

Mezzo secolo tra bit e ipertesti (Corriere Multimedia, novembre 1997)

La sua vita potrebbe essere presa come storia per la sceneggiatura di un film: «un'avventura di oltre mezzo secolo, per l'esattezza 56 anni, ad archiviare parole e frammenti di testo per poi trasformarli in bit». Stiamo parlando di Padre Roberto Busa, un energico gesuita vicentino, **classe 1913**. Lo studioso capace di recitare correttamente il rosario in sette lingue, vanta ricerche in 18 alfabeti. Greco e latino, piuttosto che ebraico o arabo per lui non nascondono segreti.

Il suo ciclopico lavoro è iniziato durante la seconda guerra mondiale scrivendo a mano migliaia di schede sull'uso delle parole nelle opere di San Tommaso. Una vita che ha seguito di pari passo l'evoluzione tecnologica dell'ultimo mezzo secolo. Dai primi computer a valvole, quelli che lui chiama scherzosamente "i dinosauri" fino ai sofisticati strumenti multimediali dell'ultima generazione.



Tanti i risultati ottenuti: la creazione dell'Index Tomisticus come indice di riferimento per i testi umanistici, la trascrizione in lemmi del Corano, e all'inizio degli anni '90 la produzione di un lucente cd rom sull'Opera Omnia di San Tommaso. Alla base dei suoi studi un'intuizione geniale, ripresa poi dagli scienziati di tutto il mondo come premessa per la linguistica computazionale: nell'informatica e nello studio delle lingue bisogna operare sull'ipertesto totale, quello che distingue il significato delle parole eliminando ad esempio gli omografi (stessa scrittura) e gli omofoni (la stessa

pronuncia).

Ma come si è sviluppata negli anni questa affascinante avventura del trasformare parole in bit? «Tutto ha avuto **inizio nel 1941** quando il Padre Provinciale mi comunicò che dovevo occuparmi di filosofia scolastica: in particolare delle recursività linguistiche presenti nelle opere di San Tommaso. Fui sorpreso perché in origine ero stato destinato ad un compito ben diverso. Infatti avrei dovuto imbarcarmi come cappellano militare in marina, sui sommergibili. Ma il destino e la regola "dell'obbedienza" decisero che il mio futuro sarebbe stato un altro. Niente fondali marini degli oceani, ma la penombra delle biblioteche e degli archivi».

E il lavoro si rivelò subito di immense proporzioni. Basta pensare che gli scritti dell'Aquinate contengono quasi 9 milioni di parole. Come paragone la Divina Commedia ne vanta "solo" 100 mila, mentre i Promessi Sposi arrivano a 230 mila. Ma dopo qualche mese ecco la scoperta. Tutto dipendeva dalla preposizione "in" che con tutte le sue recursività risultava essere il termine chiave delle opere.

«Così scrissi a mano, una per una 10 mila schede di carta esaminandone tutti i contesti. Ma per fortuna venne in mio aiuto la tecnologia. Nel 1949 feci il mio primo viaggio negli Stati Uniti. Mi recai all'Ibm di New York per chiedere di elaborare con le nuove macchine da calcolo le **concordanze linguistiche**, cioè l'elenco di tutte le frasi contenenti una determinata parola. Risposero che a loro la richiesta sembrava impossibile, tuttavia acconsentirono. Così iniziò la mia avventura a "bitizzare" parole».

Lo studioso gesuita divide le sue fatiche in tre periodi di sviluppo. Fino agli anni '60 l'obiettivo era quello di classificare 12 milioni di schede perforate corrispondenti, allora, ad un peso cartaceo di 500 tonnellate. Quel lavoro però non giunse a termine perché negli anni '70 arrivarono i nastri magnetici. Il Padre Gesuita ne impiegò ben 1800, una fila di "bit" lunga oltre 1600 chilometri. Da questi nastri uscirono per fotocomposizione le 70 mila pagine "in folio" dell'Index Tomisticus. E finalmente negli anni '90 l'ultima fase: un solo cd rom dal peso di pochi grammi contenente 1 miliardo e 600 milioni di byte compressi. Parliamo della versione ipertestuale **dell'Opera Omnia edita nel '92 dalla Editel di Milano**.

Oggi quale significato assumono gli ipertesti alla luce di Internet e della multimedialità? «In origine il termine aveva il significato di "scavalcare le frontiere di un file posto nella memoria di un computer, per richiamare informazioni presenti in altri file" - spiega Padre Busa - Nel corso degli anni il concetto si è allargato sconfinando nell'ermeneutica, la capacità cioè di interpretare e risalire da un'espressione al pensiero che l'ha generata. Si tratta di un processo paragonabile al "navigare a ritroso il fiume della creatività mentale". Il testo è un prodotto chiuso, ma il pensiero che lo ha creato è aperto e dinamico. Nel corso degli anni gli ipertesti sono poi diventati uno strumento informatico per creare i collegamenti dinamici (link) usati nella multimedialità e nei motori di ricerca su Internet».



Ma l'intensa attività dell'ottugenario gesuita non si ferma qui. E' di questi giorni la notizia della nuova edizione del cd rom sull'Opera Omnia, questa volta arricchita con un Thesaurus di riferimento. Inoltre con l'aiuto di chi gli è più vicino sta lavorando ad un progetto ambizioso. Mettere in piedi una collaborazione internazionale tra giovani studiosi di diverse confessioni per elaborare le voci di un lessico biculturale su San Tommaso. Si tratta di risalire ai concetti espressi con le parole del tredicesimo secolo, per valutarli con i termini in uso alle soglie del terzo millennio. Per gli ipertesti un dolce "ritorno alle origini".

Che cos'è un ipertesto?

Ipertesto, una parola fino a qualche anno fa sconosciuta ai più, ma entrata in uso grazie ai cd rom e alle pagine Internet. Ma qual'è il suo significato? Ecco la definizione riportata da un'enciclopedia multimediale: «documento elettronico costituito da un insieme di elementi informativi integrati (testo, grafica e suono) che dispone di una consultazione non sequenziale».

Fu **Theodore Nelson**, uno studente di informatica dell'università di Harvard, a coniare negli anni sessanta il termine "hypertext". L'ingegnoso studente voleva mettere ordine nei propri appunti, ma non trovò programmi in grado di risolvere il suo problema. Così lui stesso scrisse un software, chiamato **Xanadu**, per eseguire ricerche non sequenziali di parole o frammenti di testo. Nacque anche il termine "**link**" per indicare i collegamenti logici capaci di richiamare parti di file. Tanti bit, lasciati come i sassolini di Pollicino, attraverso i quali ritrovare le parole di un testo.

###