

# ***Informatica Personale***

## *Integrazione*

*Dante Picca*

*Anno accademico 2009-2010*

# Indice

|                                                                   |    |
|-------------------------------------------------------------------|----|
| Indice.....                                                       | 2  |
| Premessa.....                                                     | 4  |
| 1 I paradigmi.....                                                | 5  |
| 2 Introduzione all'informatica.....                               | 7  |
| Introduzione.....                                                 | 7  |
| Che cos'è l'informatica.....                                      | 7  |
| Che cosa non è l'informatica?.....                                | 8  |
| Informazioni in formato analogico o digitale.....                 | 8  |
| Il formato digitale delle informazioni.....                       | 8  |
| La tabella ASCII.....                                             | 9  |
| Il codice binario.....                                            | 12 |
| Il passaggio di basi.....                                         | 13 |
| Bit e multipli di bit.....                                        | 14 |
| Gli esecutori.....                                                | 15 |
| Gli algoritmi.....                                                | 18 |
| Esempio 1.....                                                    | 19 |
| Esempio 2.....                                                    | 19 |
| Esempio 3.....                                                    | 21 |
| La macchina di Von Neumann.....                                   | 22 |
| La memoria.....                                                   | 24 |
| Le memorie primarie.....                                          | 24 |
| Le memorie secondarie.....                                        | 25 |
| La conservazione delle informazioni nelle memorie secondarie..... | 25 |
| 3 Le reti.....                                                    | 27 |
| Presentazione.....                                                | 27 |
| Introduzione.....                                                 | 27 |
| Che cosa è una rete ?.....                                        | 28 |
| I protocolli.....                                                 | 29 |
| Tipi di reti.....                                                 | 30 |
| Reti a bus.....                                                   | 30 |
| Reti a stella.....                                                | 31 |
| Reti a maglia.....                                                | 31 |
| Reti LAN, MAN e WAN.....                                          | 32 |
| Reti a commutazione di pacchetto e di circuito.....               | 32 |
| La commutazione di circuito.....                                  | 32 |
| La commutazione di pacchetto.....                                 | 33 |
| I browser per la navigazione ipermediale.....                     | 33 |
| La ricerca di informazioni sull'internet.....                     | 34 |
| I motori web.....                                                 | 34 |
| I motori di ricerca.....                                          | 36 |
| I newsgroup.....                                                  | 37 |
| Le directory.....                                                 | 38 |
| La posta elettronica.....                                         | 40 |
| 4 Trattamento digitale delle immagini e dei suoni.....            | 41 |
| Introduzione.....                                                 | 41 |
| Le immagini.....                                                  | 41 |
| Il campionamento spaziale.....                                    | 41 |
| La quantizzazione cromatica.....                                  | 44 |
| La rappresentazione dei colori.....                               | 45 |
| Formati di conservazione delle immagini.....                      | 46 |
| Il formato bitmap.....                                            | 46 |
| Il formato GIF.....                                               | 46 |
| Il formato PNG.....                                               | 46 |
| Formati lossy e lossless.....                                     | 47 |
| Il formato JPEG.....                                              | 47 |

|                                                                    |    |
|--------------------------------------------------------------------|----|
| I suoni .....                                                      | 51 |
| Il campionamento .....                                             | 52 |
| La quantizzazione .....                                            | 54 |
| I formati di conservazione dell'audio .....                        | 55 |
| Il formato del CD audio .....                                      | 55 |
| I formati psicoacustici .....                                      | 55 |
| Il formato MP3 .....                                               | 56 |
| 5 Cittadini nella società dell'informazione .....                  | 58 |
| Che cos'è la società della conoscenza .....                        | 58 |
| La società dell'informazione.....                                  | 58 |
| La società della conoscenza.....                                   | 59 |
| Evoluzione dei mezzi di comunicazione .....                        | 61 |
| La multimedialità diffusiva .....                                  | 61 |
| La multimedialità interattiva: un nuovo ruolo per gli utenti ..... | 63 |
| L'architettura ipermediale .....                                   | 64 |
| Testo, ipertesto ed ipermedia .....                                | 64 |
| L'ipermedialità e la società della conoscenza.....                 | 67 |
| Informazioni e società della conoscenza.....                       | 67 |
| Informazioni in formato digitale .....                             | 67 |
| Il diritto d'autore .....                                          | 71 |
| Che cosa è il diritto d'autore.....                                | 71 |
| Il diritto d'autore e la società della conoscenza.....             | 73 |
| Privacy e società della conoscenza .....                           | 75 |
| Che cosa è la privacy .....                                        | 76 |
| I principi della legge sulla privacy.....                          | 79 |
| La tutela della privacy nella società della conoscenza .....       | 80 |
| Un caso di studio.....                                             | 81 |
| La vendita on-line di brani musicali .....                         | 82 |
| La vendita on-line dei giornali.....                               | 83 |
| Conclusioni .....                                                  | 84 |
| Due articoli di approfondimento .....                              | 84 |
| Indice delle figure .....                                          | 90 |
| Indice delle tabelle .....                                         | 91 |

## **Premessa**

In presente documento costituisce un approfondimento rispetto ad alcuni argomenti trattati durante il corso di informatica personale. Ho cercato, per quanto possibile, di rispettare l'ordine con il quale gli argomenti sono stati trattati a lezione ma in alcuni casi ho dovuto effettuare scelte differenti per massimizzare la comprensione del testo.

Ritengo opportuno sottolineare che il programma del corso coincide solo parzialmente con il contenuto di questo documento che, di conseguenza, non può essere l'unica fonte per lo studio.

Roma, 24 novembre 2009

Dante Picca

# 1 I paradigmi<sup>1</sup>

Poiché nello svolgimento del corso sarà spesso utilizzato un approccio paradigmatico allo studio dei diversi argomenti, è utile definire da subito che cosa sia un paradigma e quale sia il ruolo del paradigma nello studio delle scienze.

Dal punto di vista etimologico, il termine paradigma deriva dalla parola greca *paradeigma*, e significa “modello”, “progetto” o “esempio”. Con tale accezione si trova nei testi di Platone, Socrate ed Aristotele.

Con l'accezione di modello, il paradigma verbale è utilizzato nella grammatica latina per indicare l'elenco delle cinque forme fondamentali del verbo dalle quali, per i verbi regolari, possono essere ricavate tutte le altre (prima persona dell'indicativo presente, seconda persona dell'indicativo presente, prima persona dell'indicativo perfetto, supino, infinito presente).

Nel corso della seconda metà del XX secolo – grazie soprattutto ai lavori di T. S. Kuhn – il termine paradigma è stato utilizzato, nella descrizione della scienza, in una accezione più articolata.

Secondo la definizione di Kuhn, un paradigma è un “Modello o schema accettato, caratterizzato da una costellazione di conclusioni – concetti, valori, tecniche – condivise da una comunità scientifica e utilizzate dalla comunità stessa per definire problemi e soluzioni lecite”.

Il vocabolario Treccani, definisce il paradigma come un “[...] complesso di regole metodologiche, modelli esplicativi, criteri di soluzioni di problemi che caratterizza una comunità di scienziati in una fase determinata dell'evoluzione storica della loro disciplina”.

Entrambe le definizioni evidenziano ancora come il paradigma sia un modello, uno schema o un insieme di regole. Tuttavia entrambe pongono l'accento su due aspetti fondamentali che bene identificano la moderna accezione di paradigma.

In primo luogo il paradigma non è un insieme di regole “scritte nella pietra”, ma un modello che consente di identificare un modo di operare nell'affrontare la ricerca delle soluzioni ai problemi.

In secondo luogo i paradigmi utilizzati da una comunità cambiano nel tempo e possono essere diversi da comunità a comunità: per ogni fase storica o per ogni esigenza di una comunità, può essere utilizzato un paradigma in luogo di un altro.

Vediamo ora un esempio che ben chiarisce gli aspetti appena introdotti.

Sino alla rivoluzione scientifica avvenuta durante il Rinascimento, il paradigma condiviso degli scienziati nello studio delle scienze fisiche era il così detto metodo assiomatico<sup>2</sup>, derivante dall'epistemologia aristotelica.

Secondo Aristotele la matematica, la fisica e la metafisica sono scienze teoriche (cioè finalizzate alla conoscenza pura). Nel *Metafisica* viene chiarito che “Scienza si ha quando si conoscono le cause per le quali una cosa è, e si conosce che proprio per tal causa è la cosa, in modo che non possa

---

1 Per una descrizione dell'utilizzo del paradigma nella filosofia classica: <http://www.riflessioni.it/enciclopedia/paradigma.htm>

Per una descrizione di Kuhn e del suo lavoro:

[http://it.wikipedia.org/wiki/Thomas\\_Kuhn](http://it.wikipedia.org/wiki/Thomas_Kuhn)

[http://www.hermesnet.it/glossario/k/scheda\\_kunh.html](http://www.hermesnet.it/glossario/k/scheda_kunh.html)

Un testo di riferimento per l'introduzione del concetto di paradigma nello studio delle scienze è: T. S. Kuhn - *The Structure of Scientific Revolutions*, Chicago University Press, Chicago, tr. it della II ed. *La struttura delle rivoluzioni scientifiche*, Torino, Einaudi, 1979.

Nota: nel libro non si fa alcun riferimento all'approccio paradigmatico nello studio dell'informatica.

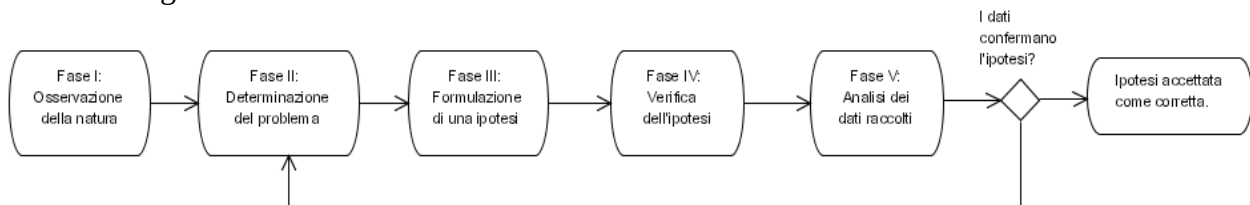
2In epistemologia, un **assioma** è una proposizione o un principio la cui verità è ritenuta talmente evidente nell'esperienza comune da non necessitare né di dimostrazione né di discussione per essere considerato ed accettato da tutti. (da <http://it.wikipedia.org/wiki/Assioma>).

darsi che sia diversamente” (da *Metafisica*, I 2). Il metodo di ricerca utilizzato dagli scienziati era di tipo rigidamente deduttivo: le asserzioni o gli eventi da giustificare vengono fatti discendere in modo deduttivo dai principi a priori (gli assiomi). Tali assiomi erano ricavati, per astrazione, dall’osservazione della realtà. Gli assiomi, avendo una valenza generale, venivano utilizzati per spiegare i singoli eventi che avvenivano nel mondo. L’approccio utilizzato viene definito di tipo top-down: partire dall’alto degli assiomi per discendere verso ciò che è da dimostrare.

Secondo la definizione data da Kuhn, quindi, il modello accettato dalla comunità ed utilizzato dalla stessa per definire problemi e soluzioni lecite è quello di fare scienza partendo dagli assiomi e deducendo di conseguenza le cause per le quali “le cose sono”.

Con il Rinascimento, cambia il paradigma utilizzato dagli scienziati per fare scienza. In primo luogo, infatti, cambia la prospettiva dell’epistemologia: non più “conoscere le cause per le quali una cosa è”, ma descrivere esattamente alcune proprietà della natura. La ricerca delle cause diviene necessaria per la comprensione dei fenomeni. Il metodo sperimentale – introdotto da Galilei – si basa sull’osservazione della realtà, da cui partire per definire le leggi che la regolano. Un approccio, quindi, di tipo bottom-up: partire dal basso (l’osservazione dei fenomeni) per arrivare alla formulazione delle leggi di valore generale.

Schematicamente il metodo sperimentale può essere riassunto nel seguente schema procedurale, visibile in Figura 1



**Figura 1: schema del metodo sperimentale.**

Nel primo passo, detto dell’*osservazione*, lo scienziato osserva un fenomeno cercando di astrarre gli aspetti salienti, utili per descriverlo e sintetizzarlo. Tale fase è profondamente influenzata dalla conoscenza e dall’esperienza dell’osservatore.

Il secondo passo, detto di *determinazione del problema*, consiste nel definire quali sono gli elementi importanti da studiare e quali variabili sono utili per studiare il fenomeno.

Successivamente viene *formulata una ipotesi*, ad esempio una legge fisica, che descrive il fenomeno per mezzo delle variabili definite precedentemente.

Tale ipotesi viene *verificata* nel quarto passo, detto della sperimentazione controllata, durante il quale vengono effettuati uno o più esperimenti mirati a verificare o confutare l’ipotesi svolta precedentemente.

Successivamente i dati vengono raccolti ed *analizzati* per verificare se l’ipotesi è corretta. In caso l’ipotesi non sia corretta si torna alla determinazione del problema che, grazie ai nuovi dati, può essere effettuata con maggiore precisione. Nel caso l’ipotesi sia corretta, questa entra di diritto all’interno di una teoria.

In questo caso, quindi, secondo la definizione di Kuhn il modello accettato dalla comunità ed utilizzato dalla stessa per definire problemi e soluzioni lecite è quello di partire dall’osservazione della natura per ricavare le leggi che le regolano attraverso un processo iterativo di formulazione di una ipotesi e sua verifica. È importante notare che questo metodo non è limitato allo studio della fisica: in molte branche delle scienze, dalla medicina alle scienze biologiche, alle scienze sociali cambia l’oggetto dello studio, ma non il metodo per affrontarlo.

Deve essere sottolineato che il metodo assiomatico, pur se superato dalla comunità degli studiosi delle scienze fisiche, viene utilizzato nello studio della matematica, dove la dimostrazione dei teoremi viene fatta discendere deduttivamente dai postulati e dalle dimostrazioni di altri teoremi.

## 2 Introduzione all'informatica

### Introduzione

Questa unità didattica contiene una introduzione all'informatica.

In particolare viene analizzato che cosa sia l'informatica partendo dalla definizione classica.

Successivamente viene analizzata la distinzione tra trattamento analogico e digitale delle informazioni.

Quindi sono introdotti gli esecutori, con particolare attenzione ai generici esecutori: quella particolare categoria di esecutori in grado di risolvere gli algoritmi.

Successivamente viene introdotto il modello della macchina di Von Neumann, come modello funzionale e strutturale del generico esecutore.

L'unità didattica si conclude con la presentazione dei diversi tipi di dispositivi di memorizzazione presenti negli elaboratori moderni (memoria primaria e memoria secondaria).

### Che cos'è l'informatica

Il termine **informatica** deriva dalla crasi dei termini **informazione** automatica ed è stato utilizzato per la prima volta nel 1962 da Philippe Dreyfus (in francese *informatique* come crasi di *information automatique*) per indicare la gestione automatica delle informazioni da parte di un elaboratore.

In un'accezione più generale l'informatica può essere definita come una **scienza** che studia i metodi e le tecniche per il **trattamento automatizzato** delle **informazioni**.

Data la centralità di questa definizione, vale la pena concentrarsi su alcune parole chiave presenti all'interno della definizione stessa.

Per **scienza** si intende il processo rivolto alla produzione di conoscenza da parte degli uomini.

In questo senso, quindi, attraverso lo studio dell'informatica definiamo e sviluppiamo i metodi e le tecniche per il trattamento automatico delle informazioni.

Con **trattamento** intendiamo, nella sua accezione più generale, tutte le attività che possono essere effettuate su dei dati, ne sono esempi la creazione, la raccolta, la memorizzazione, la trasmissione, la rappresentazione, l'elaborazione o la cancellazione. Anche se alcuni di questi trattamenti sono diventati delle scienze a se stanti (ad esempio la telematica è la scienza che studia i metodi e le tecniche per la trasmissione delle informazioni), l'informatica – nella sua accezione più generale – li comprende tutti.

Il termine **automatico** deriva dalla parola greca *automatos* e significa “di propria forza”, “indipendente”. L'indipendenza è qui riferita alla presenza di esseri umani durante il trattamento delle informazioni: compito dell'uomo è quello di trasferire al sistema informativo i dati e le istruzioni, ma non partecipa in modo continuo e diretto alle singole attività necessarie per il trattamento.

Visto il taglio delle definizioni introdotte, anche la definizione di informazione non può che essere generale: con **informazione** intendiamo tutto ciò a cui attribuiamo un valore e che, quindi, merita di essere trattata.

Il lettore poco attento può trovare che una definizione così generale di informatica, essendo nei fatti onnicomprensiva, non fornisca alcun contributo alla comprensione. In realtà si definiscono dei margini, se pure molto ampi, all'interno dei quali un'attività può essere considerata o meno informatica.

Innanzitutto chiarisce che il trattamento delle informazioni non deve necessariamente avvenire in presenza di elaboratori elettronici: il telegrafo è sicuramente un esempio calzante di informatica (o di telematica). Il primo esempio di telegrafo risale al 1837.

La definizione chiarisce, inoltre, che per avere informatica il trattamento deve avvenire in modo automatico. I segnali di fumo usati per comunicare dai nativi americani non sono un esempio di informatica poiché nessuna operazione può essere automatizzata.

La definizione evidenzia che per avere informatica non è necessario disporre di un computer: anche un semplice videoregistratore è un esempio di strumento che tratta in modo automatico le informazioni. Infatti è sufficiente inserire la cassetta (i dati da trattare) e le istruzioni con le quali operare sui dati (queste le hanno inserite i produttori del videoregistratore). All'invio del comando play, il videoregistratore tratta le informazioni presenti sulla cassetta inviando immagini e suoni al televisore. Discorso analogo vale per molti dispositivi, come i telefoni cellulari.

## Che cosa non è l'informatica?

Dall'ultimo decennio del secolo scorso, è andata maturando la convinzione che l'informatica fosse una sorta di e-magic, cioè un mezzo attraverso il quale trovare la soluzione indiscriminata a qualsiasi problema solo perché tale soluzione è mediata attraverso un computer.

Come sarà evidente in seguito, l'informatica fornisce solo gli strumenti per la soluzione automatizzata di problemi che sono stati studiati ed analizzati dall'uomo. Einstein scrisse che "Un giorno le macchine riusciranno a risolvere tutti i problemi, ma mai nessuna di esse potrà porne uno". Ciò significa che una macchina sarà in grado di svolgere qualsiasi operazione impartitagli dall'uomo, ma che nessuna macchina saprà mai porsi – *motu proprio* – una domanda a cui dover dare una risposta. Più avanti nel testo definiremo un elaboratore elettronico come uno "stupido velocissimo". Tale definizione ci consente di cogliere come l'utilità degli strumenti informatici per il trattamento delle informazioni sia legata alla velocità (ed alla precisione) con cui un elaboratore può svolgere operazioni per nostro conto, ma esclude qualsiasi possibilità di scelta autonoma da parte dell'elaboratore.

Tale caratteristica dell'informatica non deve tuttavia sminuirne l'importanza: la conoscenza di come risolvere un problema e la capacità di risolverlo sono competenze distinte. Spetta all'uomo la competenza su come risolvere un problema ed alla macchina la capacità di farlo. Per tale motivo il termine **risolutore** è spesso usato come sinonimo di uomo, ed il termine **esecutore** come sinonimo di macchina, anche se spesso l'uomo può essere anche un esecutore: il computer è solo più veloce e preciso.

## Informazioni in formato analogico o digitale

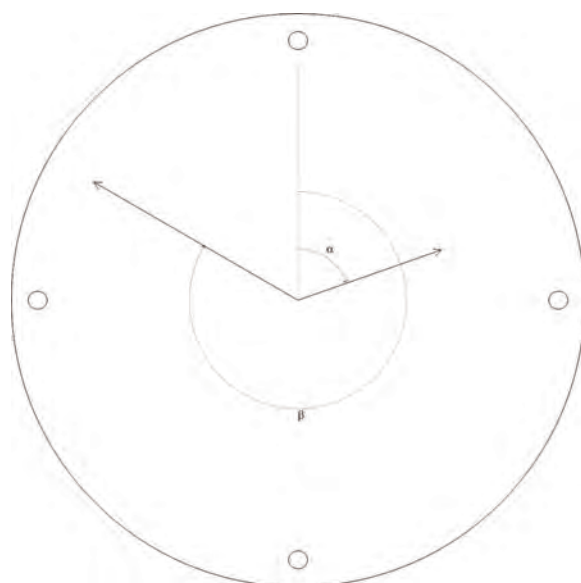
Quale tipo di informazioni sono trattate dagli elaboratori elettronici? Intorno a noi esistono moltitudini di informazioni in formato diverso: pensiamo al nome di una strada, alla temperatura dell'aria, ad un numero di telefono, al suono di una canzone o all'immagine di un paesaggio. Ognuna di queste informazioni deve, per le ragioni più diverse, di essere trattata. Per ognuna di esse il nostro cervello memorizza informazioni in formati diversi: ad esempio memorizza un suono, un insieme di colori, oppure semplicemente un numero telefonico. Se all'inizio si realizzarono strumenti appositi per il trattamento dei diversi tipi di informazioni (pensiamo alle prime agende elettroniche, in grado solo di conservare nome, cognome, telefono ed indirizzo), oggi qualsiasi telefono cellulare è in grado di trattare tipologie diverse di informazioni: non solo numeri le lettere, ma anche suoni ed immagini. La chiave di volta è stata l'introdurre un formato unico per rappresentare le informazioni, che andasse bene per un insieme eterogeneo (colori, suoni, numeri, lettere).

Che cosa c'entrano, però, i colori rappresentati all'interno di una fotografia con il nostro numero di telefono di casa? All'apparenza nulla, eppure entrambe sono conservati nella memoria del nostro cellulare, vediamo di capire come è stato risolto il problema.

## Il formato digitale delle informazioni

*Si dice che una informazione è rappresentata in formato **analogico**, quando il suo valore è ricavato per analogia (da qui analogico) con un'altra grandezza fisica.*

Prendiamo ad esempio un orologio a lancette: come facciamo a sapere che ore sono? Semplicemente misuriamo l'angolo tra la verticale e la lancetta dell'ora e quindi, effettuando una analogia tra la posizione delle lancette ed il tempo trascorso, otteniamo l'ora. La stessa operazione la effettuiamo con la lancetta dei minuti. Nell'esempio visibile in Figura 2, il valore di  $\alpha$  corrisponde all'angolo compreso tra la verticale (mezzogiorno) e la posizione della lancetta delle ore mentre il valore di  $\beta$  corrisponde all'angolo tra la verticale (zero minuti) e la posizione della lancetta dei minuti. In realtà noi effettuiamo la misura di due angoli, e dai valori ottenuti ricaviamo una misura di tempo.



**Figura 2: un orologio analogico**

Come facciamo, invece, a conoscere la temperatura dell'aria con un termometro a mercurio? Misuriamo l'altezza della colonnina di mercurio e, da quella, ricaviamo per analogia la temperatura. Anche in questo caso abbiamo utilizzato una misura analogica.

Nel caso di un termometro digitale, invece, al termine del processo di misura leggiamo semplicemente la temperatura sul display sotto forma di numeri. Lo stesso vale per la lettura dell'ora con un orologio digitale.

*Si dice che una informazione è rappresentata in formato **digitale** quando il suo valore viene espresso sotto forma di numeri (in inglese digit, da cui digitale).*

Abbiamo così imparato la differenza sostanziale tra informazioni in formato digitale ed informazioni in formato analogico. Questo, però, sembra non spiegare in che modo siamo riusciti ad insegnare al telefono cellulare come trattare informazioni tanto diverse. Se riflettiamo su quello che abbiamo appena visto, però, ci rendiamo conto che noi trattiamo la temperatura ed il tempo allo stesso modo: attraverso dei numeri. Abbiamo quindi definito un formato comune per conservare, nel nostro cervello, informazioni tanto diverse tra loro utilizzando un formato unico. Ovviamente il nostro cervello può trattare informazioni in formato analogico e digitale: spesso pur non conoscendo la temperatura di un ambiente (informazione digitale), ricordiamo una sensazione di freddo o di caldo (informazione analogica).

Gli elaboratori elettronici sono in grado di trattare informazioni solo sotto forma di numeri.

Questa è una differenza marcata rispetto a noi: quando pensiamo ad un quadro di Van Gogh, vediamo la rappresentazione che la nostra mente si è fatta del quadro stesso. Per il nostro cellulare, invece, l'immagine non è altro che una sequenza di numeri che rappresentano i diversi colori nei diversi punti dello schermo.

Abbiamo così compreso che un elaboratore elettronico tratta sotto forma di numeri l'ora, la temperatura, un numero di telefono, il nome di una strada, il suono di una canzone o una foto. Questo vuol dire che è necessario digitalizzare (cioè trasformare in numeri) le diverse informazioni che deve trattare il computer.

Esistono quattro tipologie di dati che un sistema informativo dovrebbe poter trattare e sulle quali focalizzeremo la nostra attenzione: i numeri, le lettere, le immagini ed i suoni (un filmato può essere considerato come una sequenza di immagini a cui è sincronizzato un audio). Ora vedremo come vengono trasformate le lettere in numeri, mentre nella terza unità didattica analizzeremo come vengono trattati i suoni e le immagini.

## La tabella ASCII

La trasformazione delle lettere in numeri è una procedura molto semplice: è sufficiente associare ad ogni lettera un numero diverso. In questo modo è possibile memorizzare all'interno di un elaboratore elettronico qualunque testo. Questo procedimento non è diverso da quanto fanno i bambini che, per nascondere un messaggio "segreto", trasformano ogni lettera in un numero e viceversa.

La più importante codifica dei caratteri in numeri risale al 1963, quando fu definita la tabella

ASCII (acronimo di American Standard Code for Information Interchange) che associa ad ogni carattere presente sulle tastiere americane (e quindi prive dei caratteri accentati) un numero diverso. La tabella ASCII rappresentò un'importante conquista dell'informatica: infatti sino alla sua definizione ed alla successiva elevazione a standard, ogni produttore di elaboratori elettronici utilizzava una propria tabella, e gli utenti non si potevano elaborare le informazioni trattate con elaboratori realizzati da produttori diversi.

Il lettore non deve stupirsi che nel 1963 qualcuno si ponesse il problema di scambiare le informazioni tra elaboratori elettronici: nel 1965 le idee circa la possibilità di connettere computer insieme per scambiarsi informazioni erano così diffuse che l'agenzia governativa americana ARPA (acronimo di Advanced Research Projects Agency) sponsorizzò uno studio sulle reti di computer. Lo studio portò, dopo solo tre anni alla nascita dell'arpanet, la rete che si evolse anni dopo nell'internet.

La prima versione della tabella ASCII, permetteva di rappresentare sino a 128 simboli diversi, come visibile nella Tabella 1. I codici decimali da 0 a 31 ed il codice decimale 127 non sono stampabili (ad esempio il codice 127 corrisponde alla pressione del tasto delete sulla tastiera), mentre i codici numerici da 32 a 126 sono tutti stampabili (il codice 32 corrisponde alla pressione della barra spaziatrice).

| Numero | Carattere | Numero | Carattere | Numero | Carattere | Numero | Carattere |
|--------|-----------|--------|-----------|--------|-----------|--------|-----------|
| 0      | NUL       | 32     | SP        | 64     | @         | 96     | `         |
| 1      | SOH       | 33     | !         | 65     | A         | 97     | a         |
| 2      | STX       | 34     | "         | 66     | B         | 98     | b         |
| 3      | ETX       | 35     | #         | 67     | C         | 99     | c         |
| 4      | EOT       | 36     | \$        | 68     | D         | 100    | d         |
| 5      | ENQ       | 37     | %         | 69     | E         | 101    | e         |
| 6      | ACK       | 38     | &         | 70     | F         | 102    | f         |
| 7      | BEL       | 39     | '         | 71     | G         | 103    | g         |
| 8      | BS        | 40     | (         | 72     | H         | 104    | h         |
| 9      | TAB       | 41     | )         | 73     | I         | 105    | i         |
| 10     | LF        | 42     | *         | 74     | J         | 106    | j         |
| 11     | VT        | 43     | +         | 75     | K         | 107    | k         |
| 12     | FF        | 44     | ,         | 76     | L         | 108    | l         |
| 13     | CR        | 45     | -         | 77     | M         | 109    | m         |
| 14     | SO        | 46     | .         | 78     | N         | 110    | n         |
| 15     | SI        | 47     | /         | 79     | O         | 111    | o         |
| 16     | DLE       | 48     | 0         | 80     | P         | 112    | p         |
| 17     | DC1       | 49     | 1         | 81     | Q         | 113    | q         |
| 18     | DC2       | 50     | 2         | 82     | R         | 114    | r         |
| 19     | DC3       | 51     | 3         | 83     | S         | 115    | s         |
| 20     | DC4       | 52     | 4         | 84     | T         | 116    | t         |
| 21     | NAK       | 53     | 5         | 85     | U         | 117    | u         |
| 22     | SYN       | 54     | 6         | 86     | V         | 118    | v         |
| 23     | ETB       | 55     | 7         | 87     | W         | 119    | w         |
| 24     | CAN       | 56     | 8         | 88     | X         | 120    | x         |
| 25     | EM        | 57     | 9         | 89     | Y         | 121    | y         |
| 26     | SUB       | 58     | :         | 90     | Z         | 122    | z         |
| 27     | ESC       | 59     | ;         | 91     | [         | 123    | {         |
| 28     | FS        | 60     | <         | 92     | \         | 124    |           |
| 29     | GS        | 61     | =         | 93     | ]         | 125    | }         |
| 30     | RS        | 62     | >         | 94     | ^         | 126    | ~         |
| 31     | US        | 63     | ?         | 95     | _         | 127    | DEL       |

**Tabella 1: tabella ASCII base**

Successivamente, il numero di caratteri è cresciuto a 256, attraverso diverse evoluzioni, (ad esempio per consentire il trattamento anche dei caratteri accentati presenti sulle tastiere

europee) arrivando così alla tabella ASCII estesa, implementata su tutti i computer. La parte estesa della tabella ASCII utilizzata oggi sui computer è visibile in Tabella 2.

| Numero | Carattere | Numero | Carattere | Numero | Carattere | Numero | Carattere |
|--------|-----------|--------|-----------|--------|-----------|--------|-----------|
| 128    | €         | 160    |           | 192    | À         | 224    | à         |
| 129    |           | 161    | ı         | 193    | Á         | 225    | á         |
| 130    | ,         | 162    | ¢         | 194    | Â         | 226    | â         |
| 131    | f         | 163    | £         | 195    | Ã         | 227    | ã         |
| 132    | „         | 164    | ¤         | 196    | Ä         | 228    | ä         |
| 133    | ...       | 165    | ¥         | 197    | Å         | 229    | å         |
| 134    | †         | 166    | ¦         | 198    | Æ         | 230    | æ         |
| 135    | ‡         | 167    | §         | 199    | Ç         | 231    | ç         |
| 136    | ^         | 168    | ¨         | 200    | È         | 232    | è         |
| 137    | ‰         | 169    | ©         | 201    | É         | 233    | é         |
| 138    | Š         | 170    | ª         | 202    | Ê         | 234    | ê         |
| 139    | <         | 171    | «         | 203    | Ë         | 235    | ë         |
| 140    | Œ         | 172    | ¬         | 204    | Ì         | 236    | ì         |
| 141    |           | 173    |           | 205    | Í         | 237    | í         |
| 142    | Ž         | 174    | ®         | 206    | Î         | 238    | î         |
| 143    |           | 175    | ¯         | 207    | Ï         | 239    | ï         |
| 144    |           | 176    | °         | 208    | Ð         | 240    | ð         |
| 145    | ‘         | 177    | ±         | 209    | Ñ         | 241    | ñ         |
| 146    | ’         | 178    | ²         | 210    | Ò         | 242    | ò         |
| 147    | “         | 179    | ³         | 211    | Ó         | 243    | ó         |
| 148    | ”         | 180    | ´         | 212    | Ô         | 244    | ô         |
| 149    | •         | 181    | µ         | 213    | Õ         | 245    | õ         |
| 150    | —         | 182    | ¶         | 214    | Ö         | 246    | ö         |
| 151    | —         | 183    | ·         | 215    | ×         | 247    | ÷         |
| 152    | ~         | 184    | ,         | 216    | Ø         | 248    | ø         |
| 153    | ™         | 185    | ¹         | 217    | Ù         | 249    | ù         |
| 154    | š         | 186    | º         | 218    | Ú         | 250    | ú         |
| 155    | >         | 187    | »         | 219    | Û         | 251    | û         |
| 156    | œ         | 188    | ¼         | 220    | Ü         | 252    | ü         |
| 157    |           | 189    | ½         | 221    | Ý         | 253    | ý         |
| 158    | ž         | 190    | ¾         | 222    | Þ         | 254    | þ         |
| 159    | ÿ         | 191    | ¿         | 223    | ß         | 255    | ÿ         |

**Tabella 2: tabella ASCII estesa**

Ora che disponiamo della tabella ASCII, possiamo ricavare in che modo un elaboratore memorizza una frase.

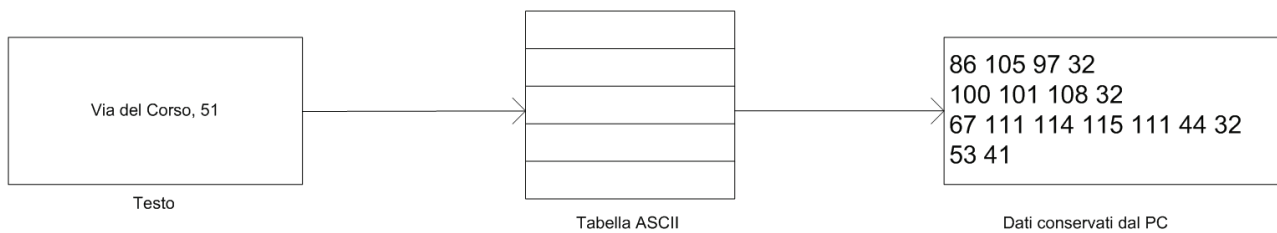
Ad esempio la frase "L'Italia è una Repubblica" nel computer è conservata secondo quanto visibile in Tabella 3.

| Testo  | L  | '  | I  | t   | a  | l   | i   | a  | è  | u   | n  | a   | R   | e  | p  | u  | b   | b   | l   | i  | c  | a   |     |    |    |
|--------|----|----|----|-----|----|-----|-----|----|----|-----|----|-----|-----|----|----|----|-----|-----|-----|----|----|-----|-----|----|----|
| Valore | 76 | 39 | 73 | 116 | 97 | 108 | 105 | 97 | 32 | 232 | 32 | 117 | 119 | 97 | 32 | 82 | 101 | 112 | 117 | 98 | 98 | 108 | 105 | 99 | 97 |

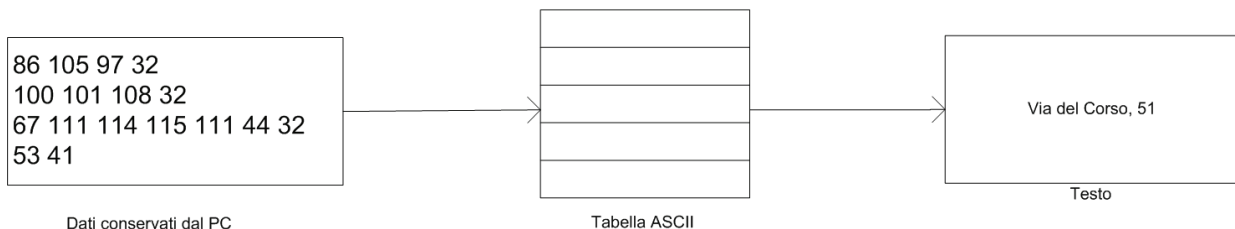
**Tabella 3: esempio di testo e sua rappresentazione in formato ASCII.**

Siamo ora in grado di comprendere cosa succede quando inseriamo o leggiamo un indirizzo dalla memoria del nostro telefono cellulare: per inserirlo, noi scriviamo le lettere attraverso la tastiera, il cellulare converte le lettere dell'indirizzo in numeri utilizzando la tabella ASCII e memorizza i numeri nella sua memoria. Per leggere i valori il cellulare preleva dalla memoria i numeri, utilizzando la tabella ASCII, trasforma i numeri in lettere e riporta le lettere sul display. Questo procedimento è esemplificato in Figura 3.

## Memorizzazione



## Riproduzione



**Figura 3: procedure di memorizzazione e riproduzione**

Può sembrare un lavoro inutile trasformare le lettere in numeri se poi è necessario fare il lavoro inverso: sicuramente è vero ma tecniche alternative, per quanto studiate, non danno ancora i medesimi vantaggi in termini di rapporto costi/prestazioni.

### Il codice binario

Ora che abbiamo chiaro che un computer tratta i dati sotto forma di numeri ci dobbiamo domandare sotto che forma questi numeri sono trattati. Non è una domanda inutile: per rispondere impareremo cosa sono i bit.

Noi uomini utilizziamo, per rappresentare i numeri, la così detta numerazione decimale: esistono dieci elementi o simboli elementari (i numeri da 0 a 9) e, combinando tali simboli elementari, è possibile rappresentare qualsiasi numero.

Un numero, comunque grande, sarà sempre una sequenza ordinata di simboli elementari. Per **ordinata** intendiamo che è importante tenere in considerazione non solo il significato del simbolo, ma anche la sua posizione relativa rispetto agli altri simboli<sup>3</sup>: l'anno 1972 è diverso dall'anno 1792 perché, seppure contengono entrambe gli stessi simboli elementari, tuttavia questi sono posizionati diversamente: sono due sequenze dei medesimi simboli elementari, ma sono ordinate in modo diverso.

Perché l'ordine dei numeri è così importante? Perché a seconda della sua posizione il simbolo elementare assume un peso differente e viene moltiplicato per potenze di 10 crescenti, partendo dalla prima posizione a destra. Ne consegue che 1792 è pari a:  $1 \cdot 10^3 + 7 \cdot 10^2 + 9 \cdot 10^1 + 2 \cdot 10^0$ , ovvero,  $1000 + 700 + 90 + 2$ , mentre 1972 è pari a:  $1 \cdot 10^3 + 9 \cdot 10^2 + 7 \cdot 10^1 + 2 \cdot 10^0$ , ovvero,  $1000 + 900 + 70 + 2$ . Diviene così evidente l'importanza dell'ordine della posizione dei numeri nella numerazione decimale.

Noi usiamo la numerazione decimale perché abbiamo dieci dita e, prima della scrittura, i simboli elementari erano proprio le dita.

Che tipo di numerazione avremmo potuto utilizzare se, invece di 10 dita, ne avessimo avute 8? Semplicemente avremmo avuto 8 simboli elementari (per convenzione i numeri da 0 a 7). Per il resto non sarebbe cambiato nulla, le operazioni sui numeri (ad esempio somma o

---

<sup>3</sup> Questo significa considerare che  $0 < 1$ ,  $1 < 2$ , ...,  $8 < 9$ .

sottrazione) sarebbero state sempre le stesse. Noi chiamiamo questa numerazione la numerazione ottale.

E se, invece di 10 dita ne avessimo avute 16? Ancora una volta avremmo avuto 16 simboli elementari (per convenzione i numeri da 0 a 9 più le lettere A, B, C, D, E, F), per il resto non sarebbe cambiato nulla. Noi chiamiamo questa numerazione la numerazione esadecimale.

Ogni numero, espresso in numerazione decimale, può essere espresso in una qualsiasi altra numerazione. La Tabella 4 evidenzia bene questo concetto.

| <i><b>Num.<br/>decimale</b></i> | <i><b>Num.<br/>ottale</b></i> | <i><b>Num.<br/>esadecimale</b></i> | <i><b>Num.<br/>decimale</b></i> | <i><b>Num.<br/>ottale</b></i> | <i><b>Num.<br/>esadecimale</b></i> |
|---------------------------------|-------------------------------|------------------------------------|---------------------------------|-------------------------------|------------------------------------|
| 0                               | 0                             | 0                                  | 10                              | 12                            | A                                  |
| 1                               | 1                             | 1                                  | 11                              | 13                            | B                                  |
| 2                               | 2                             | 2                                  | 12                              | 14                            | C                                  |
| 3                               | 3                             | 3                                  | 13                              | 15                            | D                                  |
| 4                               | 4                             | 4                                  | 14                              | 16                            | E                                  |
| 5                               | 5                             | 5                                  | 15                              | 17                            | F                                  |
| 6                               | 6                             | 6                                  | 16                              | 20                            | 10                                 |
| 7                               | 7                             | 7                                  | 17                              | 21                            | 11                                 |
| 8                               | 10                            | 8                                  | 18                              | 22                            | 12                                 |
| 9                               | 11                            | 9                                  | 19                              | 23                            | 13                                 |

**Tabella 4: rappresentazione in diverse numerazioni dei primi 20 numeri.**

Ovviamente chiunque potrebbe confondersi con le numerazioni: se dico 23, intendo 23 in decimale, in ottale, in esadecimale od in quale altra numerazione? Per evitare questa ambiguità esiste una convenzione: dopo il numero è necessario indicare il numero di simboli elementari presenti in nella numerazione utilizzata (tale valore è chiamato base). A questo punto non ci possono essere errori:  $23_8$  sarà 23 in numerazione ottale,  $23_{16}$  sarà 23 in numerazione esadecimale. Per convenzione, quando la numerazione è definita a priori, la base non viene indicata (è per questo che noi scriviamo 23 e non  $23_{10}$ ).

Tutti i numeri espressi in una numerazione qualsiasi, possono essere espressi anche nelle altre: la numerazione scelta è solo una convenzione che noi utilizziamo per rappresentare la cardinalità di un insieme. Se in una cassetta di frutta ci sono un certo numero di pere, non importa se scriviamo  $100_{10}$ ,  $144_8$  o  $64_{16}$ : il numero di pere nella cassetta non cambia. Allo stesso modo se devo dividere il contenuto della cassetta in  $5_{10}$  parti (quindi in  $5_8$  o in  $5_{16}$ ), ogni parte sarà composta da  $20_{10}$ ,  $24_8$ ,  $14_{16}$  pere. La numerazione utilizzata non cambia il risultato delle operazioni che noi eseguiamo sui numeri, cambia solo la rappresentazione con cui tali numeri ci sono presentati.

## Il passaggio di basi

Notiamo che, esattamente come avviene per la numerazione decimale, anche per gli altri sistemi di numerazione i numeri sono espressi come una sequenza ordinata di simboli elementari. Quello che cambia è che le singole cifre che compongono il numero non saranno più moltiplicate per potenze di 10, ma per potenze della basi utilizzata. Si ha quindi:

$$\begin{aligned} 100_{10} &= 1 \cdot 10^2 + 0 \cdot 10^1 + 0 \cdot 10^0 \\ 144_8 &= 1 \cdot 8^2 + 4 \cdot 8^1 + 4 \cdot 8^0 \end{aligned}$$

In una generica base  $\Psi$  si avrà:

$$64_{\Psi} = 6 \cdot \Psi^1 + 4 \cdot \Psi^0$$

Dato un numero espresso in una qualsiasi base, come si fa a ricavare il corrispondente numero in base 10?

La procedura è molto semplice, prendiamo ad esempio il numero  $3B4_{\Psi}$  e vediamo come esprimerlo in base 10. Partiamo da quanto scritto prima:

$$3B4_{\Psi} = 3 \cdot \Psi^2 + B \cdot \Psi^1 + 4 \cdot \Psi^0$$

Se la base è 16, allora riportiamo tutti i termini dalla base 16 alla base 10 abbiamo:

$$\begin{aligned}\Psi &= 16 \\ 3_{16} &= 3_{10} \\ B_{16} &= 11_{10} \\ 4_{16} &= 4_{10}\end{aligned}$$

quindi:

$$3B4_{16} = 3_{10} \cdot (16_{10})^2 + B_{10} \cdot (16_{10})^1 + 4_{10} \cdot (16_{10})^0 = 768_{10} + 176_{10} + 4_{10} = 948_{10}$$

Passare da un numero in base 10 ad un numero in una generica base utilizzando la tecnica appena vista è meno semplice, perché dovremmo saper fare le moltiplicazioni in quella base. Infatti, se volessimo trasformare il numero  $57_{10}$  in base 2, poiché:

$$\begin{aligned}10_{10} &= 1010_2 \\ 5_{10} &= 101_2 \\ 7_{10} &= 111_2\end{aligned}$$

Avremmo:

$$57_{10} = 101_2 \cdot (1010_2)^1 + 111_2 \cdot (1010_2)^0$$

Per poter risolvere questo problema dovremmo sapere come moltiplicare due numeri binari; Questa operazione è simile a quella tra due numeri decimali, ma esiste un sistema per passare dalla base 10 ad un'altra base senza bisogno di conoscere come effettuare le operazioni in basi diverse dalla base 10.

Supponiamo di voler esprimere il numero  $461_{10}$  in base 8, in generale avremo:

$$461_{10} = \alpha_8 \cdot (8)^2 + \beta_8 \cdot (8)^1 + \gamma_8 \cdot (8)^0 = \alpha_8 \cdot (8)^2 + \beta_8 \cdot 8 + \gamma_8$$

dove  $\alpha_8$ ,  $\beta_8$  e  $\gamma_8$  sono i simboli elementari in base 8 che stiamo cercando e, proprio perché sono i simboli elementari della base 8 possono assumere solo i valori 0,1,2,3,4,5,6,7 (questo è identico a quanto visto per la base 10, dove i simboli elementari sono i numeri compresi tra 0 e 9).

Se dividiamo entrambe i termini dell'equazione per 8 (cioè per la base) otteniamo:

$$461/8 = 57,625 = \alpha_8 \cdot 8 + \beta_8 + \gamma_8/8$$

Poiché  $\alpha_8 \cdot 8 + \beta_8$  è un numero intero e  $\gamma_8/8$  non è un numero intero ( $\gamma_8$  è compreso tra 0 e 7), abbiamo:

$$0,625 = \gamma_8/8$$

e

$$57 = \alpha_8 \cdot 8 + \beta_8$$

Dalla prima equazione ricaviamo:

$$\gamma_8 = 0,625 \cdot 8 = 5_8$$

A questo punto possiamo dividere l'equazione  $57 = \alpha_8 \cdot 8 + \beta_8$  per 8 ottenendo:

$$7,125 = \alpha_8 + \beta_8/8$$

procedendo similmente a prima si ricava:

$$\alpha_8 = 7_8$$

e

$$\beta_8 = 0,125 \cdot 8 = 1_8$$

quindi:

$$461_{10} = 7_8 \cdot (8)^2 + 1_8 \cdot (8)^1 + 5_8 \cdot (8)^0 = 715_8$$

Se volete esercitarvi con il passaggio di basi, potete utilizzare la calcolatrice di Windows (o del Mac o di Linux a seconda del sistema operativo che utilizzate): nella versione "scientifica", ha le funzioni per il passaggio di basi tra binario, ottale, decimale ed esadecimale.

## Bit e multipli di bit

Ora che abbiamo chiaro cosa sono le diverse numerazioni possiamo capire che tipo di dati trattano i computer.

Un computer, infatti, tratta le informazioni utilizzando la rappresentazione in base due, o binaria. Per convenzione i simboli utilizzati nella numerazione binaria sono 0 od 1. Nulla di strano, è come se un computer avesse due sole dita.

La quantità minima di informazione che può essere trattata da un computer è chiamata bit e può assumere solo i valori 0 ed 1. Bit è l'acronimo di Binary digIT (numero binario). Secondo molti, la scelta di utilizzare questo acronimo e non altri è motivata dal significato della parola bit in inglese: pezzettino. Il termine, quindi, rende immediatamente evidente che abbiamo a che fare con una quantità piccola: la minima quantità di informazione che può essere trattata da un computer. Se, in luogo di bit, si fosse usato l'acronimo di BInary diGIT bigit (in inglese

oggetto grande) questo non avrebbe reso bene l'idea.

L'informazione trasportata da un singolo bit è molto piccola (spesso oltre che 0 ed 1 si utilizzano anche altre immagini come: sì/no, acceso/spento), per tale motivo si utilizzano spesso i multipli dei bit. Anche questa non è una novità: nella nostra vita utilizziamo spesso multipli e sottomultipli delle grandezze fisiche a seconda di cosa vogliamo indicare. Indicando l'altezza di un neonato diremo che è alto 55 centimetri, indicando la distanza tra Roma e Milano parleremo di circa 700 Km. Se non avessimo voluto utilizzare i multipli (nel nostro caso i chilometri) ed i sottomultipli (nel nostro caso i centimetri) del metro, avremmo dovuto indicare l'altezza del neonato in 0,55 metri e la distanza tra Roma e Milano in circa 700.000 metri.

Anche utilizzando la numerazione in formato binario può risultare utile utilizzare i multipli dei bit. Di seguito introduciamo i più diffusi.

### Byte

Un byte corrisponde ad 8 bit. Può sembrare strano scegliere proprio 8 bit invece di 10 come siamo abituati; in realtà la lunghezza scelta per il byte è legata alla tabella ASCII. Infatti, con 8 bit si possono rappresentare tutti i numeri da 0 a 255 e la tabella ASCII ha proprio 256 valori diversi. Quindi con un byte noi possiamo memorizzare un carattere. Questo è il motivo per il quale la dimensione dei dispositivi di memorizzazione (ad esempio gli hard disk) è espressa in miliardi di byte: significa che possiamo inserire nel disco un certo numero di caratteri di testo.

### Kilo

Quando leggiamo il suffisso chilo, nella nostra testa moltiplichiamo per mille il numero cui il suffisso si riferisce: un oggetto che pesa 80 chili, vuol dire che pesa 80.000 grammi. In informatica un chilo non corrisponde esattamente a 1000, bensì a 1024. La ragione di tale apparente illogicità è legata alla scelta della numerazione utilizzata per rappresentare i numeri. Infatti nella numerazione decimale si usa 1000 perché  $1000=10^3$ , e nella numerazione binaria si usa 1024 perché  $1024=2^{10}$ .

### Mega

Per il mega valgono le stesse considerazioni del kilo: noi intendiamo un mega pari a 1000 kilo, in informatica invece un mega è pari a 1024 kilo.

### Giga

Equivale, in informatica, a 1024 mega.

### Tera

Equivale, in informatica, a 1024 giga.

Molti tendono a trascurare la differenza tra 1000 e 1024. Tuttavia la ragione alla base di tale differenza non può passare inosservata.

Poiché un bit è l'elemento minimo di informazione, non possono esistere i sottomultipli dei bit, come ad esempio il millibit.

## Gli esecutori

Vediamo ora di definire chiaramente che cosa è un elaboratore elettronico: per farlo è necessario introdurre il concetto di esecutore.

Nei fatti, per molti di noi un qualsiasi elaboratore elettronico (ad esempio lettore di CD-audio, computer, televisione o telefono cellulare) non è altro che una scatola nera, alla quale noi forniamo dei dati in ingresso e questa produce dei dati in uscita<sup>4</sup>: altro non ci interessa. Se vogliamo ascoltare la musica dobbiamo inserire nel lettore il CD musicale e premere il tasto play: in risposta avremo la musica.

Un qualsiasi entità che, presi degli input, produce degli output noi la chiamiamo **esecutore**.

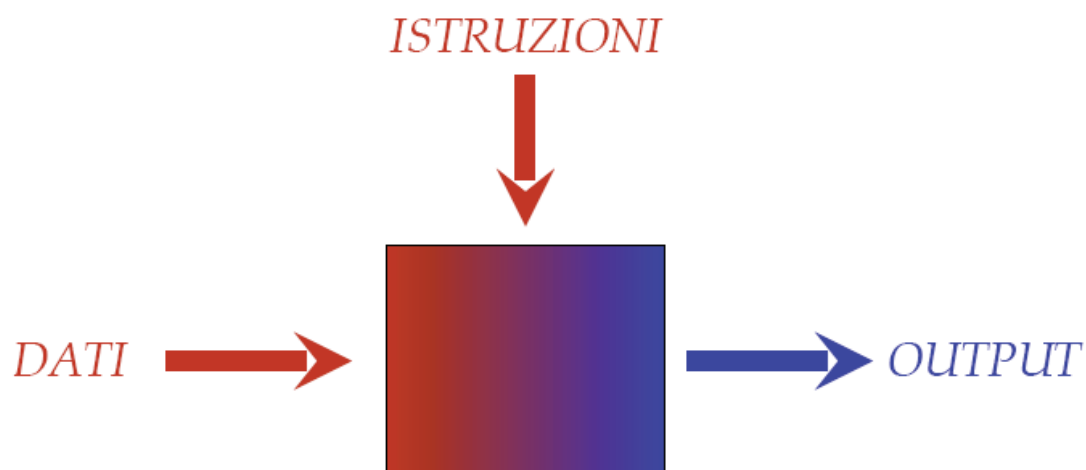
In prima istanza, quindi, un elaboratore elettronico è un esecutore.

Un esecutore, in generale, ha bisogno di due tipi di input: i **dati** e le **istruzioni**.

I **dati** sono le informazioni che vogliamo siano elaborate, le **istruzioni** sono i comandi che vogliamo vengano eseguiti sui dati per produrre gli output. Notiamo subito che esistono alcune istruzioni che siamo noi a fornire (ad esempio la pressione del tasto play per iniziare la riproduzione della musica) ed altre istruzioni che fornisce il produttore del dispositivo (tutte quelle istruzioni che consentono, ad esempio, di leggere un CD o di trasformare il suo contenuto in musica).

---

4 Nel seguito useremo il termine input come sinonimo di ingresso (ne è la traduzione inglese) ed il termine output come sinonimo di uscita (anche questo è la traduzione in inglese).



**Figura 4: un esempio di esecutore**

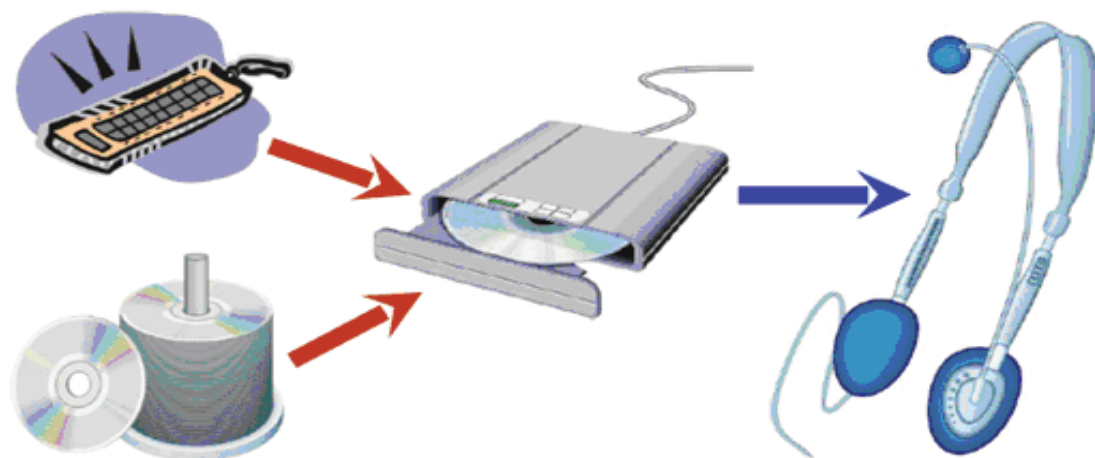
Proviamo a pensare ad un semplice esecutore: una macchina calcolatrice da tavolo. Per fare una semplice somma noi digitiamo  $3+2$  e lei risponde con 5. Per quanto detto 3, +, 2 sono gli input mentre 5 è l'output.

Tra gli input, 2 e 3 sono i dati e + è l'istruzione che la calcolatrice deve applicare sui dati per produrre l'output. Le istruzioni necessarie alla calcolatrice per operare sui numeri e calcolare la somma le ha inserite il produttore.

L'esempio della calcolatrice ci consente rispondere ad una domanda: a cosa servono i computer? Se devo dargli i dati da trattare e le istruzioni che voglio lui esegua, non posso fare tutto da solo? In teoria sì, ma il computer è molto più veloce di noi (generalmente) e molto più preciso. Potremmo così definire il computer come uno stupido velocissimo, al quale ci rivolgiamo per risolvere problemi che, altrimenti, richiederebbero troppo tempo per essere risolti (oppure che sarebbe facile, da soli, sbagliare). Se tale definizione può lasciare perplesso il lettore, vista soprattutto la pervasività dei computer nella nostra società, non si dimentichi la capacità dell'uomo di pensare, progettare, realizzare ed utilizzare tali strumenti. In questo modo sarà evidente che ogni elaboratore elettronico è solo uno strumento, nelle mani dell'uomo, per svolgere un lavoro.

Vediamo ora altri esempi di esecutori.

Nel caso in cui la scatola nera sia il lettore di CD, i dati in ingresso sono il CD-audio le istruzioni provengono, ad esempio, dalla pressione dei pulsanti per il play. Il lettore produce, in uscita, il suono della musica.



**Figura 5: il CD audio come esecutore**

Nel caso in cui l'esecutore sia il telefono cellulare il discorso non cambia, aumenta solo il numero delle funzioni e la loro complessità.

Proviamo a vedere, in generale, il funzionamento del telefono cellulare durante una telefonata; identifichiamo quindi i diversi dispositivi di input e di output. Sono dispositivi di input la tastiera, il microfono e l'antenna. I primi due servono per ricevere informazioni dall'utente, mentre l'ultimo per ricevere input dalle antenne della rete GSM. Sono invece dispositivi di output il display, la cassa e l'antenna, ancora una volta i primi due servono per dare output

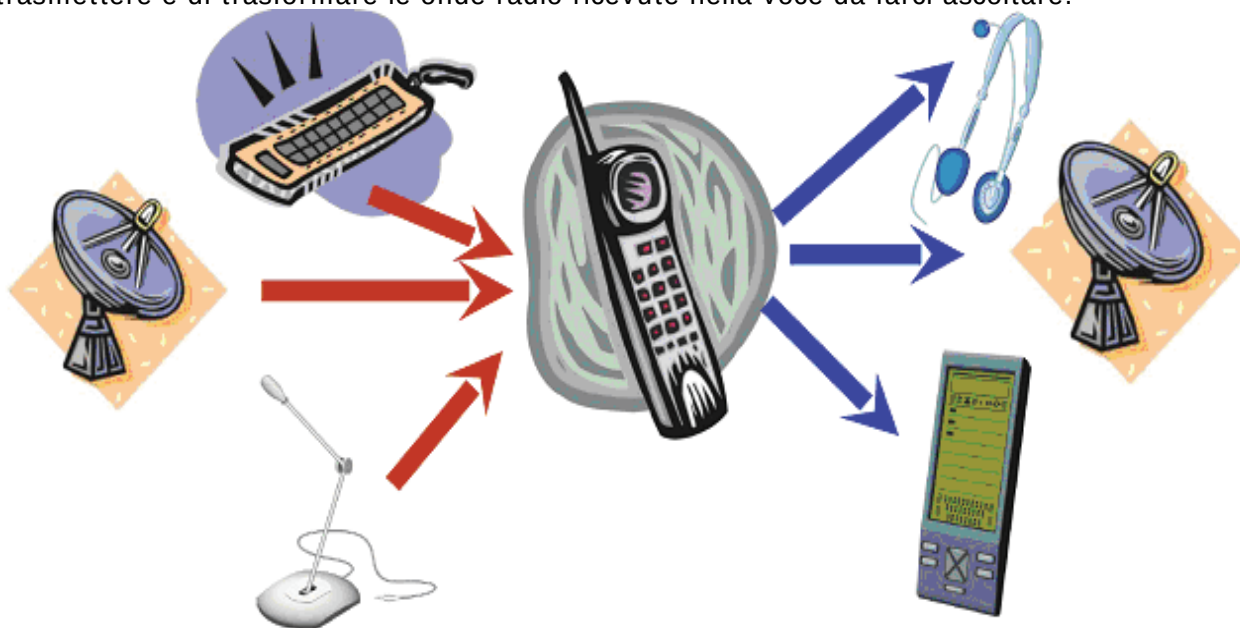
all'utente, mentre l'ultimo per trasmettere informazioni alla rete.

Mentre facciamo il numero telefonico, il cellulare accetta in ingresso i dati attraverso la pressione dei tasti corrispondenti al numero che desideriamo chiamare; in uscita mostra i numeri selezionati sul display.

Al comando di inizio della telefonata, ricevuto in ingresso dalla tastiera, il telefono inizia la comunicazione emettendo, in uscita, un segnale radio verso l'antenna GSM.

Durante la telefonata il telefono riceve in ingresso la nostra voce attraverso il microfono e le onde radio provenienti dall'antenna e che contengono la voce del nostro interlocutore; d'altronde produce in uscita la voce del nostro interlocutore dalla cassa, le onde radio verso l'antenna con la nostra voce ed eventuali messaggi sul display.

Durante la telefonata il telefono si occupa di trasformare la nostra voce in onde radio da trasmettere e di trasformare le onde radio ricevute nella voce da farci ascoltare.



**Figura 6: il telefono cellulare come esecutore**

Nel caso del PC, la complessità è ancora maggiore, poiché le funzionalità sono elevate, così come i dispositivi di ingresso ed uscita.

Nella definizione data ora di elaboratore rientra anche l'uomo. Ognuno di noi, infatti, riceve degli input attraverso i sensi, li elabora utilizzando il cervello e produce degli output, ad esempio parlando o scrivendo.

L'insieme degli esecutori è molto vasto, e contiene molti più elementi dei semplici elaboratori elettronici. È necessario, quindi, specificare bene a quale sottoinsieme di esecutori rientrano i computer.

I calcolatori sono un supporto per ricercare od esprimere le soluzioni di un problema. Sono, come detto, rapidissimi esecutori di ordini forniti dall'uomo. L'operazione di ricercare l'elenco degli ordini è una peculiarità dell'uomo.

Il calcolatore elettronico, quindi, non è necessario che conosca gli scopi per cui elabora i dati, né che abbia alcuna libertà di esecuzione: deve semplicemente limitarsi a svolgere le operazioni indicate. Se avesse la necessità di conoscere gli scopi per i quali svolge un'azione, necessariamente avrebbe le capacità di scegliere quale azione è migliore per raggiungere gli scopi: ma questo è compito del risolutore.

Tale sottoinsieme di esecutori vengono chiamati "**generici esecutori**". Osserviamo che anche un essere umano può essere un generico esecutore: chiunque batta a macchina una lettera scritta da un'altra persona è un esempio di generico esecutore. D'altronde un medico che visita i suoi pazienti non può essere un generico esecutore, poiché conosce gli scopi per i quali svolge il suo lavoro e deve interessarsi della salute del suo paziente nel complesso. Questo lo porta a decidere, sulla base della sua esperienza, la diagnosi e la cura. Ogni medico ha la libertà di esecuzione, ciò che differenzia i diversi medici è la capacità di interpretare più o meno correttamente gli input per produrre gli output (diagnosi e/o cura) più idonei.

## Gli algoritmi

Come abbiamo visto, un elaboratore elettronico è un generico esecutore che può ricevere in input i dati e le istruzioni, elaborare – sulla base delle istruzioni fornite – tali dati e produrre in uscita degli output. Per svolgere tali attività, egli non conosce gli scopi per i quali gli viene richiesto di trattare le informazioni e non ha alcuna libertà di interpretazione nel suo lavoro.

Domandiamoci quali devono essere i requisiti di un insieme di istruzioni (che chiameremo algoritmi) da fornire ad un generico esecutore per metterlo in condizione di trasformare input in output senza consentirgli alcuna libertà di interpretazione.

In generale, perché il risolutore possa descrivere all'esecutore la procedura da seguire per giungere alla soluzione di un problema, tale problema si scompone in sottoproblemi, eventualmente ancora scomponibili.

La scomposizione si ferma quando il risolutore arriva a trattare problemi elementari (o primitivi), non ulteriormente scomponibili, per la cui soluzione l'esecutore deve svolgere una o più azioni elementari.

Risolvere un problema equivale, così, ad eseguire una opportuna successione di problemi più semplici: le azioni elementari.

Nell'affermazione precedente, è stato volontariamente separato l'obiettivo della risoluzione di un problema complesso (tipico del risolutore) dall'attività di esecuzione di una opportuna sequenza di azioni elementari (compito dell'esecutore e, quindi, del computer).

**Gli algoritmi possono essere definiti come una sequenza di azioni che trasforma dei dati iniziali nel risultato finale attraverso un numero finito di passi elementari e non ambigui.**

Rispetto ad una generica sequenza di azioni l'algoritmo si distingue, già nella definizione, per la non ambiguità delle singole azioni elementari.

Quali sono i requisiti che devono avere questi algoritmi?

I primi due requisiti sono evidenti: devono operare su dei dati in ingresso e produrre dei dati in uscita. Una sequenza di istruzioni che prendesse dei dati in ingresso per non produrre nulla in uscita sarebbe inutile, come una sequenza di istruzioni che producesse degli output senza ricevere in ingresso alcun input. Il requisito di operare su dati in ingresso è detto **input**, quello di produrre dati in uscita è detto **output**.

Notiamo inoltre che, per poter produrre degli output, tutte le operazioni si devono concludere in un tempo finito poiché altrimenti il risultato, per quanto potenzialmente producibile, non sarebbe mai utilizzabile dal risolutore. Questo requisito è detto **effettività**.

Un altro requisito riguarda il numero di istruzioni, o di passi elementari, che devono essere svolti prima di poter produrre un output: il requisito della **terminazione** impone che tale numero sia finito (e non infinito) altrimenti ancora una volta, non si avrebbe alcun risultato in output. Una sequenza di istruzioni che descrivesse ad un esecutore come approssimare un cerchio utilizzando un poligono con un numero crescente di lati, non potrebbe essere un algoritmo poiché richiederebbe un tempo infinito di passi elementari per essere portata a termine.

I requisiti sino a qui introdotti sono abbastanza comuni: ogni persona di buon senso nel definire una sequenza di istruzioni si adatterebbe a questi requisiti. Tuttavia tale sequenza può essere ambigua. Il requisito che manca è l'**univocità**: non ci deve essere alcuna libertà di azione da parte dell'esecutore, altrimenti tale esecutore di istruzioni non potrebbe essere generico.

Un insieme di istruzioni che verifica questi requisiti è detto algoritmo e può essere svolto dal generico esecutore. Compito di ognuno di noi quando, come risolutore, fornisce input ad un generico esecutore, è quello di assicurarsi che le istruzioni verifichino questi semplici requisiti.

L'algoritmo non è una novità legata ai calcolatori elettronici: il termine algoritmo è legato all'ultima parte del nome del matematico persiano *Abu Ja'far Mohammed ibn Mûsâ al-Khawârizmî*, vissuto intorno al 850 dc e che, nel libro *Kitâb al-djabr wa 'l-muqâbala* (*Libro sulla ricomposizione e sulla riduzione*), sembra sia stato il primo a fare riferimento alla scomposizione dei problemi in passi elementari. Sempre da tale libro prende origine la parola algebra (dalla parola araba *al-djabr*).

Definito un algoritmo, è necessario domandarsi sotto quali ipotesi l'esecutore è in grado di eseguire algoritmi.

Affinché il generico esecutore sia in grado di svolgere correttamente l'algoritmo egli deve:  
capire le singole istruzioni fornite  
essere in grado di eseguire tali istruzioni  
essere abile a trattare i dati che riceve.

Come visto, un algoritmo è un insieme di azioni elementari, nasce quindi spontaneo domandarsi quali devono essere le proprietà di queste azioni perché siano rispettati i requisiti dell'algoritmo.

Tali proprietà sono tre.

La prima proprietà è detta **finitezza**: un'azione elementare deve poter essere conclusa in un tempo finito, altrimenti l'algoritmo richiederebbe un tempo infinito per essere completato, violando così il requisito dell'effettività.

La seconda proprietà, detta dell'**osservabilità**, richiede che ogni azione elementare deve produrre un risultato osservabile, cioè deve produrre un output parziale e funzionale al raggiungimento dell'output complessivo dell'algoritmo. Un'azione elementare che viola la proprietà dell'osservabilità è inutile e quindi non deve stare all'interno di un algoritmo.

La terza proprietà è la **riproducibilità**: se il generico esecutore ripete più volte lo svolgimento di un'azione elementare deve ottenere ogni volta, a parità di input, i medesimi output: altrimenti il generico esecutore avrebbe una qualche libertà d'azione, violando l'univocità dell'algoritmo. Se, ad esempio, chiediamo ad un generico esecutore di calcolare la somma di due numeri, diciamo 5 e 3, egli risponderà con 8. La riproducibilità impone che, chiedendogli più volte di calcolare la somma di 5 e 3, egli risponda sempre con 8: altrimenti ci sarebbe una libertà di azione da parte dell'esecutore, poiché a parità di input (3,5,+) si produrrebbero output diversi.

Vediamo ora alcuni semplici esempi di algoritmi, cercando di evidenziare le caratteristiche descritte precedentemente.

## Esempio 1

Proviamo a costruire un algoritmo che svolga una moltiplicazione tra due numeri interi, come ad esempio  $5 \cdot 8$  oppure  $132 \cdot 439$ .

Per far questo, supponiamo che il generico esecutore sappia solo fare le somme.

Sappiamo che moltiplicare due numeri  $n$  ed  $m$ , equivale a sommare  $m$  volte  $n$ . Se, ad esempio,  $m=8$  ed  $n=5$  si ha:

$$8 \cdot 5 = 8 + 8 + 8 + 8 + 8 = 40$$

Ora ci domandiamo: come può essere formalizzato l'algoritmo?

Per prima cosa il generico esecutore deve acquisire i dati in input, e verificare che i due numeri siano corretti (se i numeri fossero 4,3 e 2,4 non si potrebbe sommare 2,4 volte il numero 4,3).

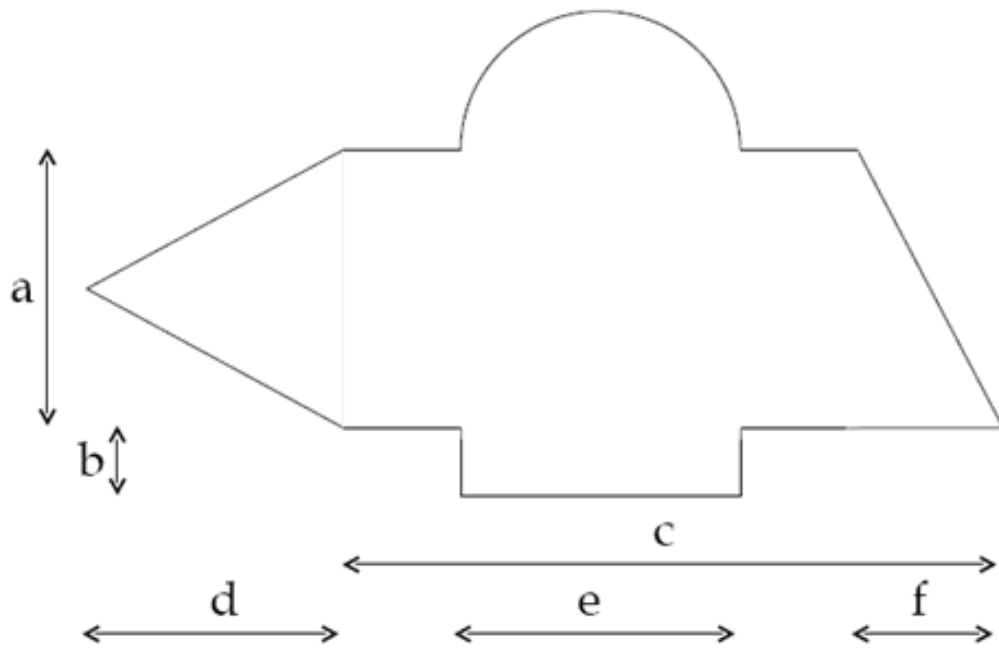
Supponiamo che i dati in input siano 8 e 5.

Il passo successivo è sommare 8 per 5 volte.

Al termine di questa operazione, il risultato della somma deve essere fornito in output.

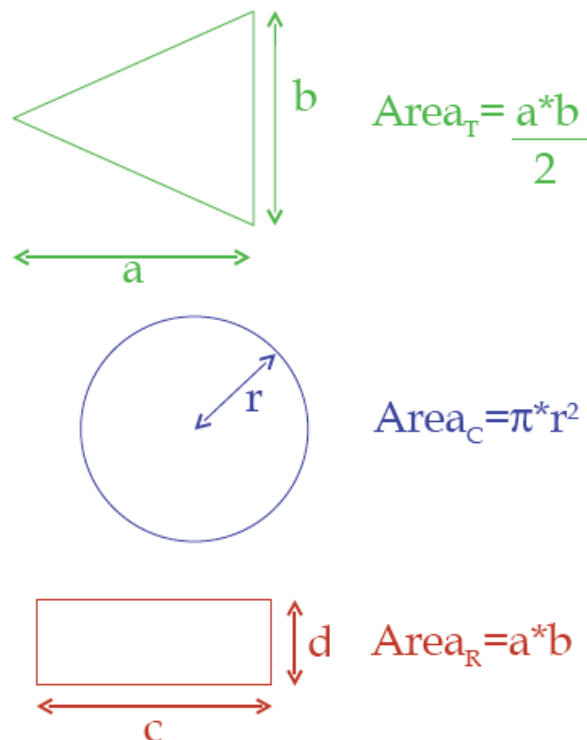
## Esempio 2

Proviamo ora a ricavare un algoritmo che ci consenta di calcolare l'area della Figura 7, supponendo che il generico esecutore sappia svolgere solo le quattro operazioni elementari.



**Figura 7: figura geometrica di cui calcolare l'area**

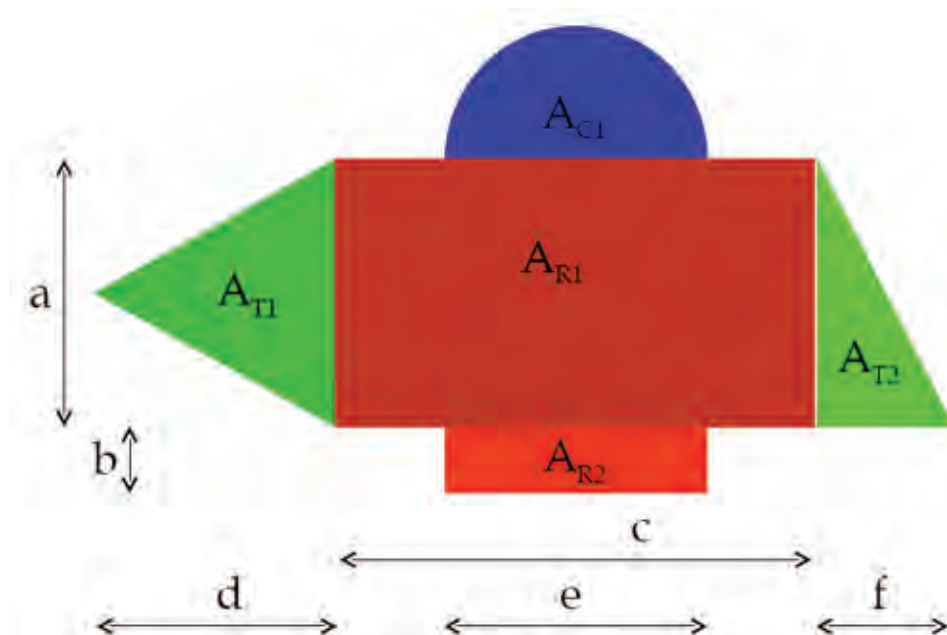
Come risolutori non conosciamo una formula per calcolare quest'area con una singola azione elementare, ma conosciamo le formule per calcolare l'area di porzioni della figura, come riportate in Figura 8.



**Figura 8: figure elementari**

Nostro compito è quello di trasformare il disegno presente in Figura 7 in un insieme di figure di cui conosciamo come calcolare l'area (quelle in Figura 8). In questo modo scomponiamo il problema (del calcolo dell'area della Figura 7), in sottoproblemi (il calcolo delle aree di Figura 8) risolvibili con semplici operazioni, note anche al generico esecutore.

La scomposizione della Figura 7 in una combinazione delle figure geometriche visibili in Figura 8 è riportata in Figura 9; non è certamente l'unica.



**Figura 9: figura base scomposta come somma di figure elementari**

Possiamo così indicare quale può essere l'algoritmo da seguire per il calcolo di quest'area.

Il primo passo consiste nel ricevere in ingresso le 6 lunghezze necessarie e verificare che tali grandezze siano corrette (ad esempio un lato deve avere necessariamente lunghezza maggiore di zero).

Successivamente, sulla base delle 6 lunghezze devono essere calcolate le aree delle singole figure elementari, nel caso in esame avremo:

$$Area_{T1} = \frac{a * d}{2}, Area_{T2} = \frac{a * f}{2}, Area_{R1} = a * c, Area_{R2} = b * e, Area_{C1} = \pi \left( \frac{e}{2} \right)^2$$

A questo punto l'esecutore può calcolare l'area della figura iniziale come somma delle singole figure elementari:

$Area_{FUGURA} = Area_{T1} + Area_{T2} + Area_{R1} + Area_{R2} + Area_{C1}$

L'algoritmo termina con la presentazione in output del valore di  $Area_{FUGURA}$ .

### Esempio 3

Sulla base di questi esempi sembra che qualsiasi sequenza di istruzioni possa essere considerata un algoritmo. In realtà la nostra vita è piena di esempi di sequenze ordinate di istruzioni che non sono algoritmi perché violano il principio dell'univocità. Addirittura il molti casi è proprio la nostra capacità di interpretare le istruzioni a dare il "colpo dell'artista".

Pensiamo ad una ricetta di cucina e vediamo perché non è un algoritmo, successivamente ci domanderemo come modificare la sequenza di istruzioni per trasformarla in un algoritmo.

Prendiamo ad esempio la ricetta della parmigiana di melanzane.

#### Ingredienti

4 melanzane grosse, olio per friggere e farina q. b., 500 gr di passata di pomodoro, mezza cipolla, un ciuffo di basilico, tre cucchiaini d'olio d'oliva, 125 gr di parmigiano grattugiato, 250 gr di mozzarella tagliata a fettine, sale e pepe q. b.

#### Esecuzione

Tagliare le melanzane a rondelle cospargerle con un pizzico di sale e lasciarle spurgare per un'ora. Asciugarle bene, infarinare, e friggerle in abbondante olio caldo. Sgocciolarle bene con carta da cucina. Intanto preparare una salsa con l'olio d'oliva, la cipolla, il pomodoro, ed il basilico. In una teglia adagiate strati di melanzane, sugo, parmigiano e mozzarella. Infornare a 180°C per 30 minuti circa.

Cominciamo a vedere quali requisiti dell'algoritmo sono verificati.

Sicuramente sono rispettati i requisiti di ingresso (dati: gli ingredienti, istruzioni: la ricetta) e di uscita (la parmigiana). Inoltre in circa due ore di lavoro l'output è prodotto, per cui anche

l'effettività è verificata. Indipendentemente dalle suddivisioni che ognuno può fare, il numero di passi elementari è sicuramente finito: nel seguito supporremo che l'algoritmo è composto di 7 passi.

Rimane solo da verificare il rispetto del requisito dell'effettività; per farlo suddividiamo l'algoritmo nei 7 passi che lo costituiscono:

1. reperire gli **ingredienti**.
2. tagliare le melanzane a **rondelle**, cospargerle con un **pizzico** di sale e lasciarle spurgare per un'ora.
3. asciugarle **bene**, **infarinarle**, e **friggerle** in **abbondante** olio **caldo**.
4. sgocciolarle **bene** con carta da cucina.
5. **preparare una salsa** con l'olio d'oliva, la cipolla, il pomodoro, ed il basilico.
6. adagiate in una teglia **strati** di melanzane, sugo, parmigiano e mozzarella.
7. infornare a 180°C per 30 minuti **circa**.

In rosso sono state riportate espressioni che indicano necessariamente una irriproducibilità delle singole azioni elementari, e quindi una violazione nel requisito dell'univocità. Tra gli ingredienti, infatti, sono riportate indicazioni come melanzane **grosse** o farina **quanto basta**, in entrambe i casi è l'esecutore che deve decidere se una melanzana è grossa oppure no, ovvero quanta farina è necessaria.

Anche nella seconda azione elementare la dimensione della rondella e la quantità di sale sono lasciati all'esperienza ed alle competenze dell'esecutore, che è quindi libero di scegliere e decidere in modo autonomo.

Notiamo, inoltre, che alcune azioni non sono elementari perché prevedono lo svolgimento di altri algoritmi. Ad esempio friggere o preparare una salsa sono algoritmi poiché per realizzarli è necessario suddividerli a loro volta in molte azioni elementari: l'esperienza dimostra, comunque, che tali sequenze di azioni violano l'univocità.

Una ricetta non è, quindi, un algoritmo. Possiamo lecitamente domandarci come deve essere modificata la ricetta per renderla un algoritmo.

Prendiamo la seconda azione elementare, dovremmo sicuramente definire lo spessore delle singole rondelle nonché la quantità di sale da utilizzare e la densità con cui deve essere cosperso. La semplice frase *tagliare le melanzane a rondelle, cospargerle con un pizzico di sale* diventerebbe così:

**tagliare le melanzane a rondelle spessore (0,75±0.05) cm e cospargere ogni rondella con 0.01 grammi di sale per grammo di melanzana. Il cospargimento deve avvenire su entrambe le superfici recise e deve avere una disomogeneità massima del 10%.**

Espressa tutta la ricetta in questi termini anche un generico esecutore sarebbe in grado di cucinare la parmigiana di melanzana per noi. Anche se una formalizzazione così spinta del problema può sembrare eccessiva, nella produzione industriale di beni alimentari non possiamo pensare ci siano batterie di cuochi pronti a cucinare per noi.

## **La macchina di Von Neumann**

Ora che abbiamo chiaro come formalizzare la soluzione di un problema per consentirne l'esecuzione anche ad un generico esecutore, possiamo domandarci - in linea assolutamente generale - quali elementi debbano comporre il generico esecutore per consentirgli di eseguire correttamente gli algoritmi.

Anche se nel corso degli ultimi decenni la tecnologia si è molto sviluppata, realizzando generici esecutori in grado di svolgere praticamente qualsiasi operazione, il modello di base non si è molto evoluto dalla sua prima definizione effettuata da John Von Neumann negli anni 40 dello scorso secolo.

Lo schema che seguiremo è detto modello della **macchina di Von Neumann**: una descrizione funzionale e strutturale del generico esecutore. Questo significa che la macchina di Von Neumann racchiude le caratteristiche essenziali di un generico calcolatore, senza entrare nel merito dei singoli aspetti implementativi: non esiste un prototipo della macchina di Von Neumann in un qualche museo, ma tutti i nostri infodomeistici - dal computer al cellulare al televisore - sono stati realizzati prendendo a modello la macchina di Von Neumann.

Sino ad ora abbiamo trattato il generico esecutore come una sorta di scatola nera: noi gli forniamo in ingresso i dati e le istruzioni e lui ci risponde con gli output. Supponiamo di poter togliere il coperchio alla scatola nera e vediamo di cosa è composta, per farlo consideriamo l'algoritmo definito nel precedente esempio 2.

Per prima cosa il risolutore deve fornire al generico esecutore i dati e le istruzioni su cui

operare: la caratteristica di lavorare su dati in ingresso è proprio una delle proprietà dell'algoritmo. Il generico esecutore, quindi, dovrà disporre di alcuni strumenti per consentire al risolutore di inserire i dati in ingresso. Questi strumenti sono detti **dispositivi di input** della macchina di Von Neumann. Questa non è una novità: quando, ad esempio, utilizziamo il telecomando del televisore stiamo fornendo dati ed istruzioni al nostro infodomeistico utilizzando un dispositivo di input.

I dispositivi di input non hanno memoria delle informazioni acquisite e, per tale motivo, devono trasferire tali input ad un altro dispositivo, che ha il compito di memorizzare i dati e le istruzioni durante lo svolgimento dell'algoritmo. Il dispositivo che ha il compito di memorizzare i dati e le istruzioni ricevuti dalle periferiche di input è la **memoria** della macchina di Von Neumann.

Al termine della fase di acquisizione dei dati, nella memoria della macchina di Von Neumann sono quindi presenti i valori di a, b, c, d ed e, nonché le istruzioni dell'algoritmo. Per poter eseguire l'algoritmo è necessario che la macchina di Von Neumann disponga di un dispositivo in grado di comprendere le istruzioni presenti nella memoria e di coordinare lo svolgimento delle diverse attività richieste per l'esecuzione dei singoli passi elementari. Questa funzione è svolta dall'**unità di controllo** (o CU: Control Unit) della macchina di Von Neumann. La CU svolge una funzione di controllo e coordinamento delle attività, ma non è in grado di eseguire le operazioni matematiche necessarie per effettuare il calcolo delle aree richiesto dall'algoritmo. Il suo compito è quello di interpretare l'istruzione

$$a * c$$

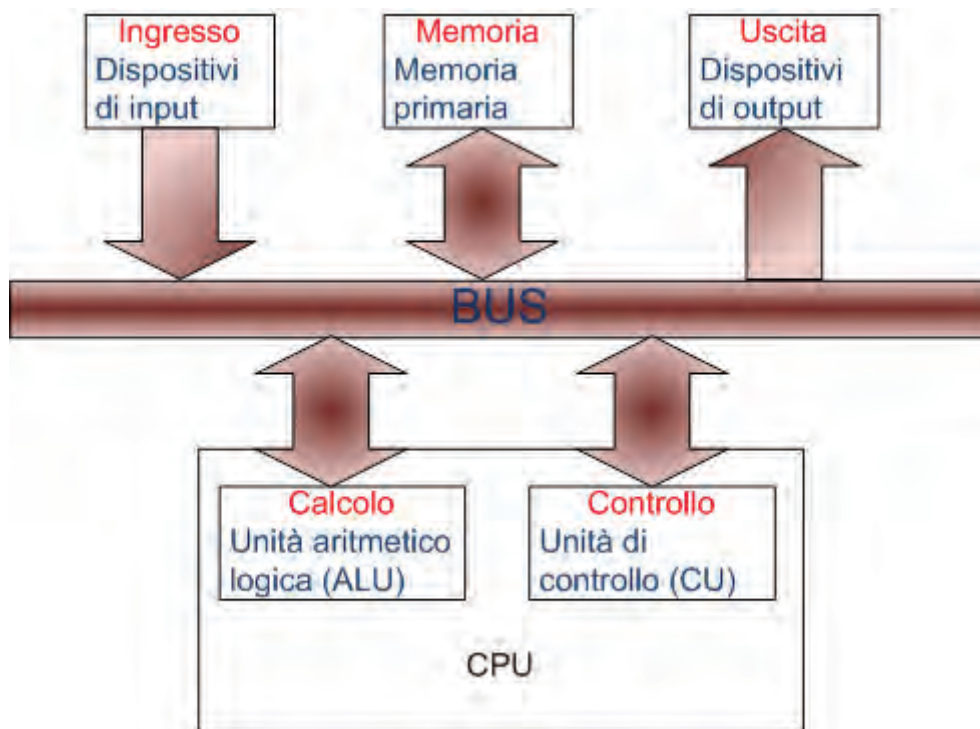
e comandare ad un dispositivo deputato al calcolo matematico di effettuare la moltiplicazione tra quei due numeri presenti in memoria. Il dispositivo deputato allo svolgimento dei calcoli matematici è l'**Unità Aritmetico Logica** (ALU: Arithmetic and Logic Unit) della macchina di Von Neumann. La UC per ogni calcolo indica quindi alla ALU quali dati prendere dalla memoria e quali operazioni svolgere su quei dati. Una volta effettuate le singole azioni elementari di calcolo delle aree, il risultato viene salvato nuovamente nella memoria, così da essere utilizzato nei successivi passi elementari, sino al calcolo di Area<sub>FUGURA</sub> come somma dei valori delle aree elementari conservati in memoria.

Al termine delle operazioni, il valore di Area<sub>FUGURA</sub> deve essere reso disponibile al risolutore attraverso un dispositivo che consente di comunicare al risolutore l'output dell'algoritmo. Tale dispositivo è detto **dispositivo di output** della macchina di Von Neumann.

Per poter interagire, i diversi dispositivi devono essere collegati tra loro attraverso un dispositivo in grado di abilitare lo scambio delle informazioni. Tale dispositivo è il **BUS**.

In Figura 10 è visibile uno schema della macchina di Von Neumann

Al termine dell'esecuzione dell'algoritmo la memoria viene svuotata, così da poter contenere gli input di un nuovo algoritmo.



**Figura 10: schema della macchina di Von Neumann**

Nel corso dei decenni, da un punto di vista tecnico sono state trovate molte soluzioni per massimizzare l'efficienza di un generico esecutore, ma lo schema generale non si è discostato molto da quello definito da Von Neumann. Ad esempio la ALU e la CU sono realizzate in un unico dispositivo fisico: la CPU (Central Control Unit) e, sono disponibili computer con molte CPU distinte; questa scelta ingegneristica migliora le prestazioni dell'esecutore ma non cambia il modello. Un'altra scelta ingegneristica finalizzata a migliorare le prestazioni è la moltiplicazione dei bus: nei computer di ultima generazione ci sono più di 10 tipi di bus diversi, ognuno finalizzato a migliorare l'efficienza di uno specifico scambio di dati.

## La memoria

Un'altra apparente differenza tra i computer moderni e la macchina di Von Neumann riguarda la memoria; parlando della memoria della macchina di Von Neumann abbiamo detto che serve per memorizzare i dati e le istruzioni necessari per lo svolgimento dell'algoritmo e che, al termine dell'esecuzione, i dati vengono cancellati dalla memoria: questo sembra in contrasto con il funzionamento dell'hard disk, il più noto dispositivo di memorizzazione utilizzato su un PC, che conserva le informazioni per un tempo arbitrariamente lungo. In realtà non c'è alcun problema, nei moderni esecutori con il termine memoria si intendono due distinte famiglie di dispositivi di memorizzazione: le **memorie primarie** e le **memorie secondarie**.

## Le memorie primarie

I dispositivi di memorizzazione previsti da Von Neumann nel suo modello, nei moderni calcolatori rientrano nella famiglia delle **memorie primarie** o **principali**. Poiché questi dispositivi servono a memorizzare i dati e le istruzioni necessarie per lo svolgimento degli algoritmi, è importante che siano dispositivi veloci nel memorizzare i dati (scrittura) e nel riprodurli (lettura), così da velocizzare lo svolgimento degli algoritmi (le moderne memorie primarie hanno velocità superiori al 1 GB/s). Inoltre, poiché in caso di assenza di corrente l'elaboratore non può funzionare, non è necessario che la memoria sia in grado di memorizzare le informazioni in modo permanente anche in assenza di corrente; le memorie che non conservano i dati in caso di mancanza della corrente sono dette memorie **volatili**. Infine, poiché la memoria primaria deve conservare solo i dati e le istruzioni per il tempo strettamente necessario allo svolgimento degli algoritmi, non è necessario che l'elaboratore disponga di una memoria di elevate dimensioni (le memorie primarie dei computer sono dell'ordine di 1 GB). Nei computer esistono due memorie primarie distinte: la prima è la RAM (Random Access Memory), la seconda è la cache. La cache è una memoria molto piccola (circa 1 MB) ma molto veloce che si trova all'interno della CPU insieme all'ALU ed alla CU: l'obiettivo di utilizzare la

cache è di velocizzare ulteriormente l'esecutore, senza modificare il modello della macchina di Von Neumann.

## Le memorie secondarie

Nel modello della macchina di Von Neumann, non viene fatta alcuna distinzione se i dati in input sono inseriti manualmente da un utente (ad esempio attraverso la tastiera) o sono acquisiti da un dispositivo (ad esempio un CD-Rom). Questa distinzione per noi utenti è fondamentale: se così non fosse, ad esempio, ogni volta che volessimo lavorare su un documento dovremmo riscriverlo da capo. Per ovviare a questo problema, esistono dei dispositivi in grado di conservare le informazioni e fornire in input all'elaboratore elettronico evitandoci l'onere di doverle inserire ogni volta. I dispositivi che svolgono questa funzione rientrano nella famiglia delle **memoria secondaria** o di **massa**. È evidente quindi che, nel modello della macchina di Von Neumann, questi dispositivi sono dispositivi di input e di output. Vediamo ora quali sono le principali caratteristiche di queste memorie. In primo luogo devono conservare le informazioni anche in assenza di corrente elettrica; le memorie che conservano i dati in caso di assenza della corrente sono dette memorie **permanenti**. In secondo luogo, poiché le informazioni sono trasferite dalla memoria secondaria a quella primaria (e viceversa) solo all'inizio ed alla fine dell'esecuzione dell'algoritmo, possono essere dispositivi lenti (le moderne memorie primarie hanno velocità inferiori al 10 MB/s). Poiché possono contenere i dati e le istruzioni relativi a molti algoritmi, oltre a conservare molte altre informazioni (ad esempio i dati prodotti dall'utente) è bene che le memorie secondarie siano di grandi dimensioni (le memorie primarie dei computer sono dell'ordine di 1000 GB).

Esempi di memorie secondarie presenti all'interno dei computer sono l'hard disk ed il DVD.

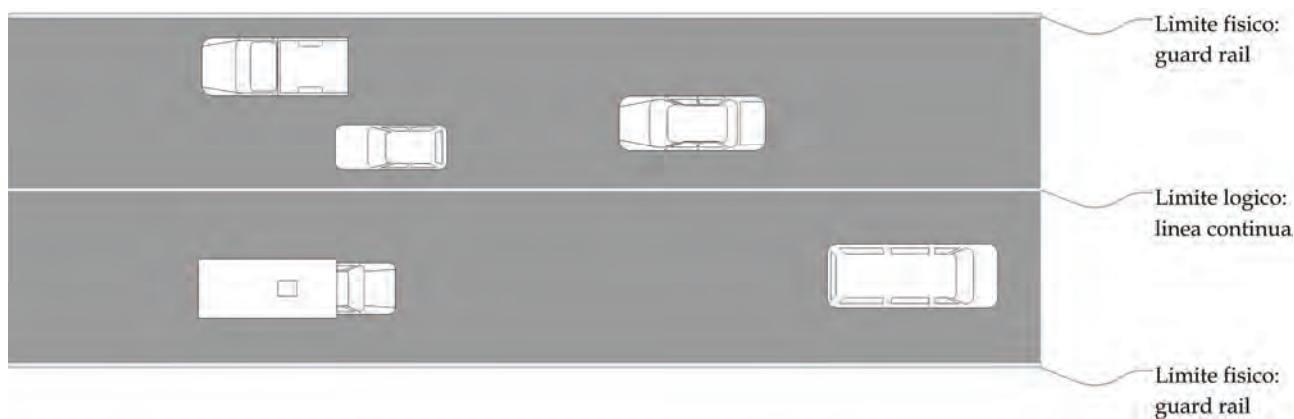
Potremmo domandarci perché non sono state realizzate delle memorie primarie veloci, permanenti e di elevate dimensioni così da rendere inutile la distinzione tra memorie primarie e secondarie. La ragione è che memorie permanenti, veloci come le primarie ma grandi come le secondarie ancora non sono state inventate. Si sta lavorando in questa direzione ma la strada sembra ancora lunga.

Inoltre, a parità di dimensione, più una memoria è veloce più costa. A parità di quantità di memoria, la RAM costa circa 1000 volte di più di un hard disk.

## La conservazione delle informazioni nelle memorie secondarie

Le memorie secondarie consentono la conservazione di dati e istruzioni per un tempo arbitrariamente lungo. In particolare conservano non solo le informazioni necessarie al computer per funzionare (ad esempio per consentire al computer di eseguire un programma), ma anche i dati che noi utenti produciamo durante il normale utilizzo come documenti, foto o canzoni. Poiché le informazioni sono conservate sui dispositivi di memorizzazione in forma binaria, è molto difficile per l'utente definire quale bit afferisce ad un particolare documento e quale ad un altro. Per tale motivo i bit afferenti allo stesso insieme di dati o istruzioni sono raggruppati in **file**. Un file (archivio in inglese) è quindi una sequenza ordinata di dati immagazzinati in formato digitale. I file sono i contenitori logici nei quali sono conservate le informazioni all'interno dei supporti di memorizzazione di massa.

Cosa significa contenitore logico? L'hard disk o il CD-Rom, così come un bicchiere od una strada, sono esempi di **contenitori fisici**: le automobili sono vincolate a restare all'interno della strada dal guard rail. Se la strada è a più carreggiate, spesso la separazione tra le due carreggiate è effettuata con una linea continua. Tale linea non costituisce un vincolo fisico ma **logico**, poiché è un vincolo che l'utente della strada sa di non dover superare anche se può agevolmente superarlo. In Figura 11 è visibile una rappresentazione grafica di quest'esempio.



**Figura 11: una strada come esempio contenitore**

Nel caso dei file presenti su un dispositivo di memorizzazione secondario vale lo stesso principio: i bit afferenti ai diversi file sono presenti sul dispositivo in modo sequenziale, ma l'elaboratore sa distinguere a quale file afferisce un singolo bit.

Per distinguere un file da un altro presente all'interno di un dispositivo di memorizzazione di massa si usa assegnare ad ogni file un nome univoco che lo distingue dagli altri.

Generalmente il nome del file ha la forma: **nomechepreferiamo.estensione**

A sinistra del punto c'è il nome del file che, a giudizio di chi lo ha creato, garantisce i migliori risultati mnemonici.

A destra invece c'è l'estensione che indica all'utente ed al elaboratore il formato dei dati contenuti nel file.

Che cosa significa il **formato** dei dati contenuti nel file? Poiché un file è una sequenza di bit, a priori non sappiamo se una certa sequenza di bit corrisponde ad una immagine, ad una canzone o ad un documento di testo.

I bit che costituiscono un file sono collegati alle informazioni in esso contenute: ad ogni bit è associato uno specifico significato, che dipende dal tipo di contenuto conservato nel file. Poiché i bit di un file possono rappresentare molti contenuti diversi, è necessario definire una convenzione che associ ad ogni bit un significato.

Il formato di un file è la convenzione usata per leggere, scrivere ed interpretare i contenuti del file.

Senza conoscere l'estensione del nome dei file, e quindi senza conoscere il formato con cui sono codificate le informazioni all'interno del file non sappiamo quale programma dobbiamo utilizzare per leggere i bit contenuti.

Attraverso l'estensione possiamo risalire a questa informazione. Se ad esempio il nome del file ha estensione txt, il contenuto sarà un testo codificato secondo la tabella ASCII; se invece il nome del file ha estensione doc, allora il testo sarà codificato per essere letto con Microsoft Word. I sistemi operativi che utilizzano interfacce grafiche (ad esempio Windows, Linux o Mac), per rendere più immediata la comprensione del formato del file, utilizzano icone diverse a seconda dell'estensione nel nome del file. Alcuni sistemi operativi tendono a nascondere l'estensione del file per aiutare gli utenti meno esperti: se non trovate l'estensione di un file vicino al suo nome, fatevi guidare dall'icona.

### 3 Le reti

#### **Presentazione**

In questa unità didattica saranno introdotti i principali concetti relativi alle reti di trasmissione dei dati, con particolare riferimento alla rete internet.

Gli obiettivi dell'unità sono da un lato quello di far apprendere i fondamenti di funzionamento di una rete, dall'altro quello di presentare i principali servizi disponibili sulla rete internet e gli strumenti per poterne fruire appieno.

Data la crescente importanza che le reti di elaboratori stanno assumendo nella nostra società, diventa sempre più rilevante comprendere in dettaglio il funzionamento di questi strumenti, per poter sfruttare tutte le possibilità offerteci.

#### **Introduzione**

I computer consentono da sempre di eseguire moltissime attività. Nel corso degli anni è emersa la necessità di scambiare informazioni tra computer diversi, ad esempio perché non tutti i computer disponevano di una stampante. Tale scambio è iniziato utilizzando dispositivi removibili di memorizzazione di massa (ad esempio il floppy disk): sul floppy venivano salvate le informazioni da trasferire ad un altro computer.

In seguito i computer sono stati connessi fisicamente l'uno all'altro usando dei cavi, in questo modo tutti i computer presenti all'interno di una stanza potevano scambiarsi dati: erano connessi all'interno di una rudimentale rete. Successivamente si cominciarono a connettere insieme i computer di uffici all'interno di enti, poi i computer degli enti tra di loro e con gli utenti, arrivando all'internet come la conosciamo oggi. Questo processo è stato così rivoluzionario che oggi è difficile pensare ad un computer non connesso ad una rete anche se in modo non permanente (ad esempio attraverso un modem).

La domanda da porsi è: perché le reti di computer sono diventate così importanti?

Le reti trasportano informazioni sotto forma di bit da un computer ad un altro. Negli ultimi anni la nostra società si sta evolvendo da una società basata sugli oggetti (tipico delle società agricole ed industriali) verso una società basata sulle informazioni: il bene primario non è più fatto di materia (come ad esempio un oggetto prodotto in fabbrica), ma di informazioni (come una notizia in televisione od una pagina web).

Le informazioni aggiornate, disponibili in ogni luogo ed in ogni momento rappresentano un bene che siamo disposti a pagare a caro prezzo: le reti di computer sono uno strumento sempre più rilevante attraverso cui queste informazioni viaggiano.

Il punti di forza delle reti di computer, per ciò che concerne lo scambio di informazioni, sono di seguito descritti.

In **primo luogo** tempi di trasferimento trascurabili; anche se – in caso di file molto pesanti o di connessioni particolarmente lente – il trasferimento dei dati può richiedere qualche minuto, comunque tale valore è trascurabile rispetto a quanto si deve aspettare utilizzando, ad esempio, la posta. Anche utilizzando il fax è possibile scambiare messaggi con tempi di trasferimento trascurabili; il fax, infatti, è basato su una delle prime reti per la trasmissione delle informazioni: la rete telefonica.

In **secondo luogo** lo scambio di informazioni avviene in formato digitale: ciò consente al destinatario di trattare le informazioni ricevute con la stessa semplicità con la quale tratterebbe informazioni da lui prodotte. Utilizzando altre reti, come ad esempio la rete telefonica quando si invia un fax, il destinatario riceve una copia fotostatica del documento inviato e – per poterla modificare – la deve ridigitare. Al contrario, se un utente invia ad un altro un file in formato digitale (ad esempio un file Word in attach via mail), il destinatario può modificarlo facilmente. Utilizzando la posta tradizionale è possibile scambiare informazioni tanto in formato digitale (ad esempio inviando un CD-Rom), quanto in formato analogico (inviando una lettera).

In **terzo luogo** il costo per inviare delle informazioni è, utilizzando una rete di computer, molto basso: infatti i costi connessione non dipendono dalla quantità di informazioni che deve essere scambiata: gli utenti pagano una quota fissa e possono restare collegati quanto vogliono e scambiare (le così dette tariffe flat). Utilizzando la posta tradizionale il pagamento viene effettuato per ogni singolo invio ed il costo, ad esempio per una lettera, è superiore all'euro. Nel caso del fax, invece, si paga per il tempo nel quale si impegna la linea: il costo dipende strettamente dalla quantità di informazioni da inviare.

In Tabella 5 è visibile una tabella riassuntiva dei punti di forza delle reti di computer rispetto ad altri tipi di reti.

|                    | Tempi di trasferimento | Formato delle informazioni | Costo dell'invio |
|--------------------|------------------------|----------------------------|------------------|
| Computer           | Trascurabili           | Digitale                   | Trascurabile     |
| Fax                | Trascurabili           | Analogico                  | Dipende          |
| Posta tradizionale | Alti                   | Analogico o digitale       | Alto             |

**Tabella 5: confronto tra diversi tipi di reti per la trasmissione delle informazioni.**

### **Che cosa è una rete ?**

Secondo “il vocabolario Treccani”: *“Una rete è un elemento o complesso di elementi simile per struttura e per funzioni ad una rete di fili intrecciati.*

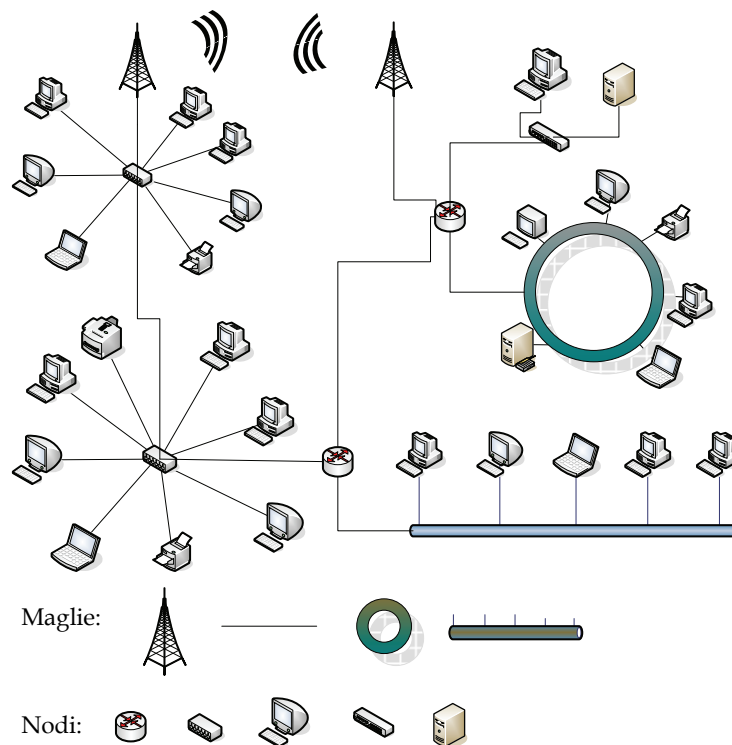
*Nelle telecomunicazioni è un complesso di apparecchi, linee, circuiti e altri impianti, per mezzo dei quali viene svolto un servizio.*

*In informatica è un sistema di interconnessioni tra calcolatori [...] che permette lo scambio di messaggi tra utenti, la trasmissione di dati o programmi e, in alcuni casi, [...] la gestione contemporanea di risorse distribuite.”.*

Esistono quindi molti tipi di reti, diverse per struttura, dimensioni e fini:

- le strade che attraversano una nazione (rete stradale)
- i cavi per il trasporto dell'energia elettrica tra le centrali ed i punti di utilizzo (rete elettrica)
- i cavi telefonici per le comunicazioni (rete telefonica)
- i nervi all'interno del nostro corpo (sistema nervoso)
- le vene/arterie all'interno del nostro corpo (sistema cardiocircolatorio).

In telematica (la branca dell'informatica che studia le tecniche per la trasmissione dei dati), tutti gli elementi che compongono una rete sono chiamati risorse trasmissive. In una rete dati le maglie della rete sono gli strumenti che hanno il compito di trasportare le informazioni (ad esempio cavi o ponti radio) da un nodo all'altro, i nodi sono gli elaboratori che hanno il compito di inviare e ricevere i dati attraverso le maglie della rete. Anche il mittente ed il destinatario sono due nodi della rete. In Figura 12 è visibile un esempio di rete dati che connette tra loro più computer. Sono stati evidenziati, per chiarezza, i nodi e le maglie.



**Figura 12: un esempio di rete in telematica, nel quale sono evidenziati i nodi e le maglie.**

## ***I protocolli***

Ogni rete è stata progettata per il trasporto di qualcosa, ad esempio le automobili nel caso della rete stradale o il sangue nel sistema cardiocircolatorio. Affinché questo trasporto avvenga senza problemi, è necessario che tutti gli elementi attivi coinvolti seguano delle regole comuni e siano in grado di scambiarsi dei messaggi chiari. Pensiamo all'esempio della rete stradale, che sicuramente conosciamo meglio: tutti gli attori devono conoscere il Codice della strada, che contiene queste regole e questi messaggi necessari al corretto transito degli utenti della strada, senza incidenti e con pochi rallentamenti. Gli attori, in questo caso, comprendono non solo gli automobilisti, ma anche i produttori di automobili ed i gestori della rete. I produttori di automobili e gli automobilisti devono condividere regole e messaggi – ad esempio per definire quale pedale serve per accelerare e quale per frenare; anche i gestori delle strade e gli automobilisti devono condividere regole e messaggi: un gestore non può utilizzare la linea orizzontale tratteggiata come separatore per due corsie se vuole inviare all'automobilista un messaggio di divieto di superamento della riga stessa. Infine tutti gli utenti di uno stesso gruppo devono condividere regole e messaggi: un automobilista di Roma ed uno di Milano devono utilizzare l'indicatore di direzione quando devono accostare o girare.

L'insieme di regole e messaggi non ambigui che governano la comunicazione tra due entità è detto **protocollo**.

Definire un protocollo significa descrivere queste regole in modo non ambiguo, specificando tutti i messaggi che le entità possono scambiarsi, il loro significato e le azioni da intraprendere in ogni situazione.

Prendiamo ad esempio il semaforo: questo strumento invia agli utenti della strada dei messaggi. Il Codice della strada indica i messaggi che il semaforo può inviarci e, a seconda del messaggio, quali regole devono essere seguite. Devono esistere regole per ogni possibile messaggio inviato dal semaforo. In Tabella 6 è riportato l'elenco dei messaggi che il semaforo può trasmettere con le relative regole da seguire. Come si vede nei casi *Verde-Giallo* e *Altro* la regola è ambigua e lasciata all'interpretazione dell'uomo: cosa impossibile nelle reti dove sono coinvolti generici esecutori!

**Messaggio**

**Regola**

|              |                                                                                                                                                                                                                                                                                                           |
|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Verde        | autorizzazione a procedere                                                                                                                                                                                                                                                                                |
| Rosso        | divieto di procedere                                                                                                                                                                                                                                                                                      |
| Verde-Giallo | non superare la striscia di arresto o la perpendicolare del segnale, a meno di non trovarsi così prossimo al segnale, quando la luce gialla si accende, da non potersi più arrestare in condizioni di sicurezza sufficienti prima di aver superato la striscia di arresto o la perpendicolare del segnale |
| Altro        | procedere con cautela dando la precedenza a chi viene da destra.                                                                                                                                                                                                                                          |

**Tabella 6: regole e messaggi del semaforo**

È importante fare una distinzione per non generare ambiguità, poiché spesso si fa confusione tra protocollo e linguaggio; come detto, il protocollo è un insieme non ambiguo di regole e di messaggi che governano la **comunicazione** tra due entità. Il linguaggio, invece è un insieme di segnali e/o simboli per mezzo dei quali entità diverse **comunicano** tra loro.

Linguaggio e protocollo riguardano lo scambio di informazioni tra entità diverse, ma operano a livello diverso: mentre attraverso il linguaggio vengono formalizzate le informazioni da trasmettere (pensiamo ad esempio alla formalizzazione di un pensiero in parole), utilizzando il protocollo si garantisce che i diversi attori siano in grado di trasmettere le informazioni dal mittente al destinatario e che il destinatario sia in grado di ricevere correttamente il contenuto della comunicazione.

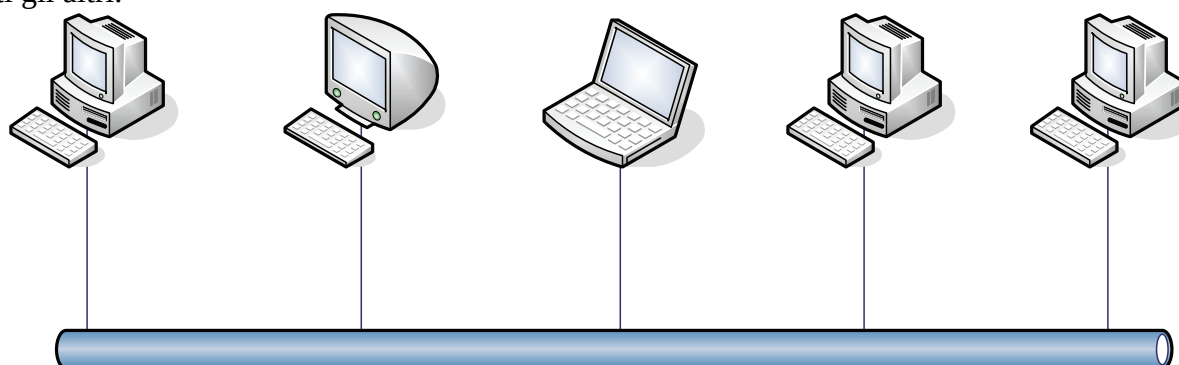
## Tipi di reti

Nel seguito sono descritti i principali tipi di reti di computer oggi disponibili.

### Reti a bus

Nelle reti a bus, ogni elaboratore o host (non solo computer, ma anche stampanti ed altri dispositivi) è collegato agli altri dispositivi in modo lineare attraverso un cavo, come visibile in Figura 13. Quando un elaboratore deve comunicare con un altro, invia i dati attraverso il cavo sino al destinatario. Opportuni protocolli devono essere definiti perché più coppie di computer possano comunicare in contemporanea.

I principali pregi di questa rete sono la semplicità di realizzazione e l'economicità. D'altronde, poiché i dati di tutti gli interlocutori transitano attraverso una sola risorsa trasmissiva, chiunque può accedere al contenuto della trasmissione, rendendo poco sicuro questo tipo di rete. Inoltre se tutti gli utenti comunicano contemporaneamente, si genera un elevato traffico. Infine, un ultimo difetto, riguarda la scarsa affidabilità: un guasto ad un computer o al cavo rende inutilizzabile la rete per tutti gli altri.



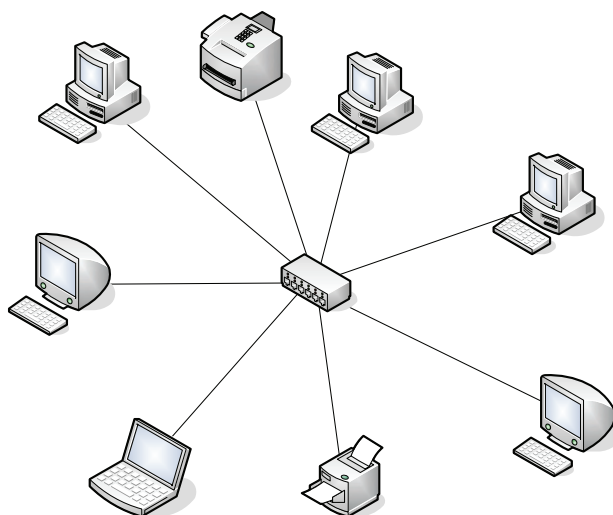
**Figura 13: esempio di rete a bus**

## Reti a stella

Le reti a stella sono evoluzioni delle reti a bus. Una rete a stella è caratterizzata da un punto centrale, detto centro stella, e diversi host connessi ad esso, come visibile in Figura 14. Quando un elaboratore deve comunicare con un altro, invia i dati al centro stella che li trasmette all'host di destinazione. Se più coppie di elaboratori devono comunicare in contemporanea, ogni mittente invierà i dati al centro stella che li ritrasmetterà al relativo destinatario.

Molti dei difetti tipici di una rete a bus sono superati dalla rete a stella. Infatti il guasto di un host o di un cavo non compromette le comunicazioni degli altri elaboratori; inoltre i dati scambiati non possono essere intercettati dagli altri host, poiché il centro stella li invierà solo al legittimo destinatario. In questa rete l'elemento critico è il centro stella, poiché un guasto del centro stella comprometterebbe il funzionamento di tutta la rete, inoltre un centro stella non molto performante potrebbe provocare un rallentamento di tutta la rete. Oggi questo rischio è molto basso: un centro stella che consente ad 8 elaboratori di scambiarsi dati a 100.000.000 di bit al secondo costa poche decine di euro, per sicurezza è sempre bene averne due !

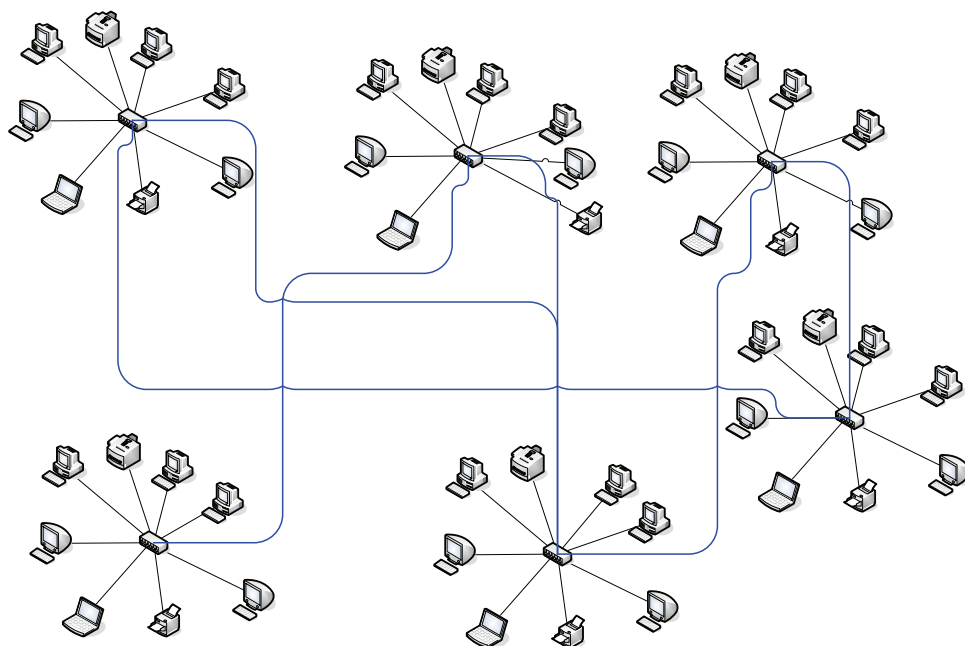
In generale i dispositivi che fungono da centro stella sono gli hub (ormai obsoleti) e gli switch.



**Figura 14: esempio di rete a stella**

## Reti a maglia

Nel caso in cui un centro stella sia connesso, oltre che ad altri host, anche ad un altro centro stella, le due reti si dicono connesse e diventano le sottoreti di una rete più grande. In questo caso il dispositivo che funge da centro stella è il router. Poiché ad una rete si possono collegare più reti, si realizza una sorta di maglia tra le reti, come visibile in Figura 15. Una rete di questo tipo è estremamente affidabile, poiché in caso di guasto ad una delle sottoreti, le altre possono continuare a comunicare senza problemi: i dati faranno un percorso diverso. L'internet, ad esempio, è una maglia composta dall'unione di migliaia di reti di dimensioni minori.



**Figura 15: esempio di rete a maglia**

## Reti LAN, MAN e WAN

Le reti **LAN** (dall'inglese Local Area Network) sono reti di elaboratori estese su aree di piccole dimensioni, come un ufficio od un palazzo.

I computer sono connessi direttamente alla rete LAN (come nelle precedenti figure) e questo consente uno scambio di dati ad elevata velocità, fino a 1.000.000.000 bit al secondo.

Un esempio di computer connessi attraverso una rete LAN sono i computer presenti nel laboratorio di informatica.

Le reti **WAN** (dall'inglese Wide Area Network) sono reti di elaboratori estese su vaste dimensioni (nazioni, continenti) e servono per connettere tra loro LAN diverse, ma geograficamente distanti, per consentire lo scambio dei dati tra gli utenti delle diverse reti. Molte WAN sono costruite per una particolare organizzazione e sono private. Una rete WAN può utilizzare una combinazione di mezzi trasmissivi diversi come ad esempio la fibra ottica, il satellite, le onde radio od i cavi in rame.

Generalmente tra le LAN e le WAN si trovano le **MAN** (dall'inglese Metropolitan Area Network), reti di connessione tra LAN di dimensione cittadina. Compito delle WAN e delle MAN è quello di trasmettere le informazioni tra i computer del mittente e del destinatario.

La città di Roma è cablata da una MAN; infatti intorno al Grande Raccordo Anulare sono depositate dei cavi in fibra ottica ai quali le diverse reti LAN o i nostri computer di casa si collegano quando vogliono navigare sull'internet.

## Reti a commutazione di pacchetto e di circuito

Le tipologie di reti che abbiamo visto sino ad adesso sono una evoluzione rispetto alle reti telefoniche. Vediamo ora come a questa evoluzione nelle reti fisiche è corrisposta una evoluzione anche nelle logiche di scambio delle informazioni.

### La commutazione di circuito

La **commutazione di circuito** è il sistema per collegare temporaneamente due utenti, come avviene ad esempio con il telefono. Durante la telefonata, infatti, si realizza un circuito temporaneo che collega i due interlocutori attraverso una serie di connessioni che si realizzano tra una centrale telefonica e la successiva.

Le risorse trasmissive (telefono, linee, porzioni di centrali) rimangono allocate ai due interlocutori sino al termine della chiamata, per tale motivo telefonando ad un utente che sta utilizzando la linea telefonica si ottiene il segnale di linea occupata.

I servizi che richiedono una connessione permanente durante il loro utilizzo sono detti **connection oriented** od orientati alla connessione.

I servizi orientati alla connessione hanno un utilizzo inefficiente delle risorse trasmissive, poiché se i due interlocutori non trasmettono alcun segnale (ad esempio sono in silenzio) le risorse non vengono sfruttate.

## **La commutazione di pacchetto**

La **commutazione di pacchetto** non stabilisce alcun percorso prestabilito tra due utenti. Le informazioni da spedire vengono organizzate in pacchetti di dati ed affidate alla rete con l'indicazione dell'indirizzo del destinatario, come avviene ad esempio per il sistema postale.

Le risorse trasmissive sono utilizzate in modo più efficiente, poiché sono allocate ad un utente per il tempo strettamente necessario alla trasmissione dei pacchetti e poi vengono allocate ad un altro utente, esattamente come un tratto di strada è allocato ad un automobilista per il tempo strettamente necessario al suo passaggio: pochi secondi dopo su quel tratto di strada possono passare altre autovetture.

A differenza della commutazione di circuito, la commutazione di pacchetto non è in grado di garantire il tempo minimo di percorrenza delle informazioni, né tanto meno un ordine esatto di arrivo dei pacchetti: sempre immaginando un parallelo con la rete stradale, poiché la rete è condivisa è possibile che un percorso sia più trafficato in alcuni momenti o che, se più macchine partono insieme da uno stesso punto, arrivino a destinazione in un ordine diverso da quello di partenza.

I servizi che non richiedono una connessione permanente durante il loro utilizzo sono detti **connectionless** o senza connessione.

La commutazione di pacchetto è adatta quando un ritardo nella ricezione dei dati può essere tollerato.

Se infatti nella ricezione di un file o di una e-mail l'utente non si rende conto di un ritardo, nel caso di un evento in tempo reale (ad esempio una telefonata) anche un secondo di ritardo viene percepito dai due interlocutori.

La rete internet è un esempio di rete connectionless, questo significa che i protocolli utilizzati non mantengono alcuna informazione di stato circa i pacchetti inoltrati. Ciascun pacchetto è trattato indipendentemente da tutti gli altri e può fare percorsi diversi (arrivando quindi in tempi diversi) rispetto agli altri.

## ***I browser per la navigazione ipermediale***

Il browser (dall'inglese *to browse*: “sfogliare un libro”) è un software applicativo che ci consente di navigare tra le risorse presenti nel web. Per consentire questa navigazione è stato necessario definire un protocollo per lo scambio delle informazioni ed un linguaggio nel quale queste informazioni devono essere scritte per poter essere comprese. Tali strumenti sono:

- l'HyperText Transfert Protocol (HTTP)
- l'HyperText Markup Language (HTML)

Il primo definisce le regole per la trasmissione dei contenuti tra il server web ed il browser, il secondo invece definisce il linguaggio con cui devono essere scritte le pagine web.

L'HyperText Markup Language è un insieme di istruzioni (i **tag**) utilizzate per specificare la struttura e la formattazione di un documento, nonché definire i collegamenti e le diverse risorse. L'autore prepara le diverse pagine web similmente a come realizzerebbe un documento con gli applicativi di videoscrittura.

Tale documento, tuttavia, è adatto ad essere distribuito attraverso rete, ad integrare contributi su media diversi (audio, filmati, immagini), nonché ad integrarsi in modo naturale con gli altri documenti presenti nella rete, diventando parte del web.

Il browser è l'applicativo software che ci consente di richiedere una particolare risorsa (ad esempio una pagina web) ad un server, utilizzando il protocollo HTTP. Il browser riceve i risultati, interpreta il linguaggio HTML e produce la pagina a video così come siamo abituati a vederle.

I browser più diffusi sono Internet Explorer e Firefox.

Inizialmente un browser doveva essere in grado di collegarsi ad un server web utilizzando il protocollo HTTP, conoscere il linguaggio HTML e saper mostrare immagini.

Oggi un browser che facesse solo questo sarebbe troppo elementare !

Infatti i browser sono molto di più che semplici visualizzatori di pagine web. Un browser, infatti, è diventato un applicativo per la riproduzione di contenuti realizzati con praticamente tutti i media digitali (ad esempio riproduce i filmati ed i brani audio). Per aumentare le possibilità del browser, i produttori hanno rilasciato delle specifiche che consentono a qualsiasi programmatore di realizzare degli applicativi che estendono le potenzialità del browser, consentendogli di svolgere nuove funzioni o integrare nuovi media. Tali applicativi sono detti plug-in. Un plug-in è per il browser qualcosa di simile a ciò che un driver è per il sistema operativo, entrambe estendono le potenzialità di comunicazione del software.

I plug-in sono usati oggi per rendere più interattive le pagine web con animazioni (plug-in per supportare i file flash), per eseguire programmi all'interno delle pagine (plug-in per supportare i programmi Java), per visualizzare documenti (plug-in per supportare la visualizzazione di file pdf) o per riprodurre brani audio o filmati.

Tutti i plug-in, ormai, si scaricano dall'internet e vengono installati facilmente sul computer.

## ***La ricerca di informazioni sull'internet***

L'internet è lo strumento attraverso il quale accedere alle informazioni presenti nei diversi siti web, così come la rete telefonica è lo strumento attraverso il quale accedete alle informazioni disponibili presso gli altri utenti. È necessario che gli utenti siano in grado di comprendere e di utilizzare questo nuovo tipo di organizzazione delle informazioni e, soprattutto, sia in grado di ricercare le informazioni a cui sono interessati riuscendo a discernere tra fonti informative di qualità diversa.

Esistono diverse tipologie di siti da cui partire quando si cerca una specifica informazione sull'internet, ognuno è stato progettato per venire incontro ad esigenze specifiche.

Come spesso succede, il confine tra le diverse tipologie di siti è sempre meno definito, poiché ogni fornitore cerca di rispondere a tutte le esigenze degli utenti per fidelizzarli e diventare il loro fornitore di riferimento.

Ormai senza un sistema di ricerca delle informazioni è molto difficile trovare ciò che si cerca. Le principali tipologie di siti per la ricerca di informazioni sull'internet sono di seguito presentati.

## **I portali web**

I **portali web** sono siti web strutturati in diverse sezioni, attraverso le quali è possibile accedere ad informazioni sui diversi argomenti trattati all'interno del portale. I portali generalisti contengono le informazioni che generalmente cercano i navigatori: ultime notizie, sport, finanza, cinema, tempo libero e così via. I portali generalisti più famosi sono Yahoo!, Excite e Virgilio.

I portali tematici, sono invece siti che contengono informazioni relative ad uno specifico argomento; ad esempio [www.uniroma2.it](http://www.uniroma2.it) è il portale di Tor Vergata, poiché fornisce informazioni sull'università ed è il punto d'accesso privilegiato alle informazioni delle diverse facoltà.



Figura 16: la homepage del portale generalista www.yahoo.it



**Figura 17: la homepage del portale tematico [www.torvergata.it](http://www.torvergata.it)**

## I motori di ricerca

I **motori di ricerca** sono servizi che consentono ad un utente di ricercare le risorse (ad esempio pagine web, o documenti) che contengono una particolare parola o frase. Ad esempio se cerchiamo tutte le risorse che contengono la frase “macchina di Von Neumann”, possiamo accedere in pochi secondi ad un archivio praticamente infinito di risorse.

Cercando su Google ([www.google.it](http://www.google.it)), il migliore motore di ricerca attualmente disponibile, la frase “macchina di Von Neumann” si ottengono circa 75.500 risorse internet che contengono questa frase (nel 2003 erano 1.250);

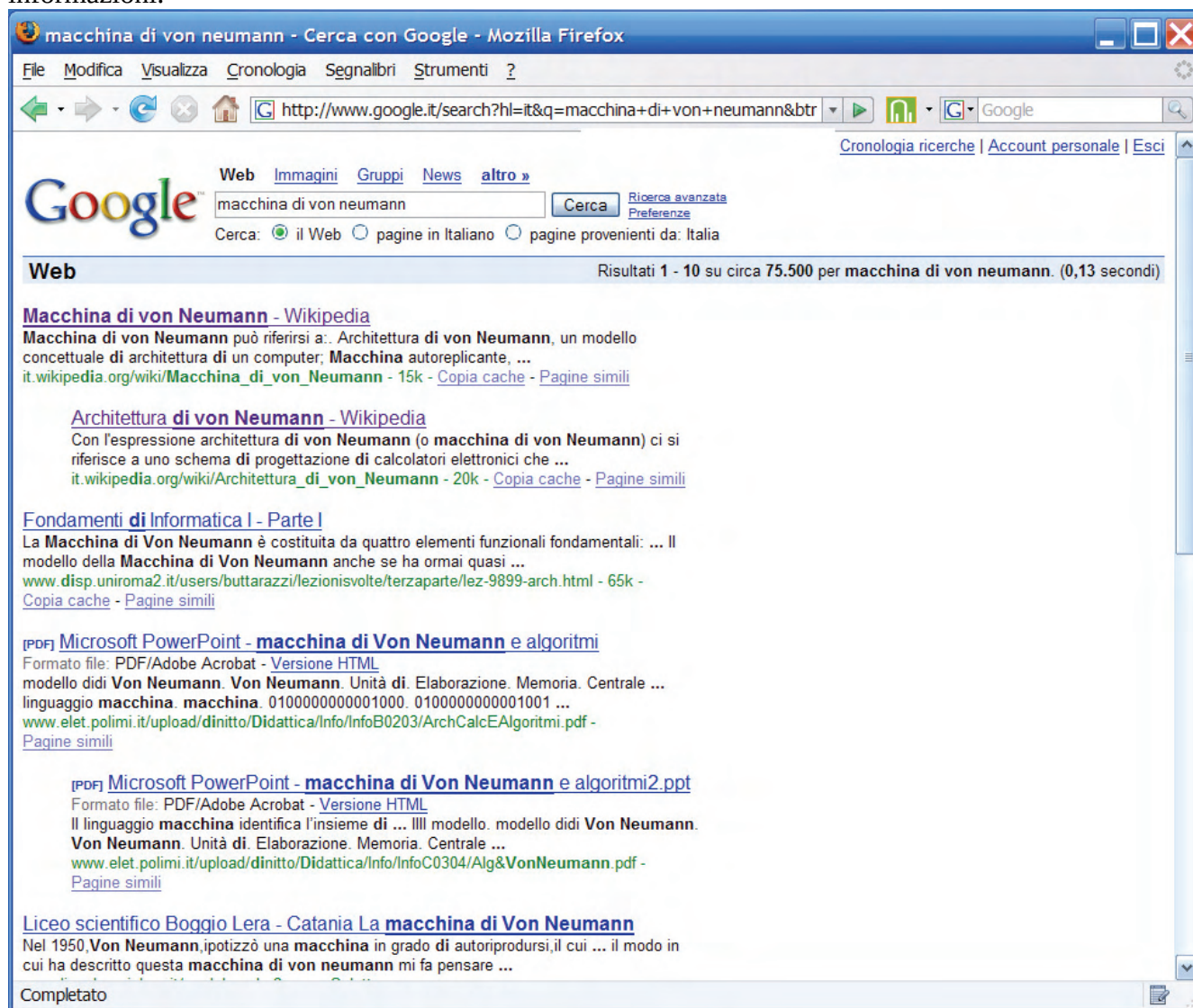
Cercando solo “Von Neumann” si trovano circa 2.460.000 documenti (nel 2003 erano 350.000) !

La grande differenza tra i valori ottenuti nel caso di cerchi “Macchina di Von Neumann” o “Von Neumann” è facilmente spiegabile: la parola “macchina” ha ristretto la ricerca alle sole pagine in italiano, mentre il nome Von Neumann è lo stesso in quasi tutte le lingue.

Il numero di risorse internet che i motori di ricerca conoscono cresce nel corso del tempo, infatti opportuni software (detti spider) indicizzano continuamente i contenuti dei siti web.

Google il 22 novembre 2003 conosceva il contenuto di 3.307.998.701 risorse web diverse, il 21 novembre 2004 questo numero era diventato 8.058.044.651: un incremento del 145% in un solo anno !!

Data l'enorme disponibilità di contenuti sull'internet, diventa centrale la capacità dell'utente di specificare in modo chiaro l'oggetto delle sue ricerche scegliendo i termini più adatti per effettuare la ricerca, nonché di saper raffinare la ricerca sempre più in dettaglio sino a selezionare un numero di siti non troppo elevato e con contenuti il più possibile coerenti con l'oggetto della sua ricerca. Ad esempio, se cerchiamo le parole macchina di Von Neumann, potremmo trovare un sito che contiene le foto della nuova automobile di un tale signor Von Neumann. Nelle librerie o sull'internet si possono trovare molti testi che aiutano gli utenti nel raffinare le proprie tecniche di ricerca delle informazioni.



**Figura 18: la pagina mostrata da Google al termine della ricerca di "macchina di Von Neumann"**

## I newsgroup

L'idea del **newsgroup** (o gruppo di discussione) è quella di condividere le proprie conoscenze ed esperienze, consentendo a chi lo desidera di esprimere la propria opinione o di porre problemi, a cui qualcuno saprà (forse) rispondere.

Sono i predecessori dei forum oggi molto diffusi. La differenza fondamentale è che, mentre i forum sono diversi per ogni sito e ad un singolo sito sono legati, i newsgroup sono trasversali: tutti i newsgroup, indipendentemente dal sito sul quale sono ospitati, sono connessi tra loro.

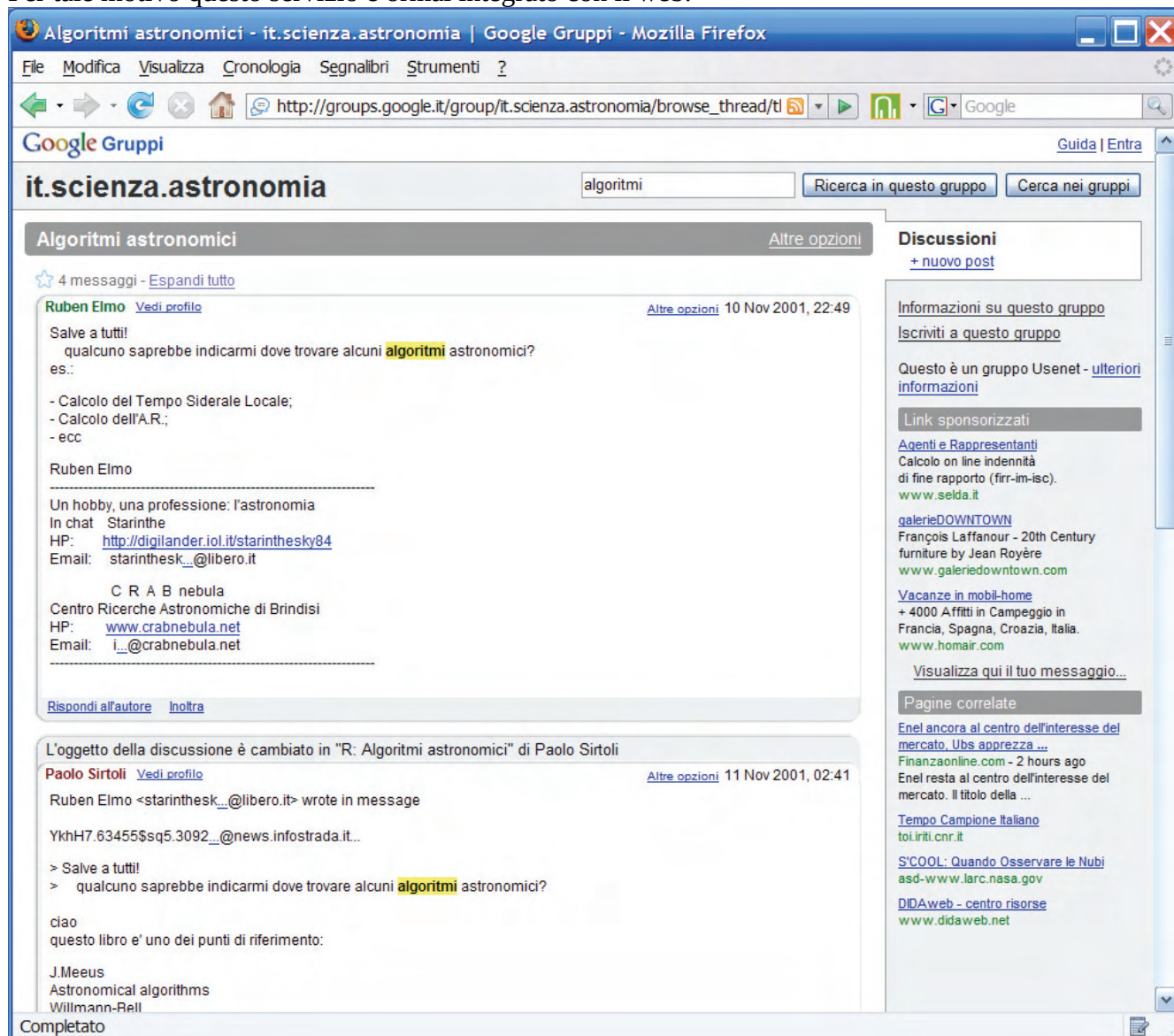
Per evitare il caos esistono newsgroup per i più diversi argomenti e chi lo desidera può aprirne di nuovi (oggi i newsgroup sono più di 300.000).

Gli argomenti sono i più disparati: esiste un gruppo italiano dedicato agli appassionati di dolci, così come uno internazionale dedicato alla fisica della relatività ed un altro dedicato alla radiologia medica.

Chiunque voglia aderire al newsgroup può inviare un messaggio ex novo oppure leggere i messaggi inviati dagli altri e decidere se eventualmente rispondere.

Inizialmente esisteva un software apposito per l'interazione con i gruppi di discussione, successivamente le funzionalità di questi client sono state integrate con i browser. Oggi molti siti offrono la possibilità di accedere ai diversi gruppi di discussione attraverso pagine web, in questo modo i messaggi possono contenere link ipertestuale ad altri messaggi a contenuti esterni.

Per tale motivo questo servizio è ormai integrato con il web.



**Figura 19: esempio di un messaggio relativo agli algoritmi nel newsgroup italiano dedicato all'astrofisica**

## Le directory

Le **directory** sono strumenti in grado di fornire agli utenti elenchi di siti web classificati per argomento. Partendo da una categoria, ad esempio "Art", è possibile trovare tutti i siti censiti che parlano del gruppo "the Eagle" selezionando le sottocategorie "Music" -> "Bands and Artists" -> "E" -> "Eagles, The". In questo modo si trovano 54 siti (nel 2003 erano 58), cercando la frase "The

Eagle” utilizzando Google come motore di ricerca si ottengono circa 82.400.000 risorse che contengono quella frase (nel 2003 erano 26.000.000).

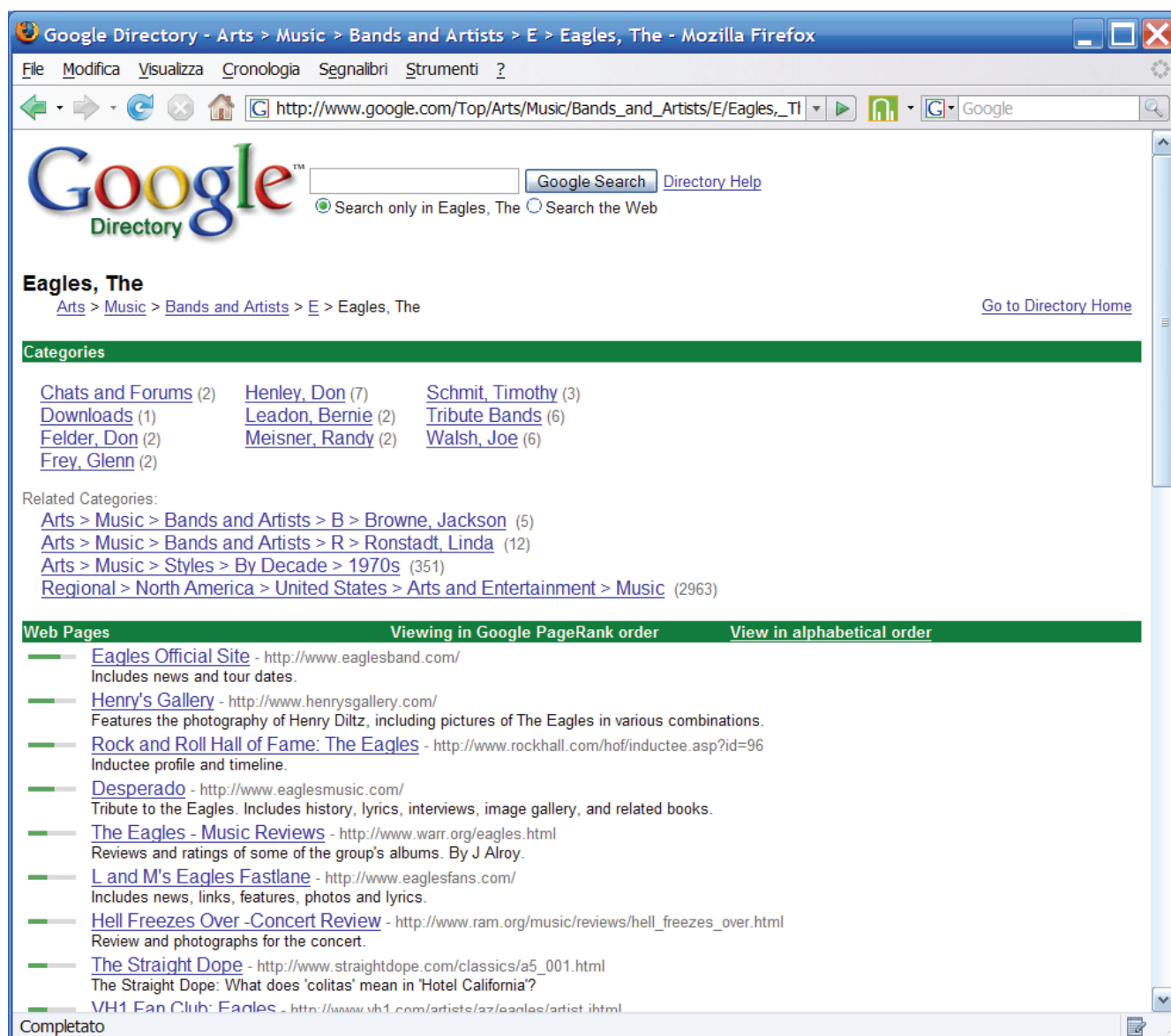
L'enorme differenza nel numero di risultati ottenuti con questi due strumenti ha varie ragioni.

In **primo luogo** i risultati ottenuti utilizzando le directory sono relativi a siti web, mentre i risultati della ricerca con Google sono risorse: un sito con 100 pagine dedicate agli Eagle conterà 1 per la directory ma 100 per Google.

In **secondo luogo** i risultati ottenuti utilizzando la directory è relativo ai soli siti con argomenti attinenti il gruppo musicale, nel caso dei risultati ottenuti con le funzioni di ricerca di Google ci saranno siti dedicati al gruppo musicale, alle aquile o semplicemente pagine web personali di utenti che scrivono la loro passione per le aquile o per il gruppo musicale ...

**Infine**, l'assegnazione di un particolare sito ad una o più categorie richiede una costante attività umana, e per tale motivo il numero di siti presenti nelle directory cresce molto lentamente rispetto ai motori di ricerca.

Utilizzando le directory, quindi, si trovano meno risorse ma, molto probabilmente, di maggiore attinenza.



**Figura 20: esempio da una directory**

## **La posta elettronica**

La posta elettronica è un servizio fruibile attraverso rete internet che consente ad un utente della rete di inviare un messaggio riservato ad un altro, come già avviene per la posta ordinaria, ma con tre importanti differenze.

In **primo luogo** le informazioni scambiate sono in formato digitale, ciò consente al destinatario di elaborare le informazioni ricevute.

In **secondo luogo** la ricezione avviene – di fatto – in tempo reale indipendentemente dalla distanza geografica tra mittente e destinatario; nella posta tradizionale una lettera impiega giorni per arrivare.

**Infine** il costo dell'invio è praticamente nullo, cosa non possibile nel caso di scambio di lettere cartacee.

Il programma a disposizione dell'utente (ad esempio Outlook o Eudora) si collega al server di posta per acquisire i nuovi messaggi o per inviare un messaggio, così come un utente si reca alla posta per inviare o ricevere un pacco.

Per poter inviare e ricevere posta elettronica ogni utente si deve munire di un indirizzo che lo identifichi in modo univoco su tutta la rete: l'indirizzo di posta elettronica.

Un esempio di indirizzo di posta elettronica è:

**mario.rossi** @ [rossimail.it](mailto:rossimail.it)

Quest'indirizzo si legge utente **mario rossi**, la cui casella di posta elettronica è presso il dominio [rossimail.it](mailto:rossimail.it)

I messaggi di posta elettronica stanno diventando uno strumento quotidiano di lavoro per molti.

Il protocollo alla base della posta elettronica, tuttavia, è molto semplice e limitato: non garantisce ad esempio al destinatario di conoscere con certezza l'identità del mittente né consente al mittente di essere informato automaticamente della presa a carico di un messaggio da parte del destinatario. A pensarci bene, questa non è una novità: anche utilizzando la posta tradizionale non abbiamo certezza del mittente (chiunque può mandarci una lettera spacciandosi per qualcun altro), né abbiamo certezza che la lettera arrivi al destinatario.

## 4 Trattamento digitale delle immagini e dei suoni

### Introduzione

Nella prima unità didattica abbiamo visto che un elaboratore elettronico deve essere in grado di trattare quattro tipi di informazioni:

- Numeri
- Caratteri
- Immagini
- Suoni.

Successivamente abbiamo visto come gli elaboratori elettronici trattano numeri e caratteri, adesso ci focalizzeremo su immagini e suoni.

Nel seguito vedremo come sono conservate le immagini ed i suoni nei file e perché esistono formati diversi.

### Le immagini

La prima domanda da porsi è semplice: Che cos'è una immagine?

Secondo il *Dizionario De Mauro della lingua italiana* una immagine è la “Raffigurazione grafica, fotografica o plastica di qualcosa o qualcuno”.

Secondo *Il vocabolario Treccani*, invece, una immagine è “Forma esteriore degli oggetti corporei, in quanto viene percepita attraverso il senso della vista [...], o rimane impressa in una lastra o pellicola o carta fotografica”.

Entrambe le definizioni sottolineano come l'immagine sia la raffigurazione di qualcosa che rimane impressa (o memorizzata) su un supporto terzo (la pellicola fotografica o il nostro hard disk). Per i nostri interessi, esistono due tipi di immagini: le immagini **analogiche** e quelle **digitali**.

Nel caso di **immagini analogiche** (ad esempio le fotografie), sulla pellicola fotografica vengono posti dei materiali fotosensibili che alterano il loro stato quando colpiti dalla luce. L'immagine rappresentata sulla carta fotografica è ottenuta per analogia con la quantità di luce che ha impresso i diversi punti della pellicola durante la fase dell'esposizione (lo scatto).

Nel caso di **immagini digitali**, i diversi colori che rappresentano l'immagine sono memorizzati come numeri. In fase di acquisizione delle immagini in formato digitale (ad esempio la macchina fotografica digitale o lo scanner) ad ogni colore è associato un numero. In fase di visualizzazione (ad esempio il monitor o la stampante) ad ogni numero è fatto corrispondere un colore.

Il processo che trasforma una immagine in una sequenza ordinata di numeri è detto **digitalizzazione**. Per comprendere questo processo è necessario introdurre due concetti: il **campionamento spaziale** e la **quantizzazione cromatica**.

### Il campionamento spaziale

Per campionamento spaziale si intende l'operazione con la quale una immagine continua viene trasformata in un insieme discreto di rettangoli più o meno grandi, i pixel. Una immagine viene così rappresentata come una matrice di pixel. Per ogni immagine deve essere definito quante righe e quante colonne di pixel la compongono. Tali valori sono generalmente indicati con la dicitura *numero\_di\_colonne x numero\_di\_righe*. Ad esempio indicando la dimensione di un'immagine con 640x480, si intende che l'immagine è composta da 480 righe e 640 colonne, per un totale di 307.200 pixel.

Il successivo processo di quantizzazione cromatica assegna ad ogni pixel è stato assegnato un colore, dato dalla media dei colori presenti all'interno della porzione di immagine sottesa al pixel.

Il pixel è, quindi, l'elemento minimo di informazione relativo ad una immagine, poiché ad ogni pixel è associato un colore.

Nasce quindi spontanea una domanda: quanti pixel servono per visualizzare bene una immagine?

Questa domanda è mal posta poiché non esiste un valore assoluto, ma tale valore dipende dalle informazioni presenti nell'immagine originale che vogliamo conservare.

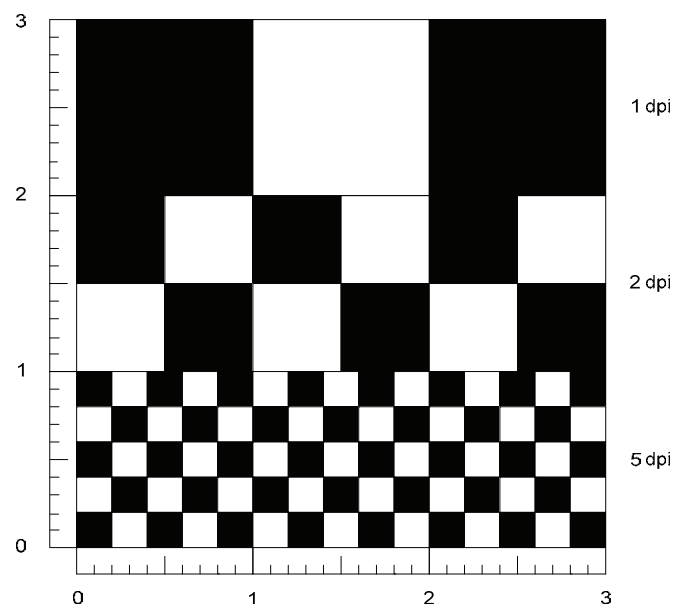
Per *risoluzione spaziale* si intende lo spazio minimo tra due oggetti per cui, all'interno dell'immagine, essi sono ancora visibili come due entità separate.

Se scelgo una matrice di pixel poco fitta (quindi un basso numero di righe e di colonne), è probabile che due oggetti che voglio memorizzare come distinti ricadano all'interno dello stesso pixel e, quindi, siano indistinguibili. Rendendo la matrice più fitta, e quindi aumentando il numero di righe e di colonne usato per memorizzare le informazioni presenti all'interno dell'immagine, un numero sempre maggiore di dettagli sarà memorizzato. Quando anche i due oggetti che intendevo memorizzare saranno distinti, la risoluzione scelta sarà quella opportuna.

In una immagine digitale più fitti sono i pixel, ovvero quanto più grande è la dimensione della matrice, più la risoluzione è alta.

Quando si acquisisce una immagine (ad esempio allo scanner), è necessario specificare la risoluzione spaziale con cui si intende acquisire l'immagine. Questa è generalmente espressa in dpi (Dot Per Inch) che corrisponde al numero di pixel per ogni pollice dell'immagine.

Cerchiamo di chiarire quanto abbiamo appreso sino ad adesso con alcuni esempi.



**Figura 21: esempio della dimensione dei pixel nel caso di 1, 2 o 5 punti per unità di lunghezza**

In Figura 21 è visibile una immagine dove, con i colori bianco e nero, sono stati rappresentati dei pixel adiacenti. Per unità di lunghezza sono stati presi 1, 2 o 5 punti. È evidente come, al crescere del numero di punti presi sia possibile riconoscere un maggiore numero di dettagli.

Tale possibilità è ancora più evidente in Figura 22 dove è indicata, nella prima riga, una ipotetica immagine lineare. Nella seconda linea, la stessa immagine viene presentata con una risoluzione che è 5 volte più grande della dimensione dei dettagli. Come si vede, l'immagine che ne risulta è una sequenza di pixel a diversa tonalità di grigio poiché il colore del pixel è dato dalla media dei colori presenti al suo interno. La risoluzione scelta non consente di distinguere alcun dettaglio informativo dell'immagine iniziale.

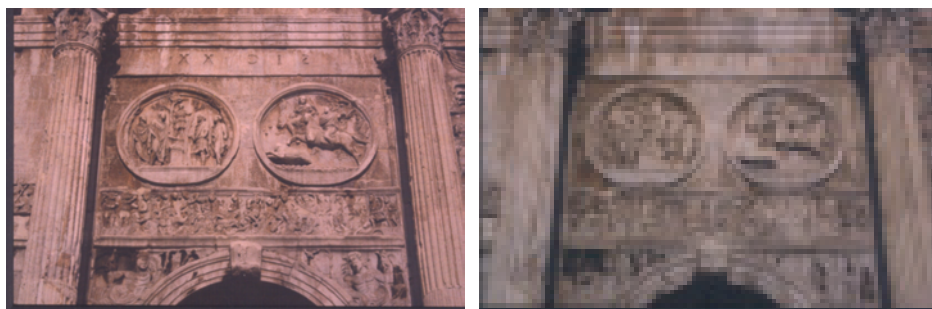
Nelle righe successive, l'immagine viene presentata con risoluzioni rispettivamente quattro, tre e due volte maggiori delle dimensioni dei dettagli. È evidente come le diverse immagini tendano a presentare un numero maggiore di informazioni e dettagli al crescere della risoluzione spaziale utilizzata. L'ultima riga presenta l'immagine rappresentata con una risoluzione pari alla dimensione dei dettagli. In questo caso tutti i dettagli presenti nell'immagine originale sono conservati nell'immagine campionata.



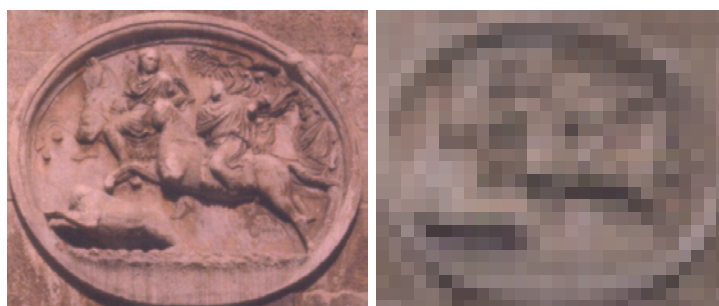
**Figura 22: effetto della risoluzione spaziale**

Passando alle immagini reali, è possibile vedere l'effetto della risoluzione spaziale scelta sulla qualità di una immagine. In Figura 23 sono visibili le scansioni di una diapositiva effettuate rispettivamente a 1200 ed a 75 dpi. Per quanto in entrambe le immagini sia possibile distinguere bene l'opera rappresentata, la seconda sembra sfuocata poiché i dettagli sono scarsamente visibili. Tale limitazione nella seconda immagine è più evidente se si effettua uno zoom di un dettaglio, poiché in questo modo gli oggetti di piccole dimensioni vengono ingranditi; è evidente come nell'immagine di sinistra sia possibile distinguere i dettagli contenuti nel rosone, mentre nell'immagine di destra non è possibile distinguere praticamente alcunché.

In Figura 24 è visibile il dettaglio del rosone di destra presente in Figura 23. La diversa qualità dell'immagine può a prima vista sembrare non giustificata, in realtà è necessario comprendere quanti pixel compongono le due immagini: poiché un pixel è il minimo elemento informativo, più pixel ci sono in una immagine più è possibile avere informazioni sul contenuto presente nell'immagine stessa. Nel caso della Figura 24, l'immagine di destra è composta da circa 16.000 pixel, l'immagine di sinistra da circa 4.147.000 pixel: l'immagine di destra contiene meno di un duecentesimo delle informazioni contenute nell'immagine di sinistra. La differenza, quindi, c'è e si vede!



**Figura 23: immagini scansionate da una diapositiva rispettivamente a 1200 ed a 75 dpi**



**Figura 24: dettaglio del rosone di destra acquisito rispettivamente a 1200 e a 75 dpi**

## La quantizzazione cromatica

I colori possono assumere infiniti valori, poiché infinite sono le sfumature, tuttavia non è possibile rappresentare in formato digitale tutti questi colori, poiché significherebbe associare ad ogni colore un numero compreso tra zero ed infinito. D'altronde memorizzare un numero troppo elevato di colori sarebbe anche inutile perché:

- l'occhio umano non è in grado di percepire la differenza tra due colori troppo simili;
- le periferiche di output (come le stampanti o gli scanner) hanno una capacità di rappresentazione dei colori molto più limitata della capacità di percezione dell'occhio umano.

Per tali motivi, dopo il campionamento spaziale, è opportuno effettuare una nuova operazione che riduce il numero di colori presenti nell'immagine memorizzata: la quantizzazione cromatica.

Con quantizzazione cromatica si intende l'operazione con la quale viene assegnato, ad ogni pixel, un numero – compreso in un opportuno intervallo - che ne identifica il colore.

Per ogni pixel, quindi, viene prima calcolato il colore medio presente nel rettangolo e, successivamente, a tale colore viene assegnato un numero compreso in un intervallo finito di possibili valori.

Per ogni immagine viene definito il numero di colori diversi possono essere rappresentati al suo interno. Tale valore viene indicato come **profondità di colore** e si esprime in bit.



**Figura 25: immagini memorizzate, rispettivamente, a 16 milioni ed a 8 colori**

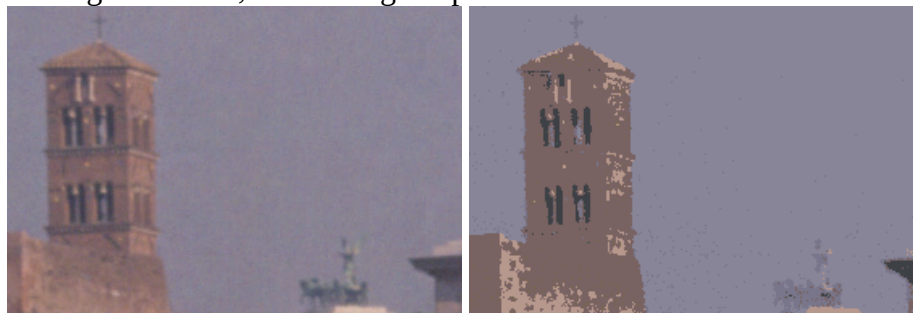
In Figura 25 sono visibili due immagini rappresentanti lo stesso panorama. Nella prima immagine i colori visibili sono scelti da un insieme di 16.777.216 di colori (24 bit di profondità), nel secondo caso da un insieme di 8 colori (3 bit di profondità).

Anche la domanda “quanti colori servono per visualizzare correttamente una immagine” è mal posta, perché dipende da ciò che vogliamo distinguere nell'immagine.

Per *risoluzione cromatica* si intende la minima differenza di colore per cui, all'interno dell'immagine, tali colori sono ancora visibili come separati.

A causa del funzionamento dei nostri occhi, la risoluzione cromatica dipende molto dalle sfumature di colore che si intendono separare. Per tale motivo, definire un valore ottimale è difficile.

In Figura 26 è visibile il dettaglio del campanile riportato in Figura 25. La riduzione della risoluzione cromatica comporta inevitabilmente una perdita nella percezione dei dettagli: sono ancora evidenti tutti gli elementi, ma i dettagli si perdono.

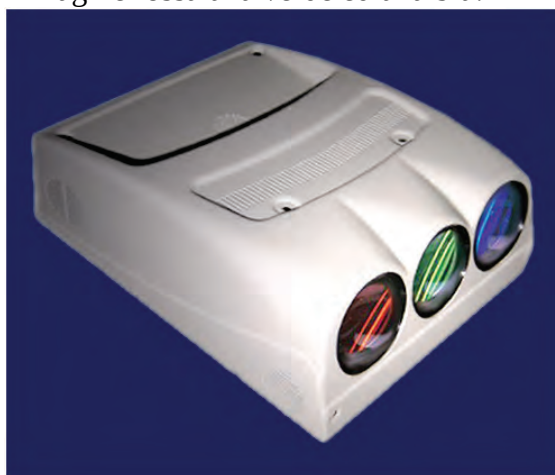


**Figura 26: dettaglio del campanile conservato, rispettivamente, a 16 milioni ed a 8 colori**

## La rappresentazione dei colori

Per rappresentare i colori il sistema scelto si chiama sistema tricromatico additivo RGB e consiste nell'indicare, per ogni pixel, la quantità di rosso (Red), di verde (Green) e di blu (Blu) che compongono il colore del pixel. Questo sistema si basa sulla sintesi additiva dei colori, cioè il nostro cervello associa ad ogni punto un colore che è dato dalla somma di opportune quantità di rosso di verde e di blu: per questo colori sono chiamati *colori primari per il sistema tricromatico RGB*.

Questo è il sistema utilizzato anche nelle televisioni, nei monitor e nei proiettori: tutti i colori che vediamo quando emessi da questi strumenti sono rappresentati come un segnale composto dalla quantità di rosso, di verde e di blu. Possiamo vedere chiaramente quanto descritto se osserviamo un proiettore a tre lampade (come quello visibile in Figura 27) all'opera: quando guardiamo su uno schermo l'immagine generata dal proiettore, di fatto stiamo vedendo la somma delle immagini prodotte con le tre lampade (una per il rosso, una per il verde ed una per il blu). L'immagine sullo schermo è la somma di una immagine rossa una verde ed una blu.



**Figura 27: videoproiettore a tre tubi**

In formato digitale i colori sono generalmente rappresentati in formato RGB, cioè specificando quanto rosso, quanto verde e quanto blu compongono il colore di ogni pixel dell'immagine.

Le quantità di rosso di verde e di blu per ogni pixel, sono definite ciascuna da un valore compreso tra 0 e 255 (un byte) e, quindi, l'immagine risultante ha 24 bit di profondità (che corrisponde a 16.777.216 colori diversi). Un numero a 24 bit può essere rappresentato, in numerazione esadecimale, con 6 simboli: i primi due per il rosso, i due centrali per il verde e gli ultimi per il blu. Nella Tabella 7 per alcuni colori è presentata la quantità di rosso, verde e blu necessari per rappresentarli.

| Rosso | Verde | Blu | Colore (hex) | Colore |
|-------|-------|-----|--------------|--------|
| 0     | 0     | 0   | 000000       |        |
| 255   | 255   | 255 | FFFFFF       |        |
| 255   | 0     | 0   | FF0000       |        |
| 0     | 255   | 0   | 00FF00       |        |
| 0     | 0     | 255 | 0000FF       |        |
| 128   | 128   | 128 | 808080       |        |
| 100   | 170   | 220 | 64AADC       |        |
| 180   | 0     | 210 | B400D2       |        |

**Tabella 7: componenti RGB di alcuni colori**

## **Formati di conservazione delle immagini**

Sono di seguito presentati alcuni dei più diffusi formati per la conservazione delle immagini.

### **Il formato bitmap**

Sulla base di quanto abbiamo imparato, le informazioni contenute in una immagine possono essere archiviate all'interno di una immagine utilizzando un formato molto semplice: per ogni pixel vengono indicati i 3 byte corrispondenti al rosso, al verde ed al blu.

Ancora una volta, come già avvenuto con la tabella ASCII, abbiamo trasformato una qualche entità (prima le lettere, ora i colori) in numeri.

Il dispositivo di input (ad esempio lo scanner) deve indicare per ogni punto le quantità di rosso, verde e blu da memorizzare. Il computer, quando invia una immagine al dispositivo di output (ad esempio il monitor), deve leggere queste quantità ed inviare gli opportuni comandi alla periferica.

Questo formato di conservazione delle immagini è detto bitmap ed i file così archiviati hanno estensione **.bmp**.

Ovviamente nel file bitmap devono essere inserite anche altre informazioni necessarie alla corretta visualizzazione delle immagini, come il numero di pixel in una riga, la risoluzione spaziale utilizzata o la profondità di colore.

Un importante vantaggio di questo formato è la semplicità: la CPU non deve svolgere particolari calcoli in fase di memorizzazione o di visualizzazione di una immagine.

Il problema fondamentale di questo formato è la dimensione del file: una immagine a tutto schermo salvata in formato bitmap pesa circa 2,5M Byte. Tale difetto rende questo formato utilizzabile solo per un ristretto insieme di applicazioni, ma lo rende inutilizzabile ad esempio per il web.

### **Il formato GIF**

Nel 1987, con l'obiettivo di definire un formato per la conservazione delle immagini che fosse sufficientemente leggero da poter essere trasferito via rete, fu sviluppato il formato GIF (Graphic Interchange Format). Le idee alla base del formato gif sono molto semplici:

- le schede video di allora non erano in grado di visualizzare molti colori (già 256 era un buon numero). L'idea fu di estrarre per ogni immagine, dai 16 milioni di colori possibili, un sottoinsieme di 256 colori che meglio degli altri rappresentano i colori presenti nell'immagine. Tali colori possono essere rappresentati con un singolo byte, riducendo la dimensione del file di un fattore 3.
- poiché l'immagine contiene solo 256 colori, è facile avere sequenze di pixel simili in punti diversi dell'immagine: è possibile effettuare una compressione dei dati, con tecniche non molto diverse da quelle che utilizziamo per gli altri file (ad esempio con winzip).

Il risultato è un file di dimensioni contenute che ben si adatta a conservare immagini artificiali (cioè realizzate completamente al computer): ne sono esempi i pulsanti, i banner o le icone. Scegliendo opportunamente i 256 colori, infatti, un'immagine artificiale può quasi sempre essere conservata in formato GIF senza perdita di informazioni. Per un'immagine foto realistica (ad esempio una fotografia) questo non sempre è vero: basta pensare alla grande varietà di sfumature che possono essere presenti in una foto di una foresta con un cielo azzurro per rendersi conto che 256 colori sono veramente pochi.

### **Il formato PNG**

La restrizione a 256 colori visibili contemporaneamente, nonché la presenza di alcuni brevetti sugli algoritmi alla base della compressione dei dati nel formato GIF, hanno spinto la comunità a sviluppare un nuovo formato più adatto alle mutate esigenze (ad esempio lo sviluppo dell'internet come rete veloce per lo scambio delle informazioni e le schede video sempre più potenti).

Per rispondere alle mutate esigenze, nel 1995 fu introdotto il formato PNG (Portable Network Graphics). Oltre alla possibilità di visualizzare immagini sino a 48 bit di profondità di colore (281.474.976.710.656 colori diversi) ed essere libero da brevetti, il formato PNG contiene molte altre caratteristiche che lo rendono adatto alla conservazione delle immagini digitali.

Utilizzando questo formato è possibile conservare immagini con qualità (e quantità di informazioni) superiore rispetto al formato bitmap, pur mantenendo le dimensioni del file contenute.

Tali caratteristiche rendono questo formato adatto all'archiviazione di immagini per le quali è richiesta una alta qualità. Questo formato è anche adatto ad essere utilizzato nei siti web.

Sfortunatamente, l'algoritmo utilizzato per la compressione dei dati nel formato PNG dà i migliori risultati di compressione con immagini non foto realistiche.

## Formati loopy e lossless

I formati visti sino a questo momento sono particolarmente adatti a conservare immagini non foto realistiche, poiché il peso di una immagine foto realistica conservata con uno dei formati precedentemente analizzati in precedenza è molto maggiore del peso di un file contenente una immagine simile ma non foto realistica. Con l'avvento e la diffusione di scanner e macchine fotografiche digitali, però, è emersa con sempre maggiore intensità la necessità di utilizzare formati per la conservazione delle immagini digitali con alta qualità e dal peso contenuto. Se non avessimo definito questi formati, le pagine dei siti web sarebbero molto più spoglie ...

I formati visti sino a questo momento utilizzano algoritmi di compressione che conservano nell'immagine digitale tutti i dati ricevuti dalla periferica di input.

Una volta definita la risoluzione e la profondità di colore, questi algoritmi cercano di comprimere i dati (riducendo quindi il peso del file) conservando tutte le informazioni, similmente agli algoritmi usati da tanti programmi (ad esempio winzip) per comprimere un generico file. I formati che utilizzano questi algoritmi di compressione sono detti **lossless** (senza perdita di informazioni).

Nel corso degli anni, molti studi sul modo in cui noi percepiamo gli stimoli visivi, hanno evidenziato che non tutte le informazioni presenti in una immagine sono indispensabili, perché molti dettagli non sono percepiti dal nostro cervello. Questa è una differenza importante tra le immagini ed un file contenente altre informazioni, come ad esempio un file word perché in questo secondo caso tutte le informazioni devono necessariamente essere conservate. Sono stati quindi definiti degli algoritmi di compressione delle informazioni presenti nelle immagini che, sulla base delle diverse sensibilità del nostro cervello, potessero eliminare le informazioni relative ad alcuni dettagli non percepite da noi fruitori.

Eliminare questi dati non comporta alcuna perdita *apparente* di informazioni da parte del fruitore ma riduce le dimensioni del file.

Gli algoritmi di compressione che riducono le dimensioni dei file contenenti immagini o suoni eliminando i dettagli scarsamente percepibili dall'uomo sono detti **lossy** (con perdita di informazioni).

L'applicazione di un metodo di compressione dati lossy comporta una seppur minima perdita di qualità.

Decomprimendo un file compresso con un algoritmo lossy la copia ottenuta sarà peggiore dell'originale, ma in genere abbastanza simile da non comportare sostanziale perdita di informazione per il fruitore. Ciò è possibile poiché i metodi di compressione lossy tendono a scartare le informazioni poco percepibili, archiviando solo quelle essenziali.

## Il formato JPEG

Il formato lossy più usato per la conservazione delle immagini è il formato JPEG (Joint Photographic Experts Group).

L'idea dell'algoritmo alla base di questo formato è quella di eliminare dall'immagine tutte le informazioni che il nostro cervello non è in grado di percepire. Come risultato si ottiene un file molto più leggero dell'originale e che archivia una immagine molto simile (per noi) all'originale.

L'algoritmo di compressione utilizzato - basato sullo studio della percezione umana rispetto agli stimoli visivi - è particolarmente adatto alla memorizzazione di fotografie e, più in generale, di immagini foto realistiche.

In fase di salvataggio dell'immagine compressa con il formato JPEG, l'utente può decidere la qualità dell'immagine archiviata: minore è la qualità dell'immagine, più leggero risulta il file che la contiene, ma minori sono i dettagli presenti nell'immagine risultante. In Figura 28 sono presenti quattro immagini Jpeg memorizzate con qualità diverse. In Tabella 8 sono riportate, per le diverse immagini, la qualità con le quali sono state salvate, la dimensione e la riduzione di dimensione rispetto ad una immagine delle stesse dimensioni salvata in formato bitmap. Sino ad una qualità del 50% l'utente difficilmente percepisce la riduzione delle informazioni memorizzate, anche se l'immagine pesa solo il 4% dell'originale immagine bitmap.

| rif. alla Figura 28 | Qualità (%) | Dimensione (kB) | Riduzione (%) |
|---------------------|-------------|-----------------|---------------|
| -                   | bmp         | 930             | -             |
| a)                  | 100         | 160             | 83            |
| b)                  | 50          | 33              | 96            |
| c)                  | 30          | 24              | 97            |
| d)                  | 1           | 6               | 99            |

**Tabella 8: dati relativi alle immagini presenti in Figura 28**



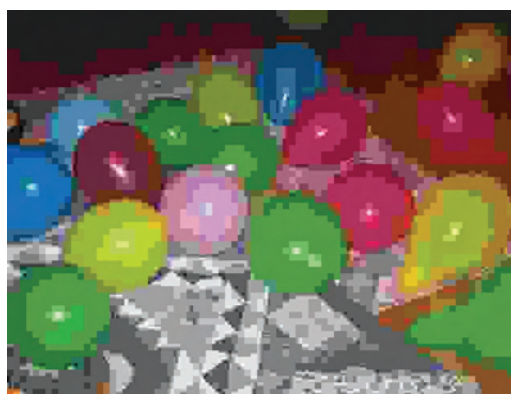
a)



b)



c)

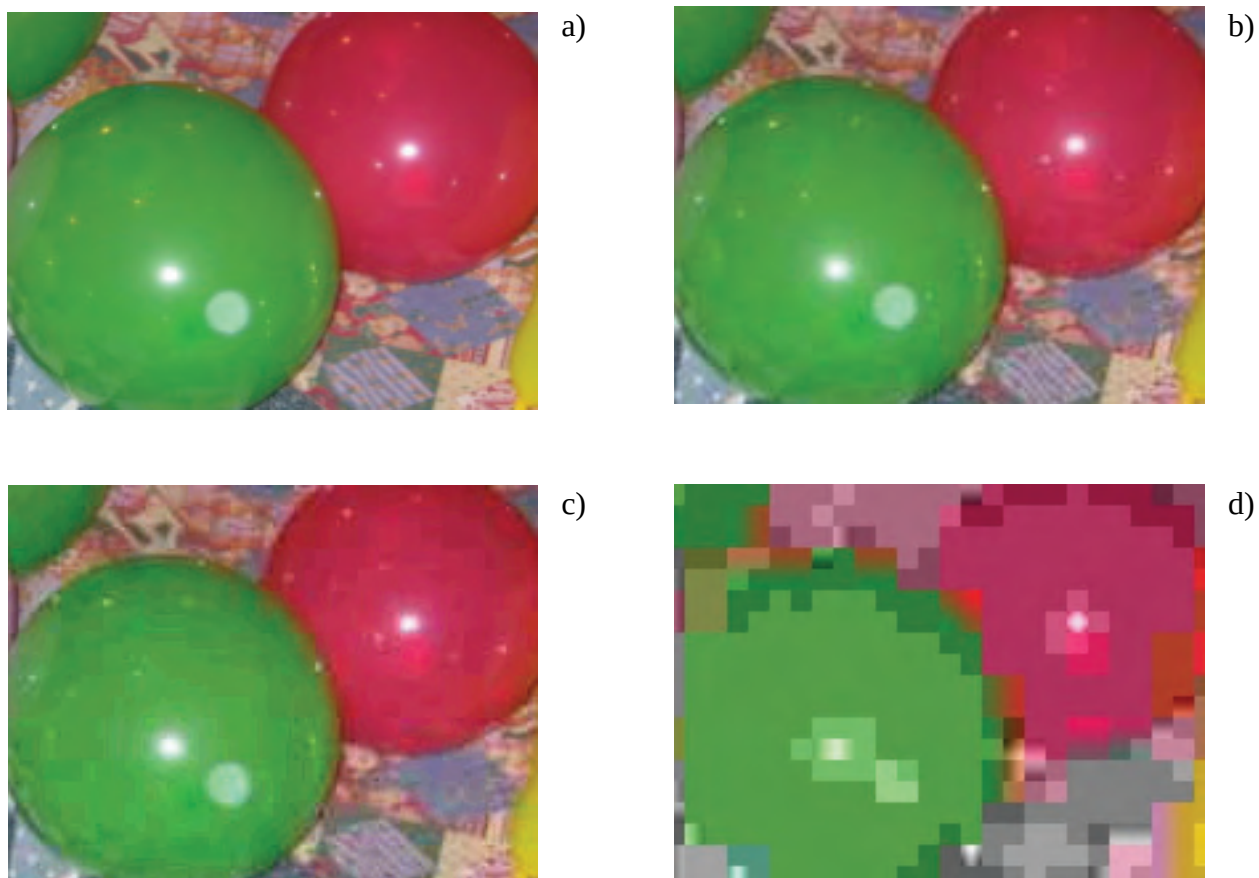


d)

**Figura 28: immagine in formato jpeg salvata con qualità differenti**

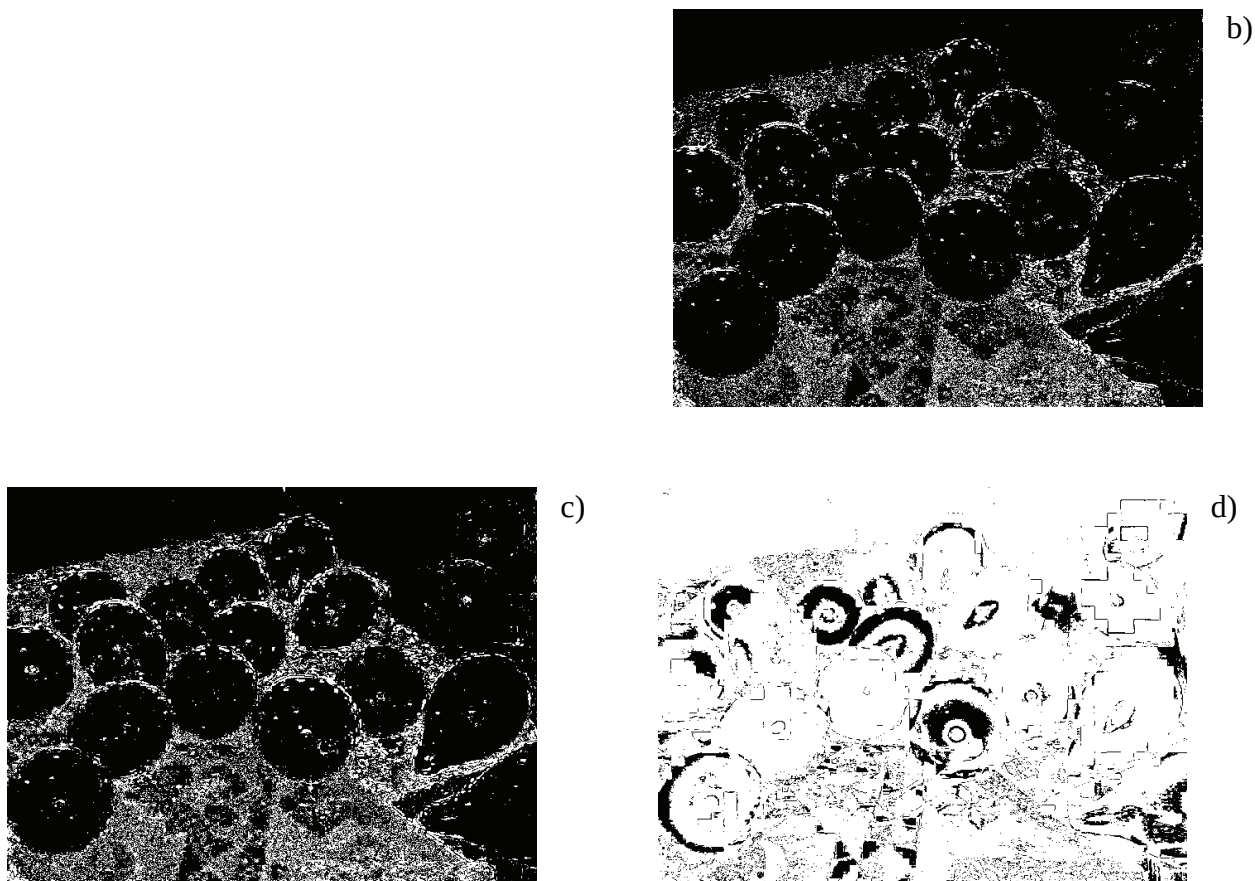
La differenza tra un file conservato con un formato lossy e quello conservato con un file lossless si evidenzia quando si effettuano degli zoom per visualizzare alcuni dettagli dell'immagine. In Figura 29 è visibile un dettaglio ingrandito delle immagini presenti in Figura 28: come visibile al

decrescere della qualità selezionata in fase di salvataggio, aumentano i dettagli e le sfumature mancanti.



**Figura 29: dettaglio delle immagini presenti in Figura 28**

Possiamo comprendere quali informazioni sono state eliminate nelle diverse compressioni confrontando, pixel per pixel, l'immagine a) con le restanti immagini presenti in Figura 28. Il risultato di questo confronto è visibile in Figura 30 dove i pixel rappresentati con il colore bianco sono i pixel che, rispettivamente per le immagini b) c) e d) differiscono sensibilmente da quelli dell'immagine a). Come si può vedere, le differenze sono più marcate nei punti di transizione tra colori molto diversi: ad esempio tra due palloncini o tra dettagli diversi della coperta. Questo conferma quantitativamente quanto visto qualitativamente in Figura 29.



**Figura 30: evidenza dei pixel differenti (in bianco) tra l'immagine a e le altre immagini presenti nella Figura 26**

La ragione di tale comportamento è legata al funzionamento dell'algoritmo di compressione lossy. Infatti, poiché in natura è raro trovare due colori molto diversi tra loro vicini, l'algoritmo utilizzato per il jpeg in caso di cambiamenti rapidi di colore non è efficace. Questa caratteristica dell'algoritmo rende il formato jpeg poco idoneo a conservare immagini non foto realistiche, dove sono possibili rapide variazioni di colore. Possiamo capire meglio questo concetto con un esempio. In Figura 31 sono riportate le immagini salvate in formato jpeg a) e gif b) di un logo. Come si vede l'immagine salvata in formato gif riproduce perfettamente il logo, mentre l'immagine salvata in formato jpeg ha una qualità molto bassa. Per quanto possa sembrare strano, ad una tale diminuzione nella qualità non corrisponde una altrettanto elevata riduzione nel peso: il file in formato gif pesa 5 volte meno del file salvato in formato jpeg! La ragione di questo risultato apparentemente incomprensibile è molto semplice: nel caso di rapidi cambiamenti di colore (in questo caso dal colore di sfondo alla scritta in bianco) l'algoritmo utilizzato per la codifica jpeg è poco efficace. Infatti, cercando di eliminare informazioni superflue nel caso di immagini foto realistiche può, nel caso di immagini non foto realistiche, produrre risultati scadenti.



a)



b)

**Figura 31: immagine non foto realistica salvata in formato jpeg a) e gif b)**

## ***I suoni***

Sino ad ora ci siamo focalizzati sulle immagini. Abbiamo quindi imparato come una qualsiasi immagine (ad esempio una fotografia) può essere trasformata – attraverso i processi di campionamento spaziale e di quantizzazione cromatica – in un insieme di bit, così da poter essere trattata da un computer.

Una cosa analoga ci aspettiamo dovrà avvenire nel caso si voglia trattare, attraverso un calcolatore elettronico, anche un contributo audio.

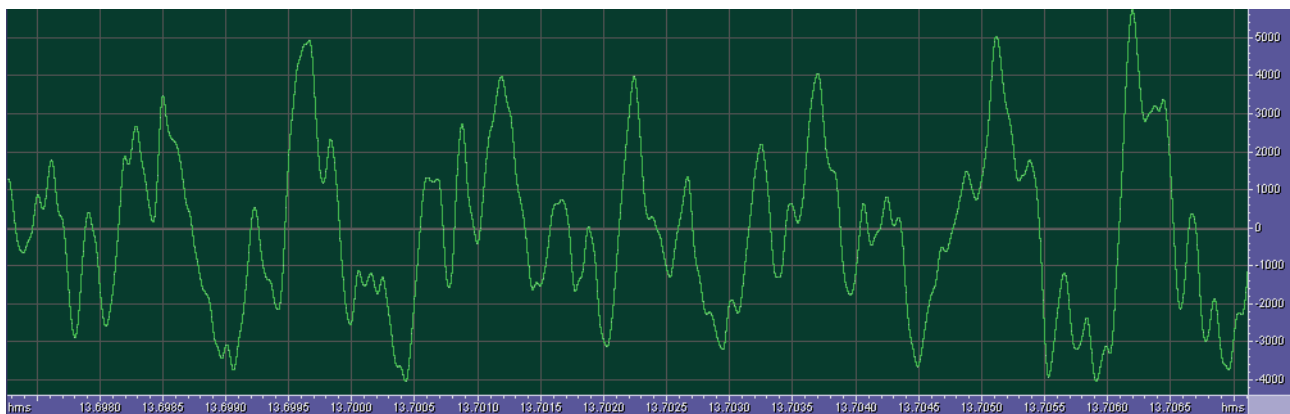
In questo caso, la prima domanda da porci è: qual è lo stimolo fisico che vogliamo digitalizzare e da cosa è prodotto?

Il suono è quello stimolo prodotto sul nostro orecchio dalla vibrazione di un corpo in oscillazione che si propaga in un mezzo (ad esempio l'aria o l'acqua).

Che cosa significa questa definizione?

Prendiamo ad esempio uno strumento musicale a corda, come una chitarra: quando pizzichiamo una corda questa comincia a vibrare e, a causa del suo moto, anche le molecole di aria intorno alla corda cominciano a muoversi in modo sincrono al movimento della corda. Il movimento di ogni molecola d'aria, spinge al movimento le altre molecole circostanti propagando il segnale; questa non è una novità: su un autobus affollato il movimento di un passeggero si propaga sui passeggeri vicini... La propagazione della vibrazione della corda attraverso l'aria, quando investe le molecole a contatto con il nostro timpano, porta alla vibrazione della membrana timpanica, questa vibrazione viene tramutata nel suono che noi percepiamo. Quando ascoltiamo un rumore, quindi, stiamo percependo la vibrazione delle molecole d'aria. Non dobbiamo pensare che il procedimento di vibrazione delle molecole d'aria richieda un tempo molto lungo per realizzarsi: un segnale acustico si propaga nell'aria a circa 1.200 Km/h!

In Figura 32 è visibile un esempio di onda che arriva ai nostri timpani (oppure sulla membrana di un microfono), come si vede il segnale ha una forma apparentemente casuale.



**Figura 32: esempio di stimolo acustico che raggiunge il nostro orecchio**

Che cosa significa digitalizzare questo segnale?

Significa trasformare la deformazione del nostro timpano (o della membrana dei un microfono) in un insieme di numeri, così da poterli trattare attraverso un computer.

Anche in questo caso, è necessario effettuare le attività di **campionamento** e di **digitalizzazione**, ma su un segnale completamente diverso rispetto a quello utilizzato per le immagini.

## Il campionamento

Nel caso delle immagini, con l'attività di campionamento spaziale, abbiamo suddiviso le due dimensioni spaziali dell'immagine (altezza e larghezza) in un insieme di intervalli; il risultato che abbiamo ottenuto sono state delle celle dell'immagine: le informazioni presenti all'interno delle singole celle le abbiamo considerate come autonome tra loro.

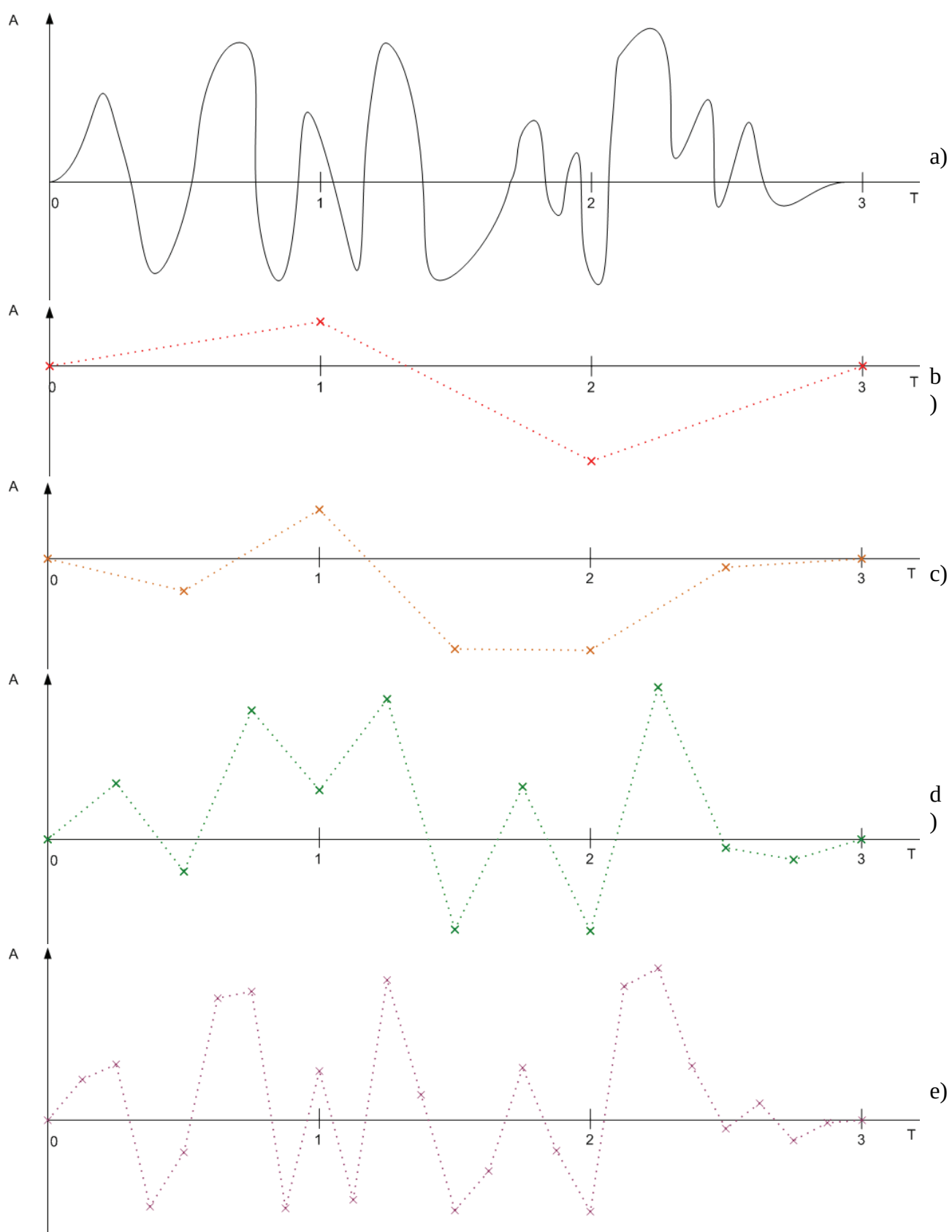
Nel caso dei suoni, l'unica dimensione nella quale il segnale si propaga è il tempo, per cui campionare un segnale sonoro significa suddividere l'asse del tempo in un insieme di intervalli regolari, dove andare a misurare l'ampiezza del segnale (che corrisponde alla deformazione della nostra membrana timpanica o alla deformazione della membrana del microfono).

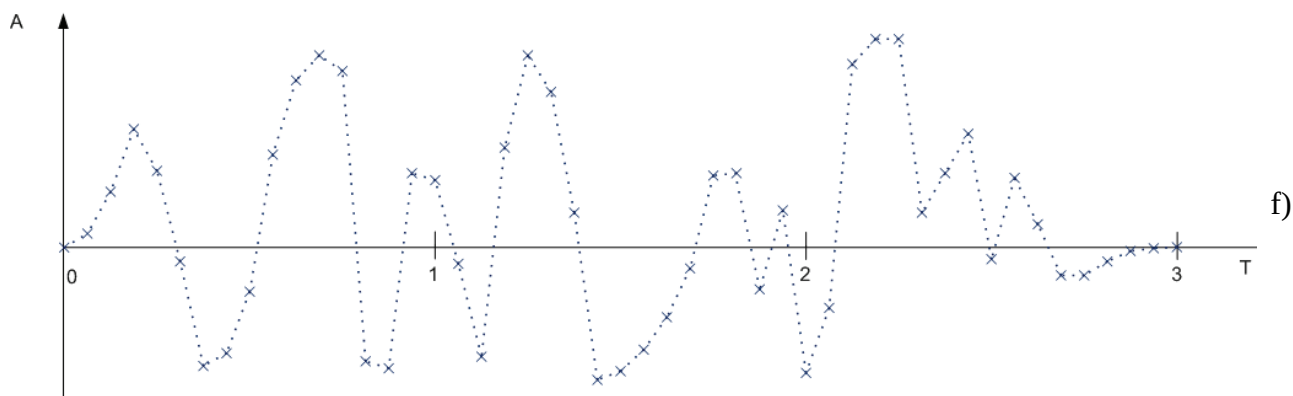
In fase di riproduzione del suono, è sufficiente inviare al dispositivo di riproduzione (ad esempio le casse audio) un impulso elettrico, con la stessa cadenza temporale con la quale è stato campionato il segnale, e di intensità proporzionale al valore dell'ampiezza del segnale.

Ancora una volta possiamo domandarci l'ampiezza degli intervalli di tempo tra i quali effettuare le misure per essere sicuri che il segnale riprodotto dalle casse sia uguale a quello che abbiamo acquisito. Anche in questo caso, tuttavia, la risposta non può che dipendere da quali informazioni presente nel suono vogliamo memorizzare: tanto minore è l'intervallo di tempo tra due misure, tanto più il segnale che conserveremo sarà simile a quello originale. Possiamo capire bene questo concetto aiutandoci con un esempio. Nell'immagine a) della Figura 33 è visibile la deformazione di una membrana che vogliamo digitalizzare. Nelle immagini b), c), d), e) ed f) sono riportate le misure (rappresentate con una x) effettuate per il campionamento ad intervalli di tempo diversi, secondo quanto riportati in Tabella 9. Come si può vedere, al crescere del numero delle misure prese (e quindi al diminuire dell'intervallo tra una misura e l'altra) il segnale campionato è più simile al segnale originale. La maggiore somiglianza tra i segnali comporta, naturalmente, una maggiore somiglianza tra il suono personale e quello percepito.

| Numero immagine | Intervallo tra le misure (in secondi) |
|-----------------|---------------------------------------|
| b)              | 1                                     |
| c)              | 0,5                                   |
| d)              | 0,25                                  |
| e)              | 0,125                                 |
| f)              | 0,0625                                |

**Tabella 9: intervalli di tempo tra le misure effettuate nelle immagini presenti in Figura 33**





**Figura 33: esempio del campionamento di un segnale ad intervalli diversi**

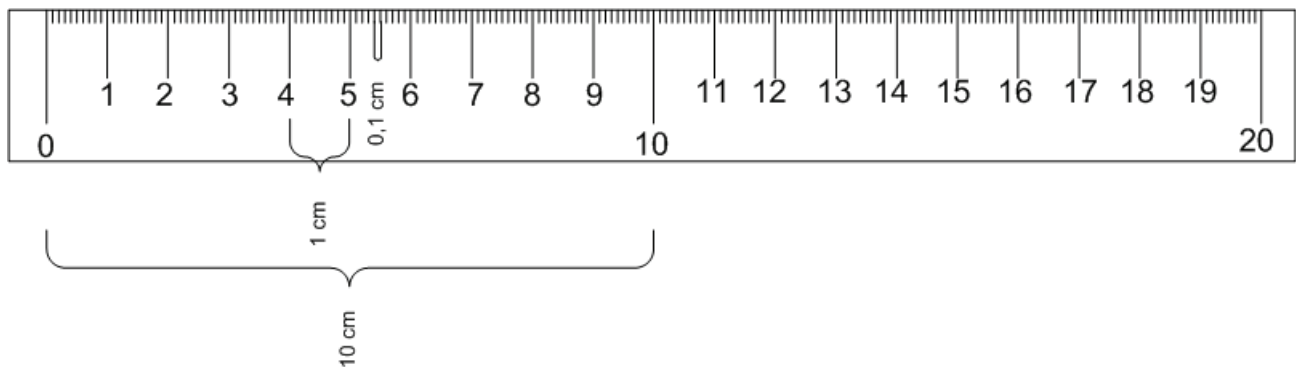
Il numero di misure da effettuare dipende dal tipo di segnale che si intende acquisire. Nel caso della musica conservata su un CD audio, il suono è stato campionato misurando l'ampiezza del segnale 44.100 volte al secondo (pari ad un intervallo tra una misura ed un'altra di 23 milionesimi di secondo)!

Ovviamente, maggiore è la frequenza di campionamento del segnale maggiori saranno i dati da conservare, diventa quindi rilevante comprendere quali informazioni portate dall'onda sonora vanno perse se il numero di misure per unità di tempo è troppo basso. Al diminuire del numero di misure effettuate in un secondo il suono risultante ha una bassa qualità nel riprodurre le componenti di alta frequenza del segnale originale, come ad esempio gli acuti le chitarre elettriche, i violini o gli organi. La scelta di 44.100 volte al secondo è legata al funzionamento dell'orecchio umano che non è in grado di percepire le differenze tra un segnale campionato 44.100 volte al secondo ed uno campionato con una frequenza più elevata.

## La quantizzazione

Una volta effettuate le misure ad intervalli di tempo regolari, i valori ottenuti devono essere memorizzati; il caso dell'audio è molto più semplice del caso delle immagini dove ad ogni pixel corrisponde un colore: per i suoni ad ogni intervallo corrisponde un numero.

Tuttavia esiste un problema comune: esattamente come ogni pixel può avere un colore tra gli infiniti possibili, ogni deformazione della membrana può avere un valore tra gli infiniti possibili; poiché memorizzare una deformazione con una precisione infinita richiederebbe un file di dimensioni infinite (esattamente come memorizzare un colore in un insieme infinito), è necessario definire un intervallo di valori all'interno dei quali scegliere quello più prossimo alla deformazione della membrana che stiamo misurando. Tanto più grande è il numero di possibili valori per la deformazione della membrana, tanto più il segnale digitalizzato sarà simile al segnale originale. Questa non è una novità: supponiamo di misurare l'altezza di un foglio di carta A4, se disponiamo di una riga che ha una i segni delle misure ogni 10 centimetri, potremmo dire che il nostro foglio è alto 30 centimetri, nulla di più. Se invece sono preseti delle tacche ogni 1 centimetro, allora la misura dell'altezza del foglio sarà ancora 30 centimetri. Se sono presenti tacche ogni 0,1 centimetri (1 millimetro), allora la misura sarà 29,7 centimetri. Se la nostra riga è lunga 50 centimetri, nel primo caso (una tacca ogni 10 centimetri) la misura dell'altezza del foglio può essere scelta tra 5 possibilità (50 centimetri / 10 centimetri), nel secondo caso (una tacca ogni centimetro), la misura può essere scelta tra 50 possibilità e nell'ultimo caso tra 500 possibilità, poiché in 50 centimetri sono presenti 500 millimetri. In Figura 34 è visibile un esempio di una riga graduata con tacche. Le tacche più alte servono per delimitare i decimetri (10 centimetri), quelle di lunghezza intermedia delimitano i centimetri e le più piccole i millimetri.



**Figura 34: esempio di una riga graduata con tacche.**

Anche in questo caso, ha poco senso cercare il valore migliore per il numero di possibili valori della deformazione della membrana da memorizzare, mentre ha senso domandarsi qual è l'effetto sul suono percepito se il numero scelto è troppo basso: se l'insieme dei valori tra cui scegliere la deformazione della membrana è troppo piccolo nel suono riprodotto è presente un rumore di fondo, una sorta di fruscio che rende meno nitido il segnale.

In generale, l'ampiezza del segnale per ogni campionamento viene rappresentata da un valore all'interno di un insieme di 65.536 valori possibili. Può sembrare bizzarro scegliere un numero così poco familiare, in realtà 65.536 è pari a  $2^{16}$ : per memorizzare l'ampiezza di una singola misura del campionamento sono necessari 16 bit. Per segnali che non necessitano di una particolare qualità per essere conservati, possono essere utilizzati 8 bit (pari a 256 livelli).

### ***I formati di conservazione dell'audio***

Nel seguito analizziamo un esempio di formato per la conservazione dei contributi audio di tipo lossless ed uno di tipo lossy.

#### **Il formato del CD audio**

Il CD audio è lo strumento per conservare i contributi audio che garantisce la migliore qualità: infatti il segnale è campionato 44.100 volte al secondo e digitalizzato a 16 bit. Inoltre il segnale è completamente stereo: i segnali per le due orecchie sono digitalizzati indipendentemente. L'eccellenza del segnale riprodotto si paga, ovviamente, in termini di spazio necessario per memorizzare tutti questi bit di informazione.

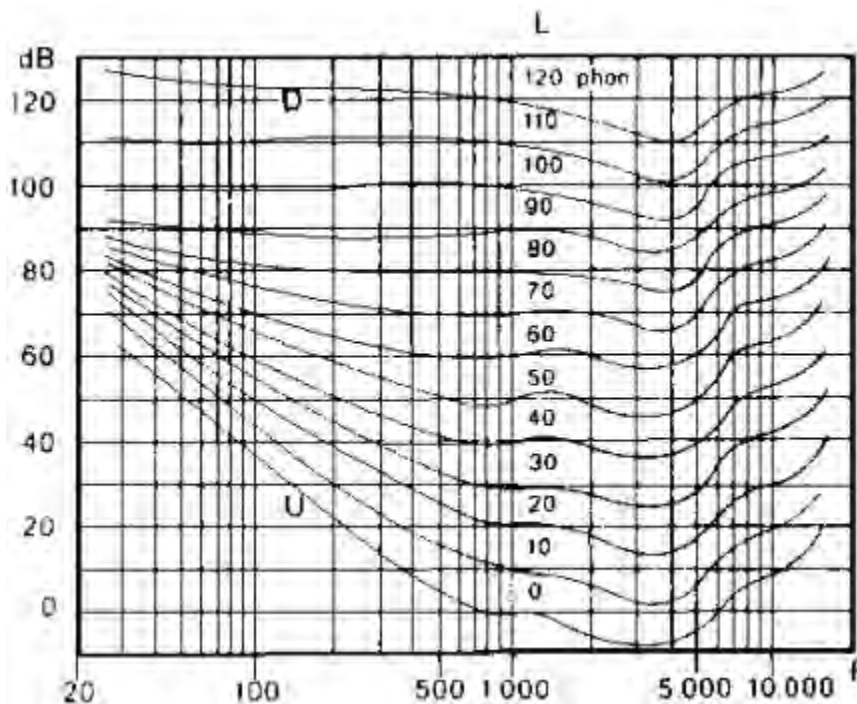
Per conservare un secondo di audio su CD sono infatti necessari:  $44.100$  (numero di campionamenti al secondo)  $\times 16$  bit  $\times 2$  (canali: destro e sinistro) = 1.411.200 bit.

Una canzone di 3 minuti pesa, così: 338.688.000 bit. Questo spiega perché su un CD musicale è possibile memorizzare poco più di un ora di musica.

Come per le immagini, un file contenente un brano audio archiviato con una qualità molto elevata richiede molto spazio e non può essere inviato via internet.

### ***I formati psicoacustici***

Per realizzare dei formati audio che consentano l'archiviazione di brani conservando una elevata qualità in fase di riproduzione ma riducendo il peso del file, sono state studiate le risposte del nostro orecchio agli stimoli acustici, così da comprendere quali informazioni sono essenziali e quali possono essere eliminate: nulla di diverso da quanto fatto per le immagini. I formati che si basano su questi studi sono detti formati psicoacustici: ovviamente sono formati lossy.



**Figura 35: rappresentazione delle curve isofoniche**

Per comprendere le linee generali di questo procedimento è possibile fare riferimento alla Figura 35. Il grafico presenta due assi: sull'asse x (quello orizzontale) sono rappresentate le diverse frequenze che possono comporre un suono, valori bassi della frequenza corrispondono a suoni simili a quelli prodotti da un tamburo o da un contrabbasso, valori alti della frequenza corrispondono a suoni simili a quelli prodotti da un violino; sull'asse y (quello verticale) è rappresentata l'intensità dei suoni che possono raggiungere il nostro orecchio. Valori bassi dell'intensità corrispondono a suoni scarsamente udibili, mentre valori alti a suoni molto intensi. Nell'immagine sono presenti 13 curve contrassegnate dai numeri da 0 a 120. Queste curve sono dette curve isofoniche, che cosa significa? Semplicemente la nostra percezione ad uno stimolo sonoro dipende dalla frequenza e dall'intensità dello stimolo stesso: un suono con valore di intensità pari a 60 dB e frequenza pari a 30 Hz, viene percepito da noi con la stessa ampiezza di un suono di intensità pari a 0 dB e frequenza pari a 1000 Hz; addirittura uno stimolo appena percettibile ad una frequenza, può essere percepito come fastidio se inviato ad una frequenza diversa. Inoltre non siamo in grado di percepire suoni che si trovano sotto la linea isofonica contrassegnata dal valore 0 (detta soglia di udibilità).

Non ha quindi senso conservare informazioni su segnali di intensità inferiore alla soglia di udibilità, eliminando molte delle informazioni presenti nel segnale. Inoltre, è possibile ridurre le informazioni memorizzate nelle zone in cui la nostra percezione è minore (ad esempio vicino alla soglia di udibilità) a favore delle informazioni presenti in zone dove il nostro sistema uditivo è molto più efficiente.

Inoltre se al nostro orecchio arrivano due suoni simili, ma uno di intensità più alta dell'altro, il nostro sistema uditivo tende a mascherare il suono di intensità più bassa. Per tale motivo, piuttosto che inviare le informazioni relative a tutte le frequenze che compongono il segnale, è sufficiente inviare informazioni relative ad alcuni segnali che mascherano gli altri.

Sempre gli studi psicoacustici hanno poi evidenziato che il nostro cervello non percepisce, per i suoni molto bassi, l'effetto stereo: è sufficiente quindi inviare alcuni segnali come stereo e gli altri come mono, risparmiando spazio.

## **Il formato MP3**

Il formato MP3 è stato il primo ad implementare un modello psicoacustico per comprimere un segnale audio.

Oltre a quanto visto il formato MP3 codifica, sulla base del tipo di segnale, pezzi diversi dello stesso brano con qualità diverse.

Infine, una volta completata la codifica, i dati vengono compressi (in modo lossless come farebbe winzip). Il risultato è un file molto più leggero ma che consente di rappresentare in modo fedele il segnale originale.

## 5 Cittadini nella società dell'informazione

Dalla seconda metà del secolo scorso si è avviata la migrazione della nostra società dalla così detta “società della materia” alla “società dell'informazione”. Questo processo migratorio terminerà quando saremo diventati una “società della conoscenza”.

Per quanto la maggior parte dei cittadini non si sia resa ancora conto delle mutazioni in atto, molti aspetti della nostra società stanno cambiando.

In questo capitolo saranno analizzate alcune di queste mutazioni.

### ***Che cos'è la società della conoscenza***

Nell'immaginario collettivo la società dell'informazione è l'idea di qualche visionario appassionato di nuove tecnologie. In realtà ogni volta che accendiamo la televisione, effettuiamo un acquisto con la carta di credito o semplicemente effettuiamo una telefonata siamo cittadini della società dell'informazione. È necessario, quindi, comprendere con esattezza che cosa si intende per società dell'informazione. Si capirà che la società dell'informazione è solo un passaggio intermedio verso la società della conoscenza.

### **La società dell'informazione**

La società nella quale viviamo oggi ha molte definizioni, ognuna scelta per evidenziare una sua caratteristica.

Si usa spesso la dizione *società post-industriale* per sottolineare la rottura con la società industriale: mentre l'espressione società industriale viene utilizzata per indicare una società creata in seguito alla rivoluzione industriale e nella quale molte delle relazioni sono condizionate dal sistema di produzione industriale, con l'espressione società post-industriale si sottolinea come il peso del mondo industriale nella società si sia fortemente ridotto a seguito dell'automazione dei processi e della delocalizzazione delle imprese industriali.

Con la dizione *società dei servizi* si intende evidenziare la prevalenza dei servizi nell'economia, a discapito dell'industria e dell'agricoltura.

Con la dizione *società dell'informazione*, infine, si intende sottolineare il valore che le informazioni ed il loro trattamento hanno assunto nella società moderna. Un esempio può aiutare a comprendere. La maggior parte delle transazioni economiche avviene ormai senza lo scambio contestuale di denaro. Quando paghiamo un acquisto con la carta di credito o con il bancomat, oppure quando effettuiamo un bonifico bancario, non stiamo scambiando denaro – e quindi materia – ma informazioni. L'esigua quantità di denaro contante nei nostri portafogli rispetto al totale delle spese che effettuiamo quotidianamente è una misura evidente della migrazione della nostra società verso una società dell'informazione.

Per realizzare una società basata sulle informazioni è necessario disporre di infrastrutture tecnologiche in grado di garantire lo scambio dei dati tra tutti gli attori. L'idea che sta alla base del progetto di costruire una società dell'informazione è ben sintetizzata dal seguente brano estratto dal rapporto ERTI (European Round table of Industrialists):

“By building these ‘**Information Highways**’ we will open the door to an information-rich society, where **every individual** can have access to **whatever information** he wants, **when and where** and in the **form that he wants** it” (Realizzando queste “autostrade dell'informazione” apriremo le porte ad una società ricca di informazioni, dove ogni individuo può avere accesso a qualsiasi informazione desideri, nei modi nei tempi e nella forma preferita).

È una definizione ambiziosa, che si basa sui principi di **indipendenza dal tempo** (in qualunque momento), **dai canali** (in qualunque forma), dal **luogo** (dovunque), e **dalla condizione sociale** (ogni individuo) per l'accesso alle informazioni.

Nell'esempio del denaro, un esempio di autostrada dell'informazione è la rete telefonica: in ogni negozio è presente un terminale (il POS – Point of Sale) che scambia informazioni attraverso la rete telefonica con la banca o con la società delle carte di credito.

Esempi di **Information Highway** sono la radio, il giornale, la televisione, la rete telefonica e la rete internet. Non tutte queste autostrade, tuttavia, verificano i requisiti indicati nella definizione. Ad esempio la televisione non rispetta il principio di indipendenza dal tempo nell'accesso alle informazioni: se il fruitore non è presente davanti allo schermo nel momento in cui una informazione viene trasmessa, questa non potrà essere più fruita; la televisione, per la grande semplicità di utilizzo, consente invece un accesso alle informazioni indipendentemente dalla condizione sociale e culturale degli individui.

Per quanto riguarda la rete internet, sicuramente l'indipendenza dal tempo è una caratteristica verificata: è possibile accedere alle informazioni presenti su un sito internet, ovvero al contenuto della propria casella di posta elettronica, in qualsiasi momento. La diffusione della rete internet, ancora molto inferiore a quella della televisione, non garantisce l'indipendenza dal luogo di accesso; è tuttavia probabile che le tecnologie basate sulla trasmissione delle informazioni senza fili dovrebbero consentire nei prossimi anni una possibilità di accesso ubiqua alla rete internet, almeno nelle nazioni più sviluppate. Deve essere sottolineato tuttavia che, i servizi offerti utilizzando la rete internet non hanno ancora raggiunto la semplicità ottenuta con i servizi radiotelevisivi, e questo rende l'indipendenza dalla condizione sociale nell'accesso alla rete internet un obiettivo lontano da raggiungere. Le problematiche legate alle difficoltà di accesso alla rete internet ed ai servizi erogati attraverso di essa vanno sotto il nome di "digital divide".

Oggi sono disponibili molteplici dispositivi che consentono il trattamento delle informazioni; tali dispositivi sono stati realizzati per meglio rispondere alle caratteristiche dell'**Information Highway** utilizzata per la trasmissione delle informazioni: la radio è stata realizzata per trasmettere suoni, la televisione suoni ed immagini, il telefono per ricevere e trasmettere suoni. La rete internet, invece, consente di ricevere e trasmettere testi, numeri, immagini e suoni: questo perché è stato definito un modo univoco per rappresentare queste tipologie di informazioni. Mentre gli altri mezzi di comunicazione trasmettono informazioni in formato analogico, la rete internet trasmette informazioni in formato digitale.

## **Che cosa è la società dell'informazione**

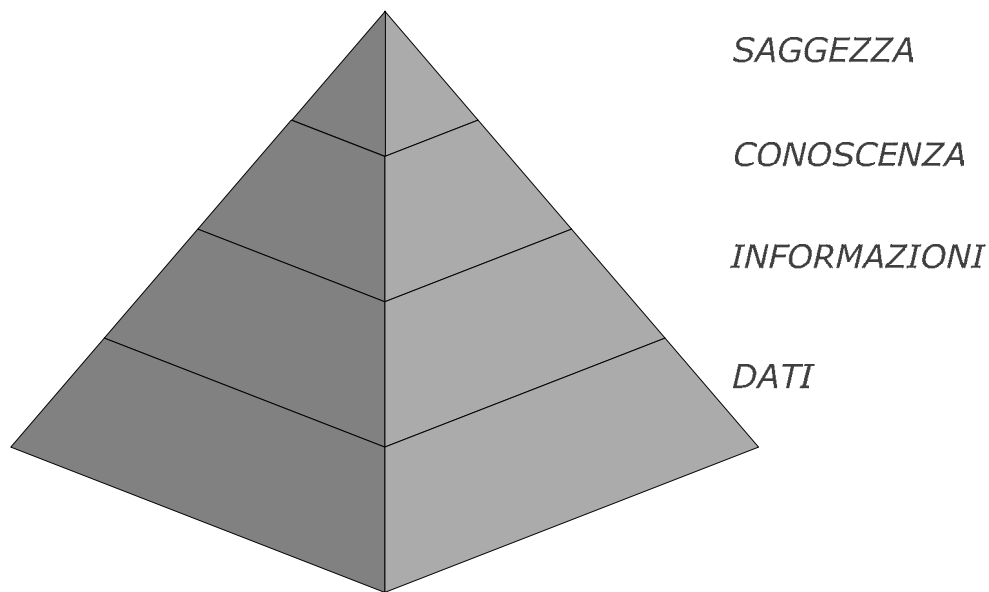
La società dell'informazione, concludendo, è una società nella quale il bene primario sono le informazioni, il loro scambio, la loro disponibilità e la capacità di trattarle. Il focus di questa società, quindi, è quello di realizzare le infrastrutture tecnologiche tanto per consentire ad ogni cittadino l'accesso quanto per fornirgli gli strumenti necessari al trattamento di queste informazioni. Lo sviluppo dei giornali, della radio, della televisione, del telefono e dell'internet sono la misura tanto dell'interesse dei cittadini in questi strumenti, quanto dell'interesse economico per investimenti del genere.

## **La società della conoscenza**

La migrazione della nostra società da società industriale a società dell'informazione non si è ancora conclusa, ed è già maturata la convinzione che l'obiettivo verso cui tendere non è la società dell'informazione ma la società della conoscenza. Per comprendere la differenza prendiamo ad esempio la piramide del knowledge management<sup>5</sup> visibile in Figura 36 ed analizzarla partendo dal basso.

---

<sup>5</sup> Con il termine knowledge management si fa riferimento a quella branca della ricerca che studia il ciclo della conoscenza all'interno di una comunità (di pratica o di apprendimento) realizzato utilizzando gli strumenti della società dell'informazione.



**Figura 36: la piramide del knowledge management**

I **dati** sono gli elementi elementari alla base della piramide; esempi di dati possono essere un numero di telefono, il valore di un titolo in borsa, la temperatura di una stanza. La quantità di dati presenti nella nostra società è immensa: basta aprire un elenco del telefono od un giornale per rendersene conto.

I dati in se non hanno né un significato né un obiettivo. Le **informazioni** (il secondo livello della piramide) sono dati che hanno un obiettivo ed un significato. Secondo Thomas Davenport e Laurence Prusak (*Working Knowledge – Harvard Business School Press, 1999*), esistono principalmente cinque modi per fornire ai dati il valore aggiunto necessario per diventare informazioni:

**contestualizzando**: cioè esplicitando il proposito per il quale i dati sono raccolti;

**classificando**: cioè evidenziando le unità o le componenti principali dei dati;

**calcolando**: cioè effettuando sui dati analisi statistiche o calcoli matematici;

**correggendo**: rimuovendo eventuali errori dai dati;

**condensando**: cioè organizzando i dati in una forma più concisa e finalizzata ad esplicitare alcune caratteristiche presenti nei dati.

Queste attività possono sembrare scontate, ma non lo sono.

Dati possono essere, ad esempio, i numeri che rappresentano le quotazioni di un titolo per ogni singolo giorno dell'anno. Si hanno informazioni quando questi numeri sono aggregati in tabelle o grafici, quando si evidenziano i guadagni o le perdite del titolo in un periodo di riferimento, ovvero quando si relazionano i dati relativi alla quotazione del titolo con i dati relativi ad altri titoli del medesimo comparto o con l'andamento della borsa nel suo insieme.

Al terzo livello della piramide troviamo la **conoscenza**. Senza la pretesa di fornire una definizione esaustiva, possiamo dire che

*la conoscenza è la consapevolezza e la comprensione di fatti, verità o informazioni ottenuti attraverso l'esperienza o l'apprendimento (a posteriori), ovvero tramite l'introspezione (a priori). La conoscenza è l'autocoscienza del possesso di informazioni connesse tra di loro, le quali, prese singolarmente, hanno un valore e un'utilità inferiori<sup>6</sup>.*

La definizione evidenzia che per avere conoscenza è necessario elaborare con la riflessione, l'analisi ed il confronto le informazioni ricevute. L'apprendimento e l'esperienza sono elementi

<sup>6</sup> La presente definizione è presa dal sito [www.wikipedia.it](http://it.wikipedia.org/wiki/Conoscenza) (<http://it.wikipedia.org/wiki/Conoscenza>).

essenziali per trasformare le informazioni in conoscenza e richiedono tanto la possibilità di accesso a molteplici fonti informative, quanto il confronto con gli altri. Solo relazionando le informazioni ricevute da molteplici fonti e confrontando le proprie idee con quelle elaborate dagli altri è possibile arrivare ad acquisire la conoscenza di un certo argomento.

Al vertice della piramide è presente la **saggezza**, nella quale la conoscenza è distillata dall'intuizione e dall'esperienza.

La migrazione dalla società dell'informazione a quella della conoscenza deve quindi focalizzarsi non più sulle infrastrutture, ma sulla possibilità di accesso alle fonti informative e sulla possibilità per ognuno diventare soggetto attivo nel processo di creazione delle informazioni, questo per fornire a tutti la possibilità di rendere disponibili le proprie idee in modo semplice ed economico.

La società dell'informazione, in cui il bene primario sono le informazioni ed il loro accesso, è solo un ponte verso la società della conoscenza, nella quale il bene primario è la conoscenza delle informazioni e non le informazioni in quanto tali.

Il Consiglio europeo riunitosi a Lisbona nel marzo 2000, ha conferito all'Unione Europea un nuovo ambizioso obiettivo: diventare entro il 2010 "l'economia basata sulla conoscenza più competitiva e dinamica del mondo, in grado di realizzare una crescita economica sostenibile con nuovi e migliori posti di lavoro e una maggiore coesione sociale."

Il rischio che si corre, nella migrazione dalla società dell'informazione alla società della conoscenza, è quello di sviluppare una società del *prêt-à-porter* dell'informazione. Tale società sarebbe un ponte verso la società della misconoscenza.

## **Evoluzione dei mezzi di comunicazione**

Una volta compresa la distinzione tra società dell'informazione e società della conoscenza analizziamo ora l'evoluzione del paradigma alla base dei principali mezzi di comunicazione e di come questa evoluzione possa abilitare l'evoluzione della nostra società verso una società della conoscenza. A questo scopo è necessario definire il concetto di multimedialità.

Si può definire la multimedialità come la compresenza e l'interazione tra più mezzi di comunicazione presenti in uno stesso supporto o contesto informativo. Esempi di multimediali sono:

- la televisione, nella quale il contenuto informativo è composto da immagini in movimento ed suoni,
- il giornale, nel quale il contenuto informativo è composto da immagini e testo scritto,
- il world wide web, nel quale il contenuto informativo è composto da testo scritto, immagini (fisse e/o in movimento) e suoni.

La radio, quindi, non è un esempio di multimedia in senso stretto poiché il messaggio informativo è composto esclusivamente da audio.

Non va dimenticato che – secondo la moderna teoria della comunicazione – la responsabilità nella trasmissione di un messaggio è esclusivamente dell'emittente che deve predisporre le informazioni da trasmettere sulla base tanto del canale di trasmissione disponibile quanto delle caratteristiche del ricevente. In questo contesto culturale e date le caratteristiche di apprendimento proprie degli esseri umani, la multimedialità si candida come modello principe nella trasmissione efficace di un contenuto informativo.

## **La multimedialità diffusiva**

Il mezzo di comunicazione di massa per eccellenza è, ancora oggi, la televisione; questo mezzo utilizza un flusso delle informazioni unidirezionale: i produttori definiscono quali informazioni trasmettere e l'utente può solo recepirle nei tempi e nei modi scelti dal produttore. L'unica libertà di scelta, per l'utente, è legata alla pluralità delle fonti di informazione: se non piace ciò che trasmette un canale televisivo è possibile cambiare canale. Il ristretto numero delle fonti informative, dovuto agli ingenti investimenti necessari ed ai limiti tecnologici degli strumenti, ha limitato la libertà degli utenti. Solo negli ultimi anni, grazie alla televisione satellitare ed al digitale terrestre, il numero di fonti è aumentato.

Alla base di strumenti quali la televisione c'è il paradigma della **multimedialità diffusiva**: l'utente riceve informazioni attraverso media diversi, ma senza poter effettuare alcuna personalizzazione o fornire feedback immediato ai produttori di contenuti.

L'elemento paradigmatico che caratterizza questo tipo di multimedialità è l'unidirezionalità del flusso delle informazioni. Il produttore invia attraverso l'etere<sup>7</sup> i propri programmi senza avere alcun riscontro sul numero di utenti che fruiscono del programma né sul livello di gradimento dello stesso; l'utente può solo fruire di uno dei programmi che riceve nel momento in cui accende l'apparecchio televisivo.

Sino a pochi anni fa le rare eccezioni a questo paradigma, che avrebbero potuto cambiare il modo di utilizzare gli strumenti, hanno avuto scarse conseguenze. Adriano Celentano, ad esempio, nel 1987 durante Fantastico 8 (programma serale del sabato) "ordinò" di spegnere ed accendere le luci in tutte le case in segno di protesta contro la caccia. Otto milioni di italiani (su 12 che in media seguivano la trasmissione) pensarono la televisione.

Questa azione di Celentano fu profondamente criticata. Molti sostengono che, alla base di tali critiche, ci fosse la constatazione del potere persuasivo che il medium ha sulle masse culturalmente non sviluppate. L'"esperimento" evidenziava in modo plateale il potere del medium, ed evidenziava possibili strade per influenzare i contenuti trasmessi. Questo evento però finì per essere ricordato come un fatto di costume realizzato da un artista fuori dal coro.

Per ovviare a questa limitazione del paradigma, è stato necessario utilizzare altri canali per fornire ai produttori dei contenuti il necessario feedback.

Per rilevare le abitudini di fruizione della televisione nel 1984 è nata l'Auditel. Questa società fornisce ad un campione di italiani rappresentativi della popolazione, dei dispositivi che trasmettono via telefono alcune informazioni come il programma in fruizione ed il numero di persone che lo stanno guardando. In questo modo è possibile per le società televisive ricevere delle informazioni medie sul numero di utenti che fruiscono dei diversi programmi televisivi.

Negli ultimi anni si sono consolidati nuovi strumenti per fornire feedback massivi alle trasmissioni televisive (o radiofoniche). Tali strumenti si basano sull'utilizzo:

- del telefono (voce o SMS) per effettuare televoti o indicare una scelta
- dell'instant pool presenti sui siti internet delle società televisive per fornire un feedback in tempo reale sul gradimento
- degli SMS per riportare commenti molto stringati
- della posta elettronica per fornire commenti più articolati.

Ovviamente per poter utilizzare gli strumenti afferenti alle ultime due voci dell'elenco è necessario disporre di una redazione che filtri ed aggrega le informazioni ricevute.

Merita di essere sottolineato che spesso questi strumenti non sono utilizzati per ricevere commenti e migliorare il prodotto televisivo realizzato, bensì per obiettivi commerciali e di immagine. La possibilità di scegliere il candidato da eliminare/premiare durante un reality show o un programma televisivo usando il televoto, nasconde spesso l'obiettivo di trarre profitti molto rilevanti senza che l'utente possa influenzare la qualità del prodotto fruito: quando il costo di una telefonata o di un SMS è – per l'utente – maggiore di quello previsto dal suo piano tariffario, la rete televisiva riceve dal gestore telefonico una percentuale molto alta del prezzo pagato dall'utente che telefona (o invia un SMS) per dare il proprio voto.

Casi di successo nel fornire un feedback alle trasmissioni si hanno spesso nella radio, dove gli utenti possono influenzare le playlist di molte radio con le proprie proposte, consentendo al produttore di

---

<sup>7</sup> Si usa il termine *etere* per indicare il mezzo attraverso il quale si propagano le onde elettromagnetiche emesse dai ripetitori delle radio o delle televisioni. In realtà tale mezzo non esiste poiché le onde elettromagnetiche possono propagarsi tanto nella materia quanto nel vuoto (ad esempio la luce solare che giunge sino a noi attraverso lo spazio vuoto). L'idea dell'etere, introdotta sul finire del XIX secolo, fu abbandonata nel 1905 con la pubblicazione della teoria della relatività ristretta, rimane ancora nel dire comune per indicare lo spazio nel quale si propagano le radiazioni elettromagnetiche.

personalizzare i contenuti erogati sulla base dei desideri degli utenti che stanno fruendo in tempo reale dei contenuti.

Nonostante i limiti sino ad ora esposti, giornali, radio e televisione sono state un elemento di rapida democratizzazione nell'accesso alle fonti informative e di crescita delle classi meno colte: probabilmente il più importante nella storia dell'uomo. La radio è stata, durante la seconda guerra mondiale, l'unico strumento per l'accesso alle informazioni (sia di propaganda che di resistenza) e la televisione, nel dopoguerra, è stato il principale strumento di aggregazione e crescita della lingua e della cultura popolare italiana.

Uno degli elementi che ha consentito il successo della radio prima e della televisione poi è stata la grande semplicità di utilizzo: è sufficiente accendere l'apparecchio e scegliere il programma per fruire dei contenuti. Inoltre la diffusione di questi mass media è stata favorita dai bassi costi di ampliamento delle reti: non richiedendo il cablaggio dei singoli appartamenti, chiunque può fruire dei programmi radio o televisivi semplicemente acquistando un apparecchio per la ricezione ed eventualmente installando un'antenna. Queste caratteristiche di semplicità d'uso e di economicità hanno reso la radio e la televisione i mass media caratterizzanti lo sviluppo della società dell'informazione.

Nel corso della seconda metà del secolo scorso, però, gli utenti hanno iniziato a percepire i limiti dei mezzi di comunicazione che utilizzano il paradigma della multimedialità diffusiva. I vincoli spazio temporali imposti dai media che si basano su questo paradigma, unito alla sempre minore qualità dei palinsesti hanno progressivamente ridotto, infatti, l'appeal di questi media in fasce sempre crescenti della popolazione: già nel 1992 Bruce Springsteen cantava *57 Channels (And Nothin' On)* [57 canali (e nulla dentro)].

### **La multimedialità interattiva: un nuovo ruolo per gli utenti**

L'avvento dell'internet ha cambiato il paradigma alla base del processo di comunicazione delle informazioni. Al paradigma della multimedialità diffusiva si sostituisce il paradigma della **multimedialità interattiva**. L'elemento paradigmatico che lo caratterizza e lo differenzia dal precedente è la bidirezionalità dei flussi informativi. Non più un produttore che eroga i contenuti ed un fruitore che li riceve nei modi e nei tempi decisi dal produttore, ma la possibilità di uno scambio bidirezionale di informazioni.

Questo nuovo tipo di multimedialità riporta l'utente al centro del processo di comunicazione, poiché egli può scegliere tanto le fonti da cui attingere le informazioni quanto le informazioni stesse. Tali input, poi, sono acquisiti in tempo reale dal produttore di contenuti che può offrire, sulla base degli input ricevuti, l'informazione desiderata.

Per avere multimedialità interattiva non è necessario che la quantità di informazioni scambiata dai due utenti sia simile: quando si fruisce di un contenuto disponibile sulla rete internet, l'utente si limita spesso ad effettuare un click su un link per fruire di un testo, una immagine o un filmato. In questo caso pochi bit sono stati inviati dal nostro computer al server web, ma questo ci ha risposto con molti bit di informazioni; l'aspetto rivoluzionario rispetto – ad esempio – alla televisione è che l'utente ha definito in modo autonomo tanto i contenuti da fruire, quanto i modi ed i tempi di fruizione.

Nella multimedialità interattiva la console di regia torna nelle mani dell'utente, poiché questo sceglie tempo, modo e forma per l'accesso alle informazioni di suo interesse.

L'internet<sup>8</sup> è, oggi, il canale principe nel quale veicolare i contenuti secondo il paradigma della multimedialità interattiva. Le caratteristiche della rete, unita al basso costo necessario alla realizzazione dei contenitori per le informazioni (i siti web) hanno consentito il rapido diffondersi delle idee e delle informazioni prodotte da ogni singolo utente.

---

<sup>8</sup> La televisione on-demand (o interattiva) offerta oggi da diversi operatori come Fastweb o Telecom Italia si basa chiaramente sulla multimedialità interattiva; questo però non deve stupire: le informazioni sono trasmesse attraverso l'internet.

Inoltre nel corso degli ultimi anni si sono affermati molteplici servizi internet quali ad esempio i blog, i forum, le chat, YouTube, MySpace o altri che consentono un elevato livello di interazione e confronto a livello di singolo utente.

L'utente smette di essere **spettatore** per diventare **spettatore**.

Anche la definizione di emittitore e di ricevitore non può più essere interpretata in senso stretto, poiché i due attori si alternano nei ruoli.

La moltitudine delle fonti informative cui accedere, unita alla possibilità per ognuno di fornire il proprio contributo e di confrontarsi con gli altri candida l'internet a diventare uno dei principali canali per trasformare la nostra società da una società dell'informazione a quella della conoscenza.

## ***L'architettura ipermediale***

Nella società della materia il libro è stato lo strumento principe per trasmettere le informazioni. La crisi di questo strumento con l'avvento dei media che implementano il paradigma della multimedialità diffusiva è sotto gli occhi di tutti. Tali media (ed in particolare la televisione) hanno, rispetto al libro, il solo vantaggio di veicolare il messaggio in forma di audio e di video ma vincolano l'utente ad una fruizione sequenziale e non consentono l'interazione e lo scambio. Nella seconda metà del secolo scorso si sono sviluppati nuovi modelli di organizzazione delle informazioni più adatti alla società della conoscenza.

## **Testo, ipertesto ed ipermedia**

In questo paragrafo analizzeremo che cos'è un testo classico e come, tale testo, può evolversi verso un ipertesto e quest'ultimo in un ipermedia.

### **Il testo**

Pensiamo a come leggiamo un libro, ad esempio un romanzo. Iniziamo dalla prima pagina e continuiamo leggendo in successione le singole pagine che lo compongono; l'autore, quando ha pensato al suo libro, ha steso lungo una sequenza narrativa tutti i contenuti che voleva trattare (ad esempio la presentazione dei personaggi, gli eventi critici o gli intrecci tra i personaggi) e successivamente li ha trattati in modo sequenziale nelle diverse pagine. Non importa se la sequenza narrativa ha dei salti di tempo e di spazio, l'unico requisito per un buon romanzo è quello di mettere il lettore sempre in grado di comprendere (o almeno intuire) cosa stia succedendo sulla base delle informazioni fornite nelle precedenti pagine.

Un romanziere che volesse – ad esempio – rendere evidente ai lettori il legame che unisce il personaggio di un suo libro alla casa nella quale ha passato l'infanzia, potrebbe far ricordare al suo personaggio eventi accaduti molti anni prima. In questo caso si ha una evidente rottura della linea temporale, poiché si passa dalla narrazione del presente a quella del passato, ma una continuità della sequenza narrativa, poiché al termine della lettura il lettore ha compreso il legame tra il personaggio e la sua casa ed è quindi in grado di comprendere meglio gli eventi seguenti.

Nel romanzo “La luna e i Falò” di Cesare Pavese si alternano tre linee temporali distinte con continui salti tra tempi e luoghi diversi, nonostante questo è presente una sola sequenza narrativa che l'autore ha definito e seguito nella stesura del romanzo.

Per violare la continuità della sequenza narrativa di un romanzo è sufficiente leggere le singole pagine che lo compongono secondo un ordine casuale e non predeterminato: il risultato che si ottiene non avrà nulla di soddisfacente, a meno che l'autore non abbia realizzato un romanzo privo della sequenza narrativa.

In realtà alcune violazioni della sequenza narrativa le commettiamo spesso: quando dimentichiamo un personaggio del romanzo dobbiamo andare a cercarlo a ritroso per ripassare chi sia.

Se invece di un romanzo pensiamo ad un manuale (ad esempio un libro di matematica o di informatica), anche in questo caso l'autore ha previsto una sorta di “sequenza narrativa”, cioè un percorso che – partendo dalla prima pagina – accompagna il lettore meno competente a comprendere i diversi contenuti presenti nel libro. Questo percorso, per quanto sia il principale, non

è l'unico; spesso infatti, nell'introduzione a molti manuali è riportata l'indicazione che – a seconda delle necessità o delle competenze del lettore – è possibile leggere solo alcuni capitoli saltandone altri. Inoltre, il lettore più esperto alla ricerca di informazioni specifiche su un certo argomento, invece di leggere tutto il libro può trovare nell'indice a quale pagina è trattato un argomento: in questo modo accede direttamente all'informazione di interesse. Nei manuali, quindi, i diversi indici presenti hanno la funzione di consentire un accesso più mirato alle singole informazioni di interesse. Spesso, poi, nei manuali è presente un piccolo vocabolario con lo scopo di fornire al lettore un approfondimento contestualizzato e di rapido accesso su alcuni concetti o parole che possono risultare meno facili da comprendere. Anche questo strumento consente un accesso alle informazioni non sequenziale.

Per rendere questi strumenti di accesso causale alle informazioni realmente utili, è necessario che l'autore del libro organizzi i suoi contenuti rendendo meno forte il legame tra i diversi concetti: altrimenti l'accesso diretto ad un argomento senza prima aver fruito dei precedenti risulterebbe poco efficace.

All'estremo opposto rispetto ai romanzi sono i vocabolari o le enciclopedie, dove ogni concetto è presentato come un'isola separata dalle altre, ma connessa attraverso riferimenti logici.

Tale modello comunicativo, non più vincolato ad una sequenza predeterminata di “prima” e “dopo”, non può essere considerata come una innovazione dei nostri tempi. La prima enciclopedia intesa come compendio dell'intero corpus delle conoscenze di un'epoca nacque infatti nel 1751: l'opera *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers* (Enciclopedia o Dizionario ragionato delle scienze, delle arti e dei mestieri) pubblicata da Diderot e D'Alambert fu realizzata da un elevato numero di esperti delle diverse discipline; ognuno produsse i contenuti relativi ai termini dei quali aveva maggiore competenza ma, poiché molti concetti richiamaavano altri concetti trattati nella stessa opera, il risultato finale fu che nell'*Encyclopédie* è presente una struttura reticolare di rimandi tra i diversi concetti. Per quest'opera non è ipotizzabile un accesso sequenziale, un accesso – cioè – nel quale il lettore inizia a leggere il contenuto della prima pagina del primo volume e prosegue sino all'ultima pagina dell'ultimo volume. Le ragioni di tale impossibilità sono principalmente due.

La prima ragione è di carattere pratico: data la mole dell'opera (17 volumi di voci, 11 di illustrazioni, 5 di materiale supplementare e 2 di indici) leggerla tutta richiederebbe un tempo sproporzionato.

La seconda ragione è di tipo cognitivo: se il modello organizzazione delle informazioni alla base dell'opera è di tipo reticolare e non prima-dopo, accedere con il modello sbagliato è inefficace come leggere un romanzo scegliendo le pagine in modo casuale. Possiamo capire bene questo concetto con un esempio: supponiamo di leggere in modo sequenziale un'enciclopedia e di arrivare alla voce **adrone**. L'autore ci spiega che “In fisica, un **adrone** è una **particella subatomica** che può interagire attraverso **interazioni forti**. Non sono particelle fondamentali, ma sono composti da **fermioni** come **quark** e **antiquark**, e **bosoni** come i **gluoni**”. A meno che il lettore non sia un esperto di fisica, troverà questa definizione incomprensibile poiché non ha idea di cosa sia – ad esempio – una particella subatomica, un quark o un gluone (in rosso sono evidenziate le parole meno comuni). Per comprendere quindi la definizione di adrone, dovrebbe aspettare di arrivare a leggere il significato della parola quark (tra le parole evidenziate in rosso è l'ultima che si incontrerebbe ad una lettura sequenziale dell'opera), sempre sperando che – nelle definizioni delle parole sconosciute – non si faccia riferimento ad altri concetti.

L'approccio migliore per comprendere che cosa sia un adrone, quindi, è quello di comprendere innanzitutto che cosa sia una particella subatomica e, successivamente, che cosa significa interazioni forti quindi continuare con gli altri concetti.

Il modello di organizzazione delle informazioni basato su una rete di relazioni tra i diversi concetti è detto causale (in inglese random).

Anche se D'Alambert non conosceva gli adroni (scoperti solo nel secolo scorso), le idee su come realizzare un'enciclopedia le aveva molto chiare. Scrisse infatti, parlando della sua opera: “una

sorta di mappa del mondo diretta a mostrare le contrade principali, la loro posizione ed interdipendenza ed il cammino che conduce direttamente dall'una all'altra".

Sulla base di quanto discusso è evidente che il libro, nella sua fisicità, è uno strumento adatto a conservare informazioni organizzate secondo una logica di tipo sequenziale ma inadatto nel consentire all'utente di fruire di contributi informativi organizzati secondo una logica di tipo random; nel corso del secolo passato diversi scienziati hanno cercato un sistema nuovo che consentisse una fruizione delle informazioni organizzate secondo una logica di tipo random in modo semplice come leggere un libro.

Perché abbandonare il libro che, per migliaia di anni, ha rappresentato il modo più veloce ed economico per conservare la conoscenza di un popolo e per consentirne la condivisione tra popoli e tra generazioni? Dalla prima stamperia di Gutenberg nel 1455 ad oggi il libro ha consentito alla nostra cultura uno sviluppo senza precedenti, perché quindi disconoscere i meriti dello strumento?

Queste domande sono sensate: nessuno pensa a sostituire il libro, ma facciamo un semplice ragionamento. I nostri processi cognitivi sono di tipo sequenziale come un libro? Sicuramente no: il pensiero procede per associazioni logiche e collegamenti a rete, spesso stiamo pensando a qualcosa e finiamo a pensare a tutt'altro perché il nostro cervello ha fatto un salto logico. Il libro non è adatto a questo, poiché vincola il pensiero ad uno svolgimento lineare e sequenziale che non gli è proprio. La ricerca di nuovi strumenti per abilitare nuove logiche di fruizione è legata, quindi, al tentativo di superare i limiti del libro. Anche in questo caso non siamo davanti ad una novità: la televisione è uno strumento che ha consentito di superare taluni limiti del libro consentendo all'utente di fruire non solo dei testi scritti, ma anche delle immagini e dei suoni.

## L'ipertesto

Una logica che si basa su un principio di accesso causale alle informazioni è la logica ipertestuale, i prodotti che sono realizzati secondo la logica ipertestuale sono detti ipertesti. Tra le possibili definizioni di ipertesto ne vengono riportate due, quella di Ted Nelson che utilizzò per la prima volta questo termine nel 1965, e quella di George P. Landow.

Iniziamo con la definizione di Nelson:

Con "ipertesto" intendo scrittura non sequenziale, testo che si dirama e consente al lettore di scegliere qualcosa che si fruisce al meglio davanti uno schermo interattivo. Così come è comunemente inteso, un ipertesto è una serie di brani di testo tra cui sono definiti legami che consentono al lettore differenti cammini

Mentre per Landow:

Possiamo definire l'ipertesto come l'uso del computer per superare le caratteristiche di linearità, limitatezza e fissità del testo scritto tradizionale. A differenza della forma statica del libro, un ipertesto può essere composto e letto in modo non sequenziale si tratta di una struttura variabile, composta da blocchi di testo [...] e da legami elettronici che li congiungono

Entrambe le definizioni fanno leva sui due aspetti che abbiamo evidenziato sino a questo momento: il primo è che un ipertesto è composto da un'insieme di unità testuali connesse tra loro, il secondo è che un libro non è adatto ad abilitare una fruizione ipertestuale di un'opera.

In cosa si discosta un testo da un ipertesto? In fondo entrambe sono sequenze ordinate di caratteri. In realtà le differenze ci sono e sono molto evidenti. Innanzitutto i contenuti di un ipertesto, così come i contenuti di un enciclopedia, non devono utilizzare la logica del prima-dopo, ma devono essere organizzati come un corpus di frammenti; ogni frammento è un nodo nella rete ipertestuale che si viene a creare. In secondo luogo, i singoli nodi di cui è composto un ipertesto non costituiscono mai elementi isolati ma fanno parte di una struttura reticolare: i singoli nodi sono collegati tra loro attraverso dei collegamenti (in inglese link): le maglie della struttura reticolare. Il link è ciò che rende l'ipertesto un mezzo di comunicazione esclusivo, "un documento che contiene non più solo il testo ma anche la rete dei suoi collegamenti intertestuali" (come indicato Bettetini, Gasparini e Vittadini in "Gli spazi dell'ipertesto").

I link tuttavia non sono visibili sullo schermo poiché costituiscono la struttura connettiva dell'ipertesto, l'architettura concettuale degli elementi. Ciò che manifesta al lettore le relazioni che

il singolo nodo intrattiene con gli altri nodi della rete ipertestuale sono le ancore (o hot word in inglese) che rappresentano i punti di partenza del legame. Selezionando un ancora l'utente può accedere al contenuto che il realizzatore dell'ipertesto ha ritenuto opportuno collegare al frammento che stavamo fruendo.

Gli unici strumenti in grado di erogare contenuti basati sulla logica ipertestuale sono, ad oggi, gli elaboratori elettronici.

## **L'ipermedia**

Nel corso degli anni si è assistito ad una evoluzione del concetto di ipertesto, poiché si sono evolute le possibilità di elaborazione dei calcolatori elettronici. Spesso, infatti, un contenuto ipertestuale non è composto solo di caratteri ma presenta anche immagini, suoni, filmati o animazioni questa evoluzione multimediale dell'ipertesto è detta ipermedia. Con questo neologismo, composto dalle parole iper e media, si intende evidenziare un accesso non lineare (il prefisso iper) a informazioni presentate in forma multimediale.

## **L'ipermedialità e la società della conoscenza**

L'ipermedia, per quanto naturale evoluzione del concetto di ipertesto, costituisce un elemento abilitante per la società della conoscenza per molteplici ragioni.

In primo luogo i nodi che costituiscono l'ipermedia possono essere realizzati da molteplici autori: se pensiamo ai contributi presenti nel web, ognuno può collegare alle pagine che costituiscono il suo sito contributi (testi, immagini, filmati) presenti su altri siti. In questo modo si favorisce la pluralità nell'accesso alle fonti informative.

In secondo luogo un ipermedia è, per sua natura, una struttura aperta alla quale ognuno può aggiungere contributi o ampliare i contributi presenti. Questo favorisce il confronto tra gli utenti e quindi lo sviluppo della conoscenza.

Nel corso degli ultimi anni si sono sviluppate alcune applicazioni web che consentono un elevato livello di interazione tra gli utenti e che, senza la necessaria massa critica di utenti, non avrebbero alcun futuro. Esempi di tali applicazioni sono i blog, i forum, le chat, YouTube, MySpace o Wikipedia. In generale le applicazioni che richiedono un elevato livello di interazione tra gli utenti sono chiamate Web 2.0.

## **Informazioni e società della conoscenza**

Come introdotto precedentemente, le informazioni veicolate dalle autostrade dell'informazione sono in formato digitale. In particolare, ogni informazione è codificata sotto forma di una sequenza ordinata di bit. Utilizzando questo formato i singoli elaboratori elettronici sono in grado di trattare i dati in modo efficiente ed efficace. Per i nostri scopi è importante evidenziare quali siano le differenze fondamentali tra una informazione in formato analogico ed una informazione in formato digitale, questo per rendere evidenti le relazioni tra le possibilità di trattamento delle informazioni nella società della conoscenza.

## **Informazioni in formato digitale**

Nel seguito sono discusse le principali differenze tra il formato analogico e quello digitale nel trattamento delle informazioni, a questo scopo ci focalizzeremo su un caso concreto. Per facilità di comprensione faremo il confronto tra un documento in formato elettronico (documento digitale) e la sua stampa su carta (documento analogico). Nel documento sono presenti informazioni otto forme di testo, immagini, tabelle e collegamenti intra-testuali. Quanto vedremo, comunque, rimane valido indipendentemente dalle informazioni conservate nel file. Non analizzeremo le differenze legate al piacere del possesso di un libro, alla facilità di lettura di un testo cartaceo rispetto ad un testo a video o al piacere del sentire il profumo della carta.

## **Formato esteriore del documento**

Una prima ed importante differenza risiede nella forma con cui sono rappresentate le informazioni. Nel caso del documento cartaceo le informazioni sono costituite dall'inchiostro presente sulla carta al termine del processo di stampa. Se la stampante utilizzata è di tipo monocromatico ogni informazione legata ai colori (ad esempio nelle immagini) va persa. Anche i collegamenti intra-testuali, per le ragioni viste in precedenza, non sono conservati nel documento analogico.

La versione digitale del documento è, invece, composta di bit racchiusi in file: il file che contiene il documento è una sequenza ordinata più o meno lunga di uni e di zeri. Questi bit rappresentano ogni informazione presente nel documento.

## **Facilità di accesso ai contenuti**

La facilità di accesso alle informazioni presenti nel documento cartaceo è evidente. È sufficiente saper leggere per poter fruire dei contenuti.

Nel caso dell'accesso alle informazioni presenti nel documento digitale la situazione è più complessa. Infatti la sequenza ordinata di bit è incomprensibile all'uomo e la fruizione può avvenire esclusivamente se mediata da un dispositivo elettronico in grado tanto di comprendere il significato dei bit che compongono il documento, quanto di presentare i contenuti all'utente. Questo comporta, necessariamente, che l'utente debba disporre tanto di un dispositivo elettronico mediatore nella fruizione dei contenuti, quanto delle conoscenze necessarie per il suo corretto utilizzo. Le problematiche relative alla difficoltà nell'accesso ai contenuti digitali per le persone meno esperte di nuove tecnologie ricadono nelle problematiche connesse al digital divide.

## **Disponibilità nel tempo dei contenuti**

La disponibilità nel tempo dei contenuti analogici è oggi consolidata per tutti: un graffito preistorico conservato all'interno di una grotta o il manoscritto originale di un'opera medioevale sono ancora accessibili ed intellegibili per ognuno di noi: la stampa del nostro documento cartaceo potrà essere fruita anche tra cento anni senza problemi.

Nel caso del documento digitale questo non è necessariamente vero: la nostra possibilità di accesso alle informazioni è legata alla possibilità che abbiamo di rendere i calcolatori elettronici in grado di comprendere e trattare i bit che compongono il documento. Questa affermazione non va sottovalutata: la velocità con la quale nuove tecnologie o nuovi formati di conservazione delle informazioni sono introdotti in commercio rendono rapidamente obsoleto quanto veniva utilizzato solo pochi anni prima. Sino a pochi anni fa, ad esempio, per realizzare il proprio archivio si utilizzavano i floppy disk. Da anni i lettori di floppy disk non sono più venduti nelle versioni base dei computer portatili e, ultimamente, stanno scomparendo anche dai computer desktop: è probabile che tra meno di dieci anni gli archivi su floppy disk non saranno più accessibili per assenza di dispositivi in grado di leggere il contenuto dei dischi. Molti sostengono che i DVD o le penne USB avranno una vita molto più breve di quella avuta dai floppy disk.

Inoltre è necessario che l'elaboratore elettronico sappia con esattezza come trasformare i singoli bit che compongono il documento digitale nei contenuti da rappresentare sullo schermo. Tale trasformazione viene effettuata seguendo delle specifiche che definiscono il formato delle informazioni archiviate nel file<sup>9</sup>. Poiché nel corso degli anni sono stati definiti molteplici formati

---

<sup>9</sup> Abbiamo precedentemente introdotto il formato analogico e quello digitale, ora introduciamo il formato delle informazioni archiviate nel file (o più semplicemente il formato del file). Per evitare ambiguità vediamo che cosa è il formato del file. Poiché un file è una sequenza di bit, a priori non sappiamo se una certa sequenza di bit corrisponde ad una immagine, ad una canzone o ad un documento di testo. I bit che costituiscono un file sono collegati alle informazioni in esso contenute: ad ogni bit è associato uno specifico significato, che dipende dal tipo di contenuto conservato nel file. Poiché i bit di un file possono rappresentare molti contenuti diversi, è necessario definire una convenzione che associ ad ogni bit un significato. Alcuni bit possono ad esempio rappresentare un colore di un pixel se il file

per conservare i documenti, nulla garantisce che tra venti anni il nostro software di videoscrittura sia ancora in grado di comprendere un documento scritto con un formato diffuso oggi. La versione 2007 di Microsoft Word, ad esempio, non è in grado di aprire i file scritti con versioni di Microsoft Word precedenti alla 6.0 (rilasciata nel 1995). Questo significa che un documento scritto con Microsoft Word 5.0 nel 1994 può non essere già oggi più leggibile per l'assenza di un software in grado di conoscere come trasformare le informazioni codificate nei singoli bit presenti nel file nelle informazioni intellegibili all'uomo. Per ovviare a questo rischio è sufficiente che gli utenti scelgano, per memorizzare le informazioni, dei formati aperti. Questo significa che esiste una specifica pubblica per la descrizione di come leggere ed interpretare il contenuto del file. Se la Microsoft avesse rilasciato le specifiche del formato utilizzato con Word 5.0, oggi chiunque potrebbe sviluppare un programma per leggere il contenuto del file prodotto nel 1994.

Una seconda famiglia di problematiche è legata alle peculiarità dell'attore che interpreta i contenuti. Nel caso dei contenuti in formato analogico l'interpretazione è affidata all'uomo che dispone di grandi capacità di interpretazione: se un'immagine o un testo sono sbiaditi spesso l'uomo è in grado di comprendere lo stesso il contenuto informativo. Nel caso in cui l'interprete sia un computer, è necessario che il flusso di dati sia corretto: se nel corso del tempo il dispositivo di memorizzazione nel quale è conservato il file si deteriora, l'elaboratore non è più in grado di comprendere il contenuto ed interrompe il caricamento impedendo l'accesso al contenuto. Nel caso in cui – ad esempio – alcuni bit del documento rappresentano il colore di un particolare pixel dell'immagine, se alcuni di questi bit si deteriorano il computer non è in grado di comprendere quale colore stampare. Un ulteriore esempio di questo problema si osserva analizzando le differenze di comportamento tra un videoregistratore (riproduttore analogico di contributi audio-video) ed un lettore DVD (riproduttore digitale di contributi audio-video) nel caso in cui il dispositivo di memorizzazione (videocassetta o DVD) presenti dei problemi. In questo caso il videoregistratore continua a mostrare i contenuti anche se con qualità degradata: è poi il nostro cervello ad elaborare le informazioni e comprendere il contenuto mentre il lettore DVD si blocca ed inibisce la prosecuzione nella fruizione. Non dobbiamo dimenticare che i dispositivi di memorizzazione delle informazioni in formato digitale a basso costo che utilizziamo quotidianamente consentono la conservazione di informazioni per tempi molto più brevi rispetto alla conservazione delle informazioni in formato cartaceo. Già oggi il contenuto di molti floppy disk non è più accessibile, il problema sta emergendo anche con i CD-Rom.

## **Conservazione nel tempo della qualità dei contenuti**

La qualità delle informazioni presenti nel documento analogico tende, nel corso degli anni, a degradarsi: i colori delle immagini si sbiadiscono, la carta diventa più opaca e spesso l'inchiostro si schiarisce sino a scomparire. Nel caso del documento digitale il valore dei singoli bit rimane inalterato nel tempo. Se ad esempio alcuni bit del documento rappresentano il colore di un particolare pixel dell'immagine, tale colore sarà stampato in modo identico oggi e tra cento anni (fatte salve le problematiche emerse nel precedente sub-paragrafo).

## **Modifica e riutilizzo dei contenuti**

Nel caso in cui emerga la necessità di estrarre una porzione delle informazioni presenti nel documento, ovvero di modificare una parte dello stesso, i vantaggi del formato digitale emergono in modo evidente.

Le informazioni presenti nel documento analogico possono essere estratte, ovvero modificate, solo immettendole *ex novo*. Se stampo il documento e poi cancello la versione elettronica, per modificare una porzione devo necessariamente digitare tutto il testo. Tale eventualità, se

---

contiene informazioni relative ad una immagine, oppure un carattere se il file contiene informazioni relative ad un testo scritto. *Il formato di un file è la convenzione usata per leggere, scrivere ed interpretare le informazioni contenute nel file.*

deprecabile nel caso di un testo scritto, diventa impraticabile quando la modifica deve essere effettuata su una immagine la cui ricostruzione può risultare impossibile.

La possibilità di modificare il contenuto di un documento in formato digitale ha favorito la diffusione dei software per la videoscrittura, poiché permette di lavorare in più riprese su un documento, modificandolo ed integrandolo sino a raggiungere una versione stabile.

Il contenuto di un documento in formato digitale può quindi essere facilmente estratto e modificato. Tale possibilità abilita processi di condivisione della conoscenza impossibili nel caso di informazioni conservate in formato analogico.

## **Dimensione del contenitore**

La conservazione dei documenti in formato analogico richiede la disponibilità di spazi rilevanti, mentre lo spazio richiesto per la conservazione dei documenti in formato digitale è estremamente ridotta. La biblioteca nazionale di Roma conserva un patrimonio di circa sei milioni di libri in formato analogico, e per questo dispone di una sede molto vasta. Per conservare sei milioni di libri nati in formato elettronico sarebbero sufficienti meno di cento hard disk, il cui volume complessivo è molto inferiore al metro cubo.

## **Equivalenza tra originale e copia**

Copiare un documento analogico, ad esempio utilizzando una fotocopiatrice, significa riprodurre su un nuovo foglio di carta le informazioni presenti nel documento originale. Poiché le informazioni in formato analogico sono codificate in punti<sup>10</sup> sul foglio, il compito della fotocopiatrice è quello di riprodurre nel modo più preciso possibile questi punti sul nuovo foglio. In generale, la copia è di qualità inferiore rispetto all'originale poiché la riproduzione non potrà mai essere identica all'originale. Di conseguenza la copia di una copia avrà una qualità inferiore alla prima copia e così via. Gli effetti di tale degradazione sono visibili a tutti: spesso negli uffici ci vengono dati dei moduli prestampati quasi illeggibili poiché, a causa del numero eccessivo di copie effettuate, la qualità è molto degradata.

Copiare i dati in formato digitale è, invece, molto più semplice: poiché il file è composto da una sequenza ordinata di uni e zeri, copiare il file significa riprodurre questa sequenza ordinata di bit. Poiché la probabilità che un elaboratore elettronico confonda un uno con uno zero (o viceversa) è nulla, la copia di un documento digitale è indistinguibile dall'originale poiché contiene la medesima sequenza ordinata di uni e di zeri. Posso quindi effettuare un numero arbitrario di copie di un documento digitale e l'ultima copia continuerà ad essere indistinguibile dall'originale.

Anche rimanendo nell'ambito delle copie di un documento analogico, disporre del file in formato digitale può risultare vantaggioso: invece di fare più fotocopie dell'originale posso stampare più copie cartacee del documento digitale. Poiché tutte le copie provengono dallo stesso file, queste saranno indistinguibili. Questa eventualità non deve stupire: in realtà utilizziamo in modo improprio il termine copia, infatti ogni stampa produce un originale identico agli altri; situazione analoga al digitare più volte la stessa lettera con una macchina da scrivere.

## **Condivisione dei contenuti**

Come visto in precedenza i documenti digitali sono composti di bit, per tale motivo il documento è facilmente condivisibile: posso mandarlo per mail, renderlo disponibile via web o rete peer to peer, posso infine distribuirlo copiandolo su un dispositivo di memorizzazione (ad esempio un CD-ROM o una penna USB), in tutti questi casi chi riceve il documento digitale possiede una copia indistinguibile dall'originale con tutti i vantaggi visti in precedenza.

---

<sup>10</sup> Il principio di funzionamento sul quale si basano quasi tutti gli strumenti di stampa è semplice: le informazioni da stampare sono suddivise in punti di dimensioni più o meno ridotte e tali punti sono stampati sulla carta. I singoli caratteri del testo, ad esempio, sono rappresentati con la presenza di molti punti vicini.

Nel caso dei documenti analogici è necessario effettuare una scelta prioritaria tra la tutela della qualità delle informazioni presenti nel documento che riceve il destinatario e il tempo necessario perché lui lo riceva. Se infatti è prioritario il tempo nel quale l'altra parte ha a disposizione il documento, è necessario inviarlo tramite fax: le informazioni presenti nel documento ricevuto dal destinatario sono di qualità inferiore alle originali. Se invece la necessità è che il destinatario riceva un documento con informazioni di qualità simile all'originale, allora devo utilizzare la posta: questa scelta aumenta sensibilmente i tempi ed i costi di trasmissione. In entrambe i casi, comunque, viene inviata una copia del documento originale, con tutte le difficoltà di trattamento viste precedentemente. Non deve essere dimenticato, infine, che un documento digitale può essere inviato a costo zero ad un numero arbitrario di persone (ad esempio via mail o via web o via pear to pear) ed ognuno avrà la sua copia indistinguibile dall'originale; per un documento analogico invece, l'invio ha un costo non nullo e deve essere inviato singolarmente ad ogni destinatario che disporrà, comunque, di una copia.

## **Conclusioni**

L'utilizzo del formato digitale per la conservazione delle informazioni abilita trattamenti inimmaginabili sino a qualche anno fa: il formato digitale per la conservazione delle informazioni costituisce un ulteriore tassello per la società della conoscenza. Il rischio, però, è una società con tante informazioni ma che non ha gli strumenti (o la cultura) per conservarle.

## **Il diritto d'autore**

Le possibilità offerte dal trattamento digitale delle informazioni abilitano la nascita di nuovi scenari: la rottura del rapporto dicotomico produttore-fruitori consente ad ognuno di noi di pubblicare sull'internet il frutto delle proprie creazioni. Tali nuovi scenari, però, comportano anche dei rischi: chiunque può modificare facilmente il lavoro effettuato da altri e pubblicare tale nuovo lavoro come suo. Questi nuovi scenari, ed i rischi connessi, richiedono un'analisi di cosa sia il diritto d'autore e quale sia il ruolo del diritto d'autore nella società della conoscenza.

## **Che cosa è il diritto d'autore**

Ogni volta che scriviamo un articolo, scattiamo una foto o comunque realizziamo un'opera di ingegno quest'opera viene protetta dal diritto d'autore. L'obiettivo di questo diritto è evidente: tutelare – anche economicamente – il nostro lavoro intellettuale e creativo nel produrre l'opera, evitando che qualcuno se ne possa approfittare. Garantendo queste tutele il legislatore intende incentivare la creatività e lo sviluppo culturale.

Poiché la cultura di un popolo si basa sul lavoro intellettuale di coloro che lo hanno preceduto, alcune delle tutele previste dal diritto d'autore deve avere una durata definita, dopo la quale le opere di ingegno diventano patrimonio della collettività. Sarebbe improponibile che, per stampare la Divina commedia o i Promessi sposi, fosse ancora necessario remunerare gli eredi di Dante Alighieri o di Alessandro Manzoni.

## **Oggetto del diritto**

Qual è l'oggetto del diritto, ovvero che cosa possiamo tutelare con l'esercizio diritto d'autore? A questa domanda risponde in modo completo la legge 633 del 1941 e le sue modifiche.

In particolare l'articolo 1 della legge definisce che:

Sono protette ai sensi di questa legge le opere dell'ingegno di carattere creativo che appartengono alla letteratura, alla musica, alle arti figurative, all'architettura, al teatro ed alla cinematografia, qualunque ne sia il modo o la forma di espressione.

Sono altresì protetti i programmi per elaboratore come opere letterarie ai sensi della convenzione di Berna sulla protezione delle opere letterarie ed artistiche ratificata e resa esecutiva con legge 20 giugno 1978, n. 399, nonché le banche di dati che per la scelta o la disposizione del materiale costituiscono una creazione intellettuale dell'autore.

Inoltre all'articolo 2 si chiarisce che:

|                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| In particolare sono comprese nella protezione:                                                                                                                                                                                                                                                                                                                                                                                              |
| 1) le opere letterarie, drammatiche, scientifiche, didattiche, religiose, tanto se in forma scritta quanto se orale;                                                                                                                                                                                                                                                                                                                        |
| 2) le opere e le composizioni musicali, con o senza parole, le opere drammatico-musicali e le variazioni musicali costituenti di per sé opera originale;                                                                                                                                                                                                                                                                                    |
| 3) le opere coreografiche e pantomimiche, delle quali sia fissata la traccia per iscritto o altrimenti;                                                                                                                                                                                                                                                                                                                                     |
| 4) le opere della scultura, della pittura, dell'arte del disegno, della incisione e delle arti figurative similari, compresa la scenografia;                                                                                                                                                                                                                                                                                                |
| 5) i disegni e le opere dell'architettura;                                                                                                                                                                                                                                                                                                                                                                                                  |
| 6) le opere dell'arte cinematografica, muta o sonora, sempreché non si tratti di semplice documentazione protetta [...];                                                                                                                                                                                                                                                                                                                    |
| 7) le opere fotografiche e quelle espresse con procedimento analogo a quello della fotografia sempre che non si tratti di semplice fotografia protetta [...];                                                                                                                                                                                                                                                                               |
| 8) i programmi per elaboratore, in qualsiasi forma espressi purché originali quale risultato di creazione intellettuale dell'autore. Restano esclusi dalla tutela accordata dalla presente legge le idee e i principi che stanno alla base di qualsiasi elemento di un programma, compresi quelli alla base delle sue interfacce. Il termine programma comprende anche il materiale preparatorio per la progettazione del programma stesso. |
| 9) le banche di dati di cui al secondo comma dell'articolo 1, intese come raccolte di opere, dati o altri elementi indipendenti sistematicamente o metodicamente disposti ed individualmente accessibili mediante mezzi elettronici o in altro modo. La tutela delle banche di dati non si estende al loro contenuto e lascia impregiudicati diritti esistenti su tale contenuto.                                                           |
| 10) Le opere del disegno industriale che presentino di per sé carattere creativo e valore artistico.                                                                                                                                                                                                                                                                                                                                        |

Ogni prodotto del nostro ingegno, quindi, è tutelato dal diritto d'autore a meno che non sia la mera rielaborazione del lavoro di qualcun altro. In tal caso è possibile che il nostro lavoro violi i diritti del vero autore: in tal caso potremmo commettere il reato di plagio.

## Diritti morali e diritti materiali

Riflettendo sulle due opere letterarie precedentemente portate ad esempio, ci rendiamo conto che esistono due famiglie di diritti che possono essere esercitati.

La tutela di tali diritti è infatti legata a due incentivi che il legislatore ha inteso attribuire all'autore. Un primo incentivo è di tipo morale e riguarda la tutela della paternità dell'opera rispetto al plagio, un secondo incentivo – questa volta di tipo economico – va inteso come incentivo della creatività e dello sviluppo culturale.

### I diritti morali

Alla prima famiglia appartengono i c.d. **diritti morali**. Lo scopo di questi diritti è quello di proteggere la personalità dell'autore come si manifesta nella sua opera.

I diritti morali previsti dalla legge sono i seguenti:

1. il diritto di paternità dell'opera;

2. il diritto a rivendicare la paternità dell'opera (in caso di autore anonimo e/o pseudonimo);
3. il diritto di opporsi a qualsiasi deformazione, mutilazione od altra modificazione, ed a ogni atto a danno dell'opera stessa, che possano essere di pregiudizio al suo onore o alla sua reputazione;
4. il diritto di determinare i tempi di pubblicazione delle opere anche dopo la sua morte;
5. il diritto di ritirare un'opera dal commercio se ricorrono gravi ragioni morali.

Le caratteristiche principali dei diritti morali sono:

- a. perpetuità: questi diritti sono acquisiti con la realizzazione dell'opera e, alla morte dell'autore, vengono gestiti dai familiari. Sulla base di tale caratteristica, oggi Dante Alighieri ha ancora il diritto di essere considerato l'autore della Divina commedia;
- b. inalienabilità: questi diritti non possono essere ceduti a terzi;
- c. indisponibilità: il titolare non può disporre liberamente di questo diritto.

L'obiettivo della tutela fornita da questa famiglia di diritti è evidente: tutelare in modo permanente il legame che si crea tra un autore e la sua opera, intesa questa come espressione del suo lavoro intellettuale.

## I diritti patrimoniali

Poiché ognuno ha diritto ad essere remunerato per il proprio lavoro, ai diritti morali si affiancano i c.d. **diritti patrimoniali** o diritti **di utilizzazione economica**. Anche questi diritti sono riconosciuti all'autore nel momento in cui realizza l'opera. Tuttavia esistono due importanti differenze. In primo luogo i diritti patrimoniali sono alienabili: l'autore può cedere in forma scritta a terzi in esclusiva uno o più dei diritti patrimoniali; ad esempio gli scrittori cedono alle case editrici il diritto patrimoniale di riproduzione dell'opera. In secondo luogo i diritti patrimoniali hanno una durata temporale precisa: per tutta la vita dell'autore e sino a tutto il settantesimo anno solare dopo la sua morte, indipendente se tali diritti sono esercitati o meno. Poiché Alessandro Manzoni è morto da più di 70 anni, nessuno può esercitare un diritto patrimoniale d'autore sui Promessi sposi.

I diritti patrimoniali esclusivi previsti dalla legge sono i seguenti:

1. il diritto di riprodurre l'opera, ad esempio stampando più copie di un libro o di un film;
2. il diritto di trascrivere un'opera ad esempio dalla forma orale alla forma scritta, ovvero riprendendo una rappresentazione dell'opera;
3. il diritto di eseguire, rappresentare o recitare in pubblico l'opera;
4. il diritto di diffusione al pubblico dell'opera, ad esempio via etere, telefono o internet;
5. il diritto di distribuzione, intesa come messa a disposizione del pubblico dell'opera;
6. il diritto di tradurre l'opera;
7. il diritto di rielaborare dell'opera, intesa come modifica e rielaborazione;
8. il diritto di pubblicare le sue opere in raccolta;
9. il diritto di noleggiare (ovvero di autorizzare il noleggio) o dare in prestito la propria opera.

Ovviamente esistono delle limitazioni all'esercizio dei diritti patrimoniali (le c.d. libere utilizzazioni); ad esempio ai giornali è riconosciuto il diritto di cronaca, nei limiti dello scopo informativo: un giornalista può riportare nel telegiornale un estratto di un concerto per informare il pubblico dell'evento.

## Il diritto d'autore e la società della conoscenza

L'origine del diritto d'autore è relativamente recente. Sino all'invenzione della stampa a caratteri mobili da parte di Johann Gutenberg avvenuta alla metà del 1400, infatti, la stampa delle copie di un'opera avveniva in modo estremamente lento e costoso. Solo agli inizi del 1700 l'Inghilterra si è dotata di una legge che introduce nel mondo anglosassone il diritto di copia (in inglese copyright). Nel seguito anche gli altri paesi hanno promulgato leggi a tutela del diritto d'autore.

## **Il diritto d'autore nella società della materia**

In Italia, se pure con molte integrazioni, la legge alla base della tutela del diritto d'autore è stata introdotta nel 1941 quando l'internet, e più in generale le possibilità di trattamento digitale delle informazioni non erano neanche ipotizzabili.

L'impostazione normativa si basa quindi sulla struttura tipica della pubblicazione libraria che prevede due attori (l'autore e l'editore) e nella quale l'autore cede i diritti patrimoniali all'editore che si occupa di stampare le copie e di venderle, remunerando l'autore per il lavoro svolto.

Vale la pena evidenziare alcune caratteristiche della società italiana nel 1941 in relazione alla società moderna:

- a. l'assenza di sistemi di copia per i libri riduceva il rischio di copie non autorizzate;
- b. i grandi investimenti necessari per la produzione di un libro rendevano facilmente identificabile il plagio;
- c. il numero di autori era estremamente esiguo;
- d. gli alti investimenti necessari per la produzione favorivano spesso accordi nei quali l'autore riservava tutti i diritti patrimoniali all'editore. Ancora oggi è frequente trovare sui libri o sui DVD la scritta "© Copyright 2008 – tutti i diritti riservati";
- e. l'esiguo numero di opere protette rispetto a quelle libere dai diritti rendeva difficile che un autore intendesse realizzare un'opera derivata da una ancora coperta dai diritti<sup>11</sup>;
- f. il numero di nuove opere pubblicate ogni anno era molto ridotto.

Nel corso degli anni la situazione è radicalmente cambiata rispetto al 1941. La rapida diffusione degli strumenti per la copia analogica delle opere (fotocopiatrice, duplicatore di cassette o videoregistratore) imposero al legislatore, già nella seconda metà del secolo scorso, di modificare la legge per guidare le mutazioni in atto. Nel corso degli ultimi anni, poi, con la diffusione dell'internet e degli strumenti per il trattamento digitale delle informazioni, la situazione si è ulteriormente evoluta con conseguente nuova modifica della legge.

## **Nuove possibilità per società della conoscenza**

Nella società della conoscenza sono presenti alcuni effetti rivoluzionari rispetto alla situazione preesistente.

In primo luogo i bassissimi costi di pubblicazione dei contenuti riducono il ruolo degli editori: chiunque può realizzare un'opera e renderla disponibile per l'acquisto via web in formato digitale. In questo modo spariscono editori, tipografi e distributori.

In secondo luogo la rete internet amplia in modo geometrico il bacino degli utenti, non più poche migliaia di autori che producono le opere e qualche milione di fruitori per tali contenuti, ma miliardi di possibili autori e miliardi di possibili fruitori. Non è raro trovare in libreria testi di nuovi autori scoperti dagli editori tra i blogger o tra gli autori di contributi presenti sulla rete.

In terzo luogo è emersa nel tempo la necessità di esplicitare in modo completo quali diritti siano riservati e quali no. L'accezione "tutti i diritti riservati" può, infatti, risultare restrittiva. Nel corso degli anni si sono sviluppate nuove forme d'accordo tra autore ed editore per adeguarsi ai nuovi modelli alla società dell'informazione; tali nuovi accordi definiscono in modo più dettagliato cosa un utente possa o non possa fare con le opere d'ingegno degli altri. Un esempio di tali nuove forme di accordo sono le c.d. licenze Creative Commons.

Infine i contenuti in formato digitale sono facilmente modificabili da chiunque. La rottura della dicotomia produttore fruitore dei contenuti rende spesso difficile determinare la paternità di un'idea che nasce proprio dal confronto di più persone che, spesso, si trovano in paesi diversi.

Un caso può aiutarci a capire cosa significhi quest'ultimo aspetto.

I White Stripes sono un gruppo rock americano composto da due artisti, la peculiarità di questo gruppo è di suonare senza un bassista. Nel 2001 i White Stripes pubblicarono l'album "White Blood Cells" senza utilizzare il basso nelle loro canzoni. Steve Mc Donald dei "Redd Kross" decise

---

<sup>11</sup> Poiché i diritti morali sono perpetui, quando si afferma che un'opera è ancora coperta (o non è più coperta) da diritto d'autore si fa riferimento esclusivamente ai diritti patrimoniali.

di provare ad inserire il basso nelle canzoni dei White Stripes. A questo scopo acquisì dal compact disk un brano dell'album, lo rielaborò inserendoci il basso e pubblicò il risultato sul suo sito in un file formato mp3. Tale azione, svolta a valle dell'autorizzazione del produttore (altrimenti si sarebbe trattato di un atto illecito), ha consentito la creazione di un'opera derivata in modo semplice rapido ed economico. Il risultato dell'opera è stato reso disponibile a tutti gli utenti della rete internet. Questo caso evidenzia come il formato digitale delle informazioni abiliti possibilità per la creazione di opere d'ingegno impensabili sino a pochi anni fa. Inoltre, l'utilizzo della rete internet abilita nuove modalità di diffusione delle opere prive dell'intermediazione.

## **Nuovi rischi dalla società dell'informazione**

Sfortunatamente non è tutto oro quello che luccica.

Riprendiamo il caso dei White Stripes, per produrre il compact disk con l'album "White Blood Cells" la Sympathy for the Record Industry (la casa discografica che ha prodotto l'album) ha dovuto investire molti soldi per remunerare gli artisti e le altre figure impegnate nella produzione, duplicare e distribuire i compact disk, nonché pubblicizzare il prodotto.

Una volta acquistato il compact disk, ognuno di noi può – con pochi euro – effettuare una copia del compact disk e rivenderlo. Poiché originale e copia sono indistinguibili, chiunque acquista il compact disk copia possiede un disco identico all'originale. Tale azione, oltre a configurarsi come un illecito penale, genera un danno economico rilevante per il produttore e, di conseguenza, per la band.

Inoltre, il file digitale può essere reso disponibile via internet su un sito web, ovvero utilizzando un software peer to peer. Anche in questo caso, chi scarica i file che compongono l'album acquisisce dei documenti che consentono una fruizione identica all'ascolto del disco originale. Il risultato è che, per ogni compact disk venduto regolarmente in negozio, migliaia di persone potrebbero fruire di una copia digitale del suo contenuto con una qualità indistinguibile dall'originale.

Anche per arginare il rischio di un danno economico rilevante, i produttori di musica sollecitano i diversi legislatori all'approvazione di norme che, per tutelare i loro legittimi interessi nei confronti di chi svolge azioni illecite, risultano particolarmente restrittive dei diritti dei fruitori con l'effetto di rallentare non solo la diffusione di materiale da parte di persone non autorizzate ma anche di frenare lo sviluppo della società della conoscenza.

## **Privacy e società della conoscenza**

"Chiunque ha diritto alla protezione dei dati personali che lo riguardano", questo è il principio sancito dal primo articolo del decreto legislativo 196/2003, noto anche come codice in materia di protezione dei dati personali (nel seguito lo indicheremo semplicemente come codice).

Questa affermazione è molto rilevante, soprattutto nella società della conoscenza, ed è quindi importante per ognuno comprenderla sino in fondo poiché una cattiva interpretazione potrebbe portare a conseguenze molto gravi. Il legislatore, infatti, ha provveduto a fornire delle indicazioni molto precise di come garantire questo diritto. All'articolo 2, infatti, chiarisce che:

Il presente testo unico, di seguito denominato "codice", garantisce che il trattamento<sup>12</sup> dei dati personali si svolga nel rispetto dei diritti e delle libertà fondamentali, nonché della dignità dell'interessato, con particolare riferimento alla riservatezza, all'identità personale e al diritto alla protezione dei dati personali.

È quindi necessario capire quali sono realmente i nostri diritti e perché questa legge sia così importante nella società della conoscenza. Nella nostra società, infatti, la mole delle informazioni

---

<sup>12</sup> Con il termine trattamento si fa riferimento a qualunque operazione o complesso di operazioni effettuate sui dati personali. Sono esempi di trattamenti la raccolta, la registrazione, l'organizzazione, la conservazione, la consultazione, l'elaborazione, la modificazione, la selezione, l'estrazione, il raffronto, l'utilizzo, l'interconnessione, il blocco, la comunicazione, la diffusione, la cancellazione e la distruzione di dati.

trattate cresce in modo esponenziale: è necessario definire chiaramente quali trattamenti di dati personali siano leciti.

## Che cosa è la privacy

Per comprendere che cosa sia la privacy<sup>13</sup> possiamo ripartire dal principio appena enunciato e porci due domande: in primo luogo ha senso domandarsi quale sia il valore (ovvero la forza) di questo diritto, se ad esempio ha più valore il mio diritto alla privacy o il diritto di una impresa che vuole sollecitare i miei acquisti con il telemarketing; la seconda domanda riguarda la definizione di dati personali, quali informazioni ricadano quindi nell'applicazione della legge. Prima di questo, però, è necessario introdurre quali sono gli attori previsti nel trattamento dei dati personali.

## Gli attori

Quali sono i principali attori che partecipano al trattamento dei dati personali?

L'attore principale è l'**interessato**: la persona a cui si riferiscono i dati. Quando forniamo i nostri dati per acquistare un biglietto aereo, ad esempio, noi siamo gli interessati.

I nostri dati sono forniti ad una terza figura che li tratta per un certo scopo, ad esempio la compagnia aerea che deve emettere il biglietto e fornire il servizio di trasporto. Quest'attore è detto **titolare** del trattamento, a lui competono le decisioni relativamente alle finalità ed alle modalità di trattamento. Ovviamente a lui competono anche le responsabilità relative a questo trattamento. Le singole persone che trattano i nostri dati (ad esempio al check in) sono dette **incaricati**. L'incaricato tratta i dati personali sulla base delle istruzioni ricevute dal titolare ed allo scopo di raggiungere le finalità del trattamento richieste.

## Il diritto alla privacy

Il diritto alla privacy viene fatto discendere dalla nostra Costituzione poiché è oggi considerato un diritto inviolabile dell'uomo. Questa discendenza è estremamente importante: infatti i diritti della persona sono di rango superiore, ad esempio, a quelli di un ente. Questo significa che un ente se tratta dati personali deve mettere in atto tutte le azioni previste dal codice (che sono molte e spesso onerose) per garantire i diritti delle persone; tra le azioni che non sono permesse ci sono le sollecitazioni commerciali agli utenti, il tracciamento dei comportamenti e la vendita dei dati senza prima aver chiesto il consenso all'interessato. Per comprendere l'importanza che viene attribuita ai dati personali, vale la pena ricordare che il trattamento dei dati personali viene equiparato allo svolgimento di attività pericolosa; questo ha una conseguenza molto importante: se una persona subisce un danno per un trattamento dei suoi dati personali, l'onere della prova per dimostrare la propria innocenza ricade sul titolare del trattamento che ha cagionato il danno. In caso di richiesta di risarcimento da parte del soggetto che si ritiene leso dalle modalità del trattamento dei propri dati, il titolare del trattamento è tenuto a provare di avere adottato tutte le misure idonee ad evitare il danno. In generale, invece, si ha l'esatto contrario: chi si ritiene leso deve dimostrare le responsabilità dell'altra parte.

La legge italiana per la tutela della privacy ricalca una direttiva europea, per tale motivo molti dei diritti e delle procedure valide per l'Italia sono valide per tutti i paesi dell'unione.

## I dati personali

Occorre ora dare una definizione di dato personale. In questo senso la legge è molto precisa. All'art. 4 del codice si afferma:

b) "dato personale", qualunque informazione relativa a persona fisica, persona giuridica, ente od associazione, identificati o identificabili, anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi compreso un numero di identificazione personale;

<sup>13</sup> Nella prassi quotidiana il termine privacy è utilizzato come sinonimo di "diritto alla riservatezza".

che cosa significa questa definizione?

La prima cosa che si evidenzia è che la tutela riguarda non solo le persone fisiche, ma anche le imprese, gli enti o le associazioni. Noi, comunque, ci focalizzeremo esclusivamente sul trattamento dei dati personali relativi alle persone fisiche.

Per comprendere completamente questa definizione è possibile focalizzarsi sui quattro concetti cardine della definizione cioè: **qualsiasi informazione, relativa a, persona fisica, identificata o identificabile**. Nel seguito il significato di questi concetti sarà chiarito partendo dal “parere 4/2007 sul concetto di dati personali” fornito dal “gruppo di lavoro per la tutela delle persone con riguardo al trattamento dei dati personali” dell’Unione Europea.

## Qualsiasi informazione

L’espressione “qualsiasi informazione” nella definizione chiarisce la volontà del legislatore di definire un ampio concetto di dati personali.

Per quanto riguarda la **natura dell’informazione**, il concetto di dati personali comprende qualsiasi tipo di affermazione su una persona; può quindi includere informazioni "oggettive" (come la presenza di una data sostanza nel sangue di una persona, la sua data di nascita o il suo indirizzo), ma anche informazioni "soggettive" come opinioni o valutazioni. Quest'ultimo tipo di informazioni costituisce un'ampia parte del trattamento dei dati personali in ambito professionale. Le seguenti affermazioni, quindi, contengono dati personali: "Tizio è un cliente affidabile", "Tizio probabilmente non morirà presto", "Tizio è un buon lavoratore e merita una promozione". Perché l'informazione diventi un dato personale non è necessario che sia vera o dimostrata.

Dal punto di vista del **contenuto dell’informazione**, il concetto di dati personali è rivolto a qualsiasi tipo d’informazioni. L’espressione "dati personali" comprende informazioni sulla vita privata e familiare in senso stretto ma anche sulle attività di qualunque tipo, come sui rapporti di lavoro o sul comportamento economico e sociale, di una persona. I dati personali comprendono quindi informazioni sulle persone, a prescindere dalla posizione o del ruolo sociale (ad esempio in quanto consumatori, pazienti, lavoratori, clienti).

Infine per quanto attiene il **formato dell’informazione** o del supporto usato, il concetto di dati personali comprende le informazioni disponibili in qualsiasi forma, alfabetica, numerica, grafica, fotografica o acustica, le informazioni registrate su carta, e le informazioni conservate nella memoria di un computer o in una videocassetta.

## Relativa a

Questo elemento fondamentale della definizione è d’importanza cruciale, poiché è molto importante stabilire con precisione tanto le relazioni ed i collegamenti che contano quanto come distinguerli.

Il linea generale, un’informazione si può considerare “concernente” una persona se la riguarda.

In molte situazioni questa relazione può essere stabilita facilmente. Per esempio, i dati registrati nel fascicolo personale di una certa persona presso l’ufficio del personale “concernono” la sua situazione in qualità di impiegato. Lo stesso avviene per i risultati di un test medico di un paziente contenuti nella sua cartella clinica, o per l’immagine di una persona filmata nel corso di un’intervista.

In alcune situazioni le informazioni concernono oggetti e non persone. Questi oggetti appartengono di solito a qualcuno o possono subire l’influenza particolare di persone o esercitare un’influenza particolare su persone, oppure possono avere una sorta di vicinanza fisica o geografica a persone o ad altri oggetti. Solo indirettamente si può quindi considerare che tali informazioni concernono quelle persone o quegli oggetti, tuttavia possono essere considerati dati personali. Per chiarire il concetto con un esempio possiamo pensare al servizio di manutenzione di un’automobile. Il registro clienti di un meccanico o di un’officina contiene informazioni sull’automobile, sul chilometraggio, sulle date dei controlli, sui problemi tecnici e sulle condizioni materiali. Queste informazioni sono associate ad un numero di targa e a un numero di motore, che a loro volta possono essere collegati al proprietario. Quando un’officina stabilisce un nesso tra veicolo e proprietario, ad esempio per la fatturazione, l’informazione sarà “relativa al” proprietario o al conducente. Se il nesso viene fatto

con il meccanico che ha lavorato sul veicolo, ad esempio per accertarne la produttività, l'informazione anche sarà "relativa al" meccanico.

## **Persona fisica**

Il diritto alla protezione dei dati personali è un diritto universale che non è limitato ai cittadini o ai residenti di un dato paese. Il concetto di persona fisica figura all'articolo 6 della Dichiarazione universale dei diritti dell'uomo, che recita "Ogni individuo ha diritto, in ogni luogo, al riconoscimento della sua personalità giuridica".

## **Identificata o identificabile**

La legge impone che l'informazione riguardi una persona fisica "identificata o identificabile". In linea generale, si può considerare "identificata" la persona fisica che, all'interno di un gruppo, è "distinta" da tutti gli altri membri. Di conseguenza, la persona fisica è "identificabile" quando, sebbene non sia stata ancora identificata, è possibile identificarla (come suggerisce il suffisso "-abile"). Questa seconda alternativa è quindi in pratica la condizione limite che determina se le informazioni rientrano nell'ambito di applicazione del terzo elemento.

L'identificazione si fonda di norma su informazioni particolari che possiamo chiamare "identificatori" e che hanno un rapporto particolarmente stretto e privilegiato con la persona interessata: ad esempio segni esterni riguardo all'aspetto, come l'altezza, il colore dei capelli, l'abbigliamento; altri indicatori possono concernere una qualità che non può essere percepita immediatamente come ad esempio la professione, una funzione, un nome. La legge fa riferimento a questi "identificatori" nella definizione di "dati personali" perché afferma che una persona fisica che può essere identificata "[...] anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi compreso un numero di identificazione personale".

Per quanto riguarda le persone identificate o identificabili **direttamente**, il nome è evidentemente l'identificatore più comune e, in pratica, la nozione di "persona identificata" implica sovente un riferimento al nome di quella persona.

Per quanto concerne le persone identificate o identificabili **indirettamente**, questa categoria rimanda tipicamente al fenomeno delle "combinazioni uniche", siano esse ampie o ridotte. Nei casi in cui, a prima vista, gli identificatori disponibili non consentono di identificare una persona particolare, si può ancora considerare quella persona "identificabile" perché quelle informazioni combinate con altre (che siano o meno conservate dal responsabile del trattamento) consentiranno di distinguerla dalle altre.

## **Tipologie di dati personali**

Con l'accezione di dato personale si fa riferimento ad una casistica molto ampia, per tale motivo il legislatore ha dettagliato alcuni sotto insiemi di dati personali.

Con **dati identificativi** si indica quel sotto insieme dei dati personali che permettono una identificazione diretta dell'interessato (ad esempio il numero di codice fiscale).

Con **dati sensibili** si intende quel sottoinsieme dei dati personali idonei a rivelare l'origine razziale ed etnica, le convinzioni religiose, filosofiche o di altro genere, le opinioni politiche, l'adesione a partiti, sindacati, associazioni od organizzazioni a carattere religioso, filosofico, politico o sindacale, nonché i dati personali idonei a rivelare lo stato di salute e la vita sessuale.

Con **dati giudiziari** si intende quel sottoinsieme dei dati personali idonei a rilevare provvedimenti in materia di casellario giudiziale, di anagrafe delle sanzioni amministrative dipendenti da reato e dei relativi carichi pendenti, o la qualità di imputato o di indagato.

Il trattamento di dati sensibili o giudiziari, ovviamente, richiede il rispetto di norme molto più stringenti rispetto ai semplici dati personali poiché un trattamento errato potrebbe avere conseguenze molto gravi: è evidente che diffondere l'indirizzo di una persona è meno grave che diffondere informazioni sul suo stato di salute e, per tale motivo. Particolari modalità di trattamento

devono essere utilizzate anche nel trattamento di altre tipologie di dati personali, ad esempio quelli che riguardano i minori.

## **I principi della legge sulla privacy**

Una volta compreso quale sia l'oggetto del diritto e quali gli attori che concorrono al trattamento, è necessario comprendere cosa preveda la legge per garantire che il trattamento di svolga in modo corretto, almeno negli aspetti più generali.

I tre cardini sui quali si basa un corretto trattamento sono: l'informativa, il consenso e il diritto di accesso ai dati.

### **L'informativa**

Prima che il titolare possa raccogliere i dati personali, è necessario che l'interessato (o eventualmente la persona presso la quale si raccolgono i dati) sia preventivamente informato in merito al trattamento dei dati che si intende effettuare. Tale informazione preventiva è effettuata fornendo all'interessato un documento redatto dal titolare contenente l'informativa sul trattamento dei dati personali. Le informazioni che devono essere presenti nell'informativa sono le seguenti:

- a) le finalità e le modalità del trattamento cui sono destinati i dati;
- b) la natura obbligatoria o facoltativa del conferimento dei dati;
- c) le conseguenze di un eventuale rifiuto ad acconsentire al trattamento;
- d) i soggetti o le categorie di soggetti ai quali i dati personali possono essere comunicati o che possono venirne a conoscenza nonché l'ambito di diffusione dei dati medesimi;
- e) i diritti che l'utente ha in relazione ai dati raccolti;
- f) gli estremi identificativi del titolare o del responsabile al quale richiedere un riscontro nell'esercizio dei propri diritti.

L'informativa quindi fornisce all'interessato tutte le informazioni necessarie per comprendere quali dati personali il titolare intende trattare, in che modo e per quali finalità; l'informativa fornisce inoltre le informazioni per esercitare i propri diritti di accesso ai dati.

### **Il consenso**

Una volta che il titolare ha fornito l'informativa all'interessato, è necessario che questi acconsenta al trattamento. Poiché il consenso è espresso a valle della comprensione dell'informativa, si dice che l'interessato fornisce un consenso informato. Solo dopo aver ricevuto il consenso, il titolare può effettuare i trattamenti nei modi e per le finalità indicati nell'informativa. Per essere valido il consenso deve essere informato, liberamente espresso e relativo ai trattamenti da eseguire. Ad eccezione dei casi nei quali il titolare tratta dati sensibili, non è necessario che il consenso abbia forma scritta.

Esistono dei casi nei quali il consenso non è richiesto: in questi casi il titolare può effettuare il trattamento dopo aver fornito l'informativa all'interessato. Vale la pena evidenziare i due casi più frequenti nei quali il consenso preventivo non è richiesto. Il primo caso riguarda l'adempimento agli obblighi di legge: se la legge obbliga un'azienda a fornire i nominativi dei suoi clienti, non deve chiederli il consenso per adempiere a quest'obbligo. Il secondo caso riguarda il trattamento di dati personali che sono necessari per adempiere agli obblighi derivanti da un contratto del quale è parte l'interessato o per adempiere, prima della conclusione del contratto, a specifiche richieste dell'interessato: se apro un conto in banca, questa può mandarmi a casa l'estratto conto senza richiedere il mio consenso preventivo; non può però inviarmi materiale pubblicitario.

### **Diritto di accesso ai dati**

Una volta che il titolare ha acquisito i dati personali e li sta trattando, l'interessato può esercitare alcuni diritti. L'esercizio di questi diritti fornisce all'interessato gli strumenti per verificare come i suoi dati sono trattati ed intervenire qualora lo ritenga opportuno.

L'interessato ha, ad esempio, il diritto a conoscere l'origine dei dati personali, le modalità e le finalità del trattamento; qualora lo desideri può ottenere la modifica, la rettifica o la cancellazione dei dati che lo riguardano e può anche opporsi per motivi legittimi al trattamento.

Il titolare è tenuto a fornire un idoneo riscontro alle richieste ricevute, anche senza formalità, da parte dell'interessato.

Garantendo questo diritto di accesso il legislatore ha inteso assicurare un ruolo attivo all'interessato che può intervenire e verificare che i trattamenti siano effettuati in modo corretto.

### **Un esempio di trattamento non corretto**

Spesso si ritiene che, rispettando quanto visto in questo paragrafo, il titolare sia in regola. Questo è vero ma solo in parte. Infatti i dati sono talvolta trattati anche al di fuori degli scopi per i quali sono stati raccolti. Un tipico esempio è il seguente: se ci registriamo al sito di una rivista per accedere ad un contenuto, acconsentiamo al trattamento esclusivamente per questa finalità. Se il titolare ci invia una e-mail per comunicarci una nuova iniziativa commerciale sta effettuando un trattamento illecito.

### **La tutela della privacy nella società della conoscenza**

Perché sino a pochi anni fa il problema della privacy era così poco sentito, mentre oggi è diventato così rilevante?

Come visto precedentemente, il costante feedback fornito dagli utenti è uno strumento chiave per consentire ai produttori di contenuti di offrire prodotto nuovi e mirati. Tuttavia, la quantità di informazioni che possono essere acquisite, conservate ed analizzate su un cittadino della società dell'informazione possono consentire se non trattate accuratamente a diversi enti di tracciare un profilo completo della persona.

A titolo esemplificativo, una società che gestisce le carte di credito ha un quadro completo dei gusti, delle abitudini e degli spostamenti dei clienti. Tali dati arrivano automaticamente al centro di calcolo e vengono memorizzati senza la necessità di un intervento umano.

Sempre più spesso, nella nostra vita quotidiana, rimane una traccia delle nostre operazioni all'interno di un qualche sistema informativo. Alcuni esempi di informazioni che vengono registrate da attori diversi nello svolgimento delle nostre attività quotidiane sono<sup>14</sup>:

- per un prelievo al bancomat vengono registrati dalla banca: la data, l'importo ed il luogo del prelievo;
- per un acquisto con la carta di credito sono registrate dal gestore del servizio la data, l'importo e l'esercizio dove è avvenuto l'acquisto;
- durante la navigazione web vengono registrati dal provider per ogni risorsa visitata: la data, il nostro indirizzo IP e l'indirizzo della risorsa;
- per ogni mail inviata vengono registrati dal provider: la data, il nostro indirizzo IP l'indirizzo del mittente e quello del destinatario;
- quando il nostro telefono cellulare è acceso vengono registrate dal nostro operatore di telefonia cellulare: la data e la posizione del nostro telefono cellulare;
- per ogni telefonata vengono registrate dal nostro operatore telefonico: la data il numero telefonico del destinatario e la durata della comunicazione;
- per ogni SMS vengono registrate dal nostro operatore telefonico: la data e il numero telefonico del destinatario.

Tutti questi dati sono come piccole briciole che abbandoniamo lungo il nostro cammino e che consentono ad altri di ricostruire il tracciato con le nostre abitudini anche se non lo sappiamo.

Se non ci fosse una legge a tutelare i cittadini della società dell'informazione, la società che gestisce le carte di credito potrebbe vendere ad una gioielleria l'informazione che il conto in banca di una

---

<sup>14</sup> Nessuno conserva (se non nei casi autorizzati dalla magistratura) il contenuto delle comunicazioni o della corrispondenza, la cui segretezza è tutelata dalla nostra costituzione.

certa persona è ricco e che a giorni è il compleanno della moglie. Un gestore telefonico di telefonia cellulare potrebbe vendere ai ristoranti il numero telefonico delle persone che, all'ora di cena, sono in loro prossimità. Una società potrebbe decidere di assumere un dipendente non sulla base delle competenze personali, ma dei risultati di analisi del sangue. Tutti questi trattamenti violerebbero, secondo il legislatore, la dignità dell'individuo e per tale motivo ha previsto una legge che ci tuteli.

## **I trattamenti effettuati con l'ausilio dei sistemi informativi**

Con l'obiettivo di garantire il massimo delle tutele, il legislatore ha previsto disposizioni aggiuntive per i trattamenti effettuati con l'ausilio degli strumenti informativi.

La facilità di acquisizione, memorizzazione e trasmissione delle informazioni in formato digitale, infatti, possono costituire un rischio per i trattamenti illeciti.

Un errore tipico che si commette spesso è quello di ritenere che, anche se i dati non servono più per gli scopi per i quali sono stati raccolti, sia meglio conservarli: non si sa mai. In realtà una volta che non sia più necessario per gli scopi per i quali sono stati raccolti, i dati non possono più essere trattati ovvero devono essere resi anonimi. Questa regola, valida indipendentemente dalle modalità di trattamento, è ancora più valida quando i dati personali sono trattati mediante l'ausilio di sistemi informativi; spesso infatti conserviamo nel nostro hard disk molti file dei quali ignoriamo l'esistenza perché non occupano uno spazio fisico: se fossimo costretti a conservare solo la stampa del contenuto del file, probabilmente la avremmo già eliminata.

Per evidenziare l'importanza che il legislatore pone nell'attenzione al trattamento dei dati personali con l'ausilio di strumenti informatici l'art. 3 del codice specifica che:

I sistemi informativi e i programmi informatici sono configurati riducendo al minimo l'utilizzazione di dati personali e di dati identificativi, in modo da escluderne il trattamento quando le finalità perseguite nei singoli casi possono essere realizzate mediante, rispettivamente, dati anonimi od opportune modalità che permettano di identificare l'interessato solo in caso di necessità.

Il legislatore ha inoltre dettagliato le disposizioni minime che devono essere rispettate dai sistemi informativi che effettuano trattamenti di dati personali. Oltre a tali disposizioni ha sancito un principio di carattere generale relativo agli obblighi di sicurezza. L'art. 31 del codice afferma che:

I dati personali oggetto di trattamento sono custoditi e controllati, anche in relazione alle conoscenze acquisite in base al progresso tecnico, alla natura dei dati e alle specifiche caratteristiche del trattamento, in modo da ridurre al minimo, mediante l'adozione di idonee e preventive misure di sicurezza, i rischi di distruzione o perdita, anche accidentale, dei dati stessi, di accesso non autorizzato o di trattamento non consentito o non conforme alle finalità della raccolta.

## **Un caso di studio**

All'inizio del 2005 sono salite agli onori della cronaca due notizie di rilievo, che dimostrano come modelli di business che sfruttano le caratteristiche della rete per svilupparsi possono imporre profondi cambiamenti anche in settori "tradizionali". Salvo dove diversamente indicato, per completezza di informazione i numeri riportati fanno riferimento ai primi mesi dell'anno 2005.

La prima notizia informa che nell'ultima settimana del 2004 - nel Regno Unito - la vendita legale di brani musicali singoli via internet ha superato la vendita di brani musicali singoli su supporto materiale (i Compact Disk).

La seconda notizia informa che i lettori del New York Times acquirenti la versione on-line del giornale hanno superato, in numero, gli acquirenti della versione cartacea.

Dopo lo sgonfiamento della bolla speculativa legata alla così detta "new economy" e lo scetticismo diffuso che l'hanno seguito, esempi di questo genere dimostrano - anche ai meno esperti - che la così detta "net economy", cioè l'economia basata sulla rete e sulle relazioni che tramite questa

possono essere instaurate, è una realtà che può portare nuovi profitti ed aprire nuovi mercati anche a imprese che investono su business “tradizionali”.

Poiché notizie di questo tipo sono sempre più frequenti, vale la pena entrare nel dettaglio per capire quali elementi caratteristici stanno decretando il successo di simili offerte.

In entrambe i casi gli acquirenti sono interessati all'acquisto di informazioni (musica o notizie) ma, prima dell'avvento dell'internet, l'acquisto delle informazioni era vincolato all'acquisto di un bene materiale (il giornale o il CD) contenitore delle informazioni di interesse. Le tecnologie dell'informazione e della comunicazione consentono il superamento di tale necessità rendendo possibile acquistare le informazioni d'interesse in formato digitale, e successivamente effettuare il loro trattamento (lettura delle notizie od ascolto della musica per i casi in esame) in autonomia.

## **La vendita on-line di brani musicali**

L'analisi dei costi di acquisto dei singoli brani musicali evidenzia immediatamente un primo elemento di attrattività per l'acquisto di brani musicali via internet: il prezzo. Un singolo brano musicale acquistato utilizzando iTunes o OD2 (i due principali negozi on-line di musica) ha un costo di 0,99 sterline, contro le circa 5,5 sterline del medesimo brano acquistato a negozio. Una componente sostanziale in tale differenza di prezzo è imputabile alla scomparsa - nel primo caso - dei costi di duplicazione, trasporto, magazzino e vendita al dettaglio dei supporti fisici di memorizzazione (parte dei così detti costi variabili). Questo non può che invogliare gli utenti ad acquistare musica in formato digitale e, successivamente, masterizzare il proprio CD in autonomia e nel rispetto della legge.

In secondo luogo l'utenza a cui si rivolgono i produttori di musica sono i giovani, per loro natura più propensi all'utilizzo dei nuovi mezzi di comunicazione e, generalmente, con un budget limitato per questo tipo di acquisti.

In terzo luogo i siti che vendono musica on-line presentano un'offerta di brani improponibile per un negozio tradizionale. A titolo esemplificativo iTunes consente di scegliere tra circa 800.000 brani musicali delle maggiori compagnie discografiche, più i brani di centinaia di etichette indipendenti. A costi estremamente contenuti è possibile acquistare in formato digitale le canzoni comperate anni fa in vinile ed ormai introvabili nei negozi.

Un ulteriore aspetto da prendere in considerazione riguarda i diritti che l'utente acquisisce nel momento in cui acquista un brano. Le tecniche utilizzate negli ultimi anni dalle case discografiche per tutelare il copyright delle proprie opere vendute su supporti materiali (le c.d. tecniche di Digital Right Management), infatti, impediscono all'acquirente qualsiasi trattamento dei dati: dalla semplice copia personale al passaggio del brano su dispositivi di memorizzazione diversi dai CD originali. Diversamente, l'utente che acquista un brano musicale, ad esempio tramite iTunes, acquisisce anche il diritto di realizzare playlist personalizzate, di masterizzare il brano su CD, nonché di trasferirlo su dispositivi portatili. I sistemi di Digital Right Management utilizzati dai distributori di musica on-line, rispondono meglio ai desideri ed agli strumenti oggi a disposizione degli utenti.

Un ultimo aspetto da prendere in considerazione riguarda il formato del prodotto acquistato: nel caso in cui la musica sia acquistata su supporti fisici (i CD), l'utente può scegliere se acquistare una singola canzone oppure l'intero album: l'esperienza mostra che troppo spesso all'interno di un album non tutti i brani sono di pari gradimento, ma tutti devono essere acquistati. Acquistando la musica via internet è possibile eleggere le singole canzoni da acquistare (dopo averne ascoltata un'anteprima), secondo il proprio gradimento. Tale ulteriore possibilità, offerta dai canali di vendita digitali, esalta il potere di scelta degli utenti ed influenza le scelte delle case discografiche, incentivate a produrre album di maggiore qualità.

Come visto, l'attrattività nel prezzo, nell'ampiezza dell'offerta e nella versatilità del prodotto sono gli elementi che stanno decretando il successo, per altro atteso, di questo nuovo modello di business. Le scelte protezionistiche svolte da molti attori (ad esempio riducendo, in modo spesso vessatorio, i diritti degli utenti che acquistano CD musicali) hanno fornito una involontaria spinta a favore di questo nuovo mercato.

## La vendita on-line dei giornali

Passando ora al caso della vendita via internet del New York Times, l'attrattività legata al costo vantaggioso è molto meno evidente: il costo di una copia cartacea è pari ad un dollaro, contro i 65 centesimi della versione digitale (per l'edizione durante i giorni lavorativi). Anche in questo caso la riduzione di prezzo è legata principalmente ai minori costi sostenuti dall'editore per pubblicare il giornale: sono infatti assenti i costi di stampa, trasporto e vendita al dettaglio, che incidono non poco sul prezzo del prodotto finale. Come per la vendita on-line di musica, anche in questo caso l'utenza di riferimento è molto specifica: il lettore medio del New York Times è infatti di estrazione culturale medio alta, categoria di utenti più abituata all'utilizzo degli strumenti della società dell'informazione e della comunicazione.

Gli elementi che costituiscono reali punti di forza per la vendita on-line di un giornale di prestigio internazionale sono principalmente la dimensione del bacino di riferimento e la rapidità nella disponibilità delle informazioni.

Utilizzando la rete internet per la distribuzione del prodotto il numero dei potenziali lettori aumenta di alcuni ordini di grandezza a parità di costo per l'editore. Inoltre gli utenti, ovunque presenti nel mondo, possono accedere all'edizione non appena questa è pubblicata senza attendere la distribuzione nei punti vendita periferici (che può richiedere anche giorni). L'effetto di questi aspetti non può essere sottovalutato: l'immaterialità del bene, e quindi la facilità con la quale può essere distribuito, è l'elemento centrale per il successo dell'iniziativa. L'alternativa per aumentare il bacino dei potenziali acquirenti, infatti, dovrebbe prevedere l'ampliamento della rete distributiva con evidenti crescite dei costi, anche organizzativi, da sostenere. Tale ampliamento riguarderebbe comunque i grossi centri urbani, dove più facilmente può essere ottenuta la massa critica di utenti necessaria a giustificare l'investimento. Per ridurre il tempo di consegna del giornale nelle sedi periferiche invece, l'alternativa consisterebbe nell'apertura di stamperie in diversi stati del mondo, così da consentire l'uscita del numero in contemporanea. Non è difficile immaginare i costi di questa operazione.

Grazie all'utilizzo dell'internet, sono gli utenti a "recarsi" presso il giornale per acquistare la loro copia: gli unici impegni della testata consistono nel garantire la continuità del servizio e nel rendere disponibile la versione digitale della rivista appena pronta. Sarà poi l'utente – sulla base delle sue specifiche esigenze – ad accedere quando lo riterrà opportuno, certo di trovare ciò che cerca.

Un'ulteriore ragione dell'interesse nella lettura della versione digitale del New York Times riguarda il formato con il quale contenuti sono proposti. La versione digitale e quella cartacea del giornale sono identiche: non quindi una versione ricca di modalità comunicative consolidate per i lettori "off-line" ed una versione più leggera (magari una semplice sequenza di pagine html) per i lettori "on-line", ma due versioni identiche. In aggiunta i sistemi di Digital Right Management utilizzati consentono all'utente di stampare le pagine, di ricercare parole all'interno del testo nonché di copiare porzioni di notizie verso altri programmi (in tal caso, però, il testo viene copiato sotto forma di immagine), abilitando molte delle possibilità offerte dal trattamento delle informazioni in formato digitale.

Un ultimo aspetto di attrattività riguarda la facile archiviabilità dei dati in formato digitale: l'abitudine dei nonni di conservare le riviste ed i giornali "a futura memoria" torna possibile anche nelle ristrette case moderne, poiché un anno di New York Times potrà essere conservato in un singolo DVD<sup>15</sup>.

Come visto, in questo caso a decretare il successo dell'iniziativa sono state principalmente le caratteristiche di penetrazione e di accessibilità proprie della rete internet che hanno consentito di aprire alle imprese mercati prima irraggiungibili. I recenti eventi di politica internazionale che pongono - nel bene e nel male - gli USA al centro dell'attenzione, hanno sicuramente aumentato il

---

<sup>15</sup> Le Scienze, edizione italiana di Scientific American, ha pubblicato nell'ottobre del 2008 la raccolta dei numeri usciti tra il 1968 ed il 2008. 40 anni di storia (circa 480 riviste) occupano solo il volume di due DVD.

desiderio dei lettori non statunitensi di accedere a prestigiose fonti d'informazione americane, favorendo il successo dell'iniziativa.

## **Conclusioni**

I due casi analizzati, per quanto simili nei prodotti venduti, evidenziano ad un'analisi più attenta scelte molto diverse. L'intuito, il coraggio e gli investimenti richiesti ai primi che si sono affacciati a questi mercati li sta ripagando in termini economici e di immagine, anche ora che altri attori stanno iniziando a competere nel loro stesso settore.

Apple e New York Times hanno dimostrato che un modello di business basato sulla rete internet che sia innovativo e guidato dagli stessi paradigmi propri dei business tradizionali ha ampie chance di riuscire. Sfortunatamente attori in possesso di idee, coraggio e disponibilità economica sono – in questo periodo – merce molto rara.

L'internet offre mercati nuovi ed abilita modelli di business sino a pochi anni fa impensabili: qualcuno l'ha capito e, operando su mercati ancora privi di concorrenti, ha iniziato a raccogliere i frutti del proprio lavoro.

## ***Due articoli di approfondimento***

Nelle prossime pagine sono riportati due articoli di approfondimento per alcune tematiche affrontate pubblicati dalla rivista Interlex (<http://www.interlex.it>).

## Attualità

### Google e la banca dati universale della conoscenza

di Dante Picca - 17.03.05

Lo scorso 14 dicembre Google ha **presentato** un nuovo ed ambizioso progetto: digitalizzare, nell'arco dei prossimi 10 anni, i testi di libera e pubblica diffusione attualmente presenti nelle biblioteche di Harvard, Stanford, Università del Michigan, Università di Oxford e la New York Public Library. Per dare un'idea dei numeri in gioco, nel 2015 Google avrà digitalizzato circa 4,5 miliardi di pagine di testo: numeri importanti se si considera che, oggi, Google indicizza circa 8 miliardi di risorse web.

Vale la pena, entrando nel dettaglio della notizia, analizzare per quali motivi questo progetto (ed i progetti simili che si stanno attuando) rappresentano un elemento abilitante per la società dell'informazione e della comunicazione.

In primo luogo una tale mole di informazioni sarà resa disponibile gratuitamente, rendendo meno remota la possibilità - per chiunque abbia una connessione all'internet - di accedere a qualsiasi informazione prodotta dall'uomo.

In secondo luogo i documenti saranno disponibili in forma testuale e non come semplice scansione delle pagine originali. Il formato testuale permette trattamenti delle informazioni (ricerca, salvataggio, modifica) impossibili nel caso di documenti memorizzati come scansioni ottiche delle pagine che li contengono. Sarà possibile trattare i testi dei libri con la stessa semplicità con la quale oggi trattiamo il contenuto delle pagine web.

Il business di Google è evidente: diventare il principale portale per l'accesso alle informazioni, indipendentemente dal formato e dal medium con cui queste sono state prodotte. Già oggi Google indicizza pagine web, immagini, newsgroup e notizie; il futuro della ricerca delle informazioni è ignoto, ma sicuramente con questo progetto il motore di ricerca americano rafforza il proprio ruolo di leader.

Al di là delle personali convinzioni sull'evoluzione della nostra società, è innegabile che la crescita della quantità dei dati disponibili imporrà strumenti di ricerca sempre migliori ed in grado di far accedere gli utenti ad un catalogo unico di informazioni, indipendentemente dal loro formato di memorizzazione.

Due osservazioni meritano di essere fatte.

La prima riguarda le dimensioni del progetto: per quanto circa 15 milioni di testi sia un numero rilevante, è praticamente trascurabile rispetto alla produzione umana.

La seconda riguarda la tipologia di testi: si tratta infatti di opere libere da copyright. Questo esclude la quasi totalità della produzione mondiale degli ultimi 70 anni.

Si tratta, comunque, di un primo ed importante passo e Google è disponibile ad ampliare il numero dei testi realizzando opportuni accordi. Il sogno di digitalizzare l'intera produzione umana e renderla disponibile gratuitamente è un progetto immenso, come mole e come spesa, ma prima o poi dovrà essere realizzato, non solo da Google: il problema sono i tempi.

Mano a mano che i testi saranno integrati all'interno del motore di Google, avremo un riscontro dell'interesse nell'iniziativa. Se tale riscontro sarà positivo, è prevedibile che nuovi attori investiranno nel progetto per ampliare il numero di documenti disponibili; è facile immaginare che saranno quindi sviluppati nuovi modelli di business che consentano l'accesso anche a testi protetti da copyright, garantendo la giusta remunerazione agli autori, come già sperimentato da iTunes per la vendita di musica on-line.

La velocità con cui la società dell'informazione si evolverà dipende in larga misura dalla quantità e dalla qualità delle informazioni che i cittadini potranno trovare, nonché dal vantaggio che si può ottenere nell'usare strumenti ed informazioni disponibili: accedere al contenuto di milioni di libri stando comodamente seduto a casa, magari davanti al camino, sarà sicuramente una forte spinta in avanti.

Questo progetto, la cui portata è facilmente comprensibile a tutti, ha consentito di mettere in luce la distanza che ancora rende molti rappresentanti della società civile pessimi cittadini della società dell'informazione. Commentando gli scenari futuribili legati a questo progetto, emergono spesso posizioni reazionarie rispetto ad un evento che dovrebbe essere salutato con piacere.

Particolarmente illuminante, in questo senso, è un articolo pubblicato su "**Le Monde**" del **24 gennaio scorso** a firma del Prof. Jean-Noël Jeanneney, presidente della Biblioteca nazionale di Francia.

Dopo una prima lode dell'iniziativa "Una libertà di scelta è finalmente possibile, grazie alla tecnologia, a vantaggio dei paesi poveri e delle popolazioni svantaggiate", l'autore evidenzia i gravi limiti dell'iniziativa, ed in particolare che "i criteri di scelta [dei testi, ndr] saranno fortemente influenzati dalla visione anglosassone e dai suoi valori." e propone quella che, secondo lui, è l'unica soluzione al problema: un'iniziativa a livello di Unione europea per opporre all'iniziativa di Google un archivio di documenti caratterizzanti della cultura europea.

Il suo appello è perentorio: "E' l'ora di un appello solenne diretto ai membri delle tre più importanti istituzioni

dell'Unione Europea perché si agisca immediatamente. Altrimenti molto presto i giochi saranno fatti, le abitudini consolidate e sarà troppo tardi per fare qualunque cosa".

Più che discutere sulla condivisibilità o meno delle affermazioni del Prof. Jeanneney e sui problemi della globalizzazione della conoscenza, vale la pena riflettere sul perché tale posizione ha sempre meno senso nella società dell'informazione.

In primo luogo Google è in assoluto il motore di ricerca più utilizzato e conosciuto. Questo ne fa un portale naturale per la ricerca e l'accesso alle informazioni. Per quanto altri attori si stiano affacciando a questo mercato, è difficile che una biblioteca europea possa divenire conosciuta come Google offrendo solo l'accesso ad informazioni di carattere culturale.

In secondo luogo un progetto di questa rilevanza richiede ingenti investimenti economici (circa 150-200 milioni di dollari il budget per Google), difficilmente i governi europei saranno disposti a finanziare progetti di tale entità senza un evidente ritorno. Google, grazie alla liquidità ottenuta dalla quotazione in borsa, dispone del capitale necessario per portare a termine l'impresa.

In terzo luogo in questa fase di evoluzione della società dell'informazione le "guerre di religione" tendono a rallentare il processo evolutivo facendo spendere, troppo spesso inutilmente, risorse.

Infine, partire oggi con un progetto del genere significherebbe essere pronti con un'offerta competitiva tra molto tempo, quando ormai le abitudini saranno consolidate.

Ciò che ha favorito lo sviluppo dell'internet e, più in generale, degli strumenti della società dell'informazione è stata la disponibilità di standard aperti e liberi, che consentissero ad attori diversi di comunicare ed interagire tra loro. Oggi la battaglia si gioca ad un livello superiore, cioè nell'integrazione tra servizi ed informazioni diverse in modo trasparente per l'utente.

La tecnologia si sta rapidamente avviando ad un ruolo di *commodity* per i servizi che sfruttano questa tecnologia per essere erogati: un motore come Google diventerà quindi una *commodity*, come oggi sono gli elenchi del telefono.

Saranno gli utenti a decretare, con il loro utilizzo, in che modo queste commodity si evolveranno e quali informazioni offriranno. Invece di investire soldi nel realizzare una piattaforma alternativa a quella proposta da Google, conviene investire per integrare nella sua offerta numerosi ed importanti contributi della cultura europea "continentale", bilanciando le informazioni che ora possiede.

Utilizzando un unico fornitore per l'accesso alle informazioni, è reale il rischio che questi possa ergersi a giudice o censore dei contenuti accessibili, come è evidente già oggi nei media tradizionali. A questo rischio, però, può essere opposta solo la pluralità delle fonti tramite le quali, sulla base delle proprie preferenze e convinzioni, gli utenti accedono alle informazioni. La sopravvivenza degli attori in un mercato aperto, libero e globalizzato è legata alla capacità di attrarre con la propria offerta gli utenti, e non a logiche statalistico-protezionistiche.

Trovo importante che una posizione così estrema come quella descritta nelle pagine di "Le Monde" sia stata espressa da uno storico dei media, quale Jeanneney è, poiché evidenzia come la velocità con cui si evolvono i paradigmi ed i metri di giudizio è molto più lenta di quella con cui la nostra società si evolve verso una società dell'informazione.

## Le relazioni - 14

### Utenti, consumatori, cittadini: gli sviluppi della società dell'informazione

di Dante Picca\* - 30.05.05

Nel corso dei secoli passati, lo sviluppo della società ha avuto tempi caratteristici sufficientemente lenti per consentire a molti attori di prendere coscienza dell'evoluzione della società e dell'affermarsi di nuove classi sociali.

La nostra società, invece, si sta evolvendo verso una società dell'informazione con tempi caratteristici così rapidi da impedire una presa di coscienza diffusa riguardo ciò che sta avvenendo.

Sfortunatamente, l'elevata velocità di questa evoluzione, unita ad una tecnofobia diffusa ed alla colpevole disattenzione degli organi d'informazione, ha come effetto un'ignoranza dei diritti e dei doveri che spettano agli attori di questa nuova società. Tale ignoranza è causa di molti dei problemi di cui si discute, anche un questo forum.

Per comprendere questa affermazione bisogna analizzare i diversi attori presenti nella società dell'informazione.

Credo che gli attuali abitanti della società dell'informazione possano essere raggruppati in tre stati: gli utenti, i consumatori ed i cittadini.

Gli utenti sono lo stato più arretrato, utilizzano gli strumenti senza rendersi conto delle peculiarità e delle potenzialità dei mezzi che hanno a disposizione.

I consumatori sono uno stato intermedio poiché, come i consumatori nella società della materia, hanno coscienza di taluni aspetti e consumano molti dei prodotti offerti.

I cittadini sono il punto di arrivo per ogni abitante: come nella società della materia, un cittadino ha una ragionevole cognizione dei suoi diritti e dei suoi doveri ed è in grado di affrontare quasi tutte le situazioni che può incontrare.

In dieci anni sono stati fatti enormi passi avanti per dotare la società dell'informazione di infrastrutture, strumenti tecnici, leggi, servizi, prodotti ed utenti: ben poco è stato fatto per rendere questi utenti dei cittadini.

Le osservazioni riportate da Maccarone nella sua relazione [L'informatica di Don Chisciotte](#) non sono, purtroppo, un'eccezione: gli stessi utenti e gli stessi problemi organizzativi e di budget si incontrano a all'università di Tor Vergata, dove sono docente del corso di informatica personale. Dai racconti di amici e colleghi la situazione è analoga in quasi tutti gli atenei che.

Come contraltare degli abitanti ci sono le aziende il cui business è, a vario titolo, legato alla società dell'informazione. Per tali aziende, i decisori interni possono essere raggruppati in due categorie: i commercianti e gli imprenditori.

Con il termine commerciante non si intende specificare una particolare tipologia d'impresa caratterizzata per dimensioni od offerta, bensì identificare un *management* interessato a tutelare, oggi, il proprio mercato da tutti gli elementi di instabilità e, quindi, spaventato dai pericoli che le nuove tecnologie pongono al loro business. Se tale miopia porterà danni nel medio e lungo periodo all'azienda, il problema sarà affrontato al momento.

In quest'ottica ritengo vada inteso il diffuso utilizzo dei sistemi di *digital restriction management* evidenziati da Giustozzi nella sua relazione [I meccanismi di DRM non funzionano e non funzioneranno mai](#). Le aspettative dei commercianti sono quelle di una società composta di utenti dai pochi diritti ed i molti doveri.

Gli imprenditori delle aziende della società dell'informazione, pur essendo consci dei pericoli presenti, investono e sviluppano la propria attività imprenditoriale cercando di ottenere vantaggi economici ed innovando prodotti e processi. Hanno coscienza che, come in tutti i progetti strategici, anche in questo caso i ritorni economici non saranno immediati, ma il mercato della rete ha dimostrato di poter remunerare generosamente chi investe in modo innovativo. Gli imprenditori hanno bisogno di consumatori o cittadini perché il loro business risulti remunerativo: tanto pochi saranno i consumatori (o i cittadini) tanto meno saranno gli imprenditori disposti ad investire nella società dell'informazione e la società rimarrà nelle mani dei commercianti.

Si dice che la classe politica sia lo specchio della società che l'ha eletta: le scelte fatte in questi anni dai governanti di molti paesi sembrano dare ragione a quest'adagio. Pochissimi dei nostri rappresentanti hanno capito la portata e la velocità dei cambiamenti in atto, irrisorio è il numero di coloro che hanno idee su cosa fare. Questa asimmetria di competenze, conoscenze ed obiettivi tra abitanti, aziende e classe politica è - come anticipato - alla base di molti degli eventi che quotidianamente ci troviamo a commentare e discutere.

La promulgazione di leggi come il famigerato "decreto Urbani", sono la dimostrazione del livello della nostra

classe politica, nonché nella dissimmetria di comprensione delle esigenze dei diversi attori. Questa non è una novità; nel corso dei secoli il legislatore - ad esempio - prima di riconoscere diritti ai lavoratori, ha atteso che la società prendesse coscienza del loro ruolo. Solo allora ha promulgato leggi che fossero un punto di equilibrio tra gli interessi dei datori di lavoro ed i diritti dei lavoratori.

Oggi siamo in una fase simile, dove il potere delle lobby dei commercianti, unito alla superficialità dei media nel trattare gli argomenti ed alla diffusa incompetenza degli utenti, sta spingendo il legislatore a segnare una via di sviluppo della società dell'informazione diversa da quella che cittadini ed imprenditori si aspettano.

Il legislatore, infatti, sembra troppo impegnato a proteggere business già consolidati dalla minaccia della net-economy per comprendere che un uso maturo delle ICT potrebbe portare tali e tanti vantaggi per le imprese da rendere i danni legati ad un uso illegittimo della rete trascurabili rispetto ai vantaggi, anche economici.

In conclusione di questa relazione, ritengo che discutere del diritto del futuro non può prescindere dal discutere come far crescere gli abitanti della società dell'informazione, siano essi utenti, commercianti, giornalisti o politici, per farne dei cittadini. Solo allora la società civile sarà in grado di influenzare, con il voto prima, e con la sua opinione poi, le scelte del legislatore. Ogni idea in questo senso è la benvenuta.

\*

- 
- ▶ **Presentazione del forum**
  - ▶ **Come partecipare al forum on line**
  - ▶ **Le relazioni**
  - ▶ **La tavola rotonda del 9 giugno a Roma - Aggiornamento del 6 giugno**
  - ▶ **Come seguire i lavori in diretta web**
  
  - ▶ **InterLex, dieci anni di lavoro di squadra**
  - ▶ **Il forum del '95**
  - ▶ **Il forum del '96**
  - ▶ **Il forum del '97**



## Indice delle figure

|                                                                                                                          |    |
|--------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1: schema del metodo sperimentale.....                                                                            | 6  |
| Figura 2: un orologio analogico .....                                                                                    | 9  |
| Figura 3: procedure di memorizzazione e riproduzione .....                                                               | 12 |
| Figura 4: un esempio di esecutore.....                                                                                   | 16 |
| Figura 5: il CD audio come esecutore.....                                                                                | 16 |
| Figura 6: il telefono cellulare come esecutore .....                                                                     | 17 |
| Figura 7: figura geometrica di cui calcolare l'area.....                                                                 | 20 |
| Figura 8: figure elementari.....                                                                                         | 20 |
| Figura 9: figura base scomposta come somma di figure elementari.....                                                     | 21 |
| Figura 10: schema della macchina di Von Neumann .....                                                                    | 24 |
| Figura 11: una strada come esempio contenitore.....                                                                      | 26 |
| Figura 12: un esempio di rete in telematica, nel quale sono evidenziati i nodi e le maglie.....                          | 29 |
| Figura 13: esempio di rete a bus .....                                                                                   | 30 |
| Figura 14: esempio di rete a stella .....                                                                                | 31 |
| Figura 15: esempio di rete a maglia .....                                                                                | 32 |
| Figura 16: la homepage del portale generalista <a href="http://www.yahoo.it">www.yahoo.it</a> .....                      | 35 |
| Figura 17: la homepage del portale tematico <a href="http://www.torvergata.it">www.torvergata.it</a> .....               | 36 |
| Figura 18: la pagina mostrata da Google al termine della ricerca di "macchina di Von Neumann" .....                      | 37 |
| Figura 19: esempio di un messaggio relativo agli algoritmi nel newsgroup italiano dedicato all'astrofisica.....          | 38 |
| Figura 20: esempio di una directory.....                                                                                 | 39 |
| Figura 21: esempio della dimensione dei pixel nel caso di 1, 2 o 5 punti per unità di lunghezza.....                     | 42 |
| Figura 22: effetto della risoluzione spaziale .....                                                                      | 43 |
| Figura 23: immagini scansionate da una diapositiva rispettivamente.....                                                  | 43 |
| Figura 24: dettaglio del rosone di destra acquisito rispettivamente a 1200 e a 75 dpi .....                              | 43 |
| Figura 25: immagini memorizzate, rispettivamente, a 16 milioni ed a 8 colori .....                                       | 44 |
| Figura 26: dettaglio del campanile conservato, rispettivamente, a 16 milioni ed a 8 colori.....                          | 44 |
| Figura 27: videoproiettore a tre tubi .....                                                                              | 45 |
| Figura 28: immagine in formato jpeg salvata con qualità differenti.....                                                  | 48 |
| Figura 29: dettaglio delle immagini presenti in Figura 28.....                                                           | 49 |
| Figura 30: evidenza dei pixel differenti (in bianco) tra l'immagine a e le altre immagini presenti nella Figura 26 ..... | 50 |
| Figura 31: immagine non foto realistica salvata in formato jpeg a) e gif b).....                                         | 51 |
| Figura 32: esempio di stimolo acustico che raggiunge il nostro orecchio .....                                            | 52 |
| Figura 33: esempio del campionamento di un segnale ad intervalli diversi .....                                           | 54 |
| Figura 34: esempio di una riga graduata con tacche.....                                                                  | 55 |
| Figura 35: rappresentazione delle curve isofoniche.....                                                                  | 56 |
| Figura 36: la piramide del knowledge management .....                                                                    | 60 |

## Indice delle tabelle

|                                                                                                   |    |
|---------------------------------------------------------------------------------------------------|----|
| Tabella 1: tabella ASCII base .....                                                               | 10 |
| Tabella 2: tabella ASCII estesa .....                                                             | 11 |
| Tabella 3: esempio di testo e sua rappresentazione in formato ASCII. ....                         | 11 |
| Tabella 4: rappresentazione in diverse numerazioni dei primi 20 numeri. ....                      | 13 |
| Tabella 5: confronto tra diversi tipi di reti per la trasmissione delle informazioni.....         | 28 |
| Tabella 6: regole e messaggi del semaforo .....                                                   | 30 |
| Tabella 7: componenti RGB di alcuni colori .....                                                  | 45 |
| Tabella 8: dati relativi alle immagini presenti in Figura 28.....                                 | 48 |
| Tabella 9: intervalli di tempo tra le misure effettuate nelle immagini presenti in Figura 33..... | 52 |