

# I media telematici come strumento per la comunicazione scientifica

Francesca Di Donato

Copyright © 2005 Francesca Di Donato

07-07-2005 20:40:30

---

## Sommario

Introduzione

Elementi di teoria dell'informazione

a. Il computer

b. Internet

c. Il *World Wide Web*

La ricerca come attività collaborativa e distribuita

Introduzione

a. Biblioteche e web: l'accesso al sapere scientifico

b. La pratica della ricerca in rete

A. Schede di lettura

"As we may think" di Vannevar Bush

1. Il problema dell'accesso

2. Il metodo della ricerca scientifica

3. Battitori di piste

"Macchine calcolatrici e intelligenza" di A. Turing

"The World Wide Web - Past, present, future. Exploring

Universality" di Tim Berners-Lee

Riferimenti utili e approfondimenti

## Introduzione

"Metodi e macchine (..) rendono facile tutto ciò che sarebbe difficile a farsi con le sole mani, senza altri strumenti." (I. Kant, *Anthropologie* B 37)

Che cosa sono i media telematici? In che misura sono e possono diventare strumenti per lo studio universitario e la ricerca scientifica? Costruire una risposta a tali questioni necessita di alcune definizioni preliminari. Il termine telematica oggi ha assunto un significato particolare e suona un po' desueto, se usato per intendere ciò a cui ci riferiremo nelle pagine che seguono. Oggetto di questa analisi è infatti

la produzione di informazione in formati digitali (tramite computer) e la condivisione del sapere attraverso la rete *Internet* e il *World Wide Web*.

Il problema della comunicazione del sapere sarà dunque affrontato da un punto di vista filosofico e politico, considerando in che modo le tecnologie della parola agiscono sulla struttura del discorso scientifico a partire da un'introduzione teorica (dunque astratta, ma il meno possibile tecnica) e storica al computer, a Internet e al Web; ricostruiremo poi, brevemente, cosa si intende con "scienza", come vi si accede e come si produce all'interno dei limiti imposti dalle nuove tecnologie, esplorandone le possibilità – esplicite e solo accennate; per indicare, infine, alcuni strumenti utili alla pratica della ricerca scientifica in rete.

I testi di riferimento di questo ipertesto sono:

1. V. Bush , "As we may think" (1945) The Atlantic Monthly, luglio 1945, vol. 176, No. 1, pp. 101-108. Il testo in lingua originale: <http://www.ps.uni-sb.de/~duchier/pub/vbush/vbush.shtml>. Una traduzione italiana: <http://www.uniopen.it/FC/ssd/ipertesto/think.html>.
2. A. Turing, *Macchine calcolatrici e intelligenza*, in V. Somenzi (a cura di), *La filosofia degli automi*, Bollati Boringhieri, Torino 1965 (*Computing Machinery and Intelligence*, Mind, London N.S. vol. 59, 433-60, 1950, disponibile in rete al seguente url: <http://www.abelard.org/turpap/turpap.htm> ).
3. T. Berners-Lee, "The World Wide Web - Past, present, future. Exploring Universality" (2002): <http://www.w3.org/2002/04/Japan/Lecture.html>.

In ausilio allo studio dei testi, sono disponibili [schede di lettura](#) e [approfondimenti](#).

## Elementi di teoria dell'informazione

### *a. Il computer*

Il modello matematico del moderno computer è stato formalizzato poco più di sessanta anni fa, negli anni Quaranta del Novecento, dal matematico inglese Alan Turing. La storia del "calcolatore",

intesa in senso ampio come storia della ricerca di procedure e strumenti in grado di aiutare l'uomo a gestire grandi quantità di dati, è assai più antica; l'origine viene fatta risalire all'invenzione dell'abaco, nel 3000 a.C., e si arricchisce nel corso del tempo, di elementi essenziali (solo per enuclearne alcuni tra i più significativi: il *sistema binario* è un'invenzione cinese del 1100 a.C.; la prima *calcolatrice meccanica* con riporto automatico fu costruita da Pascal, nel 1642; Charles Babbage, nel 1830, creò la macchina analitica, e Ada Byron, conosciuta anche con il nome da sposata di Lady Lovelace, ne descrisse il funzionamento, scrivendo il primo *programma* della storia).

Alan Turing è considerato il padre del moderno computer per aver formulato, in uno studio del 1936, il modello teorico del calcolatore a istruzioni memorizzate. Tale modello, chiamato macchina di Turing (*mdT*), è astratto e universale in quanto dimostra la possibilità di macchine (come i nostri computer) in grado di risolvere problemi diversi. Il teorema che ne deriva è che, se un problema può essere risolto con una *mdT*, è calcolabile, dunque esistono soluzioni ad esso. Un corollario importante della scoperta di Turing, tuttavia, è che esistono problemi incalcolabili (quelli che la *mdT* non può risolvere, appunto). Di particolare interesse ai fini del nostro tentativo di definire il calcolatore, è la spiegazione funzionale proposta dallo stesso Turing nel saggio "[Macchine calcolatrici e intelligenza](#)" del 1950.

Gli attuali *personal computer* sono di dimensioni assai ridotte, e si basano su una tecnologia in buona parte differente da quella in uso cinquant'anni fa. Commercializzati a partire dagli anni Ottanta dalla *Apple*, e a seguire da *IBM* e dai suoi concorrenti (che hanno offerto macchine IBM-compatibili a prezzi sempre più accessibili al pubblico), si sono rapidamente diffusi anche grazie alla propagazione del primo sistema operativo della *Microsoft*, *Windows*, che permetteva di visualizzare i dati nel computer in modalità grafica, facilitandone l'uso ai non addetti ai lavori.

Il sistema operativo è un software, cioè un programma, che si comporta da gestore delle risorse: mette a disposizione dell'utente una comoda interfaccia, tiene traccia di chi usa quale risorsa, soddisfa le richieste delle diverse risorse e ne contabilizza l'uso, mediando i conflitti provenienti da programmi e utenti diversi. Più in generale, il codice software consiste propriamente in una tavola di istruzioni, di regole che, eseguite, portano ad un risultato, producendo effetti più o meno visibili.

Per definire il calcolatore, è oggi in uso tra gli informatici un modello formale, che lo rappresenta come un'architettura di "strati" costruita secondo livelli di astrazione crescenti. Riproduciamo uno schema nella figura qui sotto, che rappresenta, semplificata, tale struttura<sup>1</sup>:



La sezione più scura dello schema, in basso, è composta dall'*hardware* (letteralmente "ferramenta", cioè i dispositivi fisici), e da un primo strato di *software*, il linguaggio macchina o *assembler* (che viene trasformato in impulsi elettrici).

Le due sezioni superiori rappresentano strati di software, sono cioè programmi, solo codice; il grigio più scuro comprende software di livello più basso: il sistema operativo e altri programmi importanti tra cui i compilatori, che trasformano codice sorgente (scritto in un linguaggio di programmazione detto di alto livello) in linguaggio macchina; colorato di grigio chiaro, lo strato più in alto contiene i programmi applicativi (i *browser* per navigare sul *web*, client di posta e editor di testo, e i molti altri che rispondono alle numerose e differenti necessità quotidiane dell'utente).

Possiamo tralasciare i dettagli tecnici, e osservare l'utilità di tale modello: la metafora degli strati mostra infatti che sotto il livello superiore, che oscura quelli inferiori alla nostra vista, si nascondono, come in una cipolla, altri strati sempre più profondi. In questo modo, ci è possibile gestire le nostre informazioni senza dover interagire con gli strati più bassi del sistema, come facevano i primi programmatori. Allo stesso modo, gli stessi programmatori hanno a disposizione strumenti che facilitano loro notevolmente il lavoro; salendo dal basso verso l'alto, i programmi vengono infatti scritti in linguaggi di livello sempre più alto, che avvicinandosi al linguaggio naturale, sono più facilmente comprensibili dagli esseri umani (si vede, in figura, un esempio di codice sorgente scritto in C++).

```
25
26 int
27 main
28 (
29     int     argc,
30     char*   argv[]
31 )
32 {
33     ParseTextOptions(argc, argv);
34
35     if (!THXMDIServer::WillBeMDIServer(THXApp::GetAppSignature(), argc, argv))
36     {
37         return 0;
38     }
39
40     THXApp* app = new THXApp(&argc, argv);
41     assert( app != NULL );
42
43     app->Run();
44     return 0;
45 }
```

Lavorando al nostro calcolatore vediamo la punta di un *iceberg*: la parte più vasta di esso è nascosta alla vista, ma sta sotto e sorregge ciò che emerge. La metafora degli strati può così essere estesa al piano dei contenuti, per mostrare che tale livello comunica ed interagisce con gli altri. Le pagine che state leggendo, ad esempio, sono state scritte con *un* particolare programma ad *un* calcolatore su cui gira *un particolare* sistema operativo, sono state codificate in *un* formato e poi pubblicate sul *Web*. Tutte queste variabili avrebbero reso un tempo impossibile leggerle da chi avesse avuto hardware e software diverso dal mio. L'avvento di *Internet* e del *World Wide Web* hanno promosso l'interoperabilità; tutt'oggi, comunque, hardware e software influenzano il modo in cui le informazioni vengono visualizzate, ma anche e soprattutto quanto sono accessibili ed esportabili, e quanto dureranno nel tempo.

## Riferimenti e link rilevanti

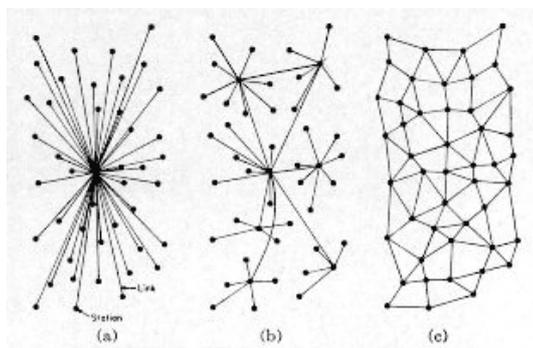
Francesca Di Donato. ["Macchine calcolatrici e intelligenza" di A. Turing \(scheda di lettura e link al testo\)](#).

Andrew S. Tanenbaum. [I moderni sistemi operativi \(cfr. nota 1\)](#).

Andrea Bedetti. [Storia del computer \(cronologia\)](#).

## b. Internet

Quando parliamo di Internet, intendiamo tanto *l'infrastruttura fisica* che collega in rete tramite cavi un insieme di calcolatori, quanto il software che permette lo scambio di informazioni tra computer, in particolare lo *stack* di *protocolli* TCP/IP. Entrambe le idee sono dovute a Paul Baran, cui, alla fine degli anni Cinquanta fu affidato il compito di studiare un sistema di comunicazione in grado di resistere ad un attacco nucleare. In un [articolo del 1964](#) Baran suggerì tre possibili [topologie](#) per una rete di comunicazione: (a) centralizzata; (b) decentralizzata; (c) distribuita.



Lo scienziato americano dimostrò che solo la terza, progettata con un'architettura a maglie analoga alla rete stradale, rispondeva all'esigenza postagli. Baran propose inoltre di dividere l'informazione in pacchetti, e di far viaggiare i pacchetti attraverso la rete indipendentemente l'uno dall'altro; questa soluzione necessitava il passaggio da un sistema [analogico](#) ad uno [digitale](#), e stentò a farsi strada.

Attualmente, la topologia di Internet è assai [più complessa](#) di quella indicata da Baran; della struttura iniziale, resta l'impossibilità di identificare al suo interno un database centrale da cui chi è connesso possa attingere informazioni; viceversa, la struttura reticolare e la sua efficienza consiste nell'assenza di un vero e proprio centro e dalla centralità dei margini. Lo *stack* di protocolli TCP/IP (*Transmission Control Protocol/Internet Protocol*) di Vinton Cerf e Robert Kahn, usa invece propriamente la tecnica di commutazione di pacchetti come fu pensata dallo scienziato americano.

Il TCP/IP non istituisce un'interazione permanente tra due macchine; uno stesso canale di rete può servire contemporaneamente più utenti, così che la rete non sia propriamente una cosa, ma piuttosto un accordo che permette la coesistenza e la cooperazione tra i nodi. Il modo in cui funziona

ricorda il sistema delle cretule di argilla, antenate della scrittura sumerica cuneiforme. Il protocollo definisce una serie di regole standardizzate e comuni che permettono la comunicazione tra differenti tipi di calcolatori che non parlano necessariamente la stessa lingua, proprio come le cretule istituivano un codice comune di scambio, astratto dalle lingue effettivamente parlate.

L'analogia con le cretule si fa calzante osservando che su Internet le informazioni sono tramesse in "piccole scatole". Il TCP suddivide il flusso cifrato dei dati in pacchetti, contrassegnando ciascuno di essi con un'intestazione particolare (composta da un'informazione che indica il loro ordine di creazione e quello che bisognerà seguire per ricomporli all'arrivo) e una somma di verifica (un numero utilizzato per accertare l'eventuale presenza di errori). Ciascun pacchetto comincia inoltre con le istruzioni d'indirizzamento previste dalla normativa IP, una notazione che corrisponde a quella utilizzata per le buste cartacee: nome mittente, nome destinatario, e indirizzi di entrambi.

È importante osservare che lo *stack* di protocolli è pubblico, cioè può e dev'essere consultato da chi intende connettere fisicamente un calcolatore alla rete Internet. È dunque un vero e proprio dominio pubblico che, lasciando i permessi fuori dal sistema, è aperto a tutti e può essere (ed è) continuamente migliorato a condizione che sia mantenuta la sua semplice struttura fondamentale.

Nonostante la completa pubblicità del protocollo che collega i singoli nodi alla rete, la gran parte dei programmi che girano sui nostri calcolatori non sono altrettanto pubblici. La chiusura delle sorgenti (ovvero la non accessibilità delle istruzioni che l'uomo scrive e può comprendere) è una prassi consolidata nel mondo del software proprietario (la Microsoft di Bill Gates in primo luogo), e genera un vero e proprio monopolio della conoscenza. Di contro, la stessa rete è stata la culla di un vasto e variegato movimento per il [software libero](#).

## Riferimenti e link rilevanti

Paul Baran. *Introduction to Distributed Communications Networks*, RM-320-PR, 1964.

Francesca Di Donato. *I ponti di Koenigsberg e l'architettura delle reti*.

[Mappe di Internet](#).

Free Software Foundation. *Che cos'è il software libero*.

Andrea Bedetti. *Storia del computer (cronologia)*.

### c. *Il World Wide Web*

Immaginate di realizzare un grande modello tridimensionale, in cui la gente compare in forma di piccole sfere, con i fili che corrono tra le varie persone aventi qualche rapporto di lavoro. Adesso immaginate di dare una scrollata alla struttura, fino a dare un senso a quel groviglio. Forse in quel punto vedrete dei gruppi intrecciati strettamente, e in altri delle aree di comunicazione debole, con poca gente. (*L'Architettura del nuovo Web*, p. 32)

E' storia nota che il *Web* è stato inventato alla fine degli anni Ottanta al CERN da Tim Berners-Lee. Al tempo, il giovane ricercatore inglese si occupava di sistemi informatici per il prestigioso istituto di fisica delle particelle con sede a Ginevra. Il suo compito richiedeva in particolare di memorizzare velocemente alcune informazioni riguardanti all'incirca 1.400 persone, che lavoravano con macchine diverse e programmi diversi. In *Weaving the Web* Berners-Lee racconta che *Enquire* fu il primo programma che anticipava il *World Wide Web*. Il nome riprendeva un'enciclopedia inglese che da bambino il programmatore amava sfogliare, dal titolo "Enquire Within Upon Everything".

*Enquire*, a differenza dei tradizionali sistemi informativi, non usava alberi ma *link* bidirezionali (all'interno del programma) e *link* monodirezionali (in direzione esterna). Ogni pagina era un nodo, e per creare nuovi nodi era necessario linkarsi a un nodo esistente; i link ai nodi si ordinavano in un elenco che, per trovare un'informazione, doveva essere sfogliato dall'inizio. Da *Enquire* al suo successore *Tangle* ("intrico") fino al *World Wide Web*, il progetto del fisico inglese prende la forma di un sistema di documentazione universale: nel suo lavoro, Berners-Lee sperimenta con difficoltà la mancanza di interoperabilità tra *hardware* e tra *software* e progetta un sistema con regole comuni, accettabili per tutti, cioè il più possibile vicino alla mancanza di regole.

A modello del suo sistema Berners-Lee sceglie l'*ipertesto*. E' necessario che il nuovo sistema funzioni tramite associazioni tra cose apparentemente scollegate. All'importanza del collegamento nei sistemi di documentazione e all'idea di ipertesto si erano già dedicati, in tempi più remoti, Vannevar Bush e Ted Nelson. Nel saggio *As we may think (1945)*, il primo aveva ipotizzato una macchina fotoelettronica, il *Memex*, in grado di attuare e seguire riferimenti incrociati su microfilm servendosi di codice binario,

fotocellule e fotografia istantanea; il secondo, autore del noto *Literary Machines* (1981) è l'inventore del termine "ipertesto", una griglia in espansione in grado, potenzialmente, di raccogliere in un unico sistema tutti i testi della letteratura mondiale, secondo il progetto dell'utopico software *Xanadu*; nel sistema ideato dal Nelson, ogni citazione sarebbe stata dotata di un *link* con la fonte, garantendo così una ricompensa agli autori citati.

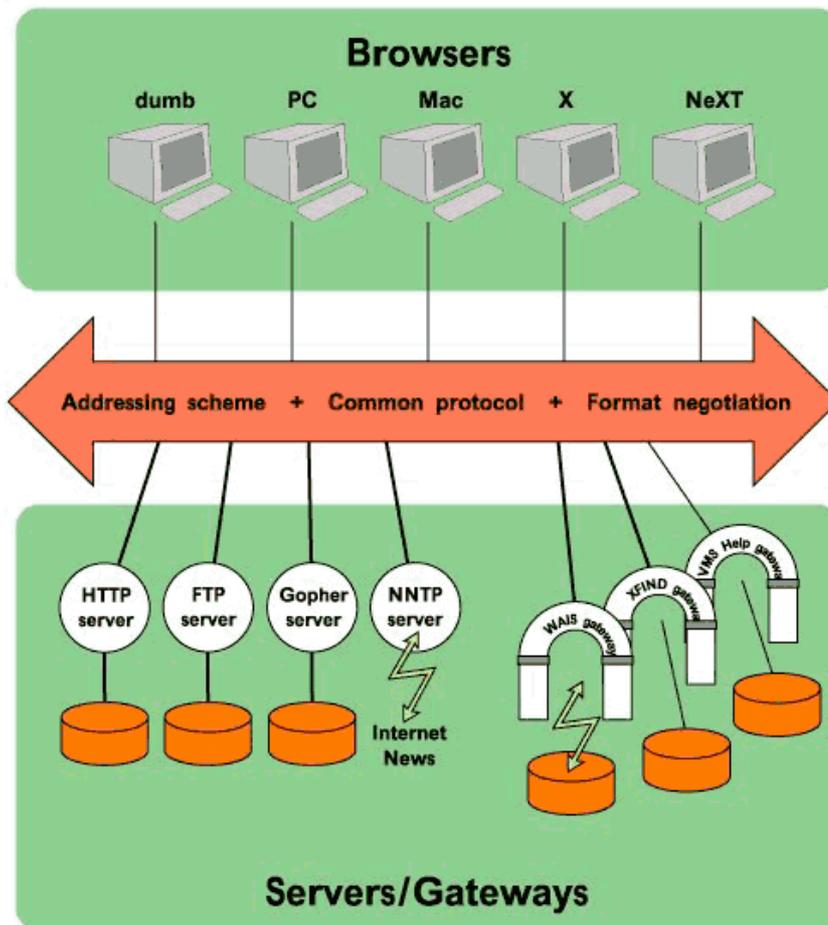
Berners-Lee si immagina l'ipertesto come uno strumento per il lavoro di gruppo e collaborativo. Nello spazio del *Web* sarebbe stato possibile leggere, muoversi e scrivere nell'ipertesto. Il sistema avrebbe dovuto però essere decentrato, in modo che ogni nodo fosse intrinsecamente equivalente agli altri; lo strumento per creare i collegamenti il *link*, che, in questo caso, è stato pensato solo monodirezionale.

Da un punto di vista tecnico, il *Web* è un architettura minima e universale a ipertesto che, per funzionare su scala planetaria, si serve di *Internet*, intesa tanto come infrastruttura fisica di comunicazione, quanto come insieme di protocolli (primo tra tutti il TCP/IP). La necessità di interoperabilità tra diversi hardware e software e il problema dell'internazionalizzazione delle codifiche sono vincoli tali da aver comportato la necessità di un principio architettonico fondante: l'apertura (universalità). Da tali presupposti sono pertanto derivate poche scelte tecniche fondamentali:

- 1. *URI o URL*: il modo in cui il web funziona richiede a ciascuno di dare un *Universal Resource Identifier* (o, più precisamente, *Uniform Resource Locator*) ai propri documenti. Ogni volta che un pezzo d'informazione viene etichettato con un indirizzo, è possibile chiedere al proprio computer, tramite un *browser*, di trovarlo. In questo modo, tutti i documenti sono uguali tra loro; la parità dei nodi è garantita mantenendo la facilità nei *link* e la possibilità di aggiungere *link* verso qualsiasi pagina.
- 2. *http*: il protocollo di scambio delle informazioni, *http*, è aperto a tutti i formati e si basa su regole minime, tali da permettere ai computer di comunicare tra loro; con un protocollo minimo e semplice, al crescere dei nodi il sistema, invece di collassare, si arricchisce.
- 3. *HTML*: il linguaggio di *mark-up* del Web è una lingua franca, comprensibile ai *browser*, che si basa su una regola fondamentale di progettazione (veicolare solo la struttura del documento ipertestuale, non il suo apparire) e permette con enorme facilità di collegare le pagine l'una all'altra.

Anche il Web dunque, è utile ripeterlo, non è un programma, ma un insieme di protocolli. Come mostra la figura in basso, le specifiche *http*, *URI* e *HTML* formano una specie di "bus" in grado di connettere i programmi che risiedono sul computer dell'utente (detti *client*) ad altri programmi installati su computer remoti che

forniscono le informazioni (cosiddetti *server*).



## Riferimenti e link rilevanti

Francesca Di Donato. *"As we may think"* di V. Bush (scheda di lettura).

Francesca Di Donato. *"The World Wide Web - Past, present, future. Exploring Universality"* di T. Berners-Lee (scheda di lettura).

## La ricerca come attività collaborativa e distribuita

## *Introduzione*

Le formulazioni teoriche alla base delle invenzioni tecnologiche trattate in precedenza, in particolare nell'[articolo di Alan Turing](#), nel saggio di [Vannevar Bush](#) e nel pensiero di Berners-Lee, condividono un importante presupposto comune: il fatto che la scienza, intesa tanto come processo di creazione del sapere quanto come il bagaglio delle conoscenze acquisite accumulate nel corso della storia, sia un insieme continuo, un'infrastruttura mobile costruita tramite collegamenti tra elementi preesistenti. Turing è scandalizzato dal fatto che, un'osservazione tanto banale, non sia ovvia: l'errore, comune ai filosofi e ai matematici, "consiste nel presupporre che appena un fatto si presenta alla mente, tutte le conseguenze di questo fatto saltino fuori simultaneamente". La ricerca scientifica, tuttavia, non funziona così; ogni scoperta poggia sulla conferma o sulla confutazione di precedenti risultati, propri o altrui. "Siamo nani sulle spalle di giganti", scriveva Newton, riprendendo un topos che, a partire da Platone, attraversa il pensiero filosofico occidentale e che ha il suo culmine nella filosofia dell'illuminismo; la stessa ipotesi che Vannevar Bush pone a fondamento dell'idea di ricerca, il fatto che la conoscenza umana sia il prodotto di un processo collaborativo, che attraversa la storia e la trascende<sup>2</sup>.

Le sezioni che seguono sono dedicate a descrivere i principali strumenti oggi disponibili per la conservazione e la comunicazione del sapere scientifico e a suggerirne possibili usi.

### *a. Biblioteche e web: l'accesso al sapere scientifico*

Tradizionalmente, i centri di archiviazione e di disseminazione del sapere sono state e sono le biblioteche (si narra che nella Biblioteca di Alessandria, la più grande dell'antichità, fossero contenuti centinaia di migliaia di "rotoli"). Oggi, la quantità e le tipologie di documenti raccolti nelle biblioteche (pubbliche, statali, universitarie, ma anche collezioni e archivi privati, etc.) sono numerosissime: oltre alle riviste, vi si trovano monografie, collane, raccolte di saggi e atti di convegni, documenti ufficiali (tra cui quelli legislativi) più la cosiddetta letteratura grigia (tesi di laurea e di dottorato, report interni e altra documentazione di ricerca). I materiali vengono archiviati e catalogati in molteplici forme generalmente standardizzate, per essere reperiti, consultati e

riutilizzati (citati, tradotti, etc.) dal pubblico dei lettori. La ricerca di standard di catalogazione comuni e condivisi (e l'adozione di formati che rendono i documenti "portabili") è un obiettivo che, ad oggi, le biblioteche continuano a perseguire.

Nell'arco degli ultimi decenni, le biblioteche si sono informatizzate ed automatizzate, mettendo in rete i propri cataloghi e offrendo strumenti di ricerca sempre più avanzati. Le informazioni sono catalogate tramite l'uso di metadati, cioè dati sui dati, quali l'autore, il titolo, o il soggetto, parole chiave che ci aiutano nella ricerca. L'articolo di Bush, più volte ricordato, ha avuto un influsso fondamentale non soltanto sullo sviluppo delle tecnologie alla base del Web, ma anche sulla messa a punto di uno strumento importantissimo nella storia recente ed attuale delle biblioteche e della conservazione della letteratura scientifica, Lo *Science Citation Index* (SCI).

### **Lo Science Citation Index**

Lo SCI è uno strumento ideato nei primi anni Cinquanta da Eugene Garfield in risposta all'esigenza di poter disporre di un sistema bibliografico per la letteratura scientifica in grado di rendere ininfluenti le citazioni non rilevanti (di dati falsi, incompleti e obsoleti). Non potendo leggere tutto, chi fa ricerca ha bisogno di criteri di qualità sui quali basarsi: lo scopo di un "indice delle citazioni scientifiche" sarebbe stato quello di rendere possibile al ricercatore la selezione delle citazioni importanti e degli articoli fondamentali per la sua ricerca.

Come Bush, Garfield pone al centro della riflessione due elementi essenziali: l'importanza delle "citazioni"; e il problema della selezione dell'informazione scientifica.

Perché un indice basato sulle citazioni? "Se uno considera i libri le macro-unità del pensiero, e gli articoli di rivista le micro-unità, l'indice delle citazioni si concentra sulle sub-micro-unità (molecolari) del pensiero", scrive Garfield. Gli scienziati, spesso, sono interessati ad un'idea particolare piuttosto che ad una concezione generale, e se appropriatamente pensati e sviluppati, indici di "pensieri" possono essere estremamente utili ai fini della selezione dell'informazione: nel processo della ricerca bibliografica, gli indici per soggetto ricoprono un ruolo minimo, anche se significativo, poiché servono semplicemente da punto di partenza della ricerca. I limiti del mezzo, secondo Garfield, stanno nel fatto che è impossibile creare indici per soggetto che coprano tutte le possibili impostazioni di cui lo scienziato ha bisogno. Gli indici a soggetto sono schematici, e la possibilità di incrementare il numero dei soggetti non elimina la rigidità del mezzo; perciò, neppure una migliore *standardizzazione* della classificazione è risolutiva.

Viceversa, sono necessari nuovi strumenti bibliografici che aiutino a colmare la distanza tra chi crea il documento, cioè l'autore, e l'impostazione dello scienziato che cerca l'informazione.

L'indice delle citazioni viene così descritto come un indice "basato sull'associazione delle idee", in grado di offrire al lettore tutta la libertà d'azione che richiede, permettendo un metodo di ricerca modellato sul soggetto.

Garfield fa riferimento ad alcuni esempi di indici che funzionano, già dalla prima metà dell'Ottocento, grazie a sistemi analoghi, e propone di sviluppare un codice di citazioni per la scienza: ad ogni articolo è associata una doppia serie di numeri, la prima corrispondente alla rivista in cui è stato pubblicato, la seconda scelta arbitrariamente, che identifica l'articolo stesso. La tabella mostra un esempio del sistema di codifica di un articolo:

**Table 1**

Index sample based on article by Hans Selye, "General adaptation syndrome" [*J. Clin. Endocrinol.* '6, 117 (1946)]

The code number for this journal in the *World List* is 11,123a; the article number is arbitrarily taken as 687 ; and the code number for the article is 11123a-687. The 23 articles that cited Selye's article are listed, followed by *A hypothetical citation index entry for Selye's article: R, review article ; A, abstract ; O, original article.*

1. Williams, R. H. : Thyroid & Adrenal Interrelations, 7: 52—57 (1947).

2. Venning, E. H. : Glycogenic Corticoids, 7: 79—101 (1947).

Nella figura l'elenco delle fonti citanti si limita alla seconda voce; per l'elenco completo e una descrizione dettagliata del meccanismo, si può [consultare](#) l'articolo di Garfield on-line.

Sotto ciascun codice numerico, per esempio 3001-6789, sarebbero ordinati altri codici numerici che rappresentano gli articoli che citano l'articolo in questione, accompagnati da un'indicazione sul tipo di fonte citante (es. articolo, recensione, traduzione, opera originale, etc..). Nei fatti, il sistema offrirebbe una lista completa, *relativa alle pubblicazioni coperte*, di tutti gli articoli originali che hanno citato l'articolo in questione.

Come lo stesso Garfield mette in luce, l'impatto di un articolo si riferisce ad un elenco di pubblicazioni limitato (il corsivo nel testo). Perché tale sistema funzioni è infatti necessario stendere un elenco delle pubblicazioni (le riviste) e catalogarle (assegnando a ciascuna

un codice, nell'esempio sotto: 11123s). L'esempio in figura (sotto) mostra l'elenco delle citazioni di un articolo, e indica le fonti citanti.

### Citation Index Entry

11123s-687

---

464-9789(R)  
 869-3366(R)  
 1105-9876(A)  
 1123-4432(R)  
 a11,123-0752(0)  
 -0779(0)  
 -7264(0)  
 -7331(0)  
 -7385(0)  
 -0866(0)  
 -8221(0)  
 -9158(0)  
 -9497(0)  
 -9529(0)

Nello stesso articolo Garfield suggerisce possibili impieghi pratici di un simile indice, che si rivelerebbe in particolare molto utile nelle ricerche storiche al fine di valutare il significato di un'opera calcolando il suo *impatto* sulla letteratura e il pensiero del tempo, e nelle scienze umane, filosofiche e sociali, nel tracciare percorsi che ricostruiscono l'origine e l'evoluzione di un'idea. Garfield aggiunge che un tale *fattore d'impatto* (*impact factor*) avrebbe potuto rivelarsi uno strumento di valutazione più indicativo e significativo rispetto al conteggio assoluto del numero delle pubblicazioni scientifiche. Infine, l'indice delle citazioni avrebbe facilitato la comunicazione tra scienziati e la nascita e lo sviluppo di nuove idee.

Garfield si pone il problema di gestire grandi quantità di dati, e propone di applicare il sistema ad un numero ristretto di riviste, poi rientrate nel numero dei "*core journals*", le riviste più prestigiose. Il sistema abbraccia, dalla sua ideazione, un'impostazione centralizzata al controllo bibliografico - e se i limiti di tale impostazione sono connaturati al mondo della stampa, e un'impostazione distribuita poteva suonare impensabile a causa della limitatezza dei mezzi a disposizione (in primo luogo dall'assenza della rete), questi limiti hanno riscontri negativi di portata ampia.

L'accentramento dell'indice:

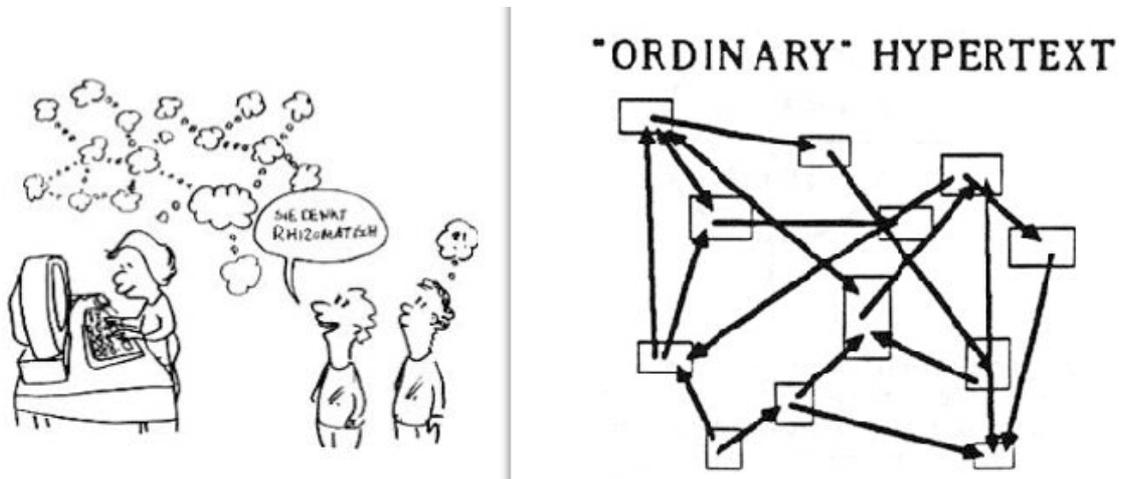
- può diventare un potente strumento di controllo dell'informazione, e creare profonde distorsioni nell'organizzazione della comunicazione scientifica. Lo SCI, ristretto ad un numero limitato di riviste fondamentali (*core journals*), si è rivelato un efficace strumento per la gestione

delle carriere universitarie ed ha attribuito molto potere ad un ristretto numero di editori, dando luogo alla cosiddetta "crisi del prezzo dei periodici" <sup>3</sup>.

- Riduce la possibilità di trovare informazioni al di fuori del nucleo fondamentale delle riviste SCI, limitando l'impatto di un sistema basato sulle associazioni; se la ricerca si nutre spesso di annotazioni secondarie e segue piste parzialmente esplorate, l'esclusione di intere aree della conoscenza umana dall'indice crea una "scienza di serie A" e una "scienza di serie B", e rende l'accesso alla seconda pressoché impossibile.

## L'ipertesto

A metà degli anni Sessanta, quando Internet non era ancora stata concepita, anche Teodor Nelson, umanista di formazione, ha avanzato l'ipotesi di costruire un sistema informativo basato sulle associazioni. A differenza del sistema di Garfield, orientato prevalentemente al lettore, il progetto di Nelson non distingue il punto di vista del ricercatore-lettore e quello dell'autore: l'"ipertesto", scrive Nelson, è la forma più generale di scrittura, una scrittura non sequenziale che riflette la struttura di ciò che scriviamo e facilita la produzione collaborativa e il riuso della conoscenza. La struttura reticolare del pensiero (figura di sinistra), viene riprodotta nell'organizzazione dell'informazione collegata attraverso associazioni (figura di destra):



In *Literary Machines 90.1*, Nelson descrive, contestualmente all'idea di "ipertesto", il progetto (oggi ancora in atto) Xanadu, "un sistema universale di editoria e di archiviazione elettronica". *Xanadu* è un programma per creare e organizzare la letteratura scientifica in forma ipertestuale che, se realizzato, incarnerebbe tanto il ruolo degli editori, quanto i compiti delle biblioteche tradizionali. Si tratterebbe infatti di una vera e propria reinvenzione della

biblioteca per costituire un unico sistema planetario integrato e decentrato, in grado di rappresentare la struttura dell'informazione tenendo traccia del [modo di operare quotidiano](#) del ricercatore, e di offrire [strumenti più adatti](#) a metterlo in pratica.

L'ipotesi alla base di *Xanadu* era la possibilità di creare un sistema che unificasse tutti gli altri, in grado di offrire tanto uno strumento di archiviazione quanto un mezzo di indicizzazione per il materiale archiviato – “una rete ipertestuale aperta e facile da usare”. Sin dalla sua prima formulazione, l'ipertesto avrebbe mantenuto una parte delle funzioni assolve dalle riviste cartacee e ne avrebbe integrate altre: “lo sviluppo della parola scritta ha perfezionato, in un continuo processo evolutivo, i meccanismi della citazione, del riferimento e della bibliografia. Le pubblicazioni scientifiche hanno sviluppato convenzioni molto precise, i giornali popolari sono molto più vaghi rispetto alle fonti, e la maggioranza di ciò che si pubblica è una via di mezzo tra questi due estremi”. La struttura dei collegamenti, indipendentemente dal rigore con cui sono codificati, scrive ancora Nelson, fornisce gli elastici invisibili che, attraverso i documenti, tengono assieme i pensieri. Così, *Xanadu* e programmi simili avrebbero svolto un'importante funzione formativa, abolendo la sequenzialità e promuovendo l'iniziativa; ognuno sarà libero di scegliere la propria spiegazione, profetizzava nel 1981 Nelson, e di tornare sui propri passi, come il lavoro creativo impone.

La previsione di Nelson suona molto simile a come oggi il Web ci appare: un enorme contenitore o una ragnatela di informazioni di tipo molteplice, scientifiche e no. La caratteristica fondamentale che accomuna l'ipertesto e il Web è la possibilità di evidenziare le relazioni tra documenti, ovvero di collegarli tra loro, citandoli. A differenza di *Xanadu*, il Web non è un programma, ma un insieme di protocolli, e la pubblicità dei protocolli rende lo spazio dell'informazione aperto ad ogni tipo di contenuti, compresi i cataloghi delle biblioteche, e le risorse tradizionalmente scientifiche. Nell'ultima sezione dell'ipertesto, che segue, confronteremo diversi strumenti pratici per la ricerca della letteratura scientifica in rete.

Anche se il Web contiene grano frammisto a loglio, le potenzialità di un sistema informativo basato su collegamenti si sono mostrate dirimpenti nello stesso ambito ristretto della comunicazione scientifica: i link ipertestuali possono essere seguiti nel giro di pochi secondi, invece che in settimane di telefonate e di inoltri della posta, e le possibilità della ricerca tramite collegamenti si stanno sviluppando in più direzioni.

Tuttavia, perché Internet e il Web possano essere a pieno titolo considerati mezzi per la comunicazione scientifica deve essere affrontato il problema, che anche Tim Berners-Lee [non ha mancato di considerare](#), relativo alla selezione delle risorse secondo criteri di qualità; è lo stesso problema affrontato da Garfield, per la cui risoluzione l'inventore del Web avanza una diversa risposta, basata

sulla decentralizzazione e su un [sistema di filtri semantici](#) scelti dal navigatore che restano all'esterno del sistema; grazie ad essi, il Web del futuro sarà probabilmente in grado di sfruttare le nuove tecnologie e di trasformare i lettori in veri e propri "battitori di piste", che sappiano scegliere le piste migliori e arricchire le precedenti, aprendo percorsi nuovi e inesplorati.

## Riferimenti e link rilevanti

Francesca Di Donato. *"As we may think" di V. Bush (scheda di lettura)*.

Francesca Di Donato. *"The World Wide Web - Past, present, future. Exploring Universality" di T. Berners-Lee (scheda di lettura)*.

Eugene Garfield. *Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas*.

Jean Claude Guédon. *La lunga ombra di Oldenburg: i bibliotecari, i ricercatori, gli editori e il controllo dell'editoria scientifica*.

*Ted Nelson su Wikipedia*.

### *b. La pratica della ricerca in rete*

Come si legge in rete <sup>4</sup> ? Banalmente, esistono tre modi per raggiungere una pagina sul Web:

- *inserendo l'indirizzo (URL) esatto nel browser*, si stabilisce una connessione diretta col punto desiderato (quando ad esempio cominciamo una ricerca, e digitiamo l'URL [www.google.it](http://www.google.it) nel browser, il protocollo http ci collega direttamente al sito di *Google*);
- *navigando*, cioè seguendo una pista tra i link che collegano il punto di partenza al punto di arrivo. Il termine "navigare" è una traduzione, impropria, dell'inglese "browsing", letteralmente sfogliare. Navigare sul web è molto diverso dallo sfogliare un libro, poiché i collegamenti tra le pagine sono ipertestuali.
- *cercandola tramite un motore di ricerca*. I motori di ricerca fanno uso di differenti algoritmi (generalmente segreti) che rispondono a criteri abbastanza noti: Peter Suber, uno dei

massimi esponenti del movimento Open Access, ha recentemente redatto, assieme con Google, un elenco di raccomandazioni da seguire per rendere i documenti depositati negli archivi aperti, più facili da indicizzare e da trovare a chi fa ricerche con Google.

**Motori di ricerca.** Un'utile pratica è provare diversi motori (si può cominciare a partire dalla pagina dell'hacker finlandese Fravia: <http://www.searchlore.org/main.htm>, dietro cui vale la pena perdere un po' di tempo), sperimentando sia i diversi motori, sia le modalità di ricerca avanzata.

**Biblioteche in rete.** Strumenti specifici per la ricerca scientifica sono gli OPAC (cataloghi on-line ad accesso pubblico delle biblioteche), tramite i quali si ottengono le informazioni bibliografiche e la collocazione nella biblioteca (es. [la Library of Congress](#); il [MetaOPAC pisano](#)). Le biblioteche offrono oggi potenti strumenti di ricerca, e -in alcuni casi - l'accesso via rete alle risorse digitali; sempre più fonti sono accessibili, anche grazie alla diffusione del movimento *Open Access*<sup>5</sup>. I portali delle biblioteche universitarie indicano generalmente un elenco delle riviste ad accesso aperto e delle riviste per cui l'Ateneo ha pagato l'abbonamento, rendendole accessibili a studenti e ricercatori. La "[Directory of Open Access Journals](#)" indicizza più di 1400 riviste che fanno parte della rete OAI.

**Altri servizi.** Per informazioni su libri recenti e stranieri che non si trovano in biblioteca, è utile consultare anche siti commerciali come *Amazon*, che offrono servizi di grande utilità anche per la ricerca bibliografica come "Customers who bought this book also bought".

**Testi e classici on-line.** In Italiano, esistono alcuni progetti degni di nota che mettono a disposizione libri, articoli e, in generale, fonti primarie (due esempi sono la [Biblioteca telematica progetto Manuzio \(Liberliber\)](#), il [Bollettino telematico di filosofia politica](#)); in inglese e in altre lingue esistono numerosi strumenti che facilitano la ricerca di letteratura scientifica, e permettono l'accesso ai contenuti. Tra questi il [Perseus](#) (particolarmente utile per leggere i classici greci e latini, in lingua originale e in traduzione inglese) e [HyperNietzsche](#) (che permette l'accesso ai manoscritti del filosofo tedesco, alle trascrizioni e alla letteratura secondaria)

**La ricerca delle citazioni.** Recentemente, sono stati sviluppati strumenti ancora non molto evoluti che rintracciano le citazioni tra i testi. Due esempi significativi sono [Citeseer](#) e [Citebase](#). Si veda anche [CiteUlike](#), che permette di visualizzare le citazioni sotto forma di grafo.

## A. Schede di lettura

### “As we may think” di Vannevar Bush

Nel 1945 Vannevar Bush, direttore capo dell'Ufficio americano per la Ricerca e lo Sviluppo scientifico dal 1941, pubblica “As we may think”<sup>6</sup>. Se durante la seconda guerra mondiale Bush aveva messo la propria opera al servizio dello Stato e, come il matematico inglese Alan Turing, aveva ricoperto un importante ruolo nella ricerca militare, il saggio del 1945 è una riflessione sul rapporto tra la tecnologia e il problema politico, assai attuale in epoca post-bellica, della felicità universale e della pace nel mondo. “Come può la tecnologia contribuire al benessere dell’umanità?” si domanda lo scienziato americano. La risposta viene costruita sciogliendo, nel corso del saggio, un’altra questione a monte: come può il bagaglio della conoscenza umana accumulatasi fino ad oggi e a venire, aiutare l’uomo a vivere in pace? Perché le tecnologie possano mettere l’uomo in grado di costruirsi “una casa in cui possa vivere in buona salute”, suggerisce con una metafora Bush, è necessaria un’indagine sul metodo della ricerca e sul processo di costruzione della scienza. Il saggio non affronta dunque una questione meramente tecnica (si veda la descrizione del *memex*); viceversa, l’argomentazione sui cui è imperniato è anche e soprattutto una riflessione filosofica e politica su come si produce e si comunica il sapere.

Vannevar Bush fonda il proprio ragionamento su una importante premessa, il fatto che la conoscenza umana è un insieme collegato in un tutto che, in quanto tale, ha una dimensione universale che non può essere limitata alla vita del singolo. Il sapere è tale in quanto frutto di un processo accumulativo, e si costruisce, grazie alla collaborazione degli scienziati, in sistemi collegati che includono l’intero patrimonio delle conoscenze umane. Per questo, l’accesso all’informazione scientifica è una condizione necessaria alla stessa possibilità della scienza.

#### *1. Il problema dell'accesso*

La conoscenza, per essere utile alla scienza, deve poter essere continuamente ampliata, in primo luogo archiviata e soprattutto consultata.

Le moderne tecnologie della parola (prima tra tutte la stampa) hanno consentito all’uomo di produrre enormi quantità di informazioni a un

ritmo sempre crescente. All'aumentare della quantità di dati e di conoscenze raccolti in libri, articoli, e forme scritte di diversa natura, non siamo altrettanto in grado di condividere tale patrimonio; la crescente specializzazione del sapere, prosegue Bush, confligge con l'esigenza di sistemi di comunicazione veloci ed efficienti.

Il principale ostacolo all'accesso della conoscenza risiede negli alti costi di riproduzione, e può essere facilmente superato con l'ausilio di tecnologie in grado di comprimere grandi quantità di libri in dispositivi di piccole dimensioni. La proposta di Bush, che nel 1945 è in anticipo sui tempi ma che nel mondo di oggi è diventata una realtà, indica che il problema dell'archiviazione può essere facilmente risolto con l'aiuto delle macchine. Viceversa, aggiunge lo scienziato americano, il problema della consultazione del sapere è assai più serio e di difficile soluzione. Esso infatti coinvolge l'intero processo grazie al quale l'uomo trae vantaggio dal suo bagaglio di conoscenze, la selezione. Bush riconosce che il cuore del problema è più profondo del semplice ritardo tecnologico, nella meccanizzazione delle biblioteche *in primis*. Il metodo della selezione funziona "come uno scalpello nelle mani di un ebanista", e pertanto dev'essere considerato e affrontato con riguardo al metodo, più generale, della ricerca scientifica, cui l'ingegnere americano dedica ampio spazio nella parte centrale dell'articolo.

## 2. Il metodo della ricerca scientifica

### a. Connettere tramite associazioni

Il cuore del problema non dipende da una incapacità dell'uomo, ma deve essere ricondotto all'artificialità dei sistemi di indicizzazione in uso negli archivi. Si tratta di sistemi gerarchici, che organizzano l'informazione secondo una struttura ad albero (un'immagine di tale struttura è l'albero delle directory del computer, in cui i file sono organizzati in catelle e sotto-cartelle). Le informazioni, a meno di duplicati, si trovano in un unico punto dell'archivio, e perciò si devono avere delle regole per decidere quale cammino ci porterà alla informazione che cerchiamo. Il problema è che queste regole sono difficili da utilizzare e da gestire. Il limite di tale struttura, è che a un documento corrisponde un unico percorso.

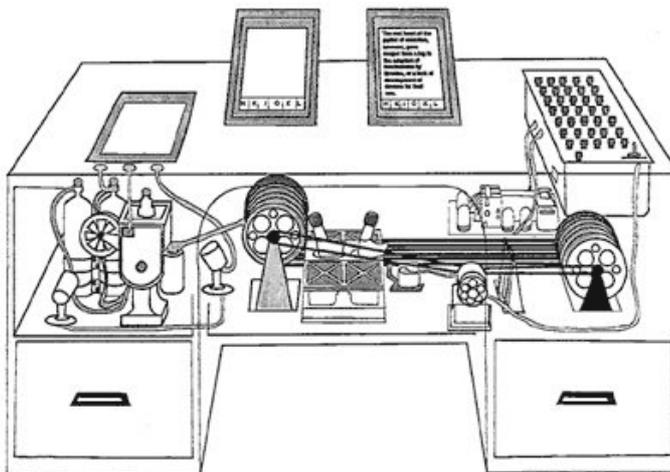
La mente umana non funziona in questo modo. Essa opera per associazioni. Una volta che essa abbia un elemento a disposizione, salta istantaneamente all'elemento successivo suggerito, in base a un intreccio di piste registrate nelle cellule del cervello, dalla associazione dei pensieri.

Bush suggerisce di ripensare la selezione dell'informazione a partire da *come possiamo pensare*, e indica, nel *memex*, un esempio di meccanizzazione della selezione per associazione, piuttosto che per

indicizzazione. L'accesso virtualmente universale al patrimonio culturale dell'umanità e l'impiego delle macchine nella ricerca hanno, ai suoi occhi, un potenziale creativo dirompente. Se infatti non esistono sostituti meccanici per il pensiero creativo, ma il pensiero creativo e il pensiero essenzialmente ripetitivo sono cose molto diverse; tuttavia ogni qualvolta si combinano e si registrano fatti sulla base di processi logici stabiliti, l'aspetto creativo "riguarda solo la scelta dei dati e del processo da impiegare, e da quel momento in poi la manipolazione è di natura ripetitiva, e dunque adatta per essere affidata alle macchine".

### b. Il *memex*

Bush immagina un dispositivo meccanico, una scrivania meccanizzata formata da schermi traslucidi inclinati, una tastiera, e gruppi di bottoni e di leve. All'interno, sono archiviati gli strumenti di lavoro del ricercatore (i libri e gli articoli su cui lavora, come gli appunti che produce), che può velocemente proiettare sugli schermi inclinati, come mostra la figura:



Il modo in cui il *memex* funziona, aiuta il ricercatore a riportare alla memoria e alla vista i dati del proprio archivio, ma fa anche qualcosa in più, in quanto agisce da estensione della memoria individuale (il nome è la contrazione di *memory extender*). Grazie ad esso infatti è possibile consultare il proprio archivio tramite lo schema di indicizzazione tradizionale: basterà digitare il codice del libro desiderato, e la copertina del libro sarà rapidamente proiettata su uno dei visori. Ma la sua caratteristica fondamentale è che permette di collegare due elementi tra loro: l'utente che si trova di fronte due schermi adiacenti, non dovrà fare altro che premere un tasto, e avrà stabilito un collegamento stabile tra i due testi. La valenza pratica della macchina ideata da Bush è descritta dall'autore nell'esempio che segue:

Il proprietario del *memex*, per esempio, è interessato alle origini e alle proprietà dell'arco. In particolare, sta cercando il motivo, nelle battaglie delle Crociate, per cui l'arco corto turco fosse superiore all'arco lungo inglese. Ha a disposizione dozzine di libri e articoli pertinenti nel suo *memex*. Dapprima sfoglia un'enciclopedia, trova un articolo interessante

ma non approfondito, lo lascia proiettato. Poi, in un resoconto storico, trova un altro argomento pertinente, e lo congiunge al precedente. Procedo in questo modo, costruendo un percorso di molti elementi. Di quando in quando inserisce un proprio commento, sia congiungendolo alla pista principale sia connettendolo come pista laterale a un elemento particolare. Quando diventa evidente che le proprietà elastiche dei materiali disponibili erano fondamentali per il rendimento dell'arco, egli devia su una pista laterale che lo porta attraverso libri di testo sull'elasticità e tavole di costanti fisiche. Inserisce una pagina di analisi scritta di suo pugno. In questo modo egli costruisce una pista della sua ricerca attraverso il labirinto del materiale che ha a disposizione.

È importante osservare che Bush anticipa, in queste pagine, una definizione della principale modalità di movimento all'interno del World Wide Web, la navigazione di pagina in pagina (nodi) tramite collegamenti (link)<sup>7</sup>. Tale sistema porta a costruire percorsi di lettura, piste di cui non restano solo i risultati, ma anche i passaggi. Così, quando numerosi elementi sono stati congiunti a formare una pista, grazie al *memex* possono essere esaminati l'uno dopo l'altro, velocemente o lentamente, muovendo una leva come quella usata per girare le pagine del libro. È come se gli elementi, parte di libri preesistenti o isolati, venissero raccolti separatamente e quindi rilegati nuovamente assieme, così da formare un nuovo libro.

### *3. Battitori di piste*

Nascerà la nuova professione di battitore di piste, persone che si ingegneranno a tracciare percorsi significativi attraverso l'immane mole dell'esperienza umana. L'eredità del maestro ai suoi discepoli non saranno più solo i suoi contributi alla conoscenza comune, ma l'intera impalcatura di conoscenze sulla quale essi sono stati costruiti.

Bush restituisce al ricercatore un ruolo attivo e fondante. L'attività della ricerca consiste nell'inserimento di ogni singolo elemento in una o più piste; tali piste restano così patrimonio dell'umanità, tanto quanto i libri, e danno vita ad enciclopedie di concezione radicalmente nuova, attraversate da una trama di percorsi associativi. Proseguendo nell'esempio sopra menzionato, lo scienziato immagina come il proprietario del *memex* potrà servirsi, in seguito, della propria ricerca sulle origini e le proprietà dell'arco: "Qualche anno dopo, durante un suo colloquio con un amico, il discorso cade sui modi strani in cui le persone resistono alle innovazioni, anche se di vitale interesse. Egli dispone di un esempio, ed esattamente di come gli europei, pur riconoscendo che la loro arma aveva una gittata minore, si rifiutarono di adottare l'arco turco: dispone addirittura di una pista al riguardo. Un tocco mostra l'elenco dei codici. La pressione di alcuni tasti porta all'inizio della pista. Una leva permette di scorrerla a piacimento, fermandosi sui punti interessanti e dipartendosi in escursioni laterali. È una pista interessante, pertinente alla discussione. Così aziona un riproduttore, fotografa l'intera

pista e la passa all'amico perché la inserisca nel proprio *memex*, congiungendola opportunamente ad altre sue piste”.

## "Macchine calcolatrici e intelligenza" di A. Turing

Nel saggio del 1950 *Macchine calcolatrici e intelligenza*<sup>8</sup>, il matematico inglese Alan Turing imposta una discussione filosofica sulla questione se le macchine possano pensare, una domanda nel seguito molto esplorata e dibattuta; lo scienziato sa bene che affrontare tale questione richiede di rispondere a due domande preliminari: che cosa significa "pensare"? Che cosa è una "macchina"?

Ricostruiamo, per cominciare, la risposta di Turing alla seconda domanda. Eviteremo in questa sede di entrare nel merito del problema, estremamente dibattuto, sull'effettiva possibilità di un'intelligenza artificiale, per concentrarci invece sulla definizione filosofica (astratta e formale) di calcolatore che propone il principale teorico della macchina calcolatrice universale<sup>9</sup>.

La strategia argomentativa del matematico inglese in via preliminare sembra aggirare il problema della definizione; Turing infatti riformula la domanda originaria (possono le macchine pensare?) nella forma del "gioco dell'imitazione". La simulazione viene giocata da tre persone, una donna (Alice, poi A), un uomo (Bob, poi B), e l'interrogante (Trudy), che può essere dell'uno e dell'altro sesso. L'interrogante viene chiuso in una stanza, separato dagli altri due. Scopo del gioco per l'interrogante è quello di determinare quale delle due persone sia la donna (X) e quale l'uomo (Y). Egli le conosce con le etichette X e Y, e alla fine del gioco darà la soluzione "X è A e Y è B" o la soluzione "X è B e Y è A". L'obiettivo dei due giocatori è differente: scopo di Bob è aiutare Trudy a trovare la giusta soluzione; scopo di Alice, al contrario, è ingannare l'interrogante. Ovviamente, Bob farà di tutto per essere d'aiuto a Trudy, dicendo la verità. Bob potrà così scrivere: "Sono io l'uomo!". Ma lo stesso potrà fare Alice: e la medesima affermazione, fatta da Bob e da Alice, non farà che aumentare la confusione dell'interrogante.

La domanda a cui Turing intende dare una risposta è "cosa accade se una macchina prenderà il posto di Alice?" La frequenza di risposte errate di Trudy sarà comparabile al caso originario, in cui una donna giochi il ruolo di Alice? (pp. 116-7, trad. it.) La definizione di macchina calcolatrice si costruisce tramite il confronto con l'essere umano razionale, un confronto per la ricerca della verità che trova luogo nel dialogo, il metodo del *logos* filosofico per eccellenza.

Turing fornisce dunque una spiegazione del funzionamento della macchina in oggetto. Le macchine interessate al gioco sono i calcolatori numerici, "macchine costruite per compiere qualsiasi operazione che possa essere compiuta da un calcolatore umano". La metafora di calcolatore proposta da Turing è costruita tramite una descrizione dei metodi che accomunano la macchina e l'essere umano nel rispondere criticamente a una domanda.

È così possibile suddividere il calcolatore numerico in tre parti:

- La prima è la memoria, un deposito di informazioni che

comprende svariati strumenti a disposizione del calcolatore umano: la carta su cui scrive i suoi conti, la carta su cui è stampato il suo "libro delle regole", vale a dire la tavola delle istruzioni da seguire (il programma), e la porzione di cervello in cui, durante il calcolo, l'uomo ritiene le informazioni.

- La seconda parte, il complesso operativo, compie le operazioni che un calcolo comporta;
- il governo, la terza componente, controlla infine che le istruzioni siano eseguite correttamente e nell'ordine giusto (pp. 121-23).

La spiegazione proposta da Turing è disegnata sulla base di un modello tripartito in analogia con il calcolatore umano; la divisione dei poteri si presta ad essere interpretata anche sul piano filosofico e politico.

Nelle pagine che seguono, lo scienziato inglese procede a dimostrare la sua ipotesi che, lo ricordiamo, è tesa a stabilire un'analogia matematica tra il calcolatore numerico e quello umano<sup>10</sup>. Turing prende dunque in esame le principali obiezioni all'ipotesi che le macchine possano pensare, e offre, al contempo una propria definizione di pensiero e di intelligenza; tale definizione nasce dalla confutazione dell'argomento proposto da Ada Byron e noto come [obiezione di Lady Lovelace](#).

Della prima programmatrice della storia, Turing riporta un'affermazione tratta da un saggio sulla macchina analitica di Babbage. "In esso, scrive, si afferma "La macchina analitica non ha la pretesa di *creare* alcunché. Può fare *qualsiasi cosa sappiamo come ordinarle di fare*" (il corsivo è nel testo)."

La macchina sarebbe dunque incapace di fare cose nuove, limitandosi ad eseguire ordini e dunque a compiere azioni le cui conseguenze sono, per l'uomo, prevedibili. Ma è veramente così? Se né Lady Lovelace né Babbage erano a conoscenza degli sviluppi e delle potenzialità della macchina analitica, si tratta comunque di un'affermazione sbagliata; di essa, tuttavia, Turing confuta due varianti per approdare ad una definizione di intelligenza.

La variante debole sostiene che una macchina non possa mai fare qualcosa di veramente nuovo. "Si può eludere per il momento l'obiezione col detto "Non c'è nulla di nuovo sotto il sole"" è la risposta ironica di Turing; che aggiunge: "chi può essere sicuro che il "lavoro originale" da lui compiuto non sia stato semplicemente la crescita di un seme gettato dall'insegnamento, o la conseguenza dell'aver seguito principi generali ben noti?"

Segue, infine, la variante forte dell'obiezione di Lady Lovelace, secondo cui una macchina non è mai in grado di cogliere un uomo alla sprovvista. Lo scienziato confuta tale tesi sulla base della propria esperienza, affermando di essere continuamente stupito dalle macchine. Egli sa bene, con questa risposta, di non portare un argomento convincente e risolutivo. Il suo critico probabilmente controbatterà che simili sorprese sono dovute a qualche atto mentale creativo da parte sua e che, in conseguenza, nessun merito deriva dalle macchine.

Tuttavia, aggiunge, "vale forse la pena di osservare che valutare qualcosa come sorprendente richiede sempre un "atto mentale creativo", tanto nel caso che ciò che sorprende sia provocato da un uomo, quanto nel caso che si tratti di un libro, di una macchina o di qualsiasi altra cosa. L'opinione che le macchine non possano far nascere sorprese è dovuta spesso a un errore cui

sono particolarmente soggetti filosofi e matematici. L'errore consiste nel presupporre che appena un fatto si presenta alla mente, tutte le conseguenze di questo fatto saltino fuori simultaneamente. È un presupposto utile in molte circostanze, ma ci si dimentica troppo facilmente che è falso. Una conseguenza naturale di questo modo di agire è che si presuppone che non ci sia alcun merito nella semplice elaborazione delle conseguenze di dati e principi generali" (A. Turing, *Macchine calcolatrici e intelligenza*, cit., pp. 140-42).

Turing indica una difficoltà nell'individuare l'autore di un pensiero creativo; la scienza è un processo e un progetto, ed è tale in quanto le conseguenze di un dato o di un'idea possono manifestarsi in tempi, in luoghi e in menti lontani dal momento e dal contesto in cui il primo pensiero, come un seme, si è impiantato. Stupisce lo scienziato il fatto che filosofi e matematici del suo tempo abbiano particolare difficoltà su questo punto; alla base di tale assunto sta infatti l'idea di scienza come attività collettiva, collaborativa e distribuita, un topos nella storia del pensiero filosofico e scientifico, scontato per Platone, per Newton ("sono un nano sulle spalle dei giganti") e per il pensiero illuminista. Infine, Turing affronta, al principio del passo citato, una questione solo apparentemente distante, affermando che l'atto mentale creativo (idea, pensiero o intelligenza) è un "prodotto" trasmesso tramite mezzi (il libro, il calcolatore) cioè è incorporato in un sostrato materiale. Suo scopo è ricordare che un libro e una macchina sono portatori di idee tanto quanto l'uomo che, ad esempio, fa un discorso in pubblico. Sembra un'affermazione banale, ma lo è meno se consideriamo che tale definizione chiama in causa il problema dell'autore di un'idea, e la questione della proprietà e del controllo della conoscenza.

## **"The World Wide Web - Past, present, future. Exploring Universality" di Tim Berners-Lee**

In una recente lezione in Giappone <sup>11</sup> Tim Berners Lee ha impostato una discussione filosofica sul principio a fondamento del [World Wide Web](#), *l'universalità*.

Il concetto del Web ha integrato molti diversi sistemi informativi disparati, creando uno spazio immaginario astratto in cui le differenze tra essi fossero ininfluenti. Il Web doveva includere ogni sorta di informazione su qualunque sistema. L'unica idea comune necessaria a mettere ciò assieme fu identificare un documento tramite lo *Universal Resource Identifier* (URI). Direttamente da ciò sono come piovute architetture (*design*) di protocollo (come http) e di formati di dati (come HTML) che hanno permesso ai computer di scambiarsi informazioni mappando i propri formati locali in *standard* in grado di offrire interoperabilità globale.

In termini semplificati, Berners-Lee spiega che il Web non è un programma ma un insieme di protocolli, e richiede a ciascuno di assegnare un nome unico ai propri documenti. Si tratta di una richiesta inaggirabile e necessaria al funzionamento del Web, ma ampia: è infatti l'unico limite che l'architettura del Web impone alla rappresentazione e alla organizzazione dei

dati.

Nella sua lezione, Berners-Lee insiste sul fatto che il passato, il presente e soprattutto il futuro del WWW dipendono dal modo in cui l'universalità di tale sistema è e sarà garantita e protetta. Sul piano tecnico, ciò è stato e sarà possibile postulando la separazione tra contenuto dell'informazione e forma in cui essa viene veicolata da due punti di vista principali: in primo luogo, mantenendo valida la condizione che ha permesso la nascita della ragnatela ipertestuale su Internet, vale a dire la sua indipendenza dal sostrato materiale (*hardware*) e dal modo in cui le idee sono codificate e veicolate (*software*); in secondo luogo, definendo con maggior rigore i confini tra "documents" e "data", i primi dedicati alla lettura da parte degli uomini, i secondi interpretabili dalle macchine.

**1. Passato e presente.** Tim Berners-Lee ricorda l'esigenza che ha visto la creazione del *World Wide Web*, quella di mettere a disposizione degli ospiti del CERN di Ginevra un comune strumento che permettesse loro di condividere documenti in rete. È noto che il primo uso delle tecnologie a fondamento del Web fu la rubrica telefonica del Cern e che l'utilità di un sistema condiviso ha stentato ad essere compresa, dapprincipio. L'idea di Berners-Lee era piuttosto semplice: creare un sistema di condivisione dei documenti "leggero", fondato sulla sottrazione, piuttosto che sull'addizione di requisiti; così operando, il sistema pensato dallo scienziato inglese sarebbe stato facilmente condivisibile dai diversi computer, sistemi operativi e programmi applicativi, e, in quanto tale, universale, cioè aperto a tutti coloro che volessero connettersi. Nella pratica, l'apertura del *World Wide Web* è garantita dalle specifiche del W3C, che raccomandano:

- l'indipendenza dall'*hardware* e dal *software*, importanti nel prevenire tanto la frammentazione in molti ipertesti sconnessi e proprietari quanto il pericolo che il *Web* finisca sotto il controllo monopolistico di una data società o di un solo stato;
- l'adozione di *standard* per la codifica dei caratteri che permettono la presenza sul *Web* alle diverse lingue del mondo;
- la creazione e la diffusione di *standard* per l'accessibilità, che tutelano chi è svantaggiato in termini di capacità (di vedere, sentire, muoversi e capire) e chi dispone di macchine con tecnologie obsolete (i paesi in via di sviluppo, ad esempio).

**2. Presente e futuro.** La riflessione di Tim Berners Lee si sposta su un *topos* della teoria informatica, la differenza tra cervello umano e macchina, tra "rima" e "ragione". Se il cervello umano funziona tramite un complesso [sistema di associazioni](#), e possiede la capacità di effettuare collegamenti che rientrino in un sistema coerente e consistente di conoscenze, le macchine, viceversa, hanno enormi capacità di calcolo, e possono processare l'informazione organizzata in forma di tabella, vale a dire gerarchicamente strutturata.

Per disporre di strumenti informativi più ricchi, l'architettura del Web dovrà definire con precisione i confini tra i "documenti" (i contenuti accessibili agli uomini) e i "dati"; una reale separazione permetterà alle macchine di processare le informazioni, una possibilità oggi resa complicata dal fatto che la gran parte dell'informazione collegata in ipertesto sulla ragnatela condivisa "*World Wide*", contiene entrambi gli elementi mescolati <sup>12</sup>

Lo scienziato inglese ha cura di precisare che non intende sostenere la possibilità di un'intelligenza artificiale - e il presupposto su cui si fonda il suo ragionamento, la netta differenza tra il funzionamento del cervello umano e della macchina, ne è una conferma. Egli vuole semplicemente sottolineare i notevoli vantaggi che la possibilità di trattare informazione semanticamente strutturata - favorendo l'interoperabilità, la portabilità e la durabilità dell'informazione (il riuso della conoscenza), e aprendo nuove possibilità di selezione del sapere - reca con sé.

Il discorso di Berners-Lee affronta la questione della selezione del sapere secondo criteri di "qualità". Se "è noto che una raccolta di testi, come un insieme di report tecnici o una biblioteca, include soltanto quegli articoli che raggiungano un certo livello di qualità" e "alcuni ritengono l'assenza di simili sistemi una carenza insita nel Web, tuttavia", precisa, è importante che il *Web* in sé non tenti di promuovere una singola nozione di qualità, ma continui a raccogliere il bello e il brutto, vero o falso che sia. Si tratta certamente di un limite, ammette, ma un'autorità centrale che esercitasse un controllo sulla qualità sarebbe assai più dannosa; e se nessuno dev'essere obbligato a leggere letteratura di bassa qualità, appunti oggi marginali potrebbero, un domani, essere a fondamento di nuove idee dalla portata rivoluzionaria. Il problema della selezione, dunque, può essere così riformulato:

Come possiamo fornire all'utente la percezione soggettiva di qualità elevata, e allo stesso tempo mantenere un Web aperto a persone i cui criteri di giudizio sono diversi?

La questione, risponde l'inventore del Web, dovrà essere affrontata dotando i lettori di strumenti di filtro in grado di sfruttare l'organizzazione semantica della conoscenza che li trasformeranno in veri e propri "battitori di piste" i quali, come nell'esempio di Bush, si scambieranno i percorsi e ne creeranno di nuovi, mantenendo comunque l'ipertesto comune svincolato da una autorità centrale che decida quale informazione è appropriata e per chi.

## Riferimenti utili e approfondimenti

F. Di Donato, *I media telematici come strumento per la comunicazione scientifica*

Sulla vita di V. Bush, *The Electronic Labirinth*, 1993-2000 (2004): <http://www.iath.virginia.edu/elab/hf10034.html>; una biografia dettagliata di Bush è *Internet Pioneers* di Scott Griffin: <http://www.ibiblio.org/pioneers/bush.html>

Sulla persona e la personalità scientifica di Alan Turing si confronti la biografia A. Hodges, *Storia di un enigma*, Bollati Boringhieri 2003, Torino; si veda anche la voce biografica della *Stanford Encyclopedia of Philosophy*: <http://plato.stanford.edu/entries/turing/>. Su Turing e il problema dell'intelligenza artificiale si confronti, come introduzione, la *Guida alla*

lettura del saggio di Turing su <http://www.swif.it/forum-scuola/forum03-04/2GLG/2glg-1inttutor.htm>; l'ipertesto raccoglie molte indicazioni bibliografiche. Per una panoramica sul dibattito generato dal test di Turing si veda J.H. Moor (a cura di), *The Turing Test*, Kluwer, Dordrecht 2003. Si confronti, infine, la confutazione di D. Harel, *Computer a responsabilità limitata. Dove le macchine non riescono ad arrivare*, Einaudi, Torino 2002.

Tim Berners Lee su *Wikipedia*: [http://en.wikipedia.org/wiki/Tim\\_Berners-Lee](http://en.wikipedia.org/wiki/Tim_Berners-Lee).

Per approfondire...

- F. Di Donato, [L'obiezione di Lady Lovelace](#)
- F. Di Donato, [I ponti di Königsberg e l'architettura delle reti](#)
- F. Di Donato, [Web Semantico: breve linkografia ragionata](#)

---

<sup>[1]</sup> Si tratta, ovviamente, di una rappresentazione semplificata; la stratificazione è, nella realtà, più varia e complessa. Cfr. A.S. Tanenbaum, *I moderni sistemi operativi*, Jackson università, Milano 2001. L'informatico olandese è inventore di *Minix*, un sistema operativo il cui sviluppo ha dato un impulso fondamentale alla ricerca e alla didattica dei sistemi operativi. È noto, tra gli *hacker*, l'aspro contrasto con Linus Torvalds, inventore di Linux, sul problema del kernel monolitico o a più strati.

<sup>[2]</sup> Che le idee siano "di tutti" suona oggi un'affermazione peregrina. Il concetto di paternità intellettuale infatti, pur di origine piuttosto recente, si è ben radicato nella cultura della carta stampata. La sua nascita è legata al modo in cui la comunità degli scienziati, al diffondersi della stampa a caratteri mobili, ha deciso di strutturare il discorso scientifico; la prima rivista scientifica (fondata nel 1665), funzionando da "ufficio brevetti delle idee scientifiche", garantiva, col conferimento da parte di pari di un titolo di "nobiltà", un particolare titolo di proprietà (quella intellettuale); nel 1710, lo *Statute of Anne* ne sanciva il diritto che si estendeva a 14 anni dalla pubblicazione. Oggi, in Italia, il diritto alla paternità intellettuale dura settant'anni dopo la morte dell'autore. Le cause di una tale estensione, e il dibattito relativo, sono questioni che esulano dallo spazio di questo ipertesto.

<sup>[3]</sup> Si confronti, su questo, l'articolo di J-C. Guédon, *La lunga ombra di Oldenburg: i bibliotecari, i ricercatori, gli editori e il controllo dell'editoria scientifica*, e in particolare il paragrafo 6: [Lo Science Citation Index e alcune delle sue conseguenze](#), BTFP 2004.

<sup>[4]</sup> La maggior parte dei documenti viene raggiunta a partire da un numero limitato di nodi, detti *hub* (connettori), che si collegano alle diverse risorse tramite *link* monodirezionali. Il *Web* è, in particolare, un grafo etichettato orientato, e la sua topologia rende il 60% dei nodi inaccessibili. Ciò nonostante, tramite la rete è possibile un'enorme quantità di informazioni in continua crescita.

<sup>[5]</sup> Negli ultimi dieci anni, la diffusione della rete degli "archivi aperti" (*Open archives*), ha aperto ulteriori possibilità, anche fuori dalle biblioteche e dalle loro reti locali (la gran parte delle risorse che le biblioteche acquistano, infatti, non sono accessibili a chi è esterno). Quale sarà l'impatto di simili iniziative è una questione delicata: l'accesso ai metadati (tramite [OAister](#), [ARC - A Cross Archive Search Service](#), e altri software analoghi) non comporta direttamente l'accesso ai contenuti; e l'accesso ai contenuti non è sufficiente alla disseminazione degli stessi tanto che il reale impatto dell'*Open Access*, come ha avuto il merito di sottolineare [Jean Claude Guédon](#), dipenderà da quanto sarà capace di far crescere l'accessibilità.

<sup>[6]</sup> Vannevar Bush, "As we may think", *The Atlantic Monthly*, luglio 1945, vol. 176, No. 1, pp. 101-108. Il testo in lingua originale:

<http://www.ps.uni-sb.de/~duchier/pub/vbush/vbush.shtml> . Una traduzione italiana: <http://www.uniopen.it/FC/ssd/ipertesto/think.html>.

[7] A differenza del Web, i collegamenti immaginati da Bush sono bidirezionali; vale a dire che ogni volta che un elemento viene proiettato, un altro elemento, ad esso collegato, può essere istantaneamente richiamato premendo un bottone, e visualizzato.

[8] A. Turing, *Macchine calcolatrici e intelligenza*, in V. Somenzi (a cura di), *La filosofia degli automi*, Bollati Boringhieri, Torino 1965 (*Computing Machinery and Intelligence*, Mind, London N.S. vol. 59, 433-60, 1950, disponibile in rete al seguente url: <http://www.abelard.org/turpap/turpap.htm>

[9] Con "Macchina di Turing" si intende una classe di macchine calcolatrici astratte, che consistono in un nastro infinito e in un'unità di calcolo con un numero finito di stati interni. L'unità di calcolo può leggere e scrivere in una casella del nastro, e spostarsi di una casella, a destra o a sinistra, lungo il nastro. L'operazione successiva viene determinata dallo stato attuale e dal simbolo letto. Previa scelta di un codice adeguato per l'interpretazione dei simboli, dunque, tali macchine possono effettuare calcoli numerici (A. Turing, *Sui numeri computabili con un'applicazione al problema della decisione*, 1936). Un importante corollario di tale teoria è che tutti i calcolatori numerici ordinari equivalgono ad una Macchina di Turing. Vale a dire che qualsiasi numero computabile con queste macchine o con qualsiasi ordinario processo di calcolo può essere computato con un'adeguata Macchina di Turing. Ciò non significa, invece, che tutti i problemi possano essere risolti. Al contrario, un ulteriore fondamentale corollario della tesi di Church-Turing è che se un problema non è risolubile con tale macchina, allora non è risolubile affatto; vale a dire che esistono problemi indecidibili (cfr. C.E. Shannon, *Calcolatori e automi*, in V. Somenzi (a cura di), *La filosofia degli automi*, pp 97-98). In rete sono presenti molti simulatori della Macchina di Turing (si veda, ad esempio, il sito del Dipartimento di Informatica dell'Università di Pisa: <http://www.di.unipi.it/settcult/turing.html>)

[10] La formulazione matematica dell'ipotesi di Turing affermava che entro 50 anni sarebbe stato possibile programmare calcolatori con una capacità di memorizzazione di circa 10 alla 9 (cosa che oggi si può effettivamente fare).

[11] ""The World Wide Web - Past, present, future. Exploring Universality"" di T. Berners-Lee (2002): <http://www.w3.org/2002/04/Japan/Lecture.html>

[12] La formattazione dei dati in HTML non permette che l'informazione sia processabile dalle macchine; si tratta di un limite di HTML, che non permette di separare pienamente il contenuto dalla sua struttura. Tuttavia, la nascita di nuovi formati di codifica (XML) e le nuove tecnologie di cui il W3C si è fatto promotore ([Web semantico](#)) stanno favorendo l'accentuarsi di questa separazione.

### Creative Commons License

This work is licensed under a [Creative Commons Attribution 2.0 Italy License](#).