

Text Mining

- **Analisi di dati testuali**
- **Consente di estrarre conoscenza da documenti non strutturati**
- **L'uso di metodi statistico-probabilistici consente di:**
 - risolvere le principali ambiguità legate alla lingua
 - indicizzare automaticamente i documenti associando loro i concetti chiave
 - riconoscere concetti anche se espressi in modo non convenzionale

Una applicazione pratica: l'informazione giuridica

- Il reperimento delle informazioni di interesse giuridico presenta molte difficoltà legate alla complessità della materia e alla abbondanza della produzione di documentazione specifica.
- A fronte di tale enorme base di informazioni, create in tempi diversi e da organi diversi occorre disporre degli strumenti informativi atti a formare un quadro normativo interpretativo coerente e pertinente alla propria esigenza d'informazione.
- Internet ha prodotto significativi cambiamenti anche nel mondo dell'informatica giuridica che ha dovuto riconsiderare sia i modelli classici di documentazione sia i sistemi di reperimento delle informazioni.
- La proliferazione dei siti Web di organismi internazionali e nazionali produttori di documentazione giuridica ha affrontato il problema della disponibilità ma non ha risolto il problema dell'accessibilità dell'informazione.

Esigenze

- Strumenti in grado di trasformare questa massa enorme di risorse in una base di conoscenza che permetta una informazione aggiornata, integrata, coerente e completa.
- Nel progettare un servizio per l'accesso in rete alle informazioni giuridiche occorre tenere presente:
 - a. le caratteristiche principali della documentazione giuridica disponibile in rete;
 - b. le caratteristiche del linguaggio del legislatore che:
 - non è necessariamente completo,
 - non è necessariamente ordinato;
 - non è sempre rigoroso.
 - c. la varietà delle esigenze informative e i differenti profili degli utenti finali.

Fonti di informazione

- In Internet è possibile reperire:
 - Siti Internet d'interesse giuridico
 - Periodici giuridici e riviste *on-line*
 - Raccolte e banche di dati giuridici
 - Cataloghi di biblioteche e di editori giuridici
 - Siti di studi legali (soprattutto internazionali)
 - Servizi vari di informazione
 - Forum di interesse giuridico
- Quasi tutta la legislazione (leggi e decreti), la giurisprudenza e la dottrina italiane e internazionali sono individuabili on line, sotto forma di riferimenti bibliografici e in gran parte anche consultabili direttamente

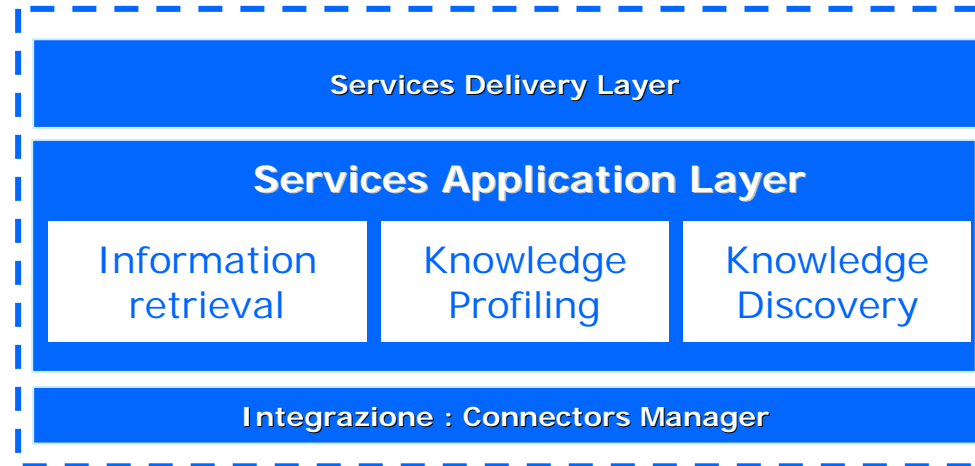
Soluzioni Volocom

- Analizzare il contenuto testuale delle leggi
- Individuare le tematiche più ricorrenti e significative (clustering)
- Raggruppare in categorie omogenee le leggi attraverso tecniche di classificazione automatica

Piattaforma di servizi

I servizi messi a disposizione dalla piattaforma VoloFrame consentono di:

- acquisire o documenti da una gamma eterogenea di flussi e formati;
- estrarre rapidamente informazioni rilevanti
- creare correlazioni tra documenti al fine di avere la visione più completa del contesto normativo che si sta analizzando (related document e related topic);
- raggruppare dinamicamente i documenti correlati (clustering) senza la necessità di definire delle categorie di consultazione predefinite
- profilare gli interessi degli utenti e proporre automaticamente (in modalità *push*) documenti che abbiano un'attinenza con i relativi profili di interesse



Le funzionalità

- Sviluppare un sistema per la gestione efficace ed informatizzata della documentazione nel rispetto delle normative vigenti in materia.
- Razionalizzare e rendere più rapidi i processi e le metodologie di ricerca e raccolta delle fonti esterne.
- Strutturare il flusso di informazioni sia all'interno sia all'esterno.

