



Franco Sivilli

*Psicologia della salute, vecchie e nuove
dipendenze*

Università degli Studi «G. d'Annunzio» Ch-Pe

Evoluzione del web, cloud e big data



“Let the web serve humanity”

„Let the web serve humanity“

Sir Tim Berners Lee, inventore
del World Wide Web



FOCUS

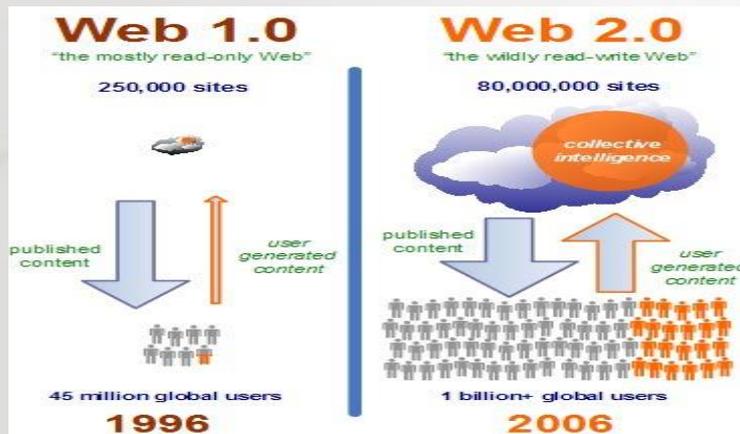
- EVOLUZIONE TECNOSOCIALE DELLA RETE: DAL WEB 1.0 AL WEB 2.0 VERSO IL WEB SEMANTICO;
- LA MASSIMIZZAZIONE DELL'UTILIZZO DELLA RETE: IL CLOUD COMPUTING E IL FENOMENO DEI BIG DATA;*
- IL FUTURO DELL' INFORMATION TECHNOLOGY: TIMORI O SPERANZE?

**EVOLUZIONE TECNOSOCIALE DELLA RETE:
DAL WEB 1.0 AL WEB 2.0 VERSO IL WEB
SEMANTICO;**



Evoluzione della Rete

1. Anni '80 → La rete come mezzo per collegare computer
2. Anni '90 → La rete come strumento per collegare documenti (**Web 1.0**)
3. 3° Millennio:
 - la Rete come strumento di connessione delle relazioni sociali (**Web 2.0**)
 - Cloud computing → ICT: da strumento di gestione a strumento di governo
4. Scenari futuri : **Web 3.0**



... e
all'orizzonte
si intravede il
Web 3.0....

Web 1.0

E' la prima fase del Web, caratterizzata dalla progressiva definizione degli standard tecnologici che hanno permesso di dare vita a questa nuova realtà e di diffonderne l'uso tra milioni di utenti.

Affermatosi e diffusosi negli anni '90, il Web 1.0 è composto prevalentemente da siti web statici, senza alcuna possibilità di interazione con l'utente eccetto la normale **navigazione** tra le pagine, l'uso delle e-mail e l'uso dei motori di ricerca.

Web 2.0

E' la fase attuale, caratterizzata da una partecipazione attiva degli utenti alla costruzione dei contenuti, alla loro classificazione e distribuzione.

Il **Web 2.0** indica genericamente uno stato di evoluzione di Internet (e in particolare del World Wide Web), rispetto alla condizione precedente. Si tende ad indicare come Web 2.0 l'insieme di tutte quelle applicazioni online che permettono uno spiccato livello di interazione sito-utente (blog, forum, chat, sistemi quali Youtube, Facebook, Twitter ecc.).

Confronti tra Web 1.0 e Web 2.0

- Da un punto di vista strettamente tecnologico, il *Web 2.0* è del tutto equivalente al *Web 1.0*, in quanto l'infrastruttura di rete continua ad essere costituita da **TCP/IP** e l'ipertesto è ancora il concetto base delle relazioni tra i contenuti.
- La differenza sta nell'**approccio con il quale gli utenti si rivolgono al Web**, che passa fundamentalmente dalla semplice consultazione (seppure supportata da efficienti strumenti di ricerca, selezione e aggregazione) alla possibilità di contribuire popolando e alimentando il Web con propri contenuti. È l'aver portato l'utente al centro della Rete la vera innovazione sancita dall'avvento del Web 2.0.

Confronti tra Web 1.0 e Web 2.0

Se è cambiato l'approccio alla Rete, non poteva che cambiare anche il modo di rappresentarsi in essa.

- Le tecniche utilizzate fino a ieri per tenere più tempo i visitatori su un sito web (*stickiness*, cioè la capacità di tenere "incollati" gli utenti ad esso) stanno lasciando il posto ad altre concezioni di contatto con il fruitore. Attraverso le tecnologie di *syndication* (RSS, Atom, tagging) chi realizza contenuti fa in modo che questi possano essere fruiti non solo sul sito, ma anche attraverso canali diversi.
- Se prima la costruzione di un sito web personale richiedeva la padronanza di elementi di HTML e programmazione, oggi con i *blog* chiunque è in grado di pubblicare e gestire i propri contenuti e questo a prescindere dal livello di conoscenza informatica del singolo utente.
- La tecnologia *Wiki* è il punto di arrivo per la gestione dei contenuti perché ne realizza tutti i paradigmi: consente ad ogni utente di aggiungere contenuti, ma anche di cambiare quelli esistenti inseriti da altri utilizzatori; con la stessa semplicità si possono aggiungere ed eliminare pagine.

<i>Web 1.0</i>	<i>Web 2.0</i>
Siti personali	Blogging
Sistemi di gestione contenuti	Wikipedia
Stickiness	Syndication

Il G.G.G. (Grafo Gigante Globale): il Web 2.0 secondo Tim Berners-Lee

Perché il GGG:

Il Grafo nasce per mettere in relazione non solo persone ma queste e i contenuti (non necessariamente i documenti in cui sono inseriti) con altri contenuti.

Obiettivo:

Riconoscere, attraverso opportune “etichette” generate dai linguaggi di marcatura, gli utenti al di là delle loro manifestazioni (i diversi profili creati on-line) allo stesso modo dei contenuti al di là delle differenti modalità di presentazione.

Il Social Networking

E' il fenomeno che sta guidando la crescita del Web 2.0.

Si tratta di piattaforme di aggregazione sulle quali le persone possono entrare in contatto, condividere contenuti, stabilire nuovi legami o riproporre quelli della vita reale esaltando la caratteristica principale del Web 2.0: la partecipazione.

Motivazioni che spingono all'uso dei SN

- Consce (es. ridurre distanze geografiche);
- Inconscie (identificazione del sé, necessità di far parte di una comunità virtuale)

Il motivo principale per cui l'utilizzo dei SN è diventato un fenomeno globale risiede nella sua peculiarità di essere diventato parte integrante della vita quotidiana di ogni utente.

Consigli utili nell'uso dei SN

- 1) Non pubblicare l'effettiva ubicazione quando si è in giro. Pubblicare ciò sui SN è una sorta di invito formale per i criminali.
- 2) Non pubblicare aspetti negativi della propria vita. Farlo è come screditare la propria immagine. I vostri Amici monitorano la vostra attività e uno sbaglio può causare danni in futuro.
- 3) Prendere distanza dal profilo dell'ex. Potrebbe sembrare crudele, ma una volta che si è deciso di uscire da una relazione non vi sono più motivazioni per visitare il profilo del proprio *ex-partner*. Se si vuole avere una vita tranquilla in futuro è consigliato cancellare il proprio ex dagli Amici.
- 4) Non sostituire gli amici reali con quelli virtuali. I SN sono un ottimo strumento per far interagire persone sparse per il mondo. Ma loro non saranno mai come i propri amici reali. Bisogna avere amici reali per accrescere se stessi e ridurre lo stress e l'ansia della vita reale.
- 5) Evitare di utilizzare i SN negli orari lavorativi. Utilizzare i SN durante il lavoro non influenza solo la propria *performance* ma aumenta anche la possibilità di essere licenziati.

Rivoluzione comunicativa

..Studio svolto a Stanford dalla Prof.ssa Lunsdorf nel 2001-2006

- La multimedialità ha provocato la nascita di nuovi gerghi all'interno del canale verbale della comunicazione.
- La tecnologia non sta uccidendo la nostra capacità di scrivere ma sta spingendo la nostra alfabetizzazione in nuove direzioni.



I risultati dello studio

- A causa dell'aumento della socializzazione che avviene *on-line* i giovani di oggi scrivono molto più di qualsiasi generazione prima di loro. Prima dell'avvento di Internet, la maggior parte degli studenti americani non aveva l'abitudine di scrivere qualcosa a meno che non fossero compiti scolastici.
- Nell'esaminare il lavoro degli studenti del primo anno non sono stati trovati esempi di prosa contaminata dal linguaggio abbreviato tipico degli SMS ed inoltre il team ha scoperto che gli studenti erano straordinariamente abili, sapevano valutare il loro pubblico ed erano bravissimi ad adattare il loro tono e la loro tecnica per ottenere i migliori risultati.
- Il fatto che gli studenti di oggi scrivano quasi sempre per un pubblico (cosa che praticamente nessuno ha mai fatto nella generazione precedente) dà loro un diverso senso di ciò che costituisce una buona prosa.
- Per loro la scrittura è l'abilità a persuadere e del saper discutere e infatti nelle interviste hanno definito buona quella prosa che ha un effetto sul mondo.

LA MASSIMIZZAZIONE DELL'UTILIZZO DELLA RETE: IL CLOUD COMPUTING E IL FENOMENO DEI BIG DATA



Cloud Computing: come nasce

Inizialmente questo nuovo paradigma computazionale è stato fatto coincidere con il concetto di *Software as a Service* ovvero l'utilizzo di software remoto, spesso attraverso un server web. In seguito, ai programmi remoti si sono aggiunti anche servizi e librerie che hanno costituito una vera e propria piattaforma software accessibile remotamente (*Platform as a Service*).

Il punto di arrivo è considerare l'intera infrastruttura come un servizio (*Infrastructure as a Service*) estendendo il concetto stesso di grid computing (calcolo distribuito) ma consentendo l'utilizzo delle risorse esclusivamente al momento in cui un utente/cliente ne ha bisogno.

Il paradigma del Cloud computing inteso come *Infrastructure as a Service* è reso possibile grazie alle standardizzazioni delle tecnologie di virtualizzazione.

Cloud Computing

Il “*cloud computing*” è un insieme di modelli di servizio che più di altri si sta diffondendo con grande rapidità tra imprese, pubbliche amministrazioni e cittadini perché incoraggia un utilizzo flessibile delle proprie risorse (infrastrutture e applicazioni) o di quelle messe a disposizione da un fornitore di servizi specializzato. Sfruttando la tecnologia del cloud computing gli utenti collegati ad un cloud provider possono svolgere tutte queste mansioni, anche tramite un semplice Internet browser.

Possono, ad esempio, utilizzare software remoti non direttamente installati sul proprio computer e salvare dati su memorie di massa on-line predisposte dal provider stesso (sfruttando sia Reti via cavo che senza fili).

Cloud Computing

Con il *cloud* diventerà sempre più sfumato il confine tra computer e Rete e sia l'infrastruttura che le applicazioni diventeranno un servizio

(*Infrastructure as a Service+Software as a Service = Cloud*).

Il computer del futuro, quindi, dovrà avere solo la potenza necessaria ad eseguire un browser come Internet Explorer (o *Firefox, Chrome, Safari*, e via discorrendo).

Quando Eric Schimdt ha detto “*il browser è il computer*”, si riferiva esattamente a questo.

Cloud Computing: criticità

Le criticità e i rischi legati al cloud sono sintetizzabili in tre macroaree:

- 1. Sicurezza;**
- 2. Privacy;**
- 3. Legale e digital forensics.**



Cloud Computing e criticità

SICUREZZA

- Difficoltà a individuare la location geografica dei miei dati
- Chi tiene i miei dati e soprattutto quali vincoli o standard di sicurezza rispetta ovvero quali vulnerabilità del detentore dei miei dati potrebbero compromettere la loro integrità o quali perdite potrei subire se ci fossero dei disastri?
- Riusciamo davvero a disporre dei nostri dati e delle nostre applicazioni come se fossero in House?
- Quali rischi corriamo condividendo le nostre risorse con altri attori?

Cloud computing e criticità

PRIVACY

Le indicazioni proposte dal Garante per la **Protezione dei Dati Personali** relative all'utilizzo consapevole dei servizi *cloud* (<http://www.garanteprivacy.it/documents/10160/10704/1819933>)

- PONDERARE PRIORITARIAMENTE RISCHI E BENEFICI DEI SERVIZI OFFERTI;
- EFFETTUARE UNA VERIFICA IN ORDINE ALL'AFFIDABILITÀ DEL FORNITORE;
- PRIVILEGIARE I SERVIZI CHE FAVORISCONO LA PORTABILITÀ DEI DATI;
- ASSICURARSI LA DISPONIBILITÀ DEI DATI IN CASO DI NECESSITÀ;
- SELEZIONARE I DATI DA INSERIRE NELLA CLOUD;
- NON PERDERE DI VISTA I DATI;
- INFORMARSI SU DOVE RISIEDERANNO, CONCRETAMENTE, I DATI;
- ATTENZIONE ALLE CLAUSOLE CONTRATTUALI;
- VERIFICARE LE POLITICHE DI PERSISTENZA DEI DATI LEGATE ALLA LORO CONSERVAZIONE;
- ESIGERE E ADOTTARE OPPORTUNE CAUTELE PER TUTELARE LA CONFIDENZIALITÀ DEI DATI;
- FORMARE ADEGUATAMENTE IL PERSONALE

Cloud Computing e criticità

LEGALE E DIGITAL FORENSICS

- Cosa potrebbe verificarsi in caso di interruzione del contratto con il fornitore?
- E' possibile riportare il servizio presso la nostra sede fisica o dislocarle presso un altro fornitore?
- È possibile accertare la prova di un reato che prevede la detenzione di materiale informativo sul cloud ed acquisirne il contenuto utile ai fini forensi?
- Qual è il foro competente alla risoluzione di eventuali controversie legali?

IL FUTURO DELL' INFORMATION TECHNOLOGY: TIMORI O SPERANZE?

01

Lo scenario futuro: il Cloud e i BIG DATA

Secondo una ricerca condotta da *Coleman Parkes Research*, il 41% delle aziende europee del settore dell'Education e **nel 34% del settore sanitario** continuano a gestire i documenti e i loro dati (aumentati in maniera esponenziale e per questo definiti BIG DATA) a livello di singolo ufficio o dipartimento e non in maniera integrata, con rischi di duplicazione e ridondanze e grosse difficoltà nell'information retrieval. Il 43% delle aziende europee coinvolte nella ricerca afferma di gestire i processi mediante attività manuali e solo il 22% ha implementato processi completamente automatizzati sfruttando anche tecnologie web.

Il white paper "*Process Efficiency Index*" mette in evidenza come i dipendenti responsabili dei processi documentali critici impieghino ogni anno per la loro gestione circa 362 milioni di ore del loro tempo.

La gestione corretta delle informazioni e dei processi attraverso tecnologie di cloud deve quindi diventare una priorità di business per riuscire a risolvere l'enigma dei "Big Data".

BIG DATA: applicazioni sanitarie

Muta concetto esattezza: il passaggio a una dimensione molto più grande modifica la capacità pratica di arrivare all'esattezza!!!

- Nel 2009 (epidemia H1N1) il sistema previsionale di Google (correlazione tra la frequenza di queries *come medicine per tosse e febbre* e la diffusione dell'influenza nel tempo e nello spazio) ha fornito in tempo reale info prezioso alle autorità sanitarie
- *Le previsioni basate sulla correlazione sono al centro dei big data*

BIG DATA: applicazioni sanitarie

- Un team IBM e Univ. Ontario coordinati dalla dott.ssa McGregor sta sperimentando un sw che analizza 16 parametri diversi di neonati prematuri generando 1250 data point al secondo prevedendo, attraverso la loro correlazione, una infezione 24 ore prima della comparsa di sintomi evidenti. E' stato scoperto (contro la logica medica) che la stabilizzazione dei prematuri precede spesso una infezione grave. I dati suggeriscono una correlazione e non un rapporto di causalità. Per fare emergere questa associazione si sono dovuti applicare metodi statistici a un'enorme quantità di dati
- *L'allarme precoce permette un trattamento più tempestivo → i big data potranno salvare delle vite!!!*

Pat Helland: if you have too much data then «good enough» is good enough

- I RDBMS non sono più attuali, sono progettati per fornire risposte a domande precise: i big data difficilmente si possono inquadrare in categorie definite e note dall'inizio. Secondo stime solo il 5% dei dati digitali sono strutturati.
- La disponibilità a usare il totale delle info anziché dei campioni ci porterà a preferire l'imprecisione alla esattezza in cambio dell'accesso all'universo delle info disponibili

BIG DATA: la fine della teoria e l'inizio della datizzazione?

- Chris Anderson (Wired): Il diluvio digitale rende il processo tradizionale su cui si fonda la scoperta scientifica (un'ipotesi che viene testata rispetto alla realtà utilizzando un modello di causalità sottostante) destinato all'estinzione per essere sostituito da un'analisi statistica di pure correlazioni del tutto svincolata dalla teoria → Google usava le key-words e non la lunghezza dei capelli delle persone come indicatore rappresentativo dell'influenza
- *Datizzazione: sarà il processo del futuro e va oltre la digitalizzazione, vuol dire convertire un fenomeno in forma quantitativa in modo da poterlo tabulare e analizzare. La rivoluzione dell'IT si è spinta finora sulla Technology, il futuro si spingerà sull'Information anche grazie **all'Internet of things***

BIG DATA: una rivoluzione che cambierà la nostra vita

- Per il 2013 la quantità di info immagazzinate nel mondo viene stimata in 1200 exabyte (meno del 2% in forma non digitale). Se venissero raccolti su cd impilati arriverebbero alla luna in 5 pile separate.
- Abbandono del campionamento statistico grazie alla datizzazione ($N=tutti$) → usare un campione statistico nell'era dei big data è come usare il cavallo nell'era dell'automobile

Lo scenario futuro: interoperabilità dei big data attraverso il cloud

E' sempre più frequente la pubblicazione da parte delle amministrazioni pubbliche europee di dati aperti (**OPEN DATA**).

Esempio: info sullo stato delle strade. Ma se prevedo di viaggiare attraverso l'Europa, ho bisogno di collegare tra loro tutti questi dati). Una base di dati unificata per tutti i paesi europei potrebbe offrire valore aggiunto all'insieme dell'economia europea.

Esempio: da un lato un comune pubblica i dati dei monumenti storici e degli hotel che si trovano nelle vicinanze di quei monumenti; dall'altra, Sovrintendenze ai Beni Culturali pubblica i dati dettagliati sui monumenti, gli artisti e i periodi storici, e sui quadri esposti nei musei o nei palazzi.

Combinare i due *dataset* potrebbe essere di grande utilità, ad esempio per offrire un servizio personalizzato sugli itinerari in base agli interessi culturali specifici di un turista.

Per fare questo, se i dati non sono "collegati" (linked) occorre in qualche modo creare questi link, processando i dati a mano o attraverso algoritmi ad hoc. Nei cosiddetti **LINKED DATA**, questi collegamenti e relazioni tra le entità descritte nei dataset sono **espliciti**.

Big data, open data e linked data ... le parole di Tim Berners Lee

L'interoperabilità di tutti questi dati è in effetti il problema principale e tutte le organizzazioni impegnate negli open data. Devono essere progressivamente adottati formati e norme comuni: i **linked data** nel loro complesso permetteranno di mettere a punto una nuova generazione di servizi web capaci di rispondere a quesiti complessi.

I **CIO**, hanno una relazione particolare con le informazioni in possesso della loro azienda. Si potrebbe quasi parlare di ossessione per il controllo. Difendono i loro database come un bene prezioso, ma dovranno cambiare il loro punto di vista sulla questione.

La diffusione di dati pubblici e privati rischia di avvenire a spese della vita privata e della privacy, di conseguenza questi organismi devono rendere anonimi questi dati per non minacciare le libertà individuali. Potremo risolvere questo problema con gli **open data** che semplificano la vita quotidiana dei cittadini.

Organismi indipendenti dovranno impegnarsi nella protezione dei dati con mezzi d'azione significativi per sorvegliare la protezione dei dati individuali. Ma questi dati sono come la dinamite e potranno scoppiare se messi nelle mani sbagliate.

Lo scenario futuro: evoluzione nel Web semantico

“il Web Semantico è un’estensione del Web corrente in cui le informazioni hanno un ben preciso significato e in cui computer e utenti lavorano in cooperazione”.

Ancor più interessanti appaiono le innovazioni associate con il Web 3.0 :

- **Un percorso evolutivo verso l’intelligenza artificiale**, capace di interagire con il Web in modo quasi umano;
- **Evoluzione verso il 3D**, che consentirebbe la trasformazione del Web in una serie di spazi in 3D condivisibili;
- **Business, intrattenimento e social networking** in qualsiasi momento e dovunque;
- **Reti sicure e affidabili**;
- **Un nuovo modo** più scorrevole che porrà fine alla divisione tra linee fisse e mobili.

Il Web semantico: come raggiungerlo

Trasformare il Web in un database

Il primo passo verso un "Web 3.0" è l'emergere del "Data Web" visto che gli archivi di dati strutturati sono pubblicati sul Web in formati riutilizzabili e "interrogabili" da remoto, come XML, RDF

Conclusioni

Internet is our shared space.

It helps us connect.

It spreads opportunity.

It enables us to learn.

It gives us a a voice.

It makes us stronger.

To keep the internet strong, we need to keep it secure... we work together to create this secure environment and make our shared space even better for the world... together, we can build a space that is greater and a more important part of the world than anything we have today...

Mark Zuckerberg 13.3.2014

“... è necessario cambiare la cornice legislativa esistente per assicurare che i nostri diritti fondamentali e le nostre libertà non ne escano indeboliti dal confronto con l’era digitale.”

Neelie Kroes, commissario per l’Agenda Digitale e vice presidente della Commissione Europea

fsivilli@unich.it

