

## eContent: tradizionale, semantico o 2.0?

Federico Meschini

Da “*The computer moves in*” a “*Yes, you. You control the information age. Welcome to your world*”. Dall’ormai celebre copertina del *Time* del 3 gennaio 1983<sup>1</sup>, che nominava il personal computer come “*machine of the year*”, a quella, quasi un quarto di secolo dopo, del 25 dicembre 2006<sup>2</sup>, in cui il titolo di “*man of the year*” va alla seconda persona, sia singolare sia plurale, della lingua inglese, scritta, o meglio, digitata al centro di un monitor. Se nel gennaio del 1983 la figura umana veniva presentata stilizzata e glaciale di fronte al computer, sul cui schermo comparivano numeri e grafici, in un ambiente cupo dove, per contrasto, solo la sedia ed il tavolo erano caratterizzati da un aspetto tradizionale e da colori caldi<sup>3</sup>, nella copertina del 2006 questa presenza è totalmente astratta, ridotta all’essenza di un pronome personale, al posto dei grafici nel monitor, ora a cristalli liquidi.

Astratta sì, ma di sicuro predominante e decisiva: ad essere celebrata come persona dell’anno è tutta quella comunità di utenti che ha cambiato la fisionomia del *WorldWideWeb*, che costantemente consuma, rielabora, produce e condivide informazione in formato digitale. Come? Con i *Blog* e i *Wiki*, con le *folksonomies*, con i *tags* ed il *social network*, con le foto condivise su *Flickr*, i video su *YouTube* e i *podcast* sparsi per l’intera rete. Con tutto quello che ora viene definito come *User Generated Content*, uno dei fattori principali dell’attuale upgrade del Web alla versione “2.0”<sup>4</sup>. Ignorare questo fenomeno comunicativo è praticamente impossibile: il *Financial Times* ha pubblicato un articolo sui *Blog* più influenti in America e in Europa<sup>5</sup>; Hillary Clinton ha reso disponibile sul suo sito un video per comunicare la discesa in campo nelle elezioni presidenziali del 2008<sup>6</sup>, e vuole continuare ad utilizzare il *Webcast*<sup>7</sup> per instaurare un dialogo virtuale con gli elettori; Antonio di Pietro, anticipando la senatrice americana, ha deciso di comunicare pubblicamente e in maniera continuativa le decisioni prese al Consiglio dei Ministri, utilizzando *YouTube*<sup>8</sup>; Microsoft ha contattato un informatico, esperto di standard, per modificare dietro pagamento la pagina di

<sup>1</sup> <<http://www.time.com/time/covers/0,16641,19830103,00.html>>

<sup>2</sup> <<http://www.time.com/time/covers/0,16641,20061225,00.html>>

<sup>3</sup> Un approccio visuale estremamente simile sarà utilizzato l’anno successivo nel famoso spot pubblicitario “1984” di Ridley Scott per l’Apple Machintosh. Il video dello spot è facilmente reperibile on-line, per maggiori informazioni vedi <[http://en.wikipedia.org/wiki/1984\\_%28television\\_commercial%29](http://en.wikipedia.org/wiki/1984_%28television_commercial%29)>.

<sup>4</sup> Per un’introduzione al Web 2.0 vedi Tim O’Reilly, *What Is Web 2.0, Design Patterns and Business Models for the Next Generation of Software* <<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1>> o la relativa voce su Wikipedia, <[http://en.wikipedia.org/wiki/Web\\_2.0](http://en.wikipedia.org/wiki/Web_2.0)>.

<sup>5</sup> *European Bloggers find their voice*, <<http://www.ft.com/cms/s/87706c4e-57be-11db-be9f-0000779e2340.html>>. Il *Blog* italiano più influente è quello di Beppe Grillo <<http://www.beppegrillo.it/>>, il quale, per il suo impegno politico e sociale, era già stato eletto nel 2005 dal *Time* come uno degli *European Heroes*, <<http://www.time.com/time/europe/hero2005/grillo.html>>.

<sup>6</sup> <<http://www.hillaryclinton.com/feature/video/?sc=8>>.

<sup>7</sup> La trasmissione di contenuti video e/o audio tramite Internet, <<http://en.wikipedia.org/wiki/Webcast>>.

<sup>8</sup> <<http://www.youtube.com/watch?v=5xRR2iDfvoE>> oppure <<http://www.youtube.com/watch?v=9xcB0Yv0I4M>>.

Wikipedia relativa al formato *Office Open XML*<sup>9</sup>, in modo da rendere, secondo loro, le informazioni presenti più obiettive, facendo così scoppiare un caso mediatico alla pubblicazione della notizia sul *Blog* dell'interessato<sup>10</sup>; la casa editrice *Penguin*, in collaborazione con un'università inglese, sperimenta il formato del "romanzo wiki"<sup>11</sup>.

Nella storia del *Time* non è stata certo la prima volta che a ricevere l'ambito titolo fosse non un singolo bensì un gruppo o un'intera categoria di persone<sup>12</sup>, ma di sicuro la scelta del *Man of Year 2006* rimane al tempo stesso sia peculiare sia estremamente significativa.

Cosa è successo in questi ventitré anni? Quali sono stati i cambiamenti e le innovazioni che hanno portato, dall'introduzione del personal computer nei primi anni '80<sup>13</sup>, alla creazione di questa *information age* distribuita e cooperativa? Di sicuro è stato un percorso non lineare, caratterizzato da numerosi fattori, che combinati hanno prodotto il risultato oggi sotto gli occhi di tutti. Tra i più importanti sicuramente la diffusione e sviluppo di sistemi operativi dotati di un'interfaccia grafica avanzata<sup>14</sup>, più semplice e immediata da utilizzare rispetto alla riga di comando, con la conseguente possibilità di manipolare facilmente non solo contenuti testuali e numerici, ma anche immagini, audio e video. Determinante invece è stato lo sviluppo delle reti telematiche, l'iniziativa che ha portato alla creazione di Internet così come la conosciamo, e, tra il 1989 e il 1991 l'invenzione dell'architettura e delle tecnologie alla base del *WorldWideWeb* da parte di Tim Berners-Lee.

---

<sup>9</sup> <<http://en.wikipedia.org/wiki/OOXML>>.

<sup>10</sup> *An interesting offer: get paid to contribute to Wikipedia*, <[http://www.oreillyn.com/xml/blog/2007/01/an\\_interesting\\_offer.html](http://www.oreillyn.com/xml/blog/2007/01/an_interesting_offer.html)>.

<sup>11</sup> <<http://www.amillionpenguins.com/>>.

<sup>12</sup> Ad esempio gli scienziati americani nel 1960 e la generazione degli under 25 nel 1966. Per un elenco completo vedi <<http://history1900s.about.com/library/weekly/aa050400a.htm>>. Secondo Valerio Evangelisti, noto scrittore di fantascienza, il titolo in realtà sarebbe dovuto andare a Hugo Chávez, vedi l'articolo

*HUGO CHÁVEZ, Uomo dell'Anno per Carmilla*,

<<http://www.carmillaonline.com/archives/2007/01/002091.html#002091>>.

<sup>13</sup> All'epoca distinto ulteriormente tra *personal* e *home computer*: il primo, spesso dotato di una maggiore potenza di calcolo, per compiti più professionali e seri; il secondo più versatile e destinato all'utilizzo da parte di tutti i membri di una famiglia (da qui il termine *home*), dalla gestione della contabilità domestica allo studio e svago dei componenti più piccoli. Due diversi spot televisivi possono rendere bene l'idea. Il primo è dell'IBM, realizzato nel 1986 con il provocatorio testo iniziale "Computers? Who's got time to figure out how to use them.", mostra come l'utilizzo di un personal computer possa aumentare la produttività e risolvere la maggior parte dei problemi in tutti i settori professionali, <<http://www.youtube.com/watch?v=RcUs35JaPPM>>. Il secondo, sempre dello stesso periodo è invece della Commodore, sul suo prodotti di punta dell'epoca, il C64 <[http://www.youtube.com/watch?v=w\\_YAbZjTBMI](http://www.youtube.com/watch?v=w_YAbZjTBMI)>. L'immagine presentata è in entrambi i casi familiare e rassicurante, in quanto lo scopo era di creare un mercato ex-novo. Per avere un'idea di quanto la situazione sia evoluta e i toni comunicativi cambiati, per non dire stravolti, basta vedere spot più recenti sia dell'IBM (<<http://www.youtube.com/watch?v=rmxPfZtV6w0>>, <<http://www.youtube.com/watch?v=2tJH5MVR2U>>, <<http://www.youtube.com/watch?v=sOtKZA9ri7M>>), destinati a un target prettamente enterprise e business, sia delle console di ultima generazione per l'*home entertainment* (<<http://www.youtube.com/watch?v=ALl6cagS2ZE>>, <<http://www.youtube.com/watch?v=n2DlGcpR2nA>>).

<sup>14</sup> Tendenza continuamente in atto, basti pensare a Vista, l'ultima versione di Windows, al Mac Os X, o Linux con OpenGL.

Nato con lo scopo di gestire l'informazione sugli esperimenti al CERN, il sistema ipertestuale progettato da Berners-Lee e denominato originariamente “*Mesh*”<sup>15</sup>, era utilizzato a malapena su scala locale<sup>16</sup>, fino a quando dall'altra parte del mondo, all'Università dell'Illinois non venne realizzato *Mosaic*, il primo browser, che diede l'impulso decisivo alla diffusione del Web così come lo conosciamo. Sin dall'inizio il linguaggio di marcatura *HTML*<sup>17</sup> è stato lo strumento principale per la pubblicazione delle informazioni in questo “docuverso ipertestuale”, caratterizzate quindi da una natura fortemente testuale. Sebbene sia possibile inserire in una pagina web anche immagini, audio e video, tutti questi elementi multimediali rimangono però, per forza di cose, incastonati nella cornice documentale sottostante la natura stessa dell'*HTML*, legata alle sue origini di applicazione *SGML*<sup>18</sup>, un metalinguaggio di marcatura estremamente potente ma di non facile implementazione, la cui stessa complessità ne ha limitato l'utilizzo ad applicazioni di grande scala e/o di nicchia<sup>19</sup>.

*HTML* è stato, e continua ad essere, il dialetto di *SGML* più diffuso e utilizzato, contribuendo così a rendere popolare il suo ingombrante genitore presso tutti quegli utenti e sviluppatori che, fino a quel momento, o non ne conoscevano l'esistenza o lo consideravano estremamente pesante ed esoterico. Si inizia presto a pensare a un modo per utilizzare le potenzialità di *SGML* sul Web, senza però dover pagare troppo in complessità e mantenendo il più possibile la semplicità d'utilizzo, uno dei fattori chiave del successo di *HTML*.

I limiti di quest'ultimo linguaggio si iniziano infatti a far sentire sin da subito. Da un lato ci sono le limitazioni intrinseche dei linguaggi di marcatura descrittivi<sup>20</sup>, il cui scopo principale è quello, per l'appunto, di descrivere un documento, lasciando quindi poco o addirittura nessuno spazio alla manipolazione dei dati e, di conseguenza, all'interattività che, in un ambiente come il *WorldWideWeb*, si rivela invece essenziale. I link ipertestuali, all'apparenza interattivi, contengono semplicemente l'informazione (statica) relativa al documento di destinazione, è poi il browser a fare tutto il lavoro, a trasportare l'utente alla destinazione prevista dal collegamento. Un'attività come la ricerca in un *OPAC*, apparentemente semplice, comporta un processo in cui il ruolo delle pagine

---

<sup>15</sup> La proposta originale del 1989 è disponibile all'url <<http://www.w3.org/History/1989/proposal.html>>. È singolare il commento di uno dei suoi capi: “*vague but exciting*”.

<sup>16</sup> Addirittura un paper sull'argomento venne rifiutato ed accettato solo come poster alla *Hypertext '91 Conference* <<http://www.w3.org/2004/Talks/w3c10-HowItAllStarted/?n=16>>.

<sup>17</sup> Inventato sempre da Tim Berners-Lee, <<http://www.w3.org/MarkUp/>>, insieme al protocollo di comunicazione *HTTP* e al meccanismo degli *URL*.

<sup>18</sup> *Standard General Markup Language*. Per ulteriori informazioni vedi; <<http://www.oasis-open.org/cover/general.html>>; <<http://www.w3.org/MarkUp/SGML/>>; <<http://etext.virginia.edu/bin/tei-tocs?div=DIV1&id=SG>>.

<sup>19</sup> In particolare nell'industria aerospaziale, nella gestione della documentazione tecnica e legale, o nelle prime versioni, per l'appunto “*SGML-based*”, dei linguaggi *DocBook* <<http://www.docbook.org/>>, o *TEI* <<http://www.tei-c.org/>>, rispettivamente per la manualistica il primo e la codifica dei testi umanistici il secondo.

<sup>20</sup> La tipologia a cui *SGML* e, logicamente, *HTML* appartengono. L'altra grande famiglia è quella dei linguaggi di marcatura procedurali, come TeX, LaTeX e Postscript. Una delle distinzioni principali è che i linguaggi di marcatura descrittivi si concentrano sulla struttura di un documento, mentre quelli procedurali sulla presentazione, e spesso arrivano ad inglobare al loro interno un linguaggio di programmazione vero e proprio.

Web si ferma allo strato presentazionale: la descrizione della maschera di ricerca e la successiva visualizzazione dei risultati. Tutte le altre operazioni, l'invio dei dati inseriti, il confronto con i dati memorizzati e l'invio dei risultati, comporta un ruolo attivo che esula dalle caratteristiche di *HTML*.

Per questo motivo è presente tutta una serie di tecnologie, dai nomi e dagli scopi più disparati, che ha lo scopo di aumentare le potenzialità e le funzionalità del Web: i primi programmi "lato server", dall'enigmatico nome di *CGI* o *Common Gateway Interface*<sup>21</sup>; il debutto di *Java* con le *Applet*<sup>22</sup>; i linguaggi di scripting da eseguire dentro il browser, *JavaScript* e *VBScript*; le animazioni in *Flash*; i linguaggi di programmazione orientati o adattati per il web, come *ASP*, *PHP*, *JSP*, che crescono fino a svilupparsi in *framework*<sup>23</sup>, ed ecco *.NET* e *J2EE*, rispettivamente di Microsoft e Sun, o i più recenti e promettenti *Django* e *Ruby on Rails*. Il panorama è estremamente variegato e multiforme, ma è sulla base di queste tecnologie, e sulla loro combinazione, che si è sviluppato, con una morfologia in continua evoluzione, il cosiddetto "*Web 2.0*", caratterizzato, tra le altre cose, dalle cosiddette *RIA*, o *Rich Internet Applications*, applicazioni web dotate di funzionalità simili a quelle dei programmi eseguiti normalmente sui personal computer<sup>24</sup>.

Questo, dicevamo, rispetto alle limitazioni dei linguaggi di marcatura in generale. Ma per quello che riguarda invece specificatamente *HTML*? Che problemi presenta, e come sono stati, o si è tentato di risolverli? Come scritto precedentemente *HTML* è un sottoinsieme di *SGML* e, per renderne l'utilizzo più semplice possibile, un sottoinsieme abbastanza limitato, cui mancano quindi alcune caratteristiche fondamentali, come ad esempio una chiara distinzione tra l'aspetto presentazionale e quello strutturale. Una prima soluzione è stata l'introduzione dei fogli di stile *CSS*<sup>25</sup>, che aumentano le possibilità presentazionali ed effettuano una prima separazione tra aspetto visivo (nel foglio di stile *CSS*) e contenuto (nel documento *HTML*), ma, per quanto utile, questo non è ancora abbastanza, in quanto il problema è di base, e necessità di un'altra soluzione.

L'idea, come scritto precedentemente, è quella di portare *SGML* sul Web, secondo però il principio "80/20": l'ottanta per cento delle possibilità, al venti per cento della complessità<sup>26</sup>. Così

<sup>21</sup> Per maggiori informazioni vedi <<http://cgi.resourceindex.com/>> o, in italiano, <<http://www.mrwebmaster.it/fareweb/cgi/>>.

<sup>22</sup> Le *Applet* sono dei programmi che possono essere inseriti in una pagina web e visualizzati ed eseguiti in un browser. Ad esempio il classico gioco "Breakout" è disponibile on line a <<http://www.geocities.com/SiliconValley/Bay/6879/Breakout.html>>. Viste con estremo interesse a metà anni '90, l'utilizzo delle *Applet* è ormai, per fortuna, estremamente limitato, e Java viene utilizzato con maggiore profitto in attività lato server, dall'altra parte del browser.

<sup>23</sup> Strutture di supporto per facilitare la la creazione, lo sviluppo e l'organizzazione dei progetti.

<sup>24</sup> Il numero, e fortunatamente anche la qualità, delle *Rich Internet Applications* cresce continuamente. Tra le principali e più note *Flickr* <<http://www.flickr.com/>>, *Google Maps* <<http://maps.google.it/>>, *Google Docs* <[docs.google.com/](http://docs.google.com/)>, *GeoBios* <[http://en.wikipedia.org/wiki/List\\_of\\_Rich\\_Internet\\_applications](http://en.wikipedia.org/wiki/List_of_Rich_Internet_applications)>. Una lista iniziale si può trovare alla relativa pagina di Wikipedia <[http://en.wikipedia.org/wiki/List\\_of\\_Rich\\_Internet\\_applications](http://en.wikipedia.org/wiki/List_of_Rich_Internet_applications)>.

<sup>25</sup> <<http://www.w3.org/Style/CSS/>>

<sup>26</sup> Noto anche in altri settori, in particolare nel management, come il "Principio di Pareto", dal nome dell'economista italiano che per primo rilevò il fenomeno. Vedi <[http://en.wikipedia.org/wiki/Pareto\\_principle](http://en.wikipedia.org/wiki/Pareto_principle)>, <<http://www.gassner.co.il/pareto/>> o, per un'applicazione alle attività quotidiane, R. KOCH, *The 80/20 Principle: The Secret to Success by Achieving More with Less*, Currency, 1999.

nel 1996 iniziano i lavori su una versione alleggerita di *SGML* e specificatamente pensata per il *WorldWideWeb*. Nel novembre dello stesso anno viene pubblicata la prima *working draft* delle specifiche del nuovo (meta)linguaggio, dal nome semplice ma efficace: *XML*, o *eXtensible Markup Language*<sup>27</sup>. Le specifiche ufficiali arriveranno due anni dopo, nel 1998, e sin dall'inizio lo scopo di *XML* è chiaro: portare ordine nel Web, all'epoca già percepito come medium fortemente entropico. Il nucleo centrale di *XML* consiste in una sintassi di base per la definizione di linguaggi di marcatura, ma questo da solo non è sufficiente per un'implementazione concreta, ed ecco allora comparire tutta una serie di tecnologie collegate, sia adattate da *SGML* sia concepite ex-novo: grammatiche per la creazione di schemi, *DTD*, *XMLSchema*, *RelaxNG*; fogli di stile per la visualizzazione dotati di capacità di selezione ed iterazione avanzate, *XSLT* e *XSL-FO*; regole di selezione di parti del documento *XML*, evolute poi in un vero e proprio linguaggio di interrogazione, *XPath* e *XQuery*; interfacce per i linguaggi di programmazione, *DOM* e *SAX* insieme a tutta una serie di strumenti per gestire e manipolare i documenti scritti utilizzando questa sintassi. A dieci anni dal rilascio delle prime bozze, il bilancio è decisamente positivo: *XML* è una tecnologia affermata ed estremamente diffusa, con un utilizzo che ha scavalcato quelli che erano i suoi confini naturali, ossia un impiego strettamente collegato alla disseminazione di informazioni strutturate e, fino ad un certo livello, semantiche sul Web<sup>28</sup>.

Non mancano naturalmente le critiche, in qualche caso neppure troppo infondate e per certi versi condivisibili, all'*hype*, all'eccessiva pubblicità cui *XML* è sottoposto, ad un utilizzo spesso non troppo appropriato, e ai difetti di fondo del modello stesso, in particolare sulla verbosità e ridondanza<sup>29</sup>.

Rimane comunque un aspetto fondamentale, *XML* è ormai dappertutto<sup>30</sup>, tranne nel posto principale in cui avrebbe dovuto essere: nel Web, dalla parte dei client. Infatti, nella visione originaria avrebbe dovuto sostituire *HTML* come formato inviato ai browser web, che avrebbero poi provveduto a visualizzare i documenti secondo delle istruzioni di formattazione separate dal contenuto stesso<sup>31</sup>. Questo non è successo, a parte qualche rara eccezione, e si è pensato fosse più conveniente impiegarlo maggiormente dall'altra parte, sul sever, continuando ad inviare al browser sempre del codice *HTML*<sup>32</sup>, ottenuto tramite interrogazioni, trasformazioni e manipolazioni dei dati

---

<sup>27</sup> <<http://www.w3.org/XML/>>

<sup>28</sup> Per dare un'idea del livello di diffusione, basti pensare che nelle suite di office automation più diffuse, *Microsoft Office* e *OpenOffice*, *XML* viene utilizzato come formato nativo di memorizzazione.

<sup>29</sup> Vedi il sito web *XMLSucks* <<http://xmlsucks.org/>> o il relativo wiki <<http://c2.com/cgi/wiki?XmlSucks>>.

<sup>30</sup> Darin McBeath, Chief Architect di Elsevier, durante il convegno <XML2006> ha dichiarato che le cinque tecnologie ad aver avuto il maggior impatto nell'industria editoriale negli ultimi dieci anni sono state rispettivamente *XML*, *XSLT*, *PDF*, *Java* e *Internet* <[http://www.oreillynet.com/xml/blog/2006/12/xml\\_conf\\_2006\\_second\\_day.html](http://www.oreillynet.com/xml/blog/2006/12/xml_conf_2006_second_day.html)>.

<sup>31</sup> Ad esempio in un file *CSS* o utilizzando tecnologie più avanzate, come *XSL-FO*.

<sup>32</sup> Sebbene nella nuova versione *XHTML* <<http://www.w3.org/TR/xhtml1/>>, rielaborata secondo le regole sintattiche del fratello maggiore, invece che quelle tradizionali di *SGML*. Nell'*XHTML* il ruolo dei *CSS* è notevolmente aumentato, ma rimane sempre un singolo linguaggio, rispetto alla possibilità di avere circolanti per il web numerose applicazioni *XML*,

in *XML*. Nel mondo digitale, la maggior parte dell'informazione strutturata è contenuta nei database relazionali, secondo un modello data-centrico, orientato verso i dati. Il Web è invece un medium documento-centrico, orientato verso i documenti. Questi due modelli sono tra di loro estremamente diversi e non isomorfici. I database relazionali sono sì enormemente utilizzati per le applicazioni Web<sup>33</sup>, ma rimangono sempre sullo sfondo, come *back-end*, per la memorizzazione dei dati.

Oltretutto i database utilizzano un metodo di memorizzazione proprietario, che rende complicata la comunicazione dei dati tra due sistemi diversi. *XML* è un passo in avanti, in quanto la sua sintassi si adatta sia al modello documento-centrico sia a quello data-centrico, ed essendo uno standard aperto, è estremamente utile per la trasmissione e l'interscambio dei dati. Da un lato quindi i database tradizionali diventano in grado di manipolare questo nuovo formato, dall'altro iniziano a comparire i primi *Native XML database*, in cui il modello interno di rappresentazione non è più quello relazionale<sup>34</sup>. Un ponte tra questi due diversi mondi è il linguaggio *XQuery*, che permette d'interrogare sia i dati *XML*, sia quelli che possono essere "visti" come tali<sup>35</sup>.

Di conseguenza, seppure lontana dall'annunciata rivoluzione, la situazione è migliorata notevolmente ed è stata introdotta una tecnologia, o meglio un insieme di tecnologie estremamente utili.

Quindi qual è il punto? Il fatto è che se osserviamo lo schema originale del sistema di *Information Management* progettato da Tim Berners-Lee<sup>36</sup>, ci si accorge facilmente di quanto sia estremamente distante dall'attuale configurazione del *WorldWideWeb*. Il Web è infatti un insieme di documenti, più o meno strutturati, collegati tra di loro in maniera ipertestuale, mentre il modello di Berners-Lee è un grafo ipertestuale sì, ma in cui ogni collegamento è fortemente tipizzato ed influenza quindi la relazione semantica di un nodo, anch'esso tipizzato, con l'altro; relazioni semantiche che nello stato attuale del Web sono molto deboli se non praticamente nulle. Evidentemente non soddisfatto dell'evoluzione della sua creatura, e con ancora il modello originario in mente, Tim Berners-Lee continua a sviluppare la sua idea, che viene esposta sia in un

---

ognuna riferita al loro particolare dominio di conoscenza.

<sup>33</sup> Nella maggior parte degli *OPAC* o dei *Content Management Systems* è presente un database in cui sono contenuti i dati, che vengono manipolati tramite un linguaggio di programmazione. Molto diffuso è l'utilizzo della coppia *PHP/MySQL*, in particolare per i software *open-source*.

<sup>34</sup> Anche se ancora distanti dalle prestazioni e dalla diffusione dei database relazionali, i *Native XML Database* hanno raggiunto una certa maturità. Degno di nota è *eXist*, *open-source* e recentemente arrivato alla versione 1.0 <[exist-sourceforge.net](http://exist-sourceforge.net)>.

<sup>35</sup> Applicazioni estremamente interessanti basate su *XQuery*, che gestiscono quantitativi notevoli di dati non omogenei tra di loro, sono *Scopus* della *Helsevier* <<http://www.scopus.com/>>, *SafariU* della *O'Reilly* <<http://www.safariu.com/>>, *Navio* <<http://www.navio.com/>>, *NASA Astronomical Data Center* <[www.cogneticsystems.com:8080](http://www.cogneticsystems.com:8080)>. Data Direct ha sviluppato un'applicazione per l'interrogazione di fonti eterogenee di dati tramite questo linguaggio, <<http://www.datadirect.com/products/xquery/index.ssp>>.

<sup>36</sup> <<http://www.w3.org/History/1989/Image1.gif>>.

libro<sup>37</sup> sia in un articolo pubblicato su *Scientific American*, con un titolo eponimo: *The Semantic Web*<sup>38</sup>.

Il Web Semantico non è, utilizzando le parole dello stesso autore, “[...] a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation”. Di conseguenza l’informazione deve in qualche modo diventare *machine-understandable*, assumere un livello di significato, o perlomeno di significato estremamente strutturato, anche per software avanzati, nell’articolo denominati “Agenti”, che siano in grado di effettuare inferenze, ottenendo così in maniera automatica nuove informazioni che verranno inviate e utilizzate dall’utente finale, o eventualmente da un altro agente software in un circolo virtuoso di creazione e utilizzo della conoscenza. Provando a utilizzare un paradigma del Web 2.0, la *cloud tags*, la nuvoletta di marcatori associati ad una risorsa o un concetto, di sicuro tra le *hot words* del *Semantic Web* troveremo etichette come *Ontology*, *Knowledge Representation*, *Knowledge Management*, *Inference Rules*, *Metadata* e *Artificial Intelligence*; infatti gli sforzi, in particolare sulla realizzazione degli standard sui cui si basa/dovrà basare, sono proprio in questa direzione<sup>39</sup>. Ricollegandosi al discorso precedente delle basi di dati, il *Semantic Web* è un enorme database interoperabile e distribuito, che utilizza la stessa infrastruttura del *WorldWideWeb*.

A più di cinque anni dall’articolo su *Scientific American*, il *Semantic Web* rimane però un qualcosa ancora in divenire, ed estremamente distante dall’utilizzo quotidiano, nonostante i lavori stiano continuando e con risultati estremamente interessanti<sup>40</sup>. I sostenitori del *Semantic Web* sono numerosi e appassionati, ed effettivamente è difficile non rimanere positivamente impressionati dall’enorme potenziale implicito. I detrattori però, seppure non criticando direttamente la visione astratta, non mancano di sottolineare come una sua realizzazione, anche parziale, sia estremamente difficile, se non impossibile, da attuare.

Sembra quasi di rivivere la stessa situazione di *XML* alla fine degli anni ’90, quando le potenzialità erano sotto gli occhi di tutti, insieme ad un discreto numero di applicazioni interessanti,

<sup>37</sup> T. BERNERS-LEE, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*, Collins, 2000. In italiano T. BERNERS-LEE, *L’architettura del nuovo Web*, Feltrinelli, Milano 2001.

<sup>38</sup> T. Berners-Lee, J. Hendler, O. Lassila, *The Semantic Web*, *Scientific American*, May 2001, pp. 34–43, <<http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&pageNumber=1&catID=2>>.

<sup>39</sup> Un’idea dei livelli progressivi su cui si appoggia il *Semantic Web* è l’ormai famosa torta: <<http://www.w3.org/2002/Talks/04-sweb/sw-stack-2002.png>>.

<sup>40</sup> In particolare le specifiche per la rappresentazione dei dati hanno raggiunto un buon livello di maturità. Principalmente viene utilizzata una struttura a livelli, in cui si parte da *RDF* <<http://www.w3.org/RDF/>> per rappresentare i dati tramite un grafo, *RDFS* <<http://www.w3.org/TR/rdf-schema/>> per strutturare in classi e proprietà le informazioni rappresentate tramite *RDF*, e infine *OWL* <<http://www.w3.org/2004/OWL/>> per la definizione di ontologie di dominio. Anche le applicazioni basate su queste tecnologie stanno raggiungendo un buon livello sia qualitativo sia quantitativo. Degni di nota sono gli italiani *Dbin* <<http://dbin.org/>>, basato su di un’architettura peer2peer per la creazione e lo scambio di ontologie, e *HyperJournal* <<http://www.hjournal.org/>>, per la pubblicazione di *eJournals open access*.

mentre ciò che in realtà mancava era un utilizzo concreto e diffuso, avvenuto in realtà solo dopo qualche anno<sup>41</sup>. Un'altra caratteristica in comune ad *XML*<sup>42</sup> è l'*hype*, l'attenzione spesso esagerata alla quale entrambi sono stati sottoposti, in particolare come risolutori di tutti i problemi del Web.

Ma mentre si teorizza, si parla, si critica, si sperimenta, si costruisce il *Semantic Web*, e *XML* è uno standard affermato, ecco all'improvviso comparire il *Web 2.0*, anche se qualche avvisaglia c'era stata, come ad esempio la diffusione ed il ruolo sempre maggiore dei *Wiki* e dei *Blog*. Quello che mancava era un'etichetta, un termine-ombrello come direbbe Umberto Eco, che potesse raccogliere e identificare le diverse metamorfosi in atto, che fosse anche un meme spendibile ed efficace a livello di marketing. Perché il *Web 2.0* è anche, e soprattutto un mercato concreto, laddove il *Semantic Web* è visto fondamentalmente come ricerca accademica, nonostante le possibilità a livello di applicazioni business siano enormi. E ancora, il *Semantic Web* nasce dall'alto, direttamente dal *W3C*; il *Web 2.0* nasce dal basso, da tutta una serie di iniziative eterogenee tra di loro con in comune determinate tecnologie o una particolare visione. Il *Semantic Web* è rappresentato dalle ontologie, un modello entità-relazioni che rappresenta una "*specification of conceptualization*", secondo la definizione data nelle scienze dell'informazione; il *Web 2.0* è caratterizzato dalle *folksonomies*, le classificazioni collaborative libere e non gerarchiche basate sui *tag*<sup>43</sup>. Da un lato abbiamo le inferenze basate su procedimenti logici rigorosi che producono nuova conoscenza, dall'altro invece il *Mash-Up*, la combinazione e il riutilizzo di informazioni e servizi<sup>44</sup>. Tra le varie tecnologie utilizzate, il *Semantic Web* ha l'enigmatico, a livello di nome, *SPARQL*<sup>45</sup>, il *Web 2.0* il più evocativo *Ajax*<sup>46</sup>.

Le differenze non finiscono certo qui e sembrerebbero suggerire una profonda differenza, se non addirittura un'incompatibilità di fondo. Durante un'intervista per l'IBM, nel luglio del 2006, Tim Berners-Lee si è espresso in proposito, dichiarando che "*Web 2.0 is of course a piece of jargon, nobody even knows what it means*"<sup>47</sup> e, in un articolo pubblicato due mesi prima confronta

---

<sup>41</sup> Senza considerare le applicazioni di nicchia, in linea di massima si può considerare il 2002 come l'anno della diffusione di *XML*, in particolare, ma non solo, con l'utilizzo dei feed *RSS* per i *Blog*.

<sup>42</sup> Oltre al fatto naturalmente che le specifiche *RDF*, *RDFS* e *OWL*, prevedono una serializzazione che utilizza questa sintassi.

<sup>43</sup> Un intervento interessante su questo argomento è quello effettuato da Gino Roncaglia, *Folksonomies e semantiche controllate: due strade conciliabili*, durante il convegno *Interoperabilità delle Biblioteche Digitali*, svoltosi all'Università di Roma Tre il 9 e 10 Ottobre 2006. Le slide sono disponibili all'url <[http://www.sba.uniroma3.it/interopbd/Roncaglia\\_Folksonomies.pdf](http://www.sba.uniroma3.it/interopbd/Roncaglia_Folksonomies.pdf)>.

<sup>44</sup> *Mix, Match, And Mutate* in *Business Week* del 25 luglio 2005. Un esempio di *MashUp* di servizi è quello dato dall'utilizzo combinato di *Google Base* <<http://base.google.com>> con *Google Maps* <<http://maps.google.com>>, per localizzare geograficamente gli annunci.

<sup>45</sup> Acronimo ricorsivo di *SPARQL Protocol and RDF Query Language*, un linguaggio d'interrogazione per *RDF*, <<http://www.w3.org/TR/rdf-sparql-query/>>.

<sup>46</sup> Acronimo di *Asynchronous JavaScript and XML*, una combinazione di *JavaScript* e *XML* che permette di effettuare chiamate asincrone al web server in modo da aggiornare interattivamente i dati presenti nella pagina web senza dover ricaricare la pagina. *Ajax*, non una tecnologia di per sé, ma un riutilizzo di tecnologie già esistenti, è al cuore di numerose applicazioni 2.0, <[http://en.wikipedia.org/wiki/Ajax\\_%28programming%29](http://en.wikipedia.org/wiki/Ajax_%28programming%29)>.

<sup>47</sup> <<http://www-128.ibm.com/developerworks/podcast/dwi/cm-int082206.txt>>.



ontologie e *folksonomies*, dichiarando che servono a due scopi fondamentalmente diversi: “*The folksonomies [...] are an interesting emergent attempt at information retrieval. But folksonomies serve very different purposes from ontologies. [...] The inferential process applied to ontologies is logic based and uses operations such as join. The inferential process used on tags is statistical in nature and employs techniques such as clustering.*”<sup>48</sup>

È realmente così? Quali sono le alternative possibili? I due modelli sono: totalmente inconciliabili, perfettamente complementari o solo parzialmente sovrapponibili? In realtà i punti di contatto sono diversi. Ad esempio i *feed Rss*, una delle caratteristiche principali dei *Blog*, nascono come applicazione di *RDF*<sup>49</sup>, iniziano a diffondersi i primi *Semantic Wiki*<sup>50</sup>, e all’inizio *Web 2.0* e *Semantic Web* venivano utilizzati come sinonimi.

Qualcuno ha anche creato, forse un po’ troppo semplicisticamente, l’equazione “*Web 2.0 + Semantic Web = Web 3.0*”<sup>51</sup>. Indubbiamente sia l’uno sia l’altro sono qualcosa di diverso rispetto al Web tradizionale (sempre che una cosa del genere sia mai esistita), ma questo è sufficiente a garantire questo risultato?<sup>52</sup>

Sin dall’inizio il *WorldWideWeb* è stato considerato come un medium estremamente democratico e innovativo, in quanto, rispetto ai media tradizionali, era relativamente facile pubblicare le proprie informazioni, a patto di avere un minimo di competenze tecniche: la conoscenza del linguaggio *HTML* e uno spazio disponibile. A prescindere dal fatto che questi requisiti sono comunque ancora selettivi, il risultato è stato la creazione di innumerevoli “*Information Silos*”, contenitori incapaci di comunicare tra di loro, laddove il valore aggiunto, in un medium decentrato come il Web, è proprio la possibilità di scambio e circolazione delle informazioni. I *personal web sites* sono stati integrati, e in certi casi sostituiti, dai più dinamici *Blog* e *Wiki*, che, una volta installata la relativa piattaforma di gestione, necessitano di conoscenze tecniche minime o addirittura nulle.

Insomma, nonostante le numerose critiche, compresa quella di trovarsi di fronte ad una nuova bolla speculativa come nel 2001<sup>53</sup>, la sensazione del *Web 2.0* è di qualcosa di nuovo, dotato di

<sup>48</sup> Tim Berners-Lee, Wendy Hall, Nigel Shadbolt, *The Semantic Web Revisited*, IEEE Intelligent Systems 21(3) pp. 96-101, May/June 2006, <[http://eprints.ecs.soton.ac.uk/12614/01/Semantic\\_Web\\_Revisted.pdf](http://eprints.ecs.soton.ac.uk/12614/01/Semantic_Web_Revisted.pdf)>.

<sup>49</sup> Vedi anche G. Roncaglia, *Blogosfera e feed RSS: una palestra per il Semantic Web?*, Networks 2, 2003, pp. 47-56. Full-text disponibile a <<http://dspace.unitus.it/handle/2067/9>>.

<sup>50</sup> *Wiki* che incorporano al loro interno un modello di gestione e rappresentazione della conoscenza, basato sui linguaggi del *Semantic Web*, in particolare *RDF*. Vedi ad esempio *OntoWorld* <[http://ontoworld.org/wiki/Main\\_Page](http://ontoworld.org/wiki/Main_Page)>. È disponibile anche una relativa estensione per *MediaWiki*, la piattaforma su cui è basata *Wikipedia* <<http://www.wikipedia.org>>: *Semantic MediaWiki* <[http://ontoworld.org/wiki/Semantic\\_MediaWiki](http://ontoworld.org/wiki/Semantic_MediaWiki)>.

<sup>51</sup> Due articoli interessanti in proposito sono B. Parrella, *Il miraggio del Web 3.0*, su Apogeeonline <<http://www.apogeeonline.com/webzine/2006/12/14/01/20061214011210>> e W. Roush, *What Comes After Web 2.0?* su Technology Review <<http://www.technologyreview.com/Infotech/17845/page2/>>.

<sup>52</sup> La relativa voce di *Wikipedia* <[http://en.wikipedia.org/wiki/Web\\_3.0](http://en.wikipedia.org/wiki/Web_3.0)> che riportava proprio questa teoria è stata cancellata e resa non accessibile. La relativa discussione è sulla pagina di *Talk* <[http://en.wikipedia.org/wiki/Talk:Web\\_3.0](http://en.wikipedia.org/wiki/Talk:Web_3.0)>.

<sup>53</sup> E naturalmente battezzata “*Bubble 2.0*”.

caratteristiche innovative ed interessanti cui, una volta abituati, si potrà difficilmente fare a meno, e verso le quali tutto il resto del Web non ancora “upgradato” verrà progressivamente paragonato, sia a livello di gestione dei contenuti (*folksonomies* e *social networks*) sia a livello di interfacce e applicazioni (*Ajax*, *Rich Internet application*, *WebServices*, *OpenApi*).

Questo dimostra, senza ombra di dubbio, come un utilizzo e uno scambio libero, fluido e continuo di metadati, seppure non estremamente rigorosi, e informazioni porta con sé un enorme valore aggiunto. Come logica conseguenza, dei contenuti strutturati e arricchiti semanticamente possono essere benissimo integrati con un’interfaccia avanzata e/o inseriti in un contesto dinamico e collaborativo, ottenendo così un beneficio reciproco e un risultato che supera la somma delle singole parti. Un caso interessante per analizzare questo rapporto, e che paradossalmente non ha nessuna delle due etichette, né semantico né 2.0, è quello di *Visual Thesaurus*<sup>54</sup>, un dizionario della lingua inglese, che utilizza il software di visualizzazione *ThinkMap*<sup>55</sup> per la creazione di grafi navigabili. Utilizzando questo thesaurus si rimane colpiti dalla ricchezza sia dell’interfaccia sia del vocabolario sottostante, il quale, per qualsiasi termine inserito, presenta tutti i vari significati sotto forma di grafo in cui i vari rami sono tipizzati a seconda della natura della relazione. Inserendo “*Information*” ad esempio, si può navigare il grafo fino ad arrivare ai termini “*Data*” e “*Knowledge*” da un lato, e “*Entropy*” dall’altro. E ancora, “*Love*” presenta numerosi nodi, che coprono tutti i vari significati del termine, dal concetto astratto ai modi di dire, passando per l’attrazione fisica e le implicazioni sessuali; è anche presente, con un rapporto di antinomia, la parola “*Hate*”. Ogni collegamento del grafo può diventare il punto di partenza per un’altra navigazione, coprendo così, virtualmente, tutto il tesoro. Dopo qualche utilizzo risulta chiaro come *Visual Thesaurus* sia in realtà basato su *WordNet*<sup>56</sup>, un database lessicale della lingua inglese in cui i termini sono organizzati tra di loro tramite i rapporti semantici tipici delle classificazioni ontologiche, come iponimia, ipernimia, ologonia e meronimia. *WordNet* è un progetto iniziato nel 1985 ed ha raggiunto una certa maturità e completezza. Il suo utilizzo con *ThinkMap* dimostra come sia possibile creare servizi avanzati, stratificati con un valore aggiunto estremamente elevato, sulla base di contenuti organizzati semanticamente. La sensazione di rigidità che spesso si prova di fronte alle applicazioni del *Semantic Web*, che viene naturale contrapporre alla fluidità del *Web 2.0*, è dovuta in molti casi non alla tecnologia o agli standard in sé, quanto alla qualità dell’ontologia sottostante, spesso non abbastanza completa o bene espressa. Una delle critiche più frequenti, e abbastanza condivisibile, è che le relazioni solitamente utilizzate tra le varie entità, come “figlio di”, “parte di”, “tipo di”, “collegato a”, non riescono a catturare quel grado minimo necessario di

<sup>54</sup> <<http://www.visualthesaurus.com/>>. Il servizio è disponibile dietro abbonamento, ma è possibile usufruire di un periodo di prova.

<sup>55</sup> Basato su tecnologia *Java*, <<http://www.thinkmap.com/>>.

<sup>56</sup> <<http://wordnet.princeton.edu/>>.

ambiguità e di sfumature insite nel linguaggio e nella trasmissione e definizione della conoscenza. D'altro canto essendo un'ontologia un modello formale, un fattore difficilmente controllabile come l'ambiguità è qualcosa che va inevitabilmente ridotto al minimo e laddove possibile controllato. Il caso di *WordNet* dimostra come la riduzione dell'ambiguità sia possibile, a fronte certo di un buon lavoro di progettazione e popolamento del modello in questione.

Un altro esempio di commistione efficace tra questi due modelli è il vincitore dell'ultima edizione del *SemanticWeb Challenge*<sup>57</sup>: il progetto olandese *MultimediaN E-Culture*<sup>58</sup>, che permette l'accesso e la navigazione di collezioni digitali distribuite di “*Cultural Heritage Items*” di diversi musei olandesi. Le varie funzioni di ricerca base, avanzata, navigazione a faccette, ricerca per relazioni e creazione di collezioni personalizzate<sup>59</sup>, utilizzano un'interfaccia basata su *Ajax*, mentre i dati sono chiaramente memorizzati utilizzando le triple di *RDF*. Così come *Visual Thesaurus*, anche *MultimediaN E-Culture*, utilizza vocabolari esterni: i thesauri del *Getty* di arte e architettura<sup>60</sup>, dei nomi di artisti<sup>61</sup> e dei luoghi geografici<sup>62</sup>, il thesaurus etnologico olandese<sup>63</sup>, e il già citato *WordNet*.

Sempre in un ambito accademico/culturale, all'interno del *NINES*, “*Networked Interface for Nineteenth-century Electronic Scholarship*”<sup>64</sup> è stata realizzata l'applicazione *Collex*, sintesi di *collect* ed *exhibit*, per l'appunto con lo scopo di collezionare, organizzare e disseminare più di 50.000 risorse digitali relative al diciannovesimo secolo. L'interfaccia principale, di sicuro perfettibile ma già estremamente interessante, è tipicamente “2.0”<sup>65</sup>, con *Cloud Tag* ed effetti *Ajax*, permette di commentare e annotare gli oggetti, creare collezioni personalizzate da utilizzare a fini didattici o di ricerca, partecipando così allo sviluppo di questa comunità scientifica on line. Tutti i dati su cui *Collex* si basa sono raccolti (o come si dice in gergo “harvestati”<sup>66</sup>) da una serie di archivi digitali già esistenti e tra di loro eterogenei, per *policies*, organizzazioni di appartenenza, implementazioni tecniche e standard. L'unico aspetto comune, e fondamentale ai fini del progetto, è l'essere incentrati su un autore o un movimento artistico relativo al diciannovesimo secolo. Naturalmente questo da solo non basta a superare le difficoltà tecniche relative alla costruzione di un accesso confederato a tutte queste risorse, seppure culturalmente affini. Come riesce *NINES* nel suo scopo? La soluzione adottata è la seguente: ognun singolo progetto, esporta i metadati relativi alle proprie risorse in *RDF*, secondo una classificazione decisa comunemente, che possa contenere

<sup>57</sup> <<http://challenge.semanticweb.org/>>.

<sup>58</sup> <<http://e-culture.multimediana.nl/>>.

<sup>59</sup> Accessibili sotto forma di demo all'indirizzo <<http://e-culture.multimediana.nl/demo/search>>.

<sup>60</sup> <[http://www.getty.edu/research/conducting\\_research/vocabularies/aat/](http://www.getty.edu/research/conducting_research/vocabularies/aat/)>.

<sup>61</sup> <[http://www.getty.edu/research/conducting\\_research/vocabularies/ulan/](http://www.getty.edu/research/conducting_research/vocabularies/ulan/)>.

<sup>62</sup> <[http://www.getty.edu/research/conducting\\_research/vocabularies/tgn/](http://www.getty.edu/research/conducting_research/vocabularies/tgn/)>.

<sup>63</sup> <<http://www.svcn.nl/index.asp>>.

<sup>64</sup> <<http://www.nines.org/>>.

<sup>65</sup> <<http://www.nines.org/collex>>. Infatti è stata implementata tramite il framework *Ruby on Rails*.

<sup>66</sup> *Da to harvest*. Questo neologismo è molto usato nell'*Open Access*, riguardo la cattura, da parte del *Service Provider*, dei metadati OAI-PMH prodotti dai vari *Data Provider*.

in maniera né troppo fluida né troppo restrittiva i vari contenuti<sup>67</sup>. In questo modo i metadati, prodotti singolarmente dai vari archivi, vengono raccolti a livello centrale costruendo così su questa base dei servizi avanzati di disseminazione e organizzazione<sup>68</sup>.

La lista di applicazioni ibride potrebbe continuare ed è in continua crescita, sia per quantità sia per qualità. Degno di nota è il progetto *Simile*<sup>69</sup>, del *M.I.T.*, nato originariamente come spin-off di *DSpace*<sup>70</sup>, con lo scopo principale di estendere le funzionalità e l'interoperabilità delle biblioteche digitali basate su questo software, proprio tramite l'utilizzo delle tecnologie del *Semantic Web*<sup>71</sup>. Gli strumenti sviluppati sino ad ora da un lato sono basati su *RDF* per la creazione, manipolazione, visualizzazione e scambio dei metadati<sup>72</sup>, mentre dall'altro consistono in applicazioni client-side basate su *AJAX*<sup>73</sup>, con dei punti di contatto tra queste due estremità<sup>74</sup> che, visti gli ultimi sviluppi, saranno sempre maggiori.

E per quanto riguarda l'editoria scientifica? Indubbiamente, tra i vari settori, questo è uno di quelli che ha più da guadagnare da un'implementazione concreta del *Semantic Web*<sup>75</sup>, e di sicuro molti lavori e sperimentazioni si stanno effettuando in tale direzione. Anche per quanto riguarda l'aspetto "2.0" qualcosa si sta muovendo, e anche in questo caso, con effetti più visibili e immediati.

Naturalmente data la complessità, la quantità di dati presenti e il modello centrale adottato, effettuare anche un semplice *restyling* o un'aggiunta di servizi di questo tipo alle piattaforme web degli editori è un'operazione né semplice né tanto meno indolore, anche se la questione non va considerata solamente dal punto di vista tecnico. Certe impostazioni di fondo, in particolare sui modelli di business e di mediazione culturale, sono ormai acquisite e consolidate, e di conseguenza

---

<sup>67</sup> Un esempio, insieme ad altre informazioni, è disponibile al relativo *Wiki* del progetto, <[http://faustroll.clas.virginia.edu/ARPwiki/index.php/RDF\\_samples](http://faustroll.clas.virginia.edu/ARPwiki/index.php/RDF_samples)>.

<sup>68</sup> Un approccio simile, seppure su più larga scala e che avrà ulteriori sviluppi, a quello adottato dal progetto *TEL - The European Library*, <<http://www.theeuropeanlibrary.org>>. Tra l'altro una certa somiglia nell'aspetto, e in particolare sui comportamenti dell'interfaccia, fa sospettare l'utilizzo dello stesso framework. *TEL* è basato sul protocollo *OAI-PMH* che fornisce uno standard comune per la trasmissione dei metadati, ma non sul formato né tanto meno sull'ontologia sottostante.

<sup>69</sup> <<http://simile.mit.edu/>>.

<sup>70</sup> La piattaforma open-source per la creazione di *repository* conformi al modello open-access, <<http://www.dspace.org>>.

<sup>71</sup> S. Mazzocchi, S. Garland, R. Lee, *SIMILE: Practical Metadata for the Semantic Web*, <<http://www.xml.com/pub/a/2005/01/26/simile.html>>.

<sup>72</sup> Come *Longwell* <<http://simile.mit.edu/wiki/Longwell>>, un browser a faccette, o *Welkin* <<http://simile.mit.edu/welkin/>>, un visualizzatore grafico.

<sup>73</sup> In particolare *Timeline* <<http://simile.mit.edu/timeline/>>, il corrispettivo di *Google Maps* per gli eventi temporali (e infatti il già citato *MultimediaN E-Culture* utilizza entrambi per contestualizzare le proprie collezioni digitali), e *Exhibit* <<http://simile.mit.edu/exhibit/>>, un framework di pubblicazione totalmente client-side, che utilizza *HTML*, *CSS* e *Javascript*. Notevole è la demo sui presidenti degli Stati Uniti <<http://simile.mit.edu/exhibit/examples/presidents/presidents.html>>, che integra al suo interno sia *Google Maps* sia *Timeline*.

<sup>74</sup> Come *Babel* <<http://simile.mit.edu/babel/>>, per la conversione dei dati nel formato supportato da *Exhibit*.

<sup>75</sup> Come il concetto di *Semantic Publishing*, T. Berners-Lee, James Hendler, *Scientific publishing on the Semantic Web*, <<http://www.nature.com/nature/debates/e-access/Articles/bernerslee.htm>>. Vedi anche <[http://en.wikipedia.org/wiki/Semantic\\_publishing](http://en.wikipedia.org/wiki/Semantic_publishing)>.

l'interesse nel mantenere l'attuale status quo è molto forte. I siti di *Elsevier*<sup>76</sup> e *Blackwell*<sup>77</sup>, ad esempio, per quanto completi, elaborati e stratificati, sono ancora decisamente "1.0", così come i principali software di gestione delle riviste elettroniche, *ScienceServer*<sup>78</sup> o *Journal OnSite*<sup>79</sup>. In ogni caso, applicazioni più recenti, e spesso posizionate ad un livello superiore o decentrato, iniziano ad implementare nuove funzioni che rientrano nei modelli del *Web 2.0*, o che perlomeno gli strizzano l'occhio. I vari portali tematici della casa editrice *O' Reilly*<sup>80</sup>, seppure non propriamente classificabili nella categoria di editoria scientifica specializzata, permettono di associare agli articoli dei *tag* tramite *del.icio.us*, potendo così vedere sia gli altri *tag* utilizzati dai vari utenti, sia gli altri articoli taggati allo stesso modo<sup>81</sup>. I titoli di questa casa editrice, inoltre, sono da sempre uno dei contenuti principali di *Safari Books OnLine*<sup>82</sup>, l'*E-library* di riferimento per il settore dell'*IT*, la cui interfaccia di navigazione si è recentemente rinnovata, utilizzando *AJAX*, in particolare per la lettura online. E ancora, *Scopus*<sup>83</sup>, database di citazioni e abstract di *Elsevier*, integra al suo interno *feed RSS*, suggerimenti contestuali per gli operatori da utilizzare nella ricerca avanzata, ed un'intera sezione denominata *ScopusLab*, con tutta una serie di strumenti di supporto. Oppure *MyAthens*<sup>84</sup>, un servizio di autenticazione e accesso unificato ai siti dei vari editori, che presenta un'interfaccia *AJAX* per elencare e selezionare le varie risorse disponibili all'utente. Questi esempi sono isolati e, rispetto a molte delle applicazioni viste precedentemente, di sicuro limitati, ma dimostrano come l'evoluzione attualmente in corso nel Web sia presente anche in questo settore, che si è accorto dei possibili utilizzi e sta iniziando a sperimentare, ben conscio delle possibilità, senza però voler stravolgere o rivoluzionare i modelli già esistenti.

Per concludere, l'interazione tra *Semantic Web* e *Web 2.0* è estremamente interessante e vantaggiosa per entrambi<sup>85</sup>, e conferma l'impressione che questi due distinti paradigmi siano, anche se non perfettamente, complementari, in quanto entrambi rispondono, con approcci diversi, all'esigenza comune di rinnovamento e aumento delle funzionalità del Web. Naturalmente non è possibile ridurre il tutto all'equazione "informazioni semantiche + interfacce avanzate", che non tiene conto né delle sfumature né delle implicazioni di entrambi i mondi, ma è un terreno abbastanza solido per una cooperazione costruttiva.

---

<sup>76</sup> <<http://www.sciencedirect.com/>>.

<sup>77</sup> <<http://www.blackwell-synergy.com/>>.

<sup>78</sup> Utilizzato in Italia sia dal Caspur <<http://periodici.caspur.it/>> sia dal Cilea <<http://scienceserver.cilea.it/>>.

<sup>79</sup> <<http://www.exlibrisgroup.com/journalsonsite.htm>>.

<sup>80</sup> Come XML.Com <<http://www.xml.com/>> o OnJava <<http://www.onjava.com/>>.

<sup>81</sup> Qualcosa di simile è presente anche nel nostrano ApogeOnLine <<http://www.apogonline.com/>>.

<sup>82</sup> <<http://proquest.safaribooksonline.com>>.

<sup>83</sup> <<http://www.scopus.com/>>.

<sup>84</sup> <<http://auth.athensams.net/my>>.

<sup>85</sup> Vedi anche C. Mitchell, *Music and Metadata* <<http://www.xml.com/pub/a/2006/11/22/music-and-metadata.html>>, che presenta un esempio di applicazione concreta dei metadata semantici, unito agli strumenti citati precedentemente, da *Google Maps* a *Piggy Bank* del progetto *Simile*.

Molto probabilmente, viste le esperienze accumulate fino ad ora, il *Semantic Web* sarà qualcosa di diverso dalla visione e idea originaria di Tim Berners-Lee. È ancora più probabile che non coprirà mai totalmente la ragnatela globale esistente, ma sarà un insieme sovrapposto al suo interno e dalla morfologia variabile. Indubbiamente però, sarà qualcosa di estremamente utile e interessante. Non ci sarebbe neppure da stupirsi se venisse influenzato a sua volta dai paradigmi e dalle applicazioni del *Web 2.0*, mettendo in discussione l'idea secondo la quale quest'ultimo è anarchico e popolare, mentre l'altro accademico e intoccabile.

*Java, Flash, XML, Semantic Web*, sono alcune delle novità, troppo spesso soggette a un *battage* mediatico eccessivo, a cui abbiamo assistito sin dalla metà degli 'anni 90. Nessuna di queste, nonostante le critiche ricevute, è mai scomparsa del tutto, ma si è sempre adattata contribuendo all'evoluzione del *WorldWideWeb*. Sarà lo stesso per il *Web 2.0*? Molto probabilmente sì, e, al di là delle questioni legate al marketing e alle inevitabili speculazioni, potrebbe essere veramente l'ingresso in una nuova fase di maturità del Web.

Perlomeno fino alla prossima copertina del *Time*.