

MOSTcold – Sistema di archiviazione documenti informatici

La piattaforma documentale MOSTdoc offre diversi moduli specializzati per affrontare temi specifici, pur conservando tutti i vantaggi offerti da un ambiente di archiviazione general purpose; uno di questi moduli permette la conservazione ed il reperimento di documenti prodotti da altri sistemi informatici. L'argomento è ben noto, ed era conosciuto negli anni passati con l'acronimo COLD (Computer Output on Laser Disc). I dischi ottici conservano una nicchia di mercato, pur essendo stati scalzati da soluzioni alternative (Nas, Centera,...), ma la problematica COLD rimane una delle più sentite nel mondo informatico, a causa dei grandi volumi di documenti, e della necessità di rendere semplice e veloce la ricerca di un documento, sia per accessi da parte dell'Help Desk, sia per richieste del cliente finale.

Per questo scopo la MOST ha sviluppato MOSTcold, modulo di gestione di file di spool, che permette l'estrazione dei documenti e dei metadati associati dagli output di stampa, e la loro indicizzazione ed archiviazione.

Sono gestiti file di spool provenienti da diversi ambienti, quali Doc1, Sidist SIAER, BPU. A seconda dei casi, i metadati per l'indicizzazione possono essere desunti dai documenti stessi, o estratti da un file di journal prodotto dalla applicazione alimentante.

Il documento può essere conservato in modalità nativa o trasformato per applicare formattazioni ad hoc, quali fincature o loghi. Il vantaggio di conservare i documenti nello stesso formato del sistema originante è sempre più limitato, vista la riduzione del costo dello storage. Per contro l'archiviazione in formato definitivo rende semplice l'adeguamento alle normative AIPA/CNIPA.

La ricerca del documento e la sua fruizione possono avvenire in vari modi:

- tramite maschere web
- mediante accesso WebDAV
- usando una modalità Web Services
- mediante estrazione di documenti su CD ROM autocontenuti

Maschera web

La “pubblicazione su web” che in molti prodotti è un modulo opzionale, in MOSTcold non ha senso di esistere, perchè tutti i documenti sono accessibili via Web. MOSTdoc è una applicazione nativamente browser-based, senza la necessità di avere programmi client specifici per l'interazione via Intranet con il server. Nel nostro caso il colloquio tra client e server avviene unicamente sulle porte http e https. La riservatezza di dati e documenti è assicurata da meccanismi di autenticazione, di encryption e di access control list, oltre alle politiche di sicurezza della rete.

WebDAV

Il medesimo archivio visto in precedenza può essere acceduto anche mediante una modalità file system; molti utenti trovano più naturale navigare in cartelle e faldoni più che effettuare delle query mediante form di ricerca. Rispetto all'accesso web le garanzie di privacy sono le stesse, i dati sono i medesimi, cambia solo lo strumento di accesso: su windows al posto di Internet Explorer si usa File Explorer.

Il percorso a cartelle che viene presentato all'utente rappresenta in modo virtuale la ricerca sull'archivio documentale, con risultati di ricerca che si autoaffinano man mano che si scende nella gerarchia delle cartelle, fino a giungere al documento interessato. Il percorso a “tree-folders” è configurabile in modo

estremamente semplice, associando a ciascun livello di cartella l'equivalente di una query sul DB. E' anche versatile in quanto è possibile personalizzare a seconda dell'utenza percorsi "logici" differenti. Ad esempio al documento di una fattura è possibile giungere tramite percorsi differenti:

- secondo una logica di protocollo:
ingresso/uscita->anno->numero di prot->documento
- secondo una logica di documentazione contabile:
bolle/fatture->anno->nome azienda->documento
- secondo una logica di aggregazione aziendale:
nome azienda->anno->tipologia documento->documento
- secondo una logica specifica :
fatture da pagare->trimestre di competenza->documento

Web Services

L'archivio può essere acceduto in modalità "program to program", mediante i web services, abilitando l'accesso ai dati e ai documenti di MOSTcold tramite una interfaccia standard (SOAP, WSDL).

Per ogni archivio documentale viene automaticamente generato dal sistema MOSTcold, il file WSDL di schema e protocollo necessario e sufficiente ai development framework degli application server a rendere veloce ed intuitiva l'integrazione di classi di oggetti-documento all'interno delle applicazioni sviluppate negli ambienti target.

CD ROM

I dati ed i documenti possono essere estratti dall'archivio e memorizzati su CD-ROM. I criteri di estrazione sono associati a query sul DB, permettendo perciò una notevole flessibilità nella selezione.

Una volta su CD, i documenti possono essere ricercati mediante maschere Web residenti su CD, e realizzate unicamente in HTML e JavaScript, rendendo quindi il CD utilizzabile su qualunque piattaforma (Windows, Mac, Unix/ Linux).

Vantaggi di MOSTcold

MOSTcold è stato progettato per installazioni di tipo enterprise, con milioni di documenti trattati e gestiti, normalmente molto omogenei tra loro. In queste realtà riteniamo più corretto puntare su requisiti quali la scalabilità, più che su meccanismi di ricerca Full Text Retrieval, adatti per numeri inferiori di documenti e poco significativi su tipologie di documento quali gli estratti conto o i fissati bollati.

MOSTcold appartiene alla famiglia MOSTdoc, e quindi trae vantaggio da tutte le features innovative di questa piattaforma:

- possibilità di utilizzare vari tipi di storage
- archiviazione a norma CNIPA
- encryption di dati e documenti
- CD autocontenuti

- varie modalità di accesso ai dati (web, webdav, web services)
- spedizioni via postel
- ambiente server multiplatforma (Linux, Solaris) e data base enterprise (Oracle, Postgres)
- client web