

La ricerca sui siti WEB

- L'accessibilità, la facile reperibilità sono caratteristiche essenziali per la vendita di prodotti non solo nel settore dei beni di consumo
- Il consumatore sempre più ha un "comportamento informato" e quindi si aspetta di avere informazioni complete e attendibili sul prodotto
- E' fondamentale conoscere il comportamento dei consumatori e proporre i prodotti secondo il loro punto di vista



L'ORÉAL
PARIS

MAYBELLINE
NEW YORK

VERSACE
VERSACE
VERSACE
VERSACE

Because for me

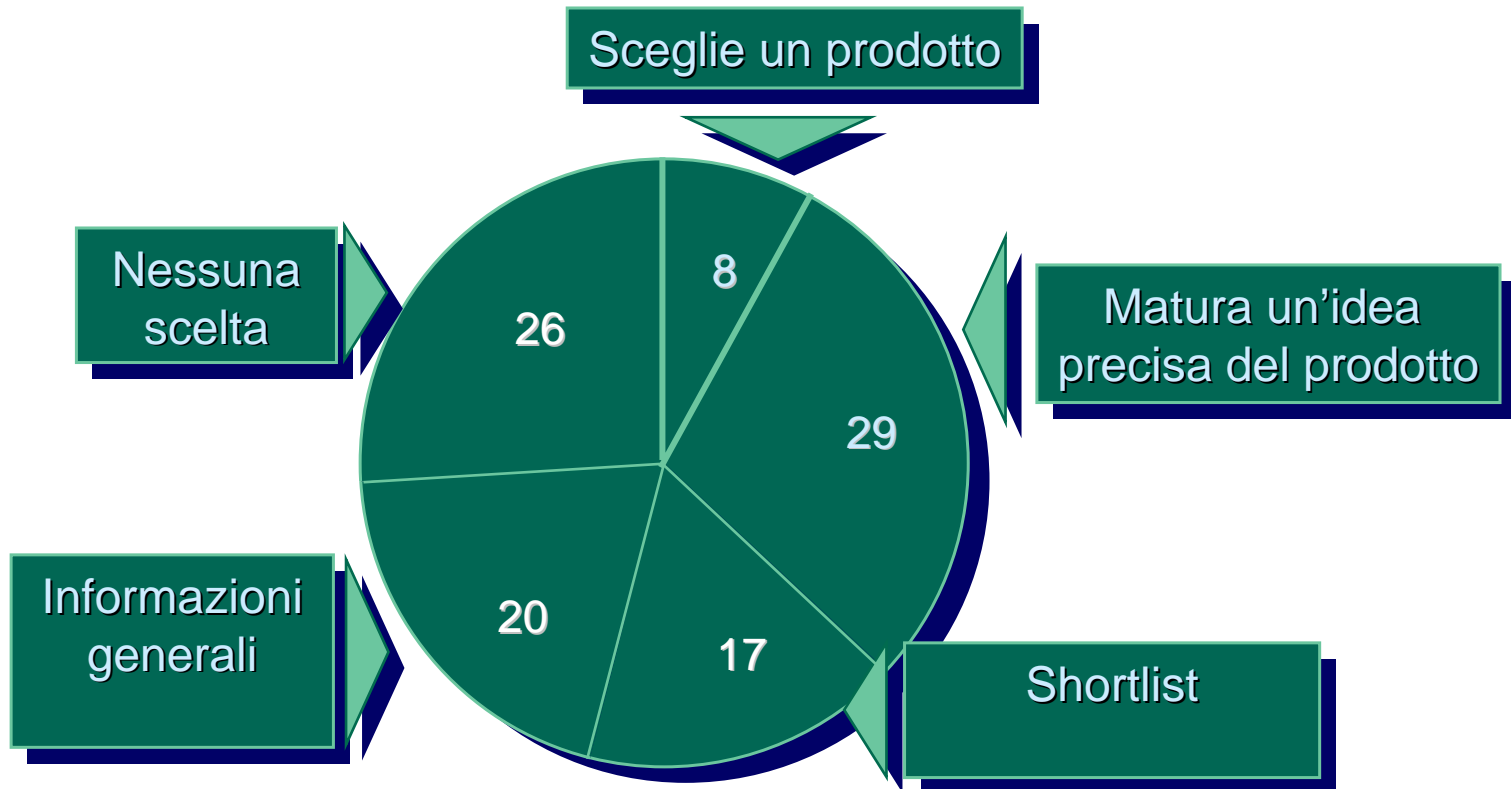
NEW! THE LATEST FOUNDATIONS

E in Internet ?

- Qual' è l'utenza di Internet?
- Nel 2010 la metà della popolazione mondiale avrà > 50 anni
- Quanta voglia avrà di "navigare" per trovare il prodotto che cerca?
- Quanto il risultato della sua ricerca deve essere preciso e immediato?



Navigare in Internet per scegliere cosa comprare in negozio



Il 60% di chi naviga in Internet lo fa per scegliere il prodotto da comprare nei tradizionali canali di distribuzione

Fonte: Boston Consulting Group

Stato dell' arte

"The quality of search on many websites is awful.

Most organizations would simply not countenance having their reception area in the mess that their search engine is in.

Yet far more people pass through their search engine every day than through their reception."

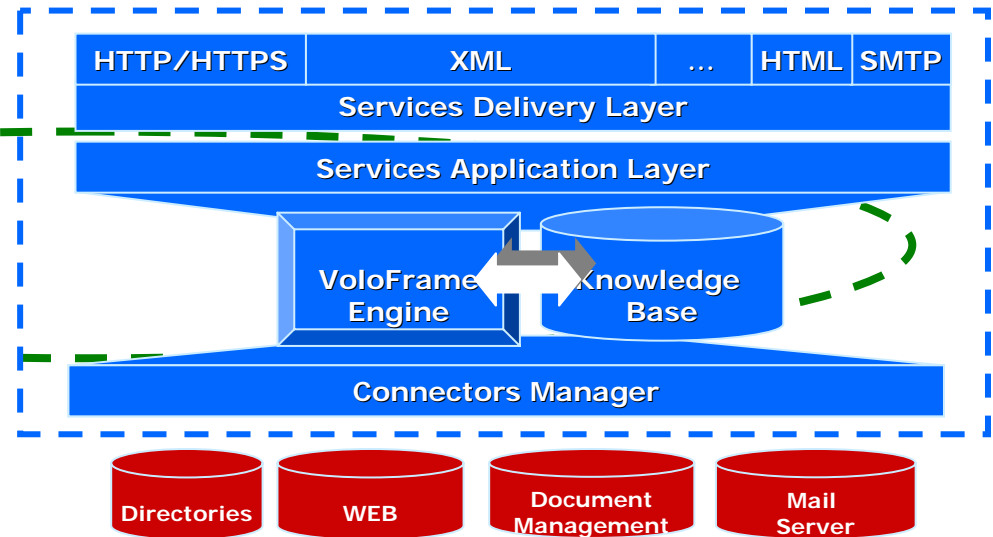
da "Content Critical" di G. McGovern



Le soluzioni Volocom

Information Retrieval

Information Retrieval



- VoloIR contiene l'insieme delle funzionalità di ricerca concettuale
- VoloIR non è solo recuperare tutti i documenti che contengono la parola ricercata, ma quella parola in quel contesto di ricerca, e si differenzia in maniera sostanziale dalle classiche metodologie di ricerca per parola chiave

Erogazione dei servizi

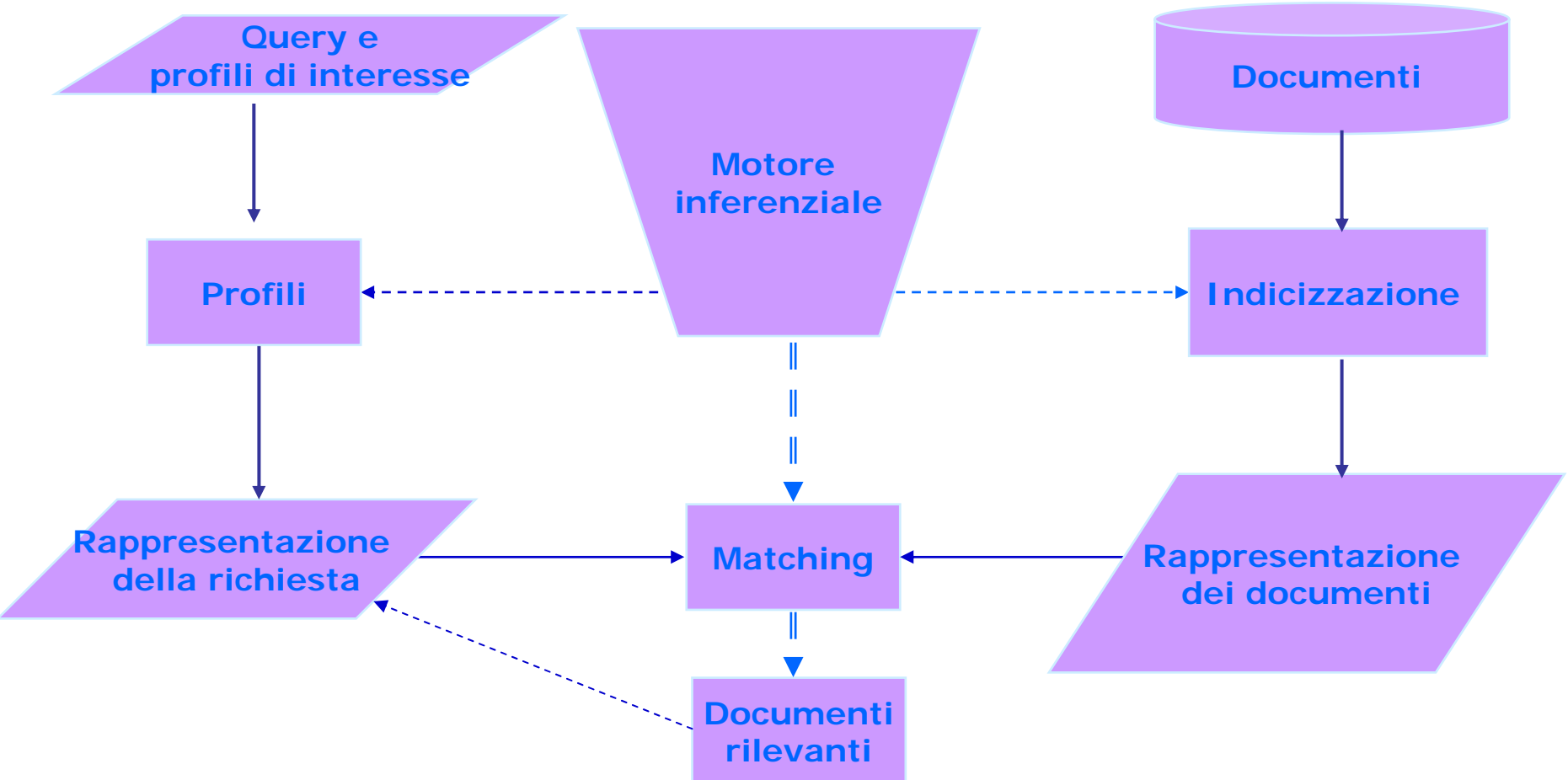
■ VoloFrame ha una struttura modulare aggregabile secondo le esigenze degli utenti

- I WEB services possono essere richiamati direttamente dalle applicazioni dell'utente
- La soluzione integrata può essere utilizzata autonomamente

Struttura del sistema IR

Ricerca

Base di conoscenza



Motore inferenziale

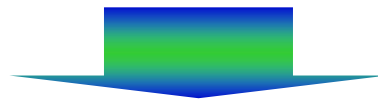
- **Interrogazione utilizzando frasi in linguaggio naturale, anziché espressioni booleane**
- **I risultati sono forniti in base alla affinità concettuale alla ricerca di partenza**
 - Identificazione dei concetti basata sulla teoria dell'informazione di Shannon
 - Analisi della rilevanza (ranking) basata su modelli probabilistici (inferenza Bayesiana)
- **Relevance feedback**
(Possibilità di utilizzare uno o più risultati per affinare la ricerca)
- **Possibilità di filtrare i risultati utilizzando dei metadati associati ai documenti (Es.: data, autore, fonte, etc.)**

Latent Semantic Indexing

- Poiché esistono diversi modi per esprimere uno stesso concetto (sinonimia), e che uno stesso termine può avere differenti significati (polisemia), la semplice ricerca “text matching” può fornire dei risultati che hanno bassa rilevanza.
- Il Latent Semantic Indexing effettua delle ricerche cercando di evidenziare una struttura semantica latente che nei differenti documenti è nascosta dalla variabilità con la quale possono essere scelte le parole.

Latent Semantic Indexing

- La tecnologia LSI esamina le parole contenute nei documenti : documenti contenenti molte parole simili vengono considerati semanticamente vicini, mentre quelli con poche parole uguali sono valutati semanticamente lontani.
- LSI quindi raggruppa i documenti semanticamente vicini e il risultato è una ricerca che trova pagine correlate anche se queste non condividono le stesse parole chiave.



Affiancando la tecnologia LSI a quelle già esistenti si possono ottenere livelli più alti di comprensione dei documenti consentendo anche la convergenza di numerosi indici di diversi database.

Abstract contestuali

- Per semplificare l'analisi dei documenti il sistema seleziona dinamicamente la parte del documento più rilevante in base alla query di partenza
- La funzionalità è disponibile per tutti i processi (ricerca, documenti correlati, profilo, agenti di ricerca, canali, etc)

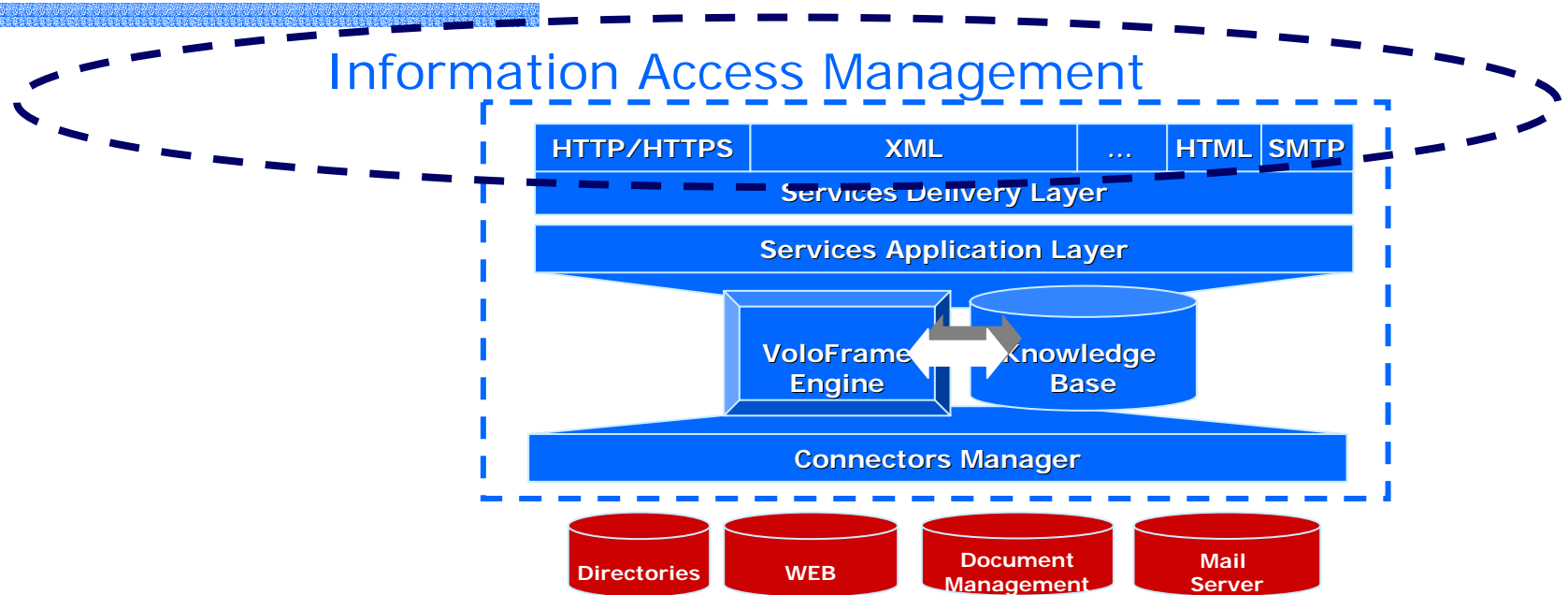
Espansione della ricerca : documenti correlati

- Creazione dinamica di collegamenti concettuali tra documenti
- Sulla base dell'impronta "concettuale" del documento aperto, il sistema è in grado di selezionare i documenti "correlati", ovvero quelli più simili concettualmente tra quelli disponibili all'interno della base di conoscenza, eventualmente distinti per tipologia.
- La qualità dei risultati non è più influenzata dalla qualità della ricerca di partenza, perché per selezionare i documenti correlati vengono utilizzati i concetti contenuti nel documento
- Il motore consente quindi di avere un risultato mirato e rilevante e contestualmente espandere la ricerca in direzioni diverse.
- Il link ai documenti è creato automaticamente e dinamicamente dal sistema e non richiede una organizzazione preventiva del corpus di documenti.
- Questa funzionalità è particolarmente rilevante quando l'utente sta effettuando una ricerca di contenuti senza avere particolari indicazioni circa il documento che li contiene.

Argomenti correlati

- **Estrazione di termini da un set di documenti in base alla loro *semantica latente* (Latent Semantic Indexing and Analysis)**
 - Poiché esistono diversi modi per esprimere uno stesso concetto (sinonimia), e che uno stesso termine può avere differenti significati (polisemia), la semplice ricerca “text matching” può fornire dei risultati che hanno bassa rilevanza.
 - Il Latent Semantic Indexing evidenzia una struttura semantica latente che nei documenti è nascosta dalla variabilità con la quale possono essere scelte le parole.

Information Access Management



- Information Access Management è l'insieme dei servizi aggiuntivi che consentono all'utente di ottimizzare l'efficienza e l'efficacia dei processi di ricerca

Accesso personalizzato

- Personalizzazione secondo i fabbisogni informativi dell'utente
- Relevance Feedback
- Creazione di home page personalizzate sulle esigenze dell'utente
- Possibilità di autenticazione dell'utente al momento dell'accesso se necessario per garantire la sicurezza
- Possibilità di associare a un utente uno o più ruoli che definiscono:
 - Le funzionalità disponibili all'utente
 - Il livello di accesso alla singola fonte

Profilazione automatica

- **Costruzione automatica di un profilo concettuale dell'utente, basato sui concetti contenuti nei documenti visualizzati e nelle ricerche effettuate**
 - L'utente interagisce con la knowledge base effettuando delle ricerche e selezionando dei documenti.
 - I concetti estratti dalle ricerche e dai documenti visualizzati sono utilizzati dal sistema per costruire un profilo "implicito" dell'utente
- **Il profilo non richiede nessuna attività esplicita all'utente**
- **Il prodotto del profilo è una selezione automatica di documenti affini al profilo concettuale dell'utente (Es.: il mio giornale personalizzato)**

- Capacità di operare in contesti poco conosciuti: l'utente è in grado di operare ricerche usando schemi concettuali e di navigazione che non richiedano necessariamente la precisa conoscenza dell'argomento;
- L'utente accede ai documenti secondo schemi liberi e non rigidamente predeterminati;
- Il sito si trasforma in una "knowledge base";
- utilizzo di processi inferenziali e metodi statistici (teoria di Shannon e statistica di Bayes) forniscono una maggiore attendibilità dei risultati delle ricerche rispetto ai tradizionali motori basati sull'utilizzo di parole chiave
- Multilinguismo: la piattaforma può operare indifferentemente su diverse lingue e gerghi senza esigenze di addestramento preliminare
- Personalizzazione sulle esigenze dei clienti: gli schemi concettuali si adattano automaticamente in funzione del fabbisogno informativo e del corpus dei documenti.