

► Internet

RSS, come ricevere da Internet informazioni su misura

Rich Site Summary è la tecnologia per tenere d'occhio gli ultimi articoli pubblicati sui siti che vi interessano, senza virus e senza spam

di Giorgio Gobbi

Definire l'acronimo RSS costringe a imbarcarsi in una spiegazione tecnica ramificata, che rischierebbe di farci perdere di vista l'obiettivo di questa tecnologia. Per qualche istante, RSS rimarrà perciò una sigla, che definiremo più avanti. Per prima cosa vediamo la funzionalità che ci viene offerta da questa tecnologia Internet, nata pochi anni fa e in rapida diffusione.

Per rimanere sempre informati

RSS risponde a un'esigenza precisa di vari strati di utenti Internet: la necessità di essere informati sui propri argomenti di interesse senza passare ore al computer per esplorare siti dopo siti che potrebbero contenere articoli e news utili per le nostre attività. Facciamo un esempio: di niente Internet parla più volentieri che di tecnologia informatica e, se il nostro lavoro richiede di essere aggiornati sugli sviluppi di hardware, software e tecnologie di base, dovremmo compilare una lunga lista di siti da tenere d'occhio quotidianamente. È vero che ci sono siti per la diffusione di comunicati stampa e di notizie informatiche, ma spesso sono troppo generici o si occupano di un settore industriale specifico: in poche parole, non soddisfano gli interessi dell'utente o del professionista informatico. Lo stesso vale negli altri campi: se volete una panoramica di come i quotidiani mondiali trattano la politica, l'economia o altri argomenti, è difficile evitare di andare a consultarne i siti, uno alla volta.

Tutto questo iniziò a cambiare alla fine degli anni '90 quando Netscape, con una tec-



AmphetaDesk è una delle applicazioni più diffuse per l'aggregazione e lettura delle news; utilizza un piccolo eseguibile multiplatforma più un normale browser

nica originata da UserLand Software, iniziò a pubblicare sul proprio portale MyNetscape dei "canali" (o newsfeed) che mantenevano l'utente aggiornato sui contenuti di varie fonti di informazioni.

Nel 2001 AOL (*America On Line*) tolse questi canali di news dal nuovo Netscape, ma la tecnologia RSS continuò ad evolversi con entusiasmo e collaborazione da più parti, tanto da rendere incerto il significato dell'acronimo RSS. Ecco qual è il concetto base: 1) i siti interessati a far conoscere gli ultimi articoli e news ne pubblicano un breve sommario in un file XML secondo il formato RSS; 2) le organizzazioni addette alla raccolta e distribuzione di queste informazioni (i sommari con i link alle pagine Web degli articoli) mantengono conti-

nuamente aggiornato un database dei siti con supporto RSS e relativi file; 3) l'utente sceglie un servizio di aggregazione e lettura delle informazioni e seleziona, tra le migliaia dei siti con supporto RSS, quelli di cui vuole essere aggiornato; 4) tramite browser, o attraverso un'applicazione desktop, l'utente riceve (per esempio ogni ora) le informazioni RSS dei siti a cui si è "abbonato" (non costa nulla), scorre rapidamente i titoli delle ultime news e può aprire le relative pagine Web nel browser o in una finestra dell'applicazione.

Nel caso di un utente informatico interessato alle novità del giorno, tutto questo significa che al mattino, anziché esplorare una decina di siti di notizie informatiche, gli basterà aprire la pagina Web o il

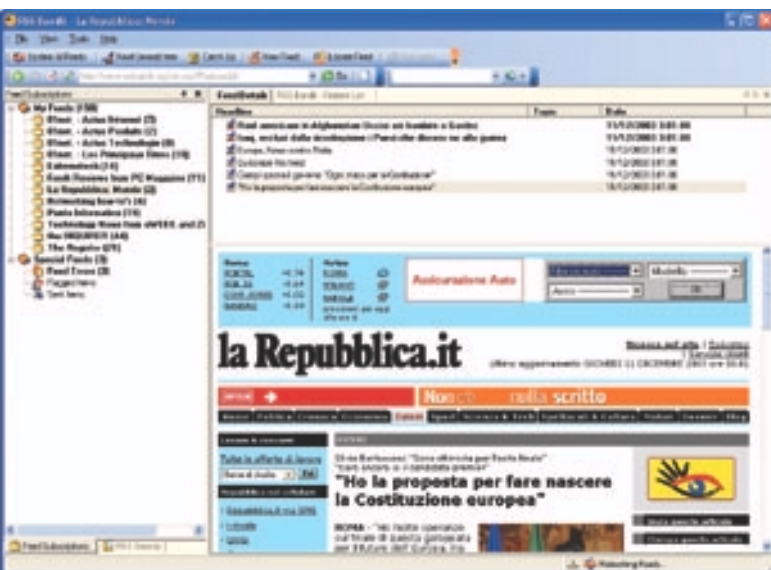
programma che ha scelto per la raccolta delle news e si troverà belli e pronti, in un unico elenco aggiornato all'ultima ora, i titoli delle news dei siti prescelti. In molti casi il servizio è gratuito; a volte, spendendo poche decine di euro, si possono avere funzionalità più estese, ad esempio per la ricerca di nuove fonti di informazioni o per l'integrazione delle news RSS con la posta.

Le tre definizioni di RSS

Finalmente spieghiamo che cosa significa RSS, avvertendo che ci sono tre correnti di pensiero, ciascuna delle quali pretende di dare la giusta definizione di RSS. I tre possibili significati sono: **Rich Site Summary**, **RDF Site Summary** e **Really Simple Syndication**.

Di queste, l'unica abbastanza intuitiva è la prima; possiamo indovinare che un "ricco sommario di un sito" sia un file contenente un elenco di articoli con i relativi link ed eventuali brevi descrizioni. Oggi l'XML (*Extensible Markup Language*) è ampiamente utilizzato per creare linguaggi specifici, quindi non stupisce che anche l'RSS sia un'applicazione dell'XML. Un sommario RSS è quindi un file XML strutturato in tag per identificare il titolo del canale (la fonte delle notizie), il link al sito che pubblica le notizie, la descrizione del canale e un elenco di item (voci), cioè gli articoli o news che periodicamente vengono aggiornati, ciascuno con link ed eventuale descrizione.

Nell'esempio di file RSS che pubblichiamo si vedono quindi i tag <channel>, <title>, <link>, <description>, <item> e alcuni altri. Pubblichiamo un esempio di file RSS nella vecchia ver-



RSS Bandit è un'applicazione .NET per l'aggregazione e lettura delle news dai siti con supporto RSS

(Sotto) Un esempio parziale di file RSS in versione 0.91, ancora molto diffusa

```
<?xml version="1.0" ?>
<!DOCTYPE rss (View Source for full doctype...)>
- <rss version="0.91">
- <channel>
  <title>the INQUIRER</title>
  <link>http://www.theinquirer.net</link>
  <description>IT news - never knowingly undersoldered</description>
  <language>en-us</language>
  <copyright>Copyright 2003, Breakthrough Publishing Limited</copyright>
  <managingEditor>mike.magee@theinquirer.net</managingEditor>
  <webMaster>webmaster@theinquirer.net</webMaster>
- <item>
  <title>Spam to email ratio soars in 2003</title>
  <link>http://www.theinquirer.net/?article=13126</link>
  </item>
- <item>
  <title>AMD claims Intel high-k not necessary for next-gen transistors</title>
  <link>http://www.theinquirer.net/?article=13125</link>
  </item>
  .....
</rss>
```

sione 0.91, tuttora ampiamente utilizzata. Come si vede, il file ha una struttura così semplice che si potrebbe scriverlo a mano con un editor di testo o HTML.

Veniamo alla seconda definizione di RSS: **RDF Site Summary**. La sostanza non cambia: si tratta sempre del sommario di un sito, salvo che si fa riferimento al formato RDF (*Resource Discovery Framework*), anch'esso di derivazione XML ed utilizzato solo nella release 1.0 di RSS. Il formato RDF del W3C (www.w3.org/RDF/) serve per numerose applicazioni, come cataloghi di biblioteche, directory a livello mondiale, aggregazione e ridistribuzione di notizie, software e contenuti e

per raccolte e collezioni personali. I file RSS 1.0 possono essere chiamati file RDF e avere estensione .rdf. Sono comunque liste di item in XML sebbene più complicate rispetto alle release 0.9x. L'estensione dei file RSS può variare, anche se di solito è .rss, .rdf, .xml o .txt.

Per proseguire nella discussione sull'RSS occorre definire alcuni termini chiave. Il primo è la parola **syndicate**, che nel contesto dell'editoria e dei media significa "organizzazione che vende materiale (come articoli, foto e filmati) per la sua diffusione simultanea su più giornali, periodici o emittenti". Questa definizione è applicabile anche all'editoria on line e al suo utilizzo ai file RSS, salvo

che in questo caso le informazioni che vengono raccolte e distribuite non sono gli articoli ma i sommari RSS che contengono il titolo, una breve descrizione (opzionale) e il link degli articoli. In altre parole, i file RSS contengono metadati, ovvero informazioni relative ai dati, dove con dati intendiamo il materiale pubblicabile. Oltre al sostantivo, c'è anche il verbo syndicate, che normalmente significa vendere materiale a un syndicate (nome) per la pubblicazione su più giornali, periodici ed emittenti.

Il termine syndication ha la stessa definizione del verbo syndicate e nel caso dei file RSS significa localizzare quanti più siti possibile con supporto RSS e prelevare continuamente i loro file RSS, in modo da fornire i sommari che vi sono contenuti a tutti gli utenti che utilizzano il servizio.

Ora siamo in grado di capire la terza definizione di RSS: **Really Simple Syndication**, ovvero syndication realmente semplice. Di solito i portali si rivolgono a un'agenzia di stampa (il syndicate) per acquistare e pubblicare le ultime notizie, magari a scorrimento continuo in una barra o in un riquadro della pagina Web. Sfruttando le informazioni di un syndicate RSS (a pagamento, per uso commerciale), un sito può incorporare le informazioni RSS dei canali di news e l'utente del sito può accedere agli articoli cliccando sul titolo.

In questo modo tutti ne traggono beneficio: l'utente ha accesso alle news, il syndicate vende le informazioni RSS a chi le usa a scopo commerciale (mentre per uso personale possono essere gratuite) e gli editori delle news ricevono traffico aggiuntivo da parte di chi è interessato alle news ma non ha intenzione di esplorare direttamente il loro sito per trovare le novità del giorno o dell'ultima ora. Come corollario ai termini syndicate e syndication, c'è anche l'aggettivo syndicated, participio passato di syndicate che, riferito a un sito, significa che esso produce e tiene aggiornato un file RSS.

I giornali on line spesso mantengono più file RSS, uno per le notizie generali e altri per le sezioni Economia, Politica, Sport, Scienza e così via. In questo modo l'utente interessato a una sezione si abbona ad essa

e riceve solo la segnalazione delle news relative. A questo punto occorre chiarire qualche altro termine del mondo RSS, dopo di che potremo passare alla pratica.

Gli aggregatori, i locator e i blog

Abbiamo visto che RSS significa *Rich Site Summary*, *RDF Site Summary* e *Really Simple Syndication* e abbiamo chiarito queste definizioni. Per non fare torto a nessuno, c'è chi propone di usare la sigla RSS come simbolo e non come acronimo, come è accaduto alla parola DVD, non più *Digital Video Disc* e neppure *Digital Versatile Disc*.

Abbiamo visto che un syndicate non ha niente a che fare con un sindacato ma indica un'agenzia di stampa o, per estensione, un'entità che raccoglie e distribuisce metadati relativi al materiale pubblicato. Da quando è nata la moda dei weblogs o blogs (siti di notizie personali e commenti in forma cronologica), si è visto un fiorire di questo tipo di syndication, che permette di trovare i blogs e di abbonarsi.

Ora introduciamo qualche altro termine che incontriamo comunemente nei documenti sull'RSS. In inglese feed, come verbo, significa alimentare o fornire il necessario per il sostentamento o per il funzionamento di una macchina. Il sostantivo feed significa quindi cibo, materiale fornito o anche il meccanismo con cui il materiale è fornito. Nel gergo RSS un FEED, chiamato anche canale, è sinonimo di un file RSS. Quindi file RSS, feed RSS, canale RSS, feed e canale sono termini equivalenti e indicano il fatto che un sito, equipaggiato di RSS, alimenta la distribuzione dei sommari (gli elenchi di news contenuti nei file RSS) tramite l'intermediazione di un syndicate. Il software del syndicate passa il tempo a leggere a rotazione tutti i feed (anche parecchie migliaia) e li mette a disposizione delle applicazioni che li prelevano e li inoltrano all'utente (saltiamo gli aspetti legali e commerciali, che trovate nei siti dei syndicate).

Weblogs e siti di news sono tipici esempi di siti attrezzati con file RSS; un sito senza continui aggiornamenti e novità non ne ha bisogno. Per scoprire quali sono i siti con RSS ci sono vari modi, ma in generale



La home page di Syndic8, uno dei principali syndicate che raccolgono i dati RSS e li mettono a disposizione di aziende e utenti

serve un **locator**, cioè una funzione software o un servizio Web che permetta di localizzare i siti con RSS che ci interessano. Faremo qualche esempio più avanti.

Per l'utente il termine chiave nel gergo RSS è "aggregatore" o aggregatore. Un **aggregatore** RSS è un programma (installato sul vostro PC o accessibile via Web) che visualizza i contenuti dei file RSS relativi ai siti che avete selezionato e che permette di leggere le news. Si usa il termine **subscribe** - abbonarsi - per selezionare i siti RSS di interesse e riceverne periodicamente i feed, cioè i titoli delle ultime novità. Un aggregatore è chiamato anche **newsreader** o news aggregator, cioè lettore di notizie o aggregatore di notizie. Due esempi di aggregatori, tra le decine disponibili, sono **AmphetaDesk**, un'applicazione che utilizza un piccolo eseguibile locale e un'interfaccia via browser, e **RSS Bandit**, un'applicazione basata sull'ambiente .Net.

Localizzare gli RSS

Prima di illustrare l'uso pratico degli aggregatori o lettori di news, spendiamo qualche parola sui criteri di scelta. Tra i tanti aggregatori che abbiamo trovato su Internet (un'ottantina), ne abbiamo provati 15 e abbiamo avuto modo di saggiarne i pro e i contro. Tra questi abbiamo scelto di presentarvi AmphetaDesk e RSS Bandit perché, oltre a essere gratuiti, sono quelli che ci sono sembrati più convincenti nel

servizio fornito all'utente. Infatti non si tratta soltanto di offrire un'interfaccia piacevole e ricca di opzioni; i punti chiave sono la facilità con cui si possono trovare i canali che ci interessano e la rapidità di consultazione.

La consultazione avviene a tre livelli: scelta del canale, scelta del titolo dell'articolo e apertura dell'articolo. In certi casi il terzo passaggio è automatico; nelle applicazioni desktop con l'interfaccia a tre pannelli (tipo Outlook Express), a volte basta selezionare un titolo nel pannello superiore per vedere l'articolo nel pannello inferiore.

Anche quando la consultazione delle news risulta comoda, resta il vero nocciolo del problema: la ricerca dei siti syndicated, cioè dotati di RSS. Di solito i lettori di news RSS hanno una funzione per l'aggiunta di nuovi canali che richiede la digitazione dell'URL del file RSS. Non abbiamo trovato programmi che attuino la ricerca automatica del file RSS sulla base dell'URL del sito, ma anche in questo caso sarebbe necessario sapere a priori che un sito supporta l'RSS.

AmphetaDesk si distingue perché, tramite la funzione **Add a channel**, permette la ricerca dei canali da un elenco di migliaia di siti suddivisi per lettera o cifra iniziale. Sono inclusi anche i siti di lingua non inglese, sebbene poco numerosi.

RSS Bandit non offre una lista di siti come AmphetaDesk, ma include una funzione di ricerca

(**Locate Feed**) che, sulla base di parole chiave immesse in una finestra, esegue una ricerca tra le migliaia di siti RSS registrati presso Syndic8.com, uno dei principali syndicate di informazioni RSS.

Se decidete di usare un lettore privo di funzione di ricerca (accettabile se non continuate a cambiare i canali di news), potete sempre cercare gli URL dei file RSS tramite la funzione Search for feed nella home page di www.syndic8.com.

Certi aggregatori hanno come interfaccia una barra lungo cui scorrono i titoli delle news (questa barra si chiama ticker, come i ricevitori telegrafici che stampavano le quotazioni di borsa su una striscia di carta un secolo fa).

Le versioni di RSS

Oggi sono in uso quasi tutte le release del formato RSS. La 0.91 di UserLand (ancora popolare sebbene in teoria superata dalla 2.0) è usata per le funzioni di base di syndication. Le 0.92, 0.93 e 0.94 di UserLand offrono maggiori informazioni rispetto alla 0.91 ma sono rese obsolete dalla 2.0.

La 1.0 del gruppo di lavoro RSS-DEV del W3C, basata su RDF, serve in generale per le applicazioni RDF ed è usata anche per syndication.

La 2.0 di UserLand offre una facile migrazione dalle 0.9x e serve per scopi generali e per

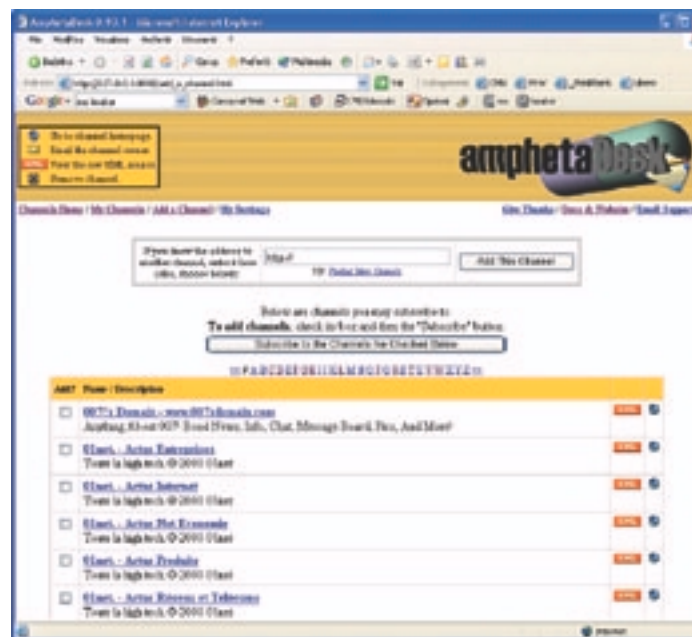
syndication a ricco contenuto di metadati. Un esempio di file RSS 2.0 è quello di UserLand (<http://productnews.userland.com/newsItems/departments/radioUserland.xml>) che contiene anche elementi grafici.

Usiamo AmphetaDesk

Presso il sito di AmphetaDesk (<http://disobey.com/amphetaDesk/>) trovate ampia documentazione e potete scaricare l'ultima versione del software per Windows, Mac e Linux.

Lo abbiamo provato con Windows XP Pro, dopo aver scaricato la versione per Windows 98 e superiore (c'è anche quella per Windows 95) e aver seguito le istruzioni di installazione (il programma è disponibile anche nel **CD Guida 1** allegato alla rivista) ed eseguire AmphetaDesk.exe. Non c'è altro, a parte l'utilità di creare un collegamento sul desktop per eseguire rapidamente il programma.

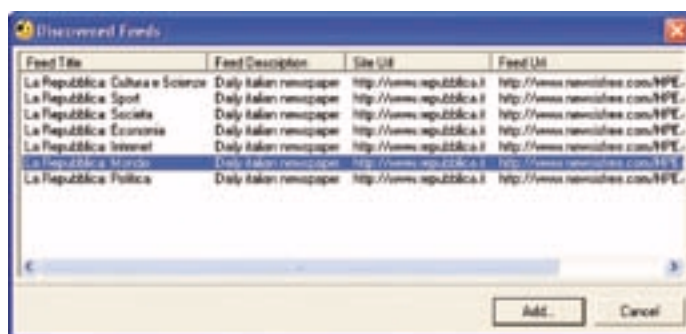
Alla prima esecuzione, AmphetaDesk verifica se ci sono nuove versioni del software e apre in Internet Explorer quella che sarà la vostra pagina Channels Home, con i canali che sceglierete. Inizialmente la pagina è già popolata con alcuni canali di notizie, ma basta un clic sul quadratino con la croce per la loro rimozione. Dopo avere rimossi i canali indesiderati (anche tutti), con un clic su **Add a channel** iniziate la ricerca



Uno dei maggiori pregi di AmphetaDesk è la facilità con cui permette di trovare e aggiungere nuovi canali di notizie



La pagina **My Channels** di AmphetamineDesk elenca i canali a cui l'utente è "abbonato"



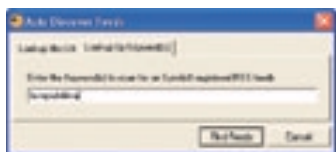
Tra i risultati della ricerca per parole chiave si possono scegliere i canali a cui "abbonarsi"

ca dei canali di vostro interesse, in base alla cifra o lettera iniziale. Se non trovate il sito che cercate, provate a cercarlo con varianti sul nome; i siti italiani con RSS sono pochi, ma ci sono diversi quotidiani, anche se non tutti hanno un file RSS corretto.

Esempi italiani funzionanti sono quelli della Gazzetta dello Sport, La Repubblica e Punto Informatico; Il Sole 24 Ore ha dato errore (anche con RSS Bandit), La Stampa ha dato errore ma ha funzionato con RSS Bandit.

Invece c'è solo l'imbarazzo della scelta con i siti di lingua inglese, che sono parecchie migliaia. Nelle prove, AmphetamineDesk ha riconosciuto correttamente file RSS nelle versioni 0.91, 1.0 e 2.0, che sono la quasi totalità.

Messe le crocette sui canali a cui volete abbonarvi, fate clic sul bottone *Subscribe* e subito trovate le relative news alla pagina Channels Home. L'elenco dei vostri canali è alla pagina My Channels, mentre alla pagina My Settings potete modificare le impostazioni del servizio, per esempio la periodicità degli aggiornamenti in minuti



La funzione **Locate Feed** di RSS Bandit permette di cercare nuovi canali anche per parole chiave, come in questo esempio

(per default 180) e il browser da usare con AmphetamineDesk.

Nel complesso, AmphetamineDesk è il servizio più snello e facile da usare tra quelli provati; è gratuito, multiplatforma e offre una facile ricerca dei canali. L'eseguibile non è piccolissimo (può anche superare i 20 MB in memoria), ma al confronto RSS Bandit usa il doppio di memoria e richiede l'installazione del .Net Framework.

Usiamo RSS Bandit

RSS Bandit ha la tipica interfaccia a tre pannelli dei lettori di news e di email. È un esempio di applicazione .NET ed è descritto sul sito di MSDN (*Microsoft Developer Network*) alla pagina <http://msdn.microsoft.com/library/default.aspx?url=/library/en-us/dnxml/html/xml02172003.asp>. L'autore dell'articolo, che è anche il principale autore del software, elenca i requisiti da cui è partito, tra cui: supporto per le release 0.91, 1.0 e 2.0 di RSS, interfaccia simile ad Outlook Express, browser incorporato per visualizzare le pagine Web, controllo della frequenza di aggiornamento, scorciatoie di tastiera per accelerare la navigazione, marcatura dei messaggi già letti in precedenza (anche se il programma è stato chiuso e riaperto), possibilità di leggere il testo XML di un canale RSS e altro ancora. Tra i requisiti non è citata la ricerca dei canali, che è comunque presente e si basa sull'accesso al database di Syndic8.

Il programma è scaricabile alla pagina www.rssbandit.org/ow.asp?Download che rimanda a www.getdotnet.com; per il download è richiesta la registrazione presso Microsoft Passport, che può avvenire al momento (è gratuita e non richiede l'immissione di un numero di carta di credito). Il file .zip supera di poco il megabyte ed espande un file .msi poco più grande che si installa con clic destro, *Installa*. Se non avete già installato .Net Framework sul PC, lo dovreste scaricare da Microsoft (lungo ma gratuito). In definitiva l'installazione può richiedere tempo (viste le grosse dimensioni di .Net Framework) ma alla fine si dispone di un'applicazione RSS efficiente e con interfaccia compatta.

Per aggiungere i canali potete usare la funzione *New Feed* se avete l'indirizzo dei file RSS, altrimenti usate *Locate Feed* per eseguire la ricerca per parole chiave (il nome della pubblicazione). Il pannello di sinistra di RSS Bandit elenca i canali (feed) che avete scelto; con clic destro sul nome di un canale, potete eliminarlo o eseguire altre funzioni, come aggiornamento, visualizzazione del sito, visualizzazione del file RSS, sua validazione e altro.

Se il file RSS del canale selezionato è valido, nel pannello superiore destro saranno elencati i titoli degli ultimi articoli. Quando selezionate un titolo nel pannello superiore, nel pannello inferiore viene aperto il documento oppure appare un link per aprire la pagina Web, che verrà visualizzata all'interno di RSS Bandit. Quando RSS Bandit funge da browser, i due pannelli di destra sono sostituiti da un unico pannello ad altezza piena dove si può navigare da pagina a pagina. Le pagine aperte in modalità browser

restano disponibili anche quando non sono più visualizzate e si possono rivedere cliccando sulle linguette che compaiono nella barra superiore man mano che si aprono nuove pagine. Sia dal pannello inferiore, sia dal pannello browser, si può stampare la pagina visualizzata con clic destro *Stampa* come in Internet Explorer; manca però la funzione di anteprima di stampa. RSS Bandit offre le funzionalità essenziali di un browser, con la finestra per gli URL dei siti o pagine da aprire e una finestra di ricerca collegata a Google, Feedster e Msn.com.

Risorse

Per iniziare la vostra esplorazione nel mondo dell'RSS potete dare un'occhiata ai siti che offrono aggregatori (vedi tabella) e ai siti dei syndicate (come www.syndic8.com e www.newsisfree.com). Se vi interessa lo sviluppo dei prodotti Microsoft potete trovare l'indirizzo dei file RSS alla pagina <http://msdn.microsoft.com/aboutmsdn/rss.asp> e un canale di news presso <http://rssnewsapps.ziffdavis.com/msw.xml>. Se accarezzate l'idea di attrezzare il vostro sito con RSS potete partire da questa pubblicazione governativa dello Utah: <http://gils.utah.gov/rss/>.

Trovate i file RSS di tre popolari notiziari informativi presso <http://punto-informatico.it/fader/pixml.txt>, www.theinquirer.net/inquirer.rss e <http://trainedmonkey.com/news/rss.php?s=11>.

Il feed di UserLand Software, culla dell'RSS, è presso <http://productnews.userland.com/newsItems/departments/radioUserland.xml>. Se tutto questo non vi basta, una ricerca su Google con RSS trova quasi cinque milioni di pagine Web. ■