

TELECOM ITALIA

FUTURE CENTRE



Ciclo di incontri

“Vita e Vite in Bit”

Quanti bit in una vita?

Memorizzare la nostra vita

I bit nella società

La Società dell'Informazione

Memorizzare la nostra vita

È possibile memorizzare la propria vita pensandola come un insieme di bit? Già circa quindici anni fa Nicolas Negroponte teorizzava, nel suo best seller “Essere Digitali”, l’avvento di un’era digitale e proponeva varie riflessioni su come le nuove tecnologie avrebbero cambiato la nostra esistenza. Questo testo visionario si presenta come una mappa per capire quali sono le vere implicazioni dell’informatica nella nostra vita quotidiana, nel lavoro, nella società, nella cultura, nella socializzazione, nell’intrattenimento e nel tempo libero. Come potremmo sfruttare i dati di tutta la nostra vita registrata in digitale? Le informazioni presenti oltre a rappresentare un diario permanente e automatico di ciò che facciamo, vediamo, sentiamo, potrebbero costituire la base per abilitare nuovi servizi e personalizzazioni dell’esperienza di ognuno di noi nella vita di tutti i giorni.

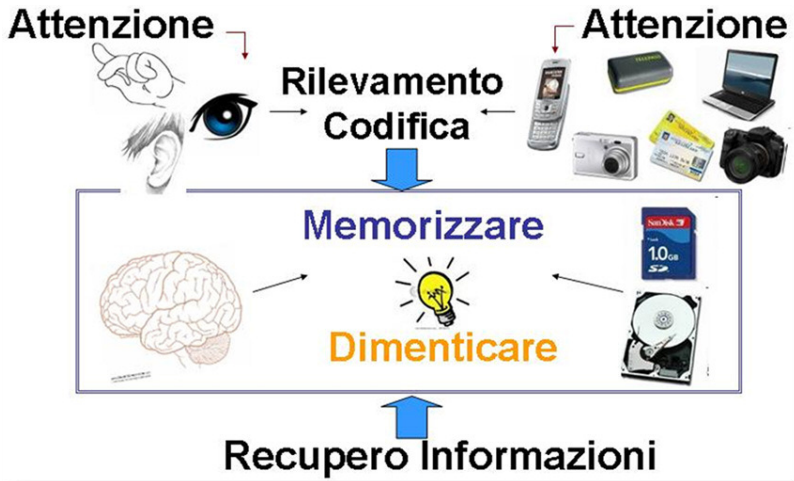
Dalla memorizzazione del nostro cervello a quella digitale

Nel primo incontro si è visto come il nostro cervello memorizza passando dalle fasi di attenzione (focus), rilevazione (cattura), contesto, memorizzazione (e oblio delle cose inutili) e recupero dell’informazione; anche il mondo digitale passa per fasi analoghe, quindi esiste un parallelismo procedurale tra i due mondi.

Attenzione è la decisione di focalizzarsi su alcuni aspetti tra i tanti segnali che arrivano. La scelta può vincolare la fase di rivelazione, infatti avviene una selezione degli aspetti di interesse tra tutto ciò che viene rilevato (mi concentro su un aspetto e trascuro tutto il resto).

Questa possibilità è fondamentale nell'ottica di memorizzare la nostra vita in bit, perché tecnicamente si potrebbe registrare tutto quanto avviene attorno a noi, ma, in realtà, a noi interessano principalmente gli aspetti su cui ci siamo concentrati e che hanno catturato la nostra attenzione (un particolare fiore, un suono, una persona). Dettagli secondari rispetto al nostro focus potrebbero però avere una grande utilità in altri ambiti in cui la ricchezza di particolari e dettagli sono essenziali (es: in tribunale, per dimostrare la nostra innocenza, sapere chi erano le persone presenti nella stanza con noi in un determinato momento, può fare la differenza per la sentenza...). Qui entrano in gioco tecnologie per catturare i movimenti degli occhi e capire la nostra visuale, sensori di movimento per sapere se siamo seduti, sdraiati, in piedi, se stiamo camminando o correndo, microfoni direzionali. Ovviamente il tutto vale più della somma delle parti, quindi unendo le informazioni derivanti da più tecnologie si ricavano maggiori informazioni rispetto a quelle di partenza: se sono in una stanza con più persone che stanno parlando e il mio sguardo è rivolto ad una in particolare, il microfono deve catturare il suono proveniente da quella persona perché si presume che io la stia ascoltando. La tecnologia può catturare tutti i suoni, ma il mio focus e interesse è su quello specifico suono.

Nella figura la freccia dell'attenzione passa prima dal dispositivo. Questo sarà vero in futuro. Ovvero, il device non rimarrà un elemento passivo che raccoglie tutto ma sarà in grado di catturare alcune cose piuttosto che altre. Alcune macchine fotografiche già oggi scattano se si sorride, il telefonino potrà rispondere in modo diverso a seconda di chi chiama e memorizzare azioni diverse a seconda della situazione senza intervento del proprietario.



Per quanto riguarda la **rilevazione** e cattura nel nostro corpo ci avvaliamo dei nostri sensi. Nel mondo digitale utilizziamo ogni dispositivo che raccoglie una piccola parte delle nostre azioni e del nostro agire: fotocamere digitali, microfoni, telefoni cellulari, telecamere di sorveglianza, GPS, sensori medicali, decoder TV digitale, ATM e POS, telepass, carte fedeltà... Ma anche navigare in Internet, chattare, inviare mail, fare acquisti sul Web e partecipare attivamente a social network e community. Il mondo digitale è così pervasivo da catturare la nostra vita senza che noi neppure ce ne accorgiamo. Tutti questi dati dove vanno a finire? Attualmente nelle banche dati di chi offre il servizio e questi dati preziosi vengono conservati per erogare il servizio o per tenere traccia delle operazioni eseguite (es: l'acquisto con carta di credito, o la registrazione della videocamera di sorveglianza in un supermercato).

Anche ogni volta che utilizziamo una carta fedeltà lasciamo una traccia del nostro agire. Se siamo al supermercato a fare acquisti e alla casa presentiamo la nostra tessera fedeltà, tutti i dati di ciò che abbia-

mo acquistato, quando e quanto abbiamo pagato diventano noti al sistema informativo del supermercato stesso per fare analisi sui nostri comportamenti di acquisto

Ironia della sorte, il vero proprietario di questi dati, la persona che li ha generati, non sempre può recuperare le sue informazioni o cancellarle o usarle come meglio crede, per lo meno in questo periodo di “vuoto” legislativo, in cui manca una vera regolamentazione sulla privacy in modo da ridare la ownership dei dati al suo legittimo proprietario. Per ora privacy è più sinonimo di chiusura dei dati, di divieto di utilizzo dei dati dell'utente da parte di terzi.



POS



ATM



PC - LAPTOP



**SENSORI
MEDICALI**



**FOTO/VIDEO
CAM DIGITALI**



**LETTORE RFID –
CODICE A BARRE**



SET TOP BOX



**CAMERE DI
SORVEGLIANZA**

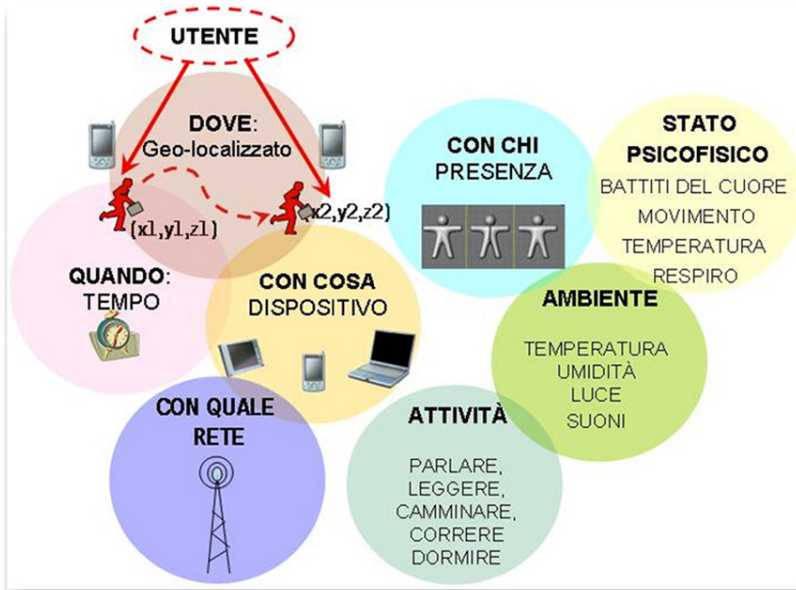


CARTE FEDELTA'

La fase di **contesto** è vista come la possibilità di dare riferimenti geografici di localizzazione e di context awareness (sapere con chi sono e quale sia la situazione) e in futuro anche sapere quali oggetti sono vicini a me. Più in particolare la contestualizzazione risponde alle

domande: dove, quando, con chi, con quale dispositivo, con quale rete, con quale stato psicofisico e anche quale attività si sta svolgendo (parlare, leggere, camminare, correre, dormire) e in quale ambiente (all'aperto, al chiuso, temperature, luce, suoni...).

La contestualizzazione delle azioni è fondamentale per arricchire le informazioni e creare metadati precisi, attendibili e di valore; per fare un esempio, se le pulsazioni del mio cuore aumentano improvvisamente forse la causa è un infarto, ma se grazie alla context awareness si capisce che sono in una stanza con una ragazza allora si può dedurre che le palpitazioni sono causate dalla sua presenza.



La **memorizzazione** avviene in apposite memorie sempre più capienti e miniaturizzate. Lo storage può essere in rete o in locale. Dipende dai dispositivi, ma tutti ormai hanno la capacità di contenere parecchie informazioni: dal media center al telefono cellulare, alla foto-

camera digitale all'mp3 player. Grazie alle memory card estraibili si può facilmente ampliare la memoria del proprio cellulare e trasferirla sul proprio PC; in ogni caso per il 2012 sono previste memory card SD da un TeraByte che eliminerebbero ogni problema di mancanza di memoria (si potrebbero davvero memorizzare tutte le telefonate, gli SMS, le mail, i video e le foto e avere un microfono sempre acceso...). Memorizzare in rete vuol dire sfruttare server remoti che offrono il servizio di storage puro, oppure servizi Web dedicati come Picasa per le foto, Youtube per i video, Box.net per i propri file, ecc. Anche gli Hard Disk ad uso domestico sono sempre più diffusi e diventano dei veri e propri "magazzini" di tutti i contenuti digitali della famiglia come nel caso dei media center.

Quando memorizziamo delle informazioni su un hard disk queste vengono fisicamente memorizzate in una certa zona del disco. Quando le cancelliamo viene rimossa dall'indice del disco l'indirizzo a cui si trovano quelle informazioni. Se in momenti successivi quella zona non viene riscritta le informazioni originali sono ancora presenti e possono essere recuperate. È come se cancellassimo l'indice di un'enciclopedia: le informazioni ci sono ancora ma senza indice è arduo trovarle. Nel nostro cervello abbiamo qualcosa di simile in termini di ricerca. Come i nostri neuroni creano collegamenti anche gli hard disk e le memorie di massa possono essere collegati con hyperlink, mashups e sincronizzazioni. Es un video che ho messo su Youtube lo posso inserire nella mia pagina di Facebook senza copiarlo, ma semplicemente inserendo un collegamento al video originale su Youtube.

Collegato al tema del ricordo troviamo quello della capacità di dimenticare. Se memorizzassimo ogni singolo dettaglio della nostra vita, come potremmo fare a recuperarlo? Dimenticare serve a dare valore a quello che più ci interessa a mettere in rilevanza gli eventi che più ci stanno a cuore. Tutto verrà registrato delle nostre azioni, anche eventi di cui noi siamo inconsapevoli, ma l'attenzione sarà su ciò che noi viviamo consapevolmente.

Il **recupero** delle informazioni è l'atto finale del processo e dati del passato si possono ricercare per timeline (linea temporale, anno, mese, giorno, ora), per parola chiave, per associazioni (luogo, persona, oggetto, evento)

Da una informazione passiamo ad un'altra con un semplice click. Ma potremo arrivare a quella informazione anche seguendo un'altra strada. Ecco allora come diventa importante nel momento in cui si voglia memorizzare la propria vita in bit stabilire varie connessioni ad una certa informazione e creare dei metadati incrociando e correlando più dati disponibili.

Le informazioni possono risiedere in data base separati e posseduti da enti diversi che però condividono i dati e i contenuti in modo che il legittimo proprietario possa avere una vista comune e associata della sua vita registrata. Nel centro ricerche Telecom Italia l'applicazione per telefoni mobili chiamata TeamLife permette di salvare contenuti (foto, video) online e aggiungere dei tag (informazioni su luogo, ora, persone presenti, commenti) in modo automatico e/o manuale e poi rivederli localizzati sulle mappe messe a disposizione da Google. Si crea il tracciamento del percorso fatto con le foto localizzate durante il tragitto



L'ambiente digitale

L'ambiente digitale (digital environment) di un individuo è sempre più ampio ed è costituito dalle tecnologie che consentono la digitalizzazione delle sue azioni e delle sue informazioni e relativa cattura e memorizzazione dei suoi dati.

Le principali tecnologie abilitanti il mondo digitale della persona sono proprio quelle che portiamo con noi, indossiamo e fanno parte della vita di tutti i giorni:

- Telefoni cellulari (comunicazione in audio, video, testo e immagini).
- Foto e video camere digitali (cattura del mondo reale tramite audio, foto, video digitale).
- Memory card (memorizzazione).
- Sensori di movimento (consapevolezza del contesto "context awareness").
- GPS (localizzazione).
- Tag RFID (consapevolezza degli oggetti che ci circondano).
- Sensori biometrici (analisi dello stato del nostro corpo, sensazioni e in futuro emozioni).

Queste tecnologie, utilizzate anche in accoppiata (es fotocamera + GPS), consentono di catturare con maggior dettaglio di informazioni le azioni e le interazioni tra persone, gli oggetti e l'ambiente.

Un esempio di tecnologie che evolvono e che creano nuovi scenari di utilizzo e raccolta dati sono i cerotti che comunicano i dati raccolti dal paziente al medico curante.

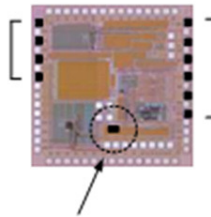
Attualmente sperimentati negli stati Uniti da Gentag, società specializzata in telemedicina, i cerotti saranno disponibili in Europa presumibilmente nel terzo trimestre del 2010. Il cerotto permette di misurare i livelli di glicemia, il battito cardiaco, la temperatura e alcune tipologie di allergie: il paziente dopo aver effettuato la rilevazione può inol-

trare le informazioni raccolte direttamente al medico di riferimento in modo tempestivo. Un dispositivo, quindi, che rientra a pieno titolo tra i sensori indossabili da una persona che catturano dati.



GENERIC RFID SENSOR PLATFORM (ILLUSTRATION ONLY)

EXTERNAL DIGITAL
SENSOR PORTS



EXTERNAL ANALOG
SENSOR PORTS

INTERNAL TEMPERATURE SENSOR

Sempre in ambito sanitario, per i possessori di iPhone, è fornito un servizio per visualizzare tutte le informazioni mediche che li riguardano.

Sono servizi di monitoraggio medico che da un lato abbattano i costi e dall'altro aumentano l'accessibilità ai servizi e ai dati sanitari che potrebbero confluire in un'unica cartella clinica personale con le informazioni storiche di salute e i dati raccolti. Correlando questi dati con altre informazioni di localizzazione, di alimentazione, di attività

sportive e di socializzazione si potrebbero potenzialmente fare diagnosi su come ci si ammala e come prevenire certe malattie sia a livello del singolo, sia a livello aggregato di società.

Dall'analogico al digitale

Memorizzare la propria vita in bit vuol dire passare da un mondo analogico ad un mondo digitale. Con digitale ci si riferisce a tutto ciò che viene gestito, visualizzato, rappresentato e analizzato con numeri. Il termine deriva dall'inglese digit, che significa cifra, che a sua volta deriva dal latino digitus, che significa dito. Digitale può essere considerato sinonimo di numerico. Se l'analogico è il mondo del continuo, nel mondo digitale domina il discreto.

Analogico si contrappone a digitale (=discreto). Analogico rappresenta il "continuo", "non discreto": dal tachimetro dell'automobile (se a lancetta), all'orologio a lancette (quando queste non si muovono a scatti).

È possibile convertire un segnale analogico nel corrispettivo digitale, costituito da una serie di numeri e questo processo è comunemente indicato come digitalizzazione.

Il "successo" del digitale è dovuto al fatto che i computer riescono a trattare ed elaborare numeri; ovvero i computer possono utilizzare e gestire solo informazioni digitali. Diventa necessario rendere tutte le informazioni di tipo digitale o, con un verbo piuttosto utilizzato, digitalizzarle. Non è quindi solo questione di qualità o di modernità ma piuttosto di una esigenza pragmatica, collegata al mondo dell'informatica.

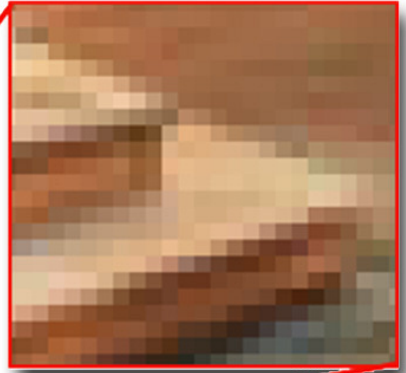
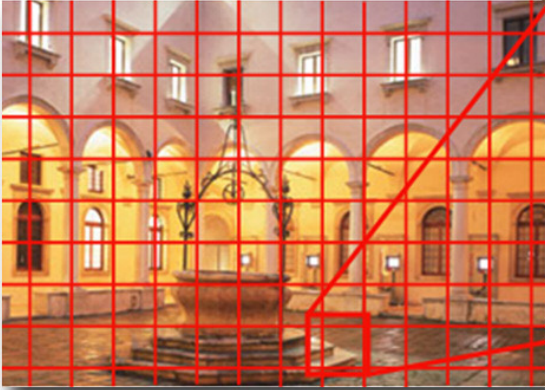
I vari tipi di informazione che devono essere digitalizzati per diventare gestibili da un computer sono solo: testo, immagini e suoni; il video può essere assimilato ad una sequenza di immagini e di suoni correlati. Ovviamente anche dati come il battito del cuore, la velocità, le coordinate geografiche possono essere memorizzati in bit; in futuro pure le emozioni saranno catturate e digitalizzate.

Per comprendere come un computer riesca a memorizzare un libro, una foto, un filmato, una canzone, il passaggio è quello di capire come le informazioni possono essere digitalizzate.

Il testo è già informazione discreta, dal momento che si basa su un numero finito di simboli (le lettere e gli altri elementi come la punteggiatura ecc.). La digitalizzazione del testo è quindi immediata: è sufficiente una tabella di conversione con la quale assegnare ad ogni lettera un numero corrispondente. Una volta stabilito che la lettera "A" corrisponde al numero 41 (è solo a titolo esemplificativo), ogni volta che chiunque preme il tasto con la "A" sulla tastiera, da qualche parte nella memoria del computer viene inserito questo numero.

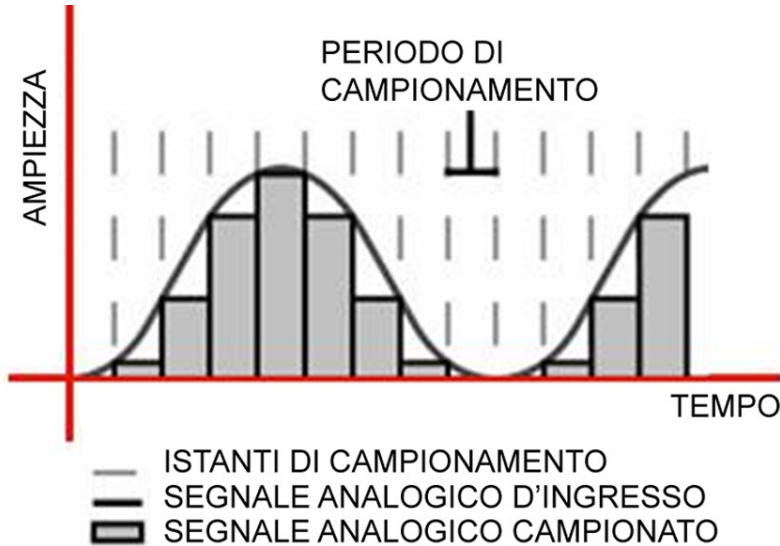
Le immagini non sono formate da elementi di base discreti come le lettere quindi si sono trovate due soluzioni di digitalizzazione: bitmap o vettoriale. Con la prima tecnica si immagina di sovrapporre ad un'immagine una griglia il più possibile fitta. Ogni elemento della griglia prende il nome di pixel (ed è l'elemento di base, simile alle lettere per il testo). Un'immagine diventa così memorizzabile sotto forma di una sequenza di numeri. Da qui il nome bitmap = mappa di bit, ovvero di numeri. La dimensione dipende dalla grandezza dell'immagine, dalla griglia più o meno fitta (risoluzione) e da quanti bit si usano per codificare il colore (profondità del colore). Più questi valori sono elevati (griglia molto fitta = alta risoluzione e molti bit utilizzati = maggiore profondità di colore) e più l'immagine digitalizzata sembrerà simile all'immagine reale. Il risultato finale è un file di dimensioni piuttosto grandi, infatti le immagini pesano molto più del testo.

Con la seconda tecnica, la grafica vettoriale, ogni immagine può essere considerata un insieme di luoghi geometrici (linee, punti, cerchi, ecc.) ed è sufficiente memorizzare le coordinate spaziali di questi elementi. Il risultato è un file molto piccolo come peso. Questo tipo di digitalizzazione non è per niente adatto a memorizzare fotografie ma è molto utilizzata per disegni tecnici, schemi, diagrammi. Ad esempio le carte geografiche sul Web o sui navigatori GPS sono realizzate con grafica vettoriale.



Suoni. La difficoltà del suono è la sua immaterialità. Il suono è costituito da vibrazioni che si propagano nell'aria e si prolungano nel tempo, infatti un suono lo possiamo rappresentare graficamente come un'onda (onde sonore). La rappresentazione numerica delle onde sonore può avvenire attraverso il "campionamento": Il suono viene "analizzato" molte volte per secondo (fino a 44.100 volte, come nel caso dei CD musicali, e allora si dice che la frequenza di campionamento è di 44.100Hz o 44,1 kHz). Ognuno di questi campioni (da cui deriva campionamento) costituisce l'elemento base al quale si può assegnare un numero. In questo caso il numero rappresenta l'ampiezza del suono nell'istante considerato. Anche per i suoni gioca un ruolo rilevante la profondità che costituisce il numero di bit utilizzati per esprimere il valore dell'ampiezza del suono.

La voce su telefono fisso è campionata a 64 kbps ma con un solo byte. Su cellulare il campionamento è a 14.400 bps, sempre con un byte; i nuovi sistemi AAC e MPEG campionato a 128, 196 kbps



Il peso di un file audio è più elevato di quello delle immagini. Il suono campionato è quello che ascoltiamo solitamente quando inseriamo un CD nello stereo e la capacità di un CD, circa 650 Megabyte, può contenere circa 74 minuti di musica campionata a 44 kHz e con 16bit di profondità (il campionamento alla qualità-CD)

Video: può essere considerato come una sequenza di immagini fisse (i fotogrammi) con audio sincronizzato. Per digitalizzare un video, ovvero "immagini in movimento con audio" si utilizza una combinazione delle tecniche bitmap e di campionamento audio. Si genera ovviamente un file molto pesante. Per risolvere questi problemi si è pensato a varie tecniche di ridimensionamento dei file, ovvero la compressione. Le tecniche di digitalizzazione sono sempre integrate da particolari algoritmi di compressione che rendono i file risultanti molto più piccoli di dimensione e più gestibili in base al formato scelto.

Cosa vuol dire bit?

“Questo file occupa 50 megabyte”, “ho un hard disk da 100 gigabyte”. Sono frasi che si sentono spesso, anche tra amici e colleghi d’ufficio...

Ma precisamente cosa sono un byte e un bit?

Partiamo dal byte: in sintesi, si tratta di un’unità di misura. Come sappiamo, il metro è l’unità di misura di lunghezza, il litro si usa per la capacità e così via. Come unità di misura per le memorie dei computer e per altre misurazioni relative all’informatica si usano il byte e i suoi multipli: è l’unità di misura del mondo digitale. Per ottenere i multipli si moltiplica per 1.024 e non per 1.000 come nel sistema metrico decimale. Questa “particolarità” è dovuta al fatto che, nel mondo dei computer, si utilizza il sistema di numerazione binario, che si basa solo sui numeri 0 e 1, a differenza del nostro mondo, dove è invece il sistema decimale (con i numeri da 1 a 9) a farla da padrone.

Comunque, nell’uso comune, 1Mbyte è equiparato a un milione di byte e 1Gbyte ad un miliardo di byte.

La parola bit è la crasi della parole binary digit e significa “cifra binaria”. Il sistema di numerazione binario è alla base dell’aritmetica interna dei computer, da cui deriva anche il nome dell’unità minimale di informazione, l’atomo sul quale è costruita l’intera tecnologia informatica, ovvero il bit.

È possibile anche vedere i bit come sottomultipli del byte: quest’ultimo è infatti costituito, per convenzione, da 8 bit, quindi $8 \text{ bit} = 1 \text{ byte}$

Anche i bit sono usati come unità di misura (per esempio della velocità di trasmissione di un collegamento Internet). Bit e byte sono anche utilizzati per dare un valore ai concetti di:

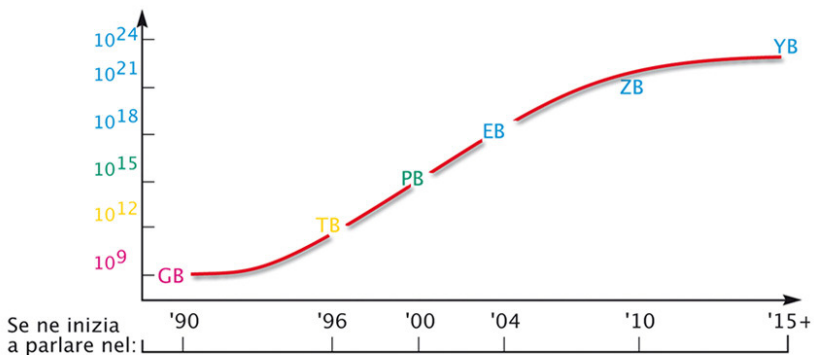
- potenza di elaborazione (ad esempio un microprocessore a 64bit è più potente e veloce di uno a 16bit).
- capacità di memoria (un hard disk da 50Gbyte è più capiente di uno da 30Gbyte).

- velocità di trasmissione (una linea ADSL da 4Megabit al secondo è molto più veloce di un modem da 56kbit al secondo).

Il “peso” degli oggetti digitali

Ogni qual volta parliamo di quantità di dati, tendiamo a farlo in astratto e diciamo kilobyte, megabyte, gigabyte senza renderci realmente conto di ciò che rappresentano. Quanto pesano un SMS, una fotografia, una parola, una canzone, un libro o un film?

Una singola foto digitale scattata da una fotocamera a 5 megapixel può essere di 40 megabyte non compressa (mentre compressa diventa 1,2 megabyte); tuttavia se cambiamo il formato e la ottimizziamo può diventare 200 kB o anche 80kB in JPEG. Quindi un oggetto digitale può avere “pesi” diversi ma rappresentare sempre la stessa cosa, questo perché si è passati da un’informazione digitale grezza ad una compressa e ottimizzata con algoritmi di compressione che cancellano il cosiddetto rumore, che consiste in quelle informazioni ridondanti che i nostri sensi non riescono a percepire e inoltre eliminano i dati ridondanti e duplicati.



Di seguito il valore dell'unità di misura digitale con l'indicazione del peso di alcuni "oggetti" digitali.

Bytes (8 bits)

- 1 byte: un carattere; ad esempio una lettera
- 10 bytes: una parola
- 100 bytes: un SMS

Kilobyte (1,000 Byte)

- 1 kilobyte: una barzelletta corta
- 2 kilobytes: una pagina scritta
- 50 kilobytes: immagine compressa di un documento

Megabyte (1,000,000 Byte)

- 1 Megabyte: una piccola storia
- 2 Megabytes: una fotografia di buona qualità
- 5 Megabytes: tutte le opere di Shakespeare; una canzone in MP3
- 10 Megabytes: un minuto di musica ad alta fedeltà
- 650-750 Megabytes: un CD-ROM

Gigabyte (1,000,000,000 Byte)

- 4 Gigabyte: un film in qualità DVD
- 20 Gigabytes: collezione audio delle opere di Beethoven;

Terabyte (1,000,000,000,000 Byte)

- 1 Terabyte: fogli di carta stampati ricavati da 50,000 alberi;

Petabyte (1,000,000,000,000,000 Byte)

- 2 Petabytes: tutte le biblioteche accademiche degli U. S. A.
- 10 PB il traffico dati nel 2008 in Italia via telefonini TIM
- 20 Petabytes: tutta la produzione di hard disk del 1995
- 200 Petabytes: tutto il materiale stampato

Exabyte (1,000,000,000,000,000,000 Byte)

- 5 Exabytes: tutte le parole dette dall'uomo
- 10 Exabytes: il numero di byte trasportati dalla rete di tlc in Italia nel 2008

Zettabyte (1,000,000,000,000,000,000,000 Byte)

Yottabyte (1,000,000,000,000,000,000,000,000 Byte)

■ **Quanti bit in una vita?**

Pensiamo di registrare ogni momento della nostra vita e delle nostre interazioni, fare della nostra memoria un database da cui attingere in ogni istante qualunque informazione generata nel corso della vita.

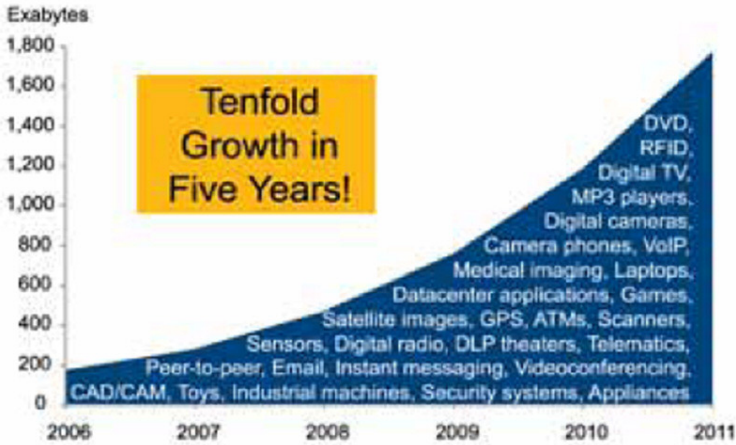
Nel report di IDC “The Diverse and Exploding Digital Universe” viene utilizzata una metodologia per dimensionare l’intero universo digitale e la stima porta alla cifra spaventosa di 281.000.000.000 Gigabyte di informazioni nel 2007, dai cui si evincono 45 Gigabyte per ogni persona sulla terra.

L’universo digitale è la somma di tutte le informazioni e i contenuti realizzati e/o conservati in digitale ed è formato da tutti i dati digitali creati, catturati, replicati nel mondo da DVD, tag RFID, TV digitale, mp3 player, videocamere digitali, fotocamere digitali, chiamate VOIP, immagini mediche, videogiochi, immagini satellitari, GPS, POS, ATM, scanner, sensori, radio digitale, peer to peer, e mail, instant messaging, videoconferenze, CAD/CAM, sistemi di sicurezza e sorveglianza, documenti digitali, contenuti sul Web ecc.

Nel 2011 viene stimato che l’ammontare delle informazioni digitali prodotte nell’anno sarà di dieci volte superiore a quelle prodotte nel 2006 per un totale di 1.800.000.000.000 Gigabyte. Dal 2007 la quantità dei dati digitali prodotti sarà maggiore della capacità di registrare e immagazzinare tutte le informazioni perché la possibilità di realizzare contenuti digitali cresce in modo pervasivo fino a coprire quasi tutti gli aspetti della vita.

Basti pensare che nel 2006 il numero delle fotocamere digitali e dei telefoni cellulari dotati di fotocamera nel mondo hanno superato la cifra di un miliardo, è facile immaginare il numero di scatti prodotti e dei momenti immortalati. Se a questo uniamo la facilità di duplicazione e condivisione a costo zero di quanto prodotto (foto o video), prendiamo consapevolezza della dimensione numerica del fenomeno.

Digital Information Created, Captured, Replicated Worldwide



Source: IDC, 2008

Chi ha i nostri dati

La società ha subito profondi cambiamenti negli ultimi venti anni e si è passati da un nucleo ristretto e localizzato di frequentazione di amici, colleghi e parenti ad un mondo aperto in cui tramite social network si riesce a tenere rapporti con un numero molto grande di individui che possono appartenere non solo a città diverse e lontane, ma anche a paesi lontani.

Se ci si sofferma a pensare chi sono gli attori che in qualche modo hanno i nostri dati e raccolgono informazioni su di noi ci accorgiamo che nuovi player si sono inseriti nella nostra vita come ad esempio i motori di ricerca o i social network. Il cambio di paradigma è che pre-

cedentemente la famigli a gli amici ristretti sapevano tutto di noi; invece, oggi, chi vive una vita digitale fornisce volontariamente o involontariamente quasi tutte le informazioni su di sé a tal punto che



Google potrebbe conoscerci meglio dei nostri genitori. In effetti ogni qual volta si effettua una ricerca, direttamente diciamo ciò a cui realmente siamo interessati e indirettamente sveliamo i nostri gusti, le nostre preferenze, i nostri dati sensibili.

Per farci un'idea di chi sono le entità che raccolgono, immagazzinano e analizzano i nostri dati sia nel mondo digitale online, sia nel mondo reale dobbiamo pensare a tutte le azioni che facciamo,

alle interazioni che abbiamo e con chi ci interfacciamo. Di seguito un elenco dei principali raccoglitori di informazioni:

- Motori di ricerca (Google, MSN Live search, Yahoo...).
- Stato / comuni.
- Banche.
- Moglie / marito / parenti.
- Siti Web/ Community / Social Network (Facebook).
- Amici.
- Operatori telefonici.
- Negozi online (Amazon...).
- Polizia.
- Scuole.
- Ospedali / medico curante.
- Supermercati (carte fedeltà).

Metadati, questi sconosciuti

Per citare Wikipedia la definizione di metadato (dal greco meta "oltre, dopo" e dal latino datum "informazione" plurale: data), letteralmente "dato su un (altro) dato", è l'informazione che descrive un insieme di dati.

Un esempio tipico di metadati è costituito dalla scheda del catalogo di una biblioteca, la quale contiene informazioni circa il contenuto e la posizione di un libro, cioè dati riguardanti i dati che si riferiscono al libro. I metadati sono una descrizione di dati.

Se, ad esempio, descriviamo una persona indichiamo caratteristiche fisiche come l'altezza, il colore degli occhi e dei capelli, o caratteristiche psicologiche, come il carattere, la simpatia, la cordialità, ecc. Queste informazioni rappresentano metadati sulla persona. In altre parole, i metadati sono dati per descrivere risorse, siano esse persone, libri, frutta, computer, pagine Web.



Quali METADATI si possono creare componendo le informazioni dell'ombra digitale?

A livello macro la scia di informazioni digitali lasciate da una persona si possono classificare in:

- SMS / MMS: testi, video e audio che l'utente scambia con la propria rete sociale ristretta.
- Voce e video chiamate: audio e video chiamate che l'utente effettua.
- Geolocalizzazione (user tracking): spostamenti dell'utente, percorsi abituali o meno; contestualizzazione delle azioni.
- Foto/Video (geolocalizzati e taggati): foto e video che l'utente scatta o riprende e può taggare e geolocalizzare, dare un valore semantico e contestualizzare.
- Sanità elettronica: cartelle elettroniche online con tutte le info storiche di salute dell'utente.
- Monitoraggio Biometrico: info di monitoring di parametri biometrici dell'utente come temperatura e battito cardiaco, glicemia e allergie.
- Documenti elettronici: documenti elettronici dell'utente come doc, fogli di calcolo, presentazioni, pdf...
- E-mail, chat / messaging: tutta la history delle conversazioni digitali.
- Social network, Blog: tutte le info presenti nei social network e nei blog dell'utente.
- Browsing (history, click stream, search): la navigazione e le ricerche effettuate sul Web.
- IPTV: cattura preferenze, utilizzo, acquisti VOD...
- Transazioni e acquisti online: tutte le informazioni di cosa si è acquistato online, quando, dove, e il costo.
- Carte di credito: tutte le informazioni di cosa si è acquistato con carta di credito, quando, dove, e il costo.
- Telepass: tutti i passaggi dai caselli autostradali, dove e quando.
- Care fedeltà: tutti i beni/servizi acquistati con sconto, quando e dove, con carte fedeltà e i punti accumulati.