

La banca nazionale delle "soluzioni"

Un progetto speciale dell'Istituto superiore per la prevenzione e la sicurezza del lavoro

di Sergio Perticaroli e Sergio Tavassi

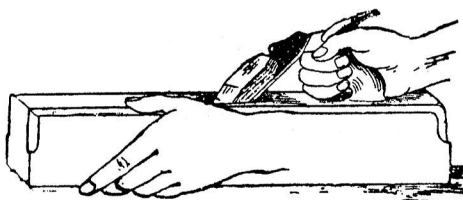
Il Dipartimento documentazione informazione e formazione dell'Ispesl (Istituto superiore per la prevenzione e la sicurezza del lavoro) aveva avviato all'inizio degli anni 1990 il progetto Sipre di collaborazione tra l'Istituto e le Regioni per la creazione di una serie di banche dati utili al Servizio sanitario nazionale.

In particolare fu avviato un progetto, in fase di completamento, di creazione di tre banche dati tra loro correlate: "Anagrafe nazionale delle ditte", "Profili di rischio dei comparti produttivi" e "Soluzioni". Quest'ultima è stata la prima banca dati delle soluzioni ad essere realizzata sia in Italia che in Europa.

Il problema

La banca dati doveva rispondere con particolari requisiti ad una serie di problemi strettamente legati sia alle considerazioni del contenuto informativo, sia alle diverse necessità di comunicazione verso gli utenti finali. Acquista in questo senso una particolare rilevanza il problema della comunicazione ad utenti finali non esperti:

— Il contenuto informativo. Il data base contiene tutte le informazioni necessarie e sufficienti per identificare una soluzione specifica per un determinato rischio professionale ed è relativa a impianti, macchi-



ne, organizzazione del lavoro, igiene ambientale e prodotti utilizzati. Il supporto informativo è rappresentato da una scheda standard (su carta o dischetto) e un corredo di disegni tecnici, schizzi esemplificatori scritti a penna, fotografie in b/n e a colori.

— La comunicazione. La banca dati doveva essere in condizioni di rispondere alle più diverse richieste ed interrogazioni da parte di utenti sia della piccola e media impresa, sia dei servizi di prevenzione delle Usl, istituti di ricerca, organizzazioni di categoria e di lavoratori. Era cioè indispensabile prevedere i più diversi livelli di evoluzione tecnologica presso l'utente e quindi una

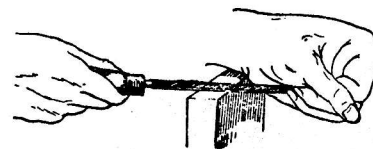
uscita in collegamento on-line, cd-rom, dischetto, stampa periodica.

La soluzione concettuale

La scelta fatta è stata quella di creare un data base testuale (documenti e immagini) da gestire con criteri documentalistici in maniera omogenea a tutte le banche dati documentalistiche del dipartimento. La ricerca doveva basarsi su full-test o su ricerca di descrittori controllati dal thesaurus Ilo-Cis. È utile accennare al fatto che tale scelta non è ovvia, dato che il corrispondente istituto olandese, su commissione Cee, è orientato a creare un data base con codifica dettagliata di tutti gli aspetti informativi, in altre parole a creare un sistema di classificazione. È noto che i sistemi di classificazione sono sistemi gerarchici rigidi percorribili solo nei due sensi possibili, dalla codifica più generale a quella più specifica e viceversa. Quando il numero degli "oggetti" da classificare diventa elevato si determina massicciamente una perdita di informazione conseguente al fatto che la griglia di classificazione è un sistema discreto e non un sistema continuo. Inoltre una struttura di classificazione presenta un ritardo sistematico di aggiornamento rispetto alla evoluzione tecnologica.

Per superare questa perdita di informazione sono state perfezionate strutture di tipo combinatorio (thesauri) che partono dal documento e ne individuano i contenuti attraverso termini (indicizzatori) preordinati.

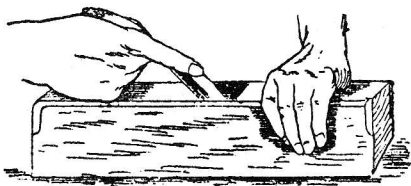
La struttura di termini preordinati contiene sia i termini concettualmente vicini, sia i termini cor- ➤



relati, sia i sinonimi e tutti sono ricercabili e gestibili con software di information retrieval.

La soluzione tecnica

La gestione dell'archivio avviene attraverso un software personalizzato, che si basa sulla piattaforma di sviluppo multimediale Fdm (Flat Document Management) con configurazione client-server: le stazioni client sono costituite da pc Ms-Windows, il server è di tipo Unix, dovendo servire accessi on-line multiutenza in rete Internet.



A questa è stato applicato un secondo server Bull dpx/20 (power-computer a tecnologia Risc) su cui è collocato un data base relazionale per la classificazione dei documenti (Oracle) e il software di information retrieval Textworks (Fulcrum Technologies Inc.).

Fdm offre tutte le funzionalità in ambiente Ms-Windows, che sono progettate in modo da essere facilmente utilizzabili da linguaggi di alto livello (visual Basic). Fdm si basa su alcune interfaccia standard: — Interfaccia Odbc (Open Data Base Connected) che permette a Fdm l'accesso al data base Oracle.

Si sottolinea che qualsiasi data base va bene (Access o Dbiv in ambiente Dos) purché supporti l'interfaccia Odbc; Fdm è quindi indipendente dal motore di classificazione.

— Utilizzo di file system per l'archiviazione del contenuto dei documenti. Lo scopo è dare ad Fdm solo una gestione logica dei supporti su cui vengono archiviati i contenuti dei documenti, rendendo

indipendente dal formato fisico (dischi ottici o magnetici).

— Per ricercare i documenti in base a operatori logici che agiscono sulle sequenze di parole, sono state integrate le conversioni di immagini in testo (Ocr).

Infine è stato attivato il trasferimento di files dal sistema operativo Unix al sistema operativo Ms-Windows al fine di distribuire il data base su dischetti, o su cd-rom con le funzioni solo di ricerca e stampa, per diffondere le informazioni a quegli utenti che non sono in grado di collegarsi via Internet.

Funzionalità

Il documento arriva al Centro accompagnato da un corredo di foto e disegni. Il documento viene scritto su pagine in formato modulo con Winword 6.0 e collegato automaticamente ad Fdm, che acquisisce chiavi di ricerca predefinite nel testo grazie al motore relazionale Oracle. Contemporaneamente Fdm integrato con Text-search indicizza tutto il documento e rende subito disponibile la ricerca full-text. Successivamente sempre Fdm

permette il collegamento al documento di un numero illimitato di immagini (foto e disegni) acquisiti via scanner.

La ricerca di un documento è possibile perciò sia utilizzando le chiavi predefinite offerte dal menù di Fdm sia attraverso la ricerca full-text di tutto il documento. In entrambi i casi le immagini sono immediatamente visualizzabili. La potenzialità di gestire film e suoni, prevista dal prodotto Fdm, non è utilizzata nell'applicazione descritta. ■

