uno dei vantaggi che non deve essere dimenticato è l'uso di aggregatori¹² che evitano di visionare centinaia di siti o di rieseguire la ricerca per ottenere la stessa informazione.

Il compito dell'aggregatore è dunque quello di trasformare l'XML in qualcosa di leggibile, di controllare periodicamente la ricerca, di

informarci quando nuovi elementi o notizie si sono aggiunti all'informazione originale, fornirci l'iperlink al testo originale (articolo) o ai testi correlati così come forniti dall'autore.

Le figure 2 e 3 che seguono mostrano gli aggiornamenti della nostra ricerca PubMed su "cancer" così come si presentano dal Reader di Google.

Alla base dell'innovazione

del Web 2.0 sta l'idea di poter aggiungere in maniera semplice, precisa e sempre più standardizzata un'indicizzazione "intelligente" a tutta l'informazione primaria inserita nella rete.

Il progetto rappresenta una duplice sfida nelle abilità: per la comunità degli utenti riguarda l'abilità d'interrogazione e per noi addetti ai lavori l'abilità di programmazione logica e di formazio-

ne degli utenti alla meta-informazione descrittiva, costruendo una rete di relazioni semantiche. Sfida che proietta i documentalisti nel futuro stesso del web, nel Web 3.0.

Questa evoluzione in web semantico, capace di costruire, ricercare e mettere in connessione logica le relazioni semantiche, produrrà nel mondo della ricerca dell'informazione biomedica lo

stesso effetto che ebbe l'introduzione dell'*Index Medicus* da parte del dottor Billings nel lontano inizio del XIX secolo.

Francesca Gualtieri

GIDIF, RBM

francesca.gualtieri@rottapharm.com

Fig. 1 – Esempio di feed tratto da PubMed

```
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0">
  <channel>
    <title>PubMed: cancer</title>
    <link>http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=pubmed&term=cancer</link>
    <description>NCBI: db=PubMed; Term=cancer</description>
    <language>en-us</language>
    <docs>http://blogs.law.harvard.edu/tech/rss/docs</docs>
    <ttl>1440</ttl>
    <image>
      <title>NCBI PubMed</title>
      <url>http://www.ncbi.nlm.nih.gov/entrez/query/static/gifs/iconsm1.gif</url>
      <link>http://www.ncbi.nlm.nih.gov/entrez/query.fcgi</link>
      <description>PubMed, a service of the National Library of Medicine, developed by the National Center for Biotechnology Information (NCBI) includes citations for biomedical articles from MEDLINE and additional life science journals.</description>
    </image>
    <item>
      <title>Methylphenidate use in children and risk of cancer at 18 sites: results of surveillance analyses.</title>
      <link>http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?tbl=NoSideBarFile&db=PubMed&cmd=Retrieve&list_uids=18041106&dopt=Abstract</link>
      <description>
        <![CDATA[
          <table border="0" width="100%">
            <tr>
              <td align="left"><td align="right"><a href="http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed&cmd=Display&db=PubMed&from_uid=18041106">Related Articles</a></td></tr>
            <tr>
              <td colspan="2"><p><b>Methylphenidate use in children and risk of cancer at 18 sites: results of surveillance analyses.</b></p>
                <p><b>Pharmacoeconomic Drug Saf. 2007 Nov 27;16(12):1268-1272</b></p>
                <p><b>Authors: Oestreicher N, Friedman GD, Jiang SF, Chan J, Quesenberry C, Habel LA</b></p>
                <p><b>PURPOSE: A recent report linked methylphenidate (MPH) use in children to cytologic abnormalities in plasma lymphocytes, a possible</b></p>
                <p><b>PMID: 18041106 [PubMed - as supplied by publisher]</b></p>
              </td></tr>
          </table>
        </description>
    </item>
  </channel>
</rss>
```

Fig. 2 – Aggiornamento della ricerca PubMed in formato elenco

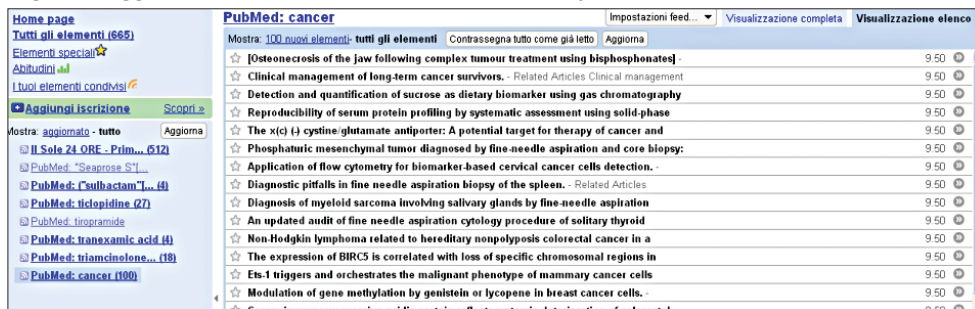
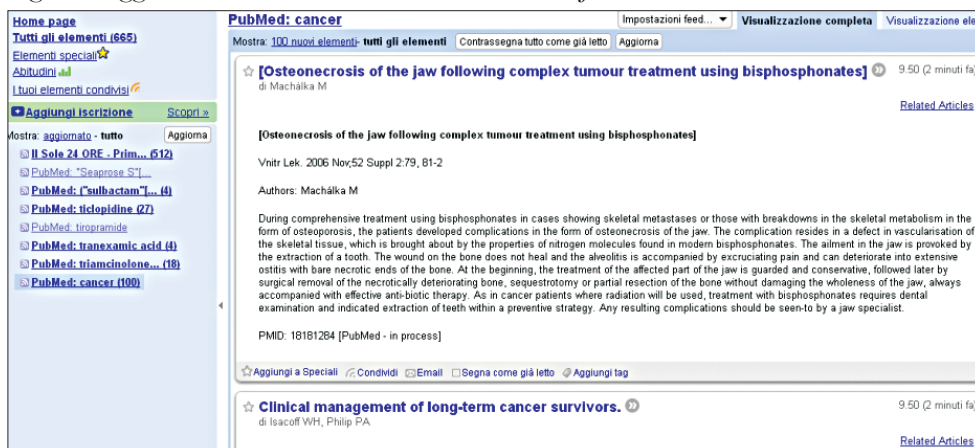


Fig. 3 – Aggiornamento della ricerca PubMed in formato abstract



Note

- <http://it.wikipedia.org/wiki/world_wide_web>.
- DANIELA CANALI, *Web semantico e ontologie*, "Biblioteche oggi", 23 (2005), 5, p. 50-58.
- PAOLO MORATI, *Rivoluzione architeturale in arrivo sul fronte del Web*, "ComputerWorld Italia", (2007), 28, p. 12.
- TONY HAMMOND ET AL., *The role of RSS in science publishing: syndication and annotation on the web*, "D-Lib Magazine", 10 (2004), 12.
- MICHAEL P. SAUERS, *Blogging and RSS. A librarian's guide*, Information Today, Medford NJ, 2006, cap. 5 *An introduction to RSS*, p. 123.
- GINO RONCAGLIA, *Blogosfera e feed RSS: una palestra per il Semantic Web?*, (2003), <http://www.swif.uniba.it/lei/ai/networks/03-2/>.
- <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/erss.cgi?rss_guid=0-U1gyUXk3Mf6Cxy0inIgcqolul33Tg0HwMn9dlo9WXF> da PubMed, RSS Feed.
- ANTONELLA DE ROBBIO, *Servizi bibliotecari personalizzati basati su RSS feeds: i diversi volti di un nuovo canale di comunicazione*, in *La biblioteca su misura. Verso la personalizzazione del servizio*, a cura di Claudio Gamba e Maria Laura Trapletti, Milano, Editrice Bibliografica, 2007, p. 184-206.
- EUGENIO SANTORO, "Recenti progressi in medicina", (2007), 98, p. 484-494.
- ANTONELLA DE ROBBIO, *Servizi bibliotecari personalizzati...*, cit.
- WILLIAM MILLER – RITA M. PELLEN, *Libraries and Google*, Binghampton, The Haworth Information Press, 2005.
- MICHAEL P. SAUERS, *Blogging and RSS...*, cit., cap. 6 *Using an aggregator*, p. 143.
- DEAN GIUSTINI, "British Medical Journal", (2007), 335, p. 1273; <http://www.bmj.com/cgi/content/full/335/7633/1273>.