

Il World Wide Web: una meravigliosa avventura scientifica, tecnologica e umana

Oreste Signore

Email: oreste.signore@gmail.com - Tel. 339-3082198
personal home page: <http://www.orestesignore.eu/weblab/people/oreste/>

Sommario - Il Web è il risultato di una sintesi di una serie di idee e di tecnologie (Internet, SGML, ipertesto) e nasce in un ambiente di ricerca di punta, grazie all'intuizione geniale di Tim Berners-Lee e alla sua perseveranza nel difenderne i principi base. L'esplosione del Web porta alla nascita del World Wide Web Consortium (W3C), fondato per regolarne lo sviluppo e portarlo al massimo del suo potenziale, definendone gli standard tecnici con un processo aperto che coinvolge l'intera comunità scientifica. Gli standard di interoperabilità definiti dal W3C consentono il superamento delle barriere tecnologiche, economiche, culturali e fisiche, valorizzando così l'enorme valenza sociale del Web e consentendone la diffusione anche nei paesi più poveri.

I principi e le motivazioni che hanno portato all'invenzione del Web, che si presenta oggi come una piattaforma unificante per tutte le applicazioni (One Web Platform, HTML5) sono ancora rispettati nei suoi sviluppi più avanzati e interessanti (Linked Open Data, Semantic Web).

Il Web, che costituisce una ricchezza per il mondo intero, viene ormai ritenuto un servizio garantito, ma occorre un impegno costante da parte di tutti per evitarne la frammentazione e mantenere intatti i suoi principi di universalità, indipendenza, rispetto dei diritti umani.

Introduzione

La combinazione di una serie di condizioni tecnologiche e applicative ha portato Tim Berners-Lee ad inventare il Web, che oggi è parte integrante della nostra vita quotidiana. Se alcuni anni fa si diceva: "*L'ho sentito in televisione*", ora si dice: "*L'ho trovato sul Web*", oppure "*Lo cerco sul Web*".

Lo sviluppo delle tecnologie ha consentito un atteggiamento più partecipativo da parte degli utenti, e l'esplosione del fenomeno noto come "Web 2.0" ha portato alla diffusione delle reti sociali. Una frase ricorrente adesso è: "*Lo metto sul Web*".

Non è però facile sfruttare adeguatamente l'enorme patrimonio informativo esistente sul web, perché le informazioni sono archiviate in maniera eterogenea. Il Semantic Web si propone come l'ambiente in cui superare queste difficoltà, consentendo il reperimento, l'integrazione e l'utilizzo delle informazioni.



Le radici del Web

L'invenzione del Web è il risultato di un lungo percorso, cui hanno contribuito molti ricercatori, favorito dalla definizione dello standard SGML e dagli sviluppi delle reti, partendo dalle idee pionieristiche di Vannevar Bush sull'ipertesto, e passando attraverso l'invenzione del mouse e le visioni affascinanti di Ted Nelson. Non è stata un'illuminazione improvvisa, come la mitica "mela di Newton", ma un'idea nuova, quella di poter *combinare in modo libero* le idee, stimolata da una sfida posta da un ambiente di ricerca di altissimo livello (il CERN¹).

Internet

L'elemento scatenante dello sviluppo di Internet si può far risalire ai primi successi spaziali della Russia, quando in piena guerra fredda era viva la lotta tra gli USA e la Russia per la conquista dello spazio. L'orgoglio americano viene ferito profondamente quando il 4 ottobre 1957 i russi lanciano lo Sputnik, primo satellite artificiale costruito dall'uomo. Per rispondere a questo successo dei russi il Dipartimento della Difesa americano avvia, il 7 febbraio 1958, il progetto ARPA (Advanced Research Project Agency), con lo specifico scopo² di: *"provide within the Department of Defense an agency for the direction and performance of certain advanced research and development projects"*. Si arriva così al primo collegamento in rete (UCLA-Stanford) nel 1969, poi compaiono i primi collegamenti in LAN (la rete ethernet a Xerox PARC), e infine la prima dimostrazione pubblica di ARPANET (1972). ARPANET è la prima rete operativa basata sulla tecnologia della commutazione di pacchetto (packet switching³) le cui basi risalgono all'inizio degli anni 1960 e che è ora la tecnologia adottata per tutte le reti di trasmissione dati e voce, avendo soppiantato la tecnologia precedente (commutazione di circuito) in cui veniva stabilita una connessione dedicata per la durata della trasmissione. La rete a commutazione di pacchetto, invece, ottimizza l'utilizzo delle risorse di rete, anche se non garantisce la trasmissione del messaggio, che viene scorporato in una serie di pacchetti che vengono poi istradati sulla rete scegliendo per ognuno di essi il percorso migliore, e alla fine vengono ricombinati, in modo da ricomporre il messaggio originale.

Dopo i primi esperimenti⁴ con ARPANET, DARPA inizia a lavorare su altre tecnologie di trasmissione dati, e questi studi prendono impulso dall'assunzione in DARPA di Robert E. Kahn, che aveva individuato l'importanza di comunicare tra reti terrestri e reti via satellite. Nel 1973 Vinton Gray Cerf, che già aveva partecipato allo sviluppo di ARPANET, si unisce a Kahn per lavorare su *modelli aperti per l'interconnessione*, con l'obiettivo di progettare la generazione successiva dei protocolli per ARPANET. Nel 1973 Kahn e Cerf, riconoscendo l'importanza dei contributi di altri ricercatori, presentano la loro proposta, in cui le differenze tra i protocolli di rete vengono mascherate usando un *internetwork protocol* comune, e la responsabilità dell'affidabilità della trasmissione dei messaggi viene demandata agli host, e non più alla rete. Riducendo così al minimo il ruolo della rete⁵, diventa possibile interconnettere facilmente reti eterogenee.

Successivamente il gruppo di ricerca di Cerf continua l'attività, giungendo nel 1973-74 alla definizione della prima specifica del protocollo TCP (Transmission Control Protocol). Nel 1975 viene data una prima dimostrazione di interconnessione tra due reti, e le sperimentazioni si susseguono negli anni successivi, fino al completamento della transizione di ARPANET al protocollo TCP/IP (1° gennaio 1983)⁶.

Nel frattempo, soprattutto in Europa si lavora alla definizione del modello ISO/OSI (Open Systems Interconnection), standard stabilito nel 1978 dall'ISO (International Organization for Standardization), che ha avuto una funzione sostanzialmente culturale, dato che non sono mai state realizzate implementazioni complete, ma solo proprietarie o in ambito accademico.

SGML

L'invenzione della stampa ad opera di Gutenberg (circa 1439) rivoluziona il metodo di produzione di documenti scritti, permettendo di passare dai manoscritti alla produzione a stampa, con una

¹ <http://public.web.cern.ch/public/Welcome.html>

² http://www.arpa.mil/Docs/Intro_-_Van_Atta_200807180920581.pdf

³ http://en.wikipedia.org/wiki/Packet_switching

⁴ Risale a questo periodo, ed esattamente al 1971, ad opera di Ray Tomlinson, la nascita del primo programma di email e successivamente, nel 1972, l'estensione agli utenti di ARPANET, grazie all'utilizzo del carattere '@' (<http://openmap.bbn.com/~tomlinso/ray/home.html>)

⁵ Si veda l'umoristica proposta IETF RFC 1149 del 1 Aprile 1990, per l'implementazione utilizzando i piccioni viaggiatori (<http://tools.ietf.org/html/rfc1149>).

⁶ Il primo nodo italiano di ARPANET fu installato nel 1986 al CNUCE (Pisa).

diffusione molto più ampia delle opere. Questo determina la nascita di nuove professioni e del *markup*, cioè del linguaggio di marcatura dei testi usato da personale specializzato per annotare sui manoscritti le informazioni (scelta del carattere, marginatura, spaziatura, etc.) da passare ai tipografi per la composizione manuale dei caratteri.

Una svolta tecnologica importante è la comparsa delle macchine da scrivere con caratteri intercambiabili, prima fra tutte (non storicamente, ma per diffusione e impatto) nel 1961 la IBM Selectric, con la “pallina” dei caratteri⁷. Tutti possono ora in preparare documenti di buona qualità e, con l'avvento dei sistemi in time sharing, documentazione in formato elettronico, utilizzando sistemi di markup proprietari, il cui limite principale è costituito dal fatto che il markup è mescolato al contenuto del documento. Nel 1967 William Tunnicliffe, che come compagnia privata provvedeva alla pubblicazione delle pagine gialle, propone di separare il contenuto informativo dei documenti dal loro formato di presentazione. Prende così avvio il progetto GenCode⁸, con l'obiettivo di adottare un markup generalizzato basato sulla struttura gerarchica del documento, tentando di standardizzare un insieme di elementi comuni. Nel 1960 Charles G. Golfarb, Edward M. Mosher e Raymond L. Lorie avevano sviluppato il GML (acronimo dei loro cognomi, ma anche di **Generalized Markup Language**). Il GML era un insieme di macro che consentiva di implementare i tag del formattatore di testi IBM (lo Script). Caratteristiche importanti erano, tra le altre, la possibilità di definire un markup personalizzato e di validare il documento rispetto a una Document Type Definition (DTD). Nel 1978 l'ANSI inizia il lavoro per rilasciare uno standard per la definizione di testi, basata su GML. Il Draft Information Standard (DIS) viene rilasciato nel 1985, e Anders Berglund al CERN implementa il DIS e usa SGML per scrivere il testo dell'Information Standard (ISO 8879), rilasciato definitivamente nel 1986.

Hypertext

Nel 1945 Vannevar Bush pubblica sulla rivista Atlantic Monthly un articolo⁹ intitolato “As we may think”, nel quale ipotizza una macchina, denominata Memex (per Memory Extension):

A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.

[...] It affords an immediate step, however, to *associative indexing*, the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another. This is the essential feature of the memex. The *process of tying two items together is the important thing*.

Il Memex, concepito con le tecnologie dell'epoca, non fu mai realizzato, e l'idea fu dimenticata. Douglas Englebart nel 1960 inventa il mouse, e tra il 1960 e il 1970 realizza NLS (o **NLine System**) che costituisce la prima implementazione dell'ipertesto, definito (1965) da Theodor Holm (Ted) Nelson come:

A body of written or pictorial material interconnected in a complex way that it could not be conveniently represented on paper. It may contain summaries or maps of its contents and their interrelations; it may contain annotations, additions and footnotes from scholars who have examined it.

Ted Nelson inizia quindi il progetto Xanadu (1965-1999) caratterizzato dalla presenza di link bidirezionali nell'universo dei documenti, prevedendo anche un meccanismo di “pay per link”. Il suo libro “*Literary Machines*”, pubblicato nel 1981, contiene una infinità di idee brillanti. A metà degli anni 1980 l'ipertesto attira l'attenzione di una vasta comunità di ricercatori.

Tim Berners-Lee e la nascita del World Wide Web

Negli anni 1980 la tecnologia delle reti si diffonde ampiamente nel CERN, utilizzando protocolli proprietari (CERNET), sperimentando collegamenti IP via satellite, utilizzando Ethernet (1983) e poi con una implementazione pilota di TCP/IP (1984). Infine, nel 1989 il CERN viene connesso ad Internet e diventa, nel 1990, il più grosso sito internet in Europa.

Così maturano alcune delle condizioni tecnologiche che portano all'invenzione del Web. I gruppi di ricerca del CERN sono composti spesso da centinaia di persone, appartenenti a decine di

⁷ http://en.wikipedia.org/wiki/IBM_Selectric_typewriter

⁸ http://en.wikipedia.org/wiki/Markup_language#GenCode

⁹ A: http://www.multimedialab.be/doc/citations/bush_aswemaythink.pdf una riproduzione dell'articolo.

istituzioni di varie nazionalità. Per la riuscita degli esperimenti è di importanza vitale che tutti i partecipanti siano aggiornati sugli eventi, e la natura intrinsecamente distribuita dei gruppi richiede che centinaia di computer siano connessi in rete e collaborino per la raccolta e l'elaborazione dei dati. Queste esigenze, già ampiamente presenti negli anni 1980, costituiscono la premessa per l'invenzione del World Wide Web.

Tim Berners-Lee fin da piccolo aveva coltivato l'idea di rendere più intuitivo il computer, mettendolo in grado di realizzare le associazioni tipiche della mente umana. Dopo la laurea in fisica ha la possibilità di un lavoro a tempo determinato (programmatore a contratto) presso il CERN. Ed è così che già nel 1980, in occasione del suo primo soggiorno al CERN, comincia ad immaginare il web (*"Suppose all the information stored on computers everywhere were linked. Suppose I could program my computer to create a space in which anything could be linked to anything."*) e realizza ENQUIRE¹⁰, successivamente andato perduto, che prevede un insieme di nodi e di link, in cui ogni nodo ha un nome, un tipo e una lista di link tipati bidirezionali. Nel 1989 propone la condivisione di documenti e definisce una DTD SGML, che segna la nascita di HTML. È del marzo 1989 *"Information Management: A Proposal"*¹¹, che delinea l'architettura di quello che diventerà il web. Questa proposta, che il suo capo Mike Sendall giudica *"Vague, but exciting"*¹² rimane inizialmente senza risposta, e viene ripresentata, di nuovo senza successo, nel maggio 1990. Nel frattempo compare un nuovo personal computer, chiamato NeXT¹³, che sembra avere caratteristiche molto interessanti per sviluppare il progetto. Tim Berners-Lee chiede a Mike Sendall di acquistargli una di queste macchine, e questi acconsente, suggerendogli anche di utilizzarla per sviluppare il suo progetto di ipertesto.

Nella proposta del 1989 emergono già i tre elementi fondamentali del web: Universal Resource Identifier, protocollo per il fetch dei documenti e document mark-up (URI, HTTP e HTML), ma non esiste ancora un nome per il sistema, chiamato semplicemente "Mesh". Una volta iniziata la realizzazione si pone anche il problema di trovare un nome per il progetto. Considera, e scarta, varie alternative:

- "Mesh" o "Information Mesh" perché troppo simili a "mess";
- "Mine Of Information" perché l'acronimo MOI in francese significa "io"
- "The Information Mine" perché l'acronimo sarebbe TIM.

Alla fine decide di chiamarlo World Wide Web, per indicare un insieme di nodi e archi in cui ogni nodo può essere collegato con qualunque altro¹⁴.

Negli anni successivi (1991-93) continua nel suo lavoro di progettazione e sviluppo del Web al CERN, raffinando le specifiche iniziali di URI, HTTP e HTML. Il primo browser realizzato è un *browser/editor*, che consente la visualizzazione e l'aggiornamento dei contenuti, rispettando così il principio di fondo che aveva portato all'invenzione del Web:

The Web is more a *social* creation than a technical one. I designed it for a social effect - to *help people work together* - and not as a technical toy. The ultimate goal of the Web is to *support and improve our weblike existence* in the world."

(Tim Berners-Lee - Weaving the Web, p. 123)

In questa fase Robert Cailliau è di aiuto prezioso, per il contributo di entusiasmo e per la diffusione dell'iniziativa. Come nota curiosa, vale la pena di ricordare che nel 1992 Tim Berners-Lee incontra Ted Nelson e ne approfitta per saldare il conto, pagandogli la copia di *Literary Machines* ricevuta qualche anno prima.

Nell'ottobre 1994 Tim Berners-Lee fonda il World Wide Web Consortium presso il *Massachusetts Institute of Technology*, Laboratory for Computer Science, in collaborazione con il CERN, e con il supporto da parte di DARPA e della *Commissione Europea*.

¹⁰ ENQUIRE è il titolo di un libro di epoca vittoriana, contenente informazioni su una grande varietà di argomenti, dalle norme di galateo alle modalità da seguire per l'impiccagione.

¹¹ <http://www.w3.org/History/1989/proposal.html>

¹² <http://info.cern.ch/Proposal.html>

¹³ Prodotto dalla NeXT Inc. fondata da Steve Jobs.

¹⁴ Tim Berners-Lee dichiara di aver preso il termine web dalla matematica.

Il World Wide Web Consortium (W3C)

Obiettivi e organizzazione

Il W3C è un consorzio internazionale di organizzazioni, guidato da Tim Berners-Lee, con lo scopo di portare il Web al massimo del suo potenziale. Il W3C persegue questo obiettivo definendo protocolli comuni che favoriscano l' *evoluzione* e assicurino l' *interoperabilità* del Web.

Nell' aprile 1995 l' INRIA (*Institut National de Recherche en Informatique et Automatique*) diventa l'host europeo del W3C, seguito poi nel 1996 da *Keio University of Japan (Shonan Fujisawa Campus)* in Asia. Nel 2003 l' ERCIM (*European Research Consortium in Informatics and Mathematics*) subentra all' INRIA nel ruolo di host europeo. Infine, nel 2013 L'Università di Beihang in Cina diventa il quarto host.

Il W3C, finanziato dai membri, è *neutrale* rispetto ai venditori, opera con tutta la comunità per produrre specifiche e software di riferimento resi disponibili gratuitamente in tutto il mondo, ed è impegnato nel guidare l' *evoluzione* tecnologica del Web, promuovendo l' *interoperabilità* e favorendo la discussione tecnica. In pochi anni ha definito un gran numero¹⁵ di specifiche tecniche (*Recommendation*) che regolamentano l' infrastruttura tecnologica del Web, al fine di renderla robusta, scalabile e adattiva. Gli obiettivi a lungo termine possono essere sommariamente espressi come:

- *Web for Everyone*: accesso per tutti, indipendentemente da hardware, software, infrastruttura di rete, lingua madre, cultura, posizione geografica, capacità fisiche o mentali;
- *Web on Everything*: accesso con qualunque dispositivo;
- *Knowledge Base*: possibilità sofisticate di ricerca e condivisione dell'informazione, da parte di esseri umani e *anche* macchine;
- *Trust and Confidence*: tecnologie per un ambiente collaborativo, per un Web con identificazione delle responsabilità, sicurezza, fiducia, riservatezza.

Le Recommendation

Le W3C Recommendation sono il risultato di un processo cooperativo, regolato dal *Process Document*, che prevede una serie di passi e di documenti, alcuni riservati ai partecipanti ai gruppi di lavoro, altri disponibili per i membri, che votano per approvarli o modificarli, e altri ancora pubblici.

L' *intera* procedura prevista dalla *Recommendation Track* è stata progettata per ottenere il massimo consenso sul contenuto, e assicurare il pieno appoggio da parte del W3C e dell' *intera* comunità web. Le specifiche passano attraverso vari stadi, denominati *maturity level*, i più rilevanti dei quali sono:

- *Working Draft*
- *Candidate Recommendation*
- *Proposed Recommendation*
- *W3C Recommendation*

Il passaggio da uno stato all' altro avviene mediante votazione da parte dei membri. Il passaggio dallo stato di "*Last Call Working Draft*" a quello di "*Candidate Recommendation*" comporta una "*Call for implementations*", e il livello di "*Proposed Recommendation*" viene raggiunto solo dopo aver maturato una soddisfacente esperienza implementativa¹⁶. Le W3C Recommendation, perciò, non sono meri documenti cartacei, ma specifiche di cui è stata dimostrata l' efficacia e che sono implementabili con uno sforzo ragionevole.

Dal punto di vista formale il W3C non è un organo di standardizzazione, per cui le W3C Recommendation non sono degli standard in senso proprio, e vengono spesso citate come "*standard de facto*". Però non sono originate da posizioni dominanti del mercato, ma sono specifiche tecniche sulle quali è stato raggiunto un accordo da parte di tutta la comunità del Web.

L'evoluzione del Web

Nel corso degli anni il Web ha subito una significativa evoluzione, rimanendo però fedele ai suoi principi e all'architettura di base.

¹⁵ Vedi <http://www.w3.org/TR/> per l'elenco completo

¹⁶ Almeno due implementazioni interoperabili e su due piattaforme diverse.

Da HTML a XML

HTML è un linguaggio di markup sostanzialmente descrittivo, che adotta le parentesi angolari per individuare i tag, può essere scritto utilizzando un qualsiasi editore di testi, e ha un numero fisso di elementi definiti. Si è giunti a una sua definizione rigorosa dopo vari anni, quando ormai il numero di pagine web esistenti era enorme, e dopo la sua definizione formale moltissime pagine Web sono divenute non valide. Indubbiamente, l'adozione di un markup descrittivo e l'assenza di meccanismi rigidi per il controllo della validità del markup sono stati fattori di successo del Web.

HTML costituisce un compromesso tra potenza e semplicità d'uso, per cui risulta adatto per applicazioni semplici, ma si presta con difficoltà per applicazioni dai requisiti più complessi.

La comparsa di nuove applicazioni centrate sui dati, con la conseguente necessità di gestire transazioni sul Web, scambiarsi informazioni tra database, avere viste diverse degli stessi dati, e consentire la presentazione personalizzata delle informazioni, ha portato alla definizione di un nuovo linguaggio di markup che costituisce la tecnologia chiave del Web: XML (EXtensible Markup Language), definito dal W3C nel 1998.

Dal Web al Web Semantico

Il Web è nato per consentire la condivisione di documenti. Il Semantic Web si pone l'obiettivo di consentire agli utenti di rendere disponibili i loro dati agli altri e di aggiungere dei link per renderli accessibili seguendo tali link, per cui può essere visto come un'estensione dei principi del Web dai documenti ai dati. Il Semantic Web fornisce una infrastruttura comune che consente la condivisione e il riutilizzo dei dati tra applicazioni, imprese e comunità. È uno sforzo cooperativo guidato dal W3C, al quale partecipano molti ricercatori e partner industriali. Il Semantic Web è basato su RDF (Resource Description Framework) e talvolta viene chiamato "web of data" proprio per sottolineare il fatto che sul web è disponibile un'enorme quantità di dati, ma, dato che questi sono controllati dai programmi applicativi, e ogni applicazione li gestisce in maniera diversa, non è possibile combinare le varie informazioni, se non con un intervento manuale.

Il Semantic Web prevede tra l'altro la definizione di formati comuni per poter integrare e combinare dati provenienti da diverse fonti. Questo consente a una persona, o a una macchina, di iniziare le sue ricerche su un database, e da lì accedere alle informazioni contenute in un numero indefinito di altri database, che sono connessi per il fatto di contenere informazioni sulla stessa cosa.

Il Semantic Web si basa sull'ipotesi che le macchine possano accedere ad un *insieme strutturato di informazioni* e ad un *insieme di regole di inferenza* da utilizzare per il ragionamento automatico. La sfida è perciò fornire un linguaggio per esprimere *dati* e *regole* per ragionare sui dati, che consenta l'*esportazione* sul web delle regole da qualunque sistema di rappresentazione della conoscenza, con l'obiettivo di consentire alle macchine di estrarre la conoscenza disponibile sul Web, spesso disponibile in formati eterogenei, e combinarla per poter estrarre nuova conoscenza. Questo richiede di poter rappresentare, esportare e condividere la conoscenza, in modo indipendente dallo specifico ambiente operativo. Le tecnologie del W3C (RDF, RDFS, OWL, SPARQL, etc.) consentono appunto di raggiungere questo obiettivo¹⁷. È importante sottolineare che *condividere i dati e la conoscenza non comporta tradurre tutte le informazioni in RDF*.

Le tecnologie del Semantic Web hanno ormai raggiunto un buon livello di maturità, e sono ampiamente utilizzate in molte applicazioni¹⁸.

I Linked Open Data^{19,20}, che stanno attirando l'attenzione di molti operatori e sono diventati un obbligo per le Pubbliche Amministrazioni, utilizzano in maniera sostanziale le tecnologie del Semantic Web, in particolare RDF.

Il Web 2.0: un fenomeno sociale

La fisionomia dei siti web sta cambiando: gli utenti non sono più meri fruitori passivi dell'informazione, ma diventano utenti attivi che creano e aggiungono contenuti, sviluppando l'abitudine e le competenze per il lavoro di gruppo. Cresce così il numero di reti sociali (*social networking*), che sono la caratteristica fondamentale del Web 2.0²¹, anche se non sono una

¹⁷ Oreste Signore: *Introduzione al Semantic Web*. <http://www.w3c.it/papers/wsb08.pdf>

¹⁸ <http://www.w3.org/2001/sw/sweo/public/UseCases/slides/HTML/Slides.html>

¹⁹ Oreste Signore: *Un'introduzione al Semantic Web e ai Linked Data*. <http://www.w3c.it/talks/2010/smau2010/>

²⁰ Oreste Signore: *I principi di base dei LOD e del Web semantico*. <http://www.orestesignore.eu/w3c/talks/2013/lod/slides.html>

²¹ <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

novità assoluta, ma un'evoluzione di alcune forme di interazione sociale che il web ha sempre supportato (computer conferencing, email, liste di distribuzione, etc.).

Alcuni studiosi segnalano che la diffusione dei contenuti generati dagli utenti fa venir meno le fonti tradizionali del sapere, e il fatto che questi contenuti siano creati dagli utenti utilizzando sistemi diversi (podcast, blog, wiki, sistemi di chat, e altro software per il social networking), rende difficile tener traccia di dove si trovi l'informazione, e problematico accedere ad essa, sia per utenti abituali che casuali.

Web 2.0 o Semantic Web?

I vantaggi dell'approccio Web 2.0 sono essenzialmente legati alla crescita di un senso "sociale", ma vanno evidenziati alcuni rischi intrinseci. Proprio il fatto di contribuire in maniera così consistente al contenuto di un sito, con un evidente impegno e dispendio di energie, rende l'utente "dipendente" dal sito medesimo, legandolo definitivamente al formato dati adottato, per cui eventuali cambi di ambiente saranno inevitabilmente onerosi. Analogamente, se l'utente decide di partecipare a più reti sociali, in assenza di standard condivisi potrebbe essere costretto a ripetere il lavoro varie volte.

Le tecnologie semantiche sono in una fase di diffusione anche nella realtà industriale.

Web 2.0 e Semantic Web (o Web 3.0) sono da considerare due approcci complementari, piuttosto che alternativi. Il Web 2.0 ha un livello d'ingresso più basso (è molto facile utilizzarlo), ma anche orizzonti abbastanza limitati (in particolare, l'approccio delle folksonomy ha dei limiti intrinseci). D'altro canto, il Web 3.0 richiede investimenti iniziali più rilevanti, e quindi presenta un livello d'ingresso più alto, ma ha un potenziale nettamente superiore.

Il futuro del Web può essere in pericolo?

Per molti di noi il Web è ormai una realtà acquisita, probabilmente non sarebbero in grado di immaginare una vita senza la disponibilità del Web. Ma proprio per questo motivo, perché il Web viene ritenuto un'infrastruttura sempre disponibile, come l'elettricità o l'acqua, si rischia di perdere di vista la necessità di garantirne l'evoluzione nel rispetto dei principi fondanti.

In un articolo abbastanza recente²² Tim Berners-Lee sottolinea i pericoli incombenti a chiama tutti vigilare ed eventualmente a far sentire la propria voce.

Si rimanda all'articolo citato per un approfondimento, mentre qui ci si limita a ricordare che i temi chiave sono l'universalità, gli open standard, la separazione tra Web e Internet, i diritti umani elettronici, il divieto di intercettazioni e l'aggancio con il futuro.

Sono interessanti, oltre alle considerazioni sulla necessità che il Web sia accessibile a tutti, indipendentemente da limitazioni tecniche o disabilità, quelle sull'importanza degli URI come base per l'*universalità*, evidenziando come di fatto non sia possibile fare riferimento a informazioni non associate ad un URI, ma accessibili solo all'interno di specifici contesti applicativi, con il rischio che si vengano a creare regimi di monopolio. Gli URI sono un componente fondamentale del Web, perché permettono di collegare tra di loro due informazioni qualunque. Si ricordi anche che la gestione degli URI costituisce la base di un altro importantissimo elemento del Web: la *decentralizzazione*. Nell'architettura del Web, chiunque può pubblicare qualunque cosa, senza chiedere autorizzazioni a nessuna autorità centrale, semplicemente creando una pagina HTML, assegnandole un URI e rendendola accessibile sul Web con il protocollo HTTP.

Gli *open standard* sono poi la base per lo sviluppo del web: devono essere disponibili gratuitamente e condivisi dalla platea dei tecnici del web. Sarà poi libera determinazione delle singole imprese utilizzarli per realizzare eventualmente applicazioni dalle quali trarre profitti. Gli open standard sono anche un eccezionale motore di innovazione e produttività. Non mancano esempi di imprese che preferiscono sviluppare senza adeguarsi agli open standard, ma anche esempi di come tali scelte abbiano portato poi le imprese addirittura al fallimento, magari anche dopo aver conquistato fette rilevanti di mercato.

Dal punto di vista tecnico, è importante che Internet e il Web continuino a svilupparsi in maniera *indipendente*. Il Web è un'applicazione sviluppata su Internet, e come tale può trarre vantaggio dai suoi sviluppi, ma può a sua volta evolvere e ampliare il suo campo d'azione semplicemente facendo affidamento sui protocolli di Internet (TCP, IP)

²² [Long Live the Web: A Call for Continued Open Standards and Neutrality](#)", Scientific American Magazine, December 2010

L'aspetto dei *diritti umani elettronici* porta a considerazioni relative alla net neutrality, e quindi alla posizione di alcuni Internet Service Provider (ISP) che vorrebbero concedere ai loro utenti la larghezza di banda garantita per contratto solo se utilizzano servizi per i quali l'ISP ha stipulato accordi con i relativi fornitori. Il mancato rispetto della net neutrality può costituire un grosso ostacolo allo sviluppo delle piccole e medie imprese.

Intercettare e analizzare i comportamenti degli utenti, tenendo traccia degli URI acceduti, costituisce una vera e propria *intercettazione non autorizzata*, e come tale una violazione di legge. Questo tipo di "spionaggio" purtroppo non è appannaggio solo dei regimi totalitari, e, se scoperto, richiederebbe una chiara e ferma presa di posizione da parte dei cittadini.

Infine, è opportuno restare agganciati agli sviluppi *futuri*, che non possono e non sono nelle mani di una singola persona o organizzazione. L'importante è mantenere saldi i principi fondanti, perché il Web possa darci ancora ulteriori vantaggi. Solo a titolo di esempio, la connettività wireless può essere fondamentale in alcuni paesi ancora in via di sviluppo, in cui non sono disponibili connessioni via cavo. In alcuni settori, come la medicina o il governo dei cittadini, la possibilità di collegare tra di loro i dati esistenti presso istituzioni diverse apre prospettive molto interessanti, per la cura di alcune malattie o un maggior coinvolgimento dei cittadini nelle scelte politiche. Per questo motivo l'attenzione di ricercatori e sviluppatori si sta concentrando sui Linked Open Data.

Conclusioni

Il Web nasce in un ambiente di ricerca di altissimo livello, che poneva sfide impegnative, e già nella sua proposta originaria contiene idee non ancora completamente realizzate. Alla sua nascita hanno contribuito molti ricercatori, che operavano in settori diversi, ma la sua concretizzazione si deve all'intuizione di Tim Berners-Lee, alla sua capacità di individuare gli elementi essenziali dell'architettura, e alla sua costanza nel perseguirne l'implementazione e nel mantenerlo un ambiente libero da protocolli proprietari²³.

Il World Wide Web Consortium opera per far raggiungere al Web il suo massimo potenziale, e definisce gli standard web con un processo aperto e collaborativo.

Tuttavia, il futuro del Web potrebbe essere in pericolo se l'intera comunità internazionale non vigilerà per mantenere saldi i principi che hanno portato a questa meravigliosa invenzione, che non è solo scientifica e tecnologica, ma ha importanti risvolti e motivazioni ideali.



Nato nel 1946, laureato in Fisica nel 1970, Oreste Signore è stato dipendente CNR dal 1973 al 30 giugno 2012, concludendo la sua carriera con la qualifica di Dirigente Tecnologo presso l'Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" (ISTI), nato dalla fusione tra [CNUCE](#) e [I.E.I.](#) Attualmente è associato alla ricerca presso l'ISTI. È il responsabile dell' [Ufficio Italiano W3C](#) presso il C.N.R. fin dalla sua costituzione nell'aprile 1999.

I suoi interessi scientifici sono nell'area dei Sistemi Informativi, nel cui ambito si è occupato di basi di dati, progettazione concettuale, ipermedia, interfacce utente, applicazioni web, information retrieval, ingegneria del software (in particolare di reverse engineering). I suoi interessi scientifici attuali sono orientati alle tecnologie Web, e in particolare al Semantic Web.

È autore di più di 300 tra libri, lavori e presentazioni in riviste e atti di convegni nazionali e internazionali. Ha ricoperto un ruolo importante nella definizione degli standard per il catalogo nazionale dei beni culturali e delle regole tecniche per l'accessibilità dei siti Web

²³ "If someone tries to monopolize the Web - by, for example, pushing a proprietary variation of network protocols - they're in for fight." (da Weaving the Web, p. 107-108)