

**11 e 12 settembre 2008** c/o Biblioteca delle Oblate - via dell'Oriuolo, 26 - FIRENZE



## L'informazione utile: come e dove trovarla

*Dott. Ing. Nicola Baldini*  
*focuseek*



# Dove?

- Nella propria rete “aziendale”
- Nelle fonti pubbliche aperte

# Fonti pubbliche aperte

- Siti Web
- Forum
- Blog

**Motori di ricerca (o quasi)**

- FTP repos
- Canali chat
- Database

???

# Intranet

- Documentazione su file system (-)
- Portali interni aziendali (~)
- Email (~)
- Database (+)
- Ecc.

# Informazione (non strutturata)

Utile



“Potenzialmente Utile”

# Come?

- Un agente software simula un utente che in maniera sistematica e periodica scandisce le fonti di potenziale interesse
- L'agente crea un indice analitico di tutto quello che trova e si salva una copia locale delle informazioni non persistenti

## Crawler (o spider)

# I motori di ricerca (pubblici)

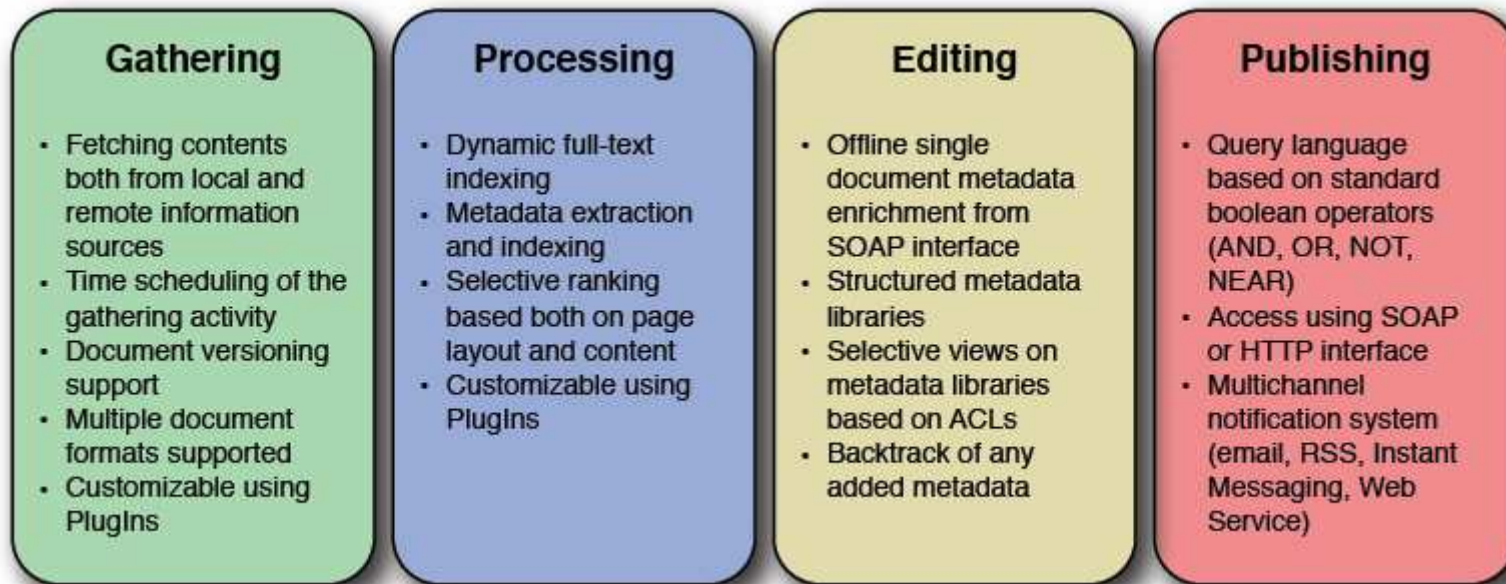
I vantaggi li conosciamo, ma...

- Non possono indicizzare la tua intranet
- Indicizzano solo il Web pubblico che interessa loro
- Nessuna garanzia sulla frequenza di refresh degli indici

# Quindi...

- Dotarsi di un proprio sistema di Crawling per le fonti interne e quelle esterne di specifico interesse che lavori in affiancamento ai motori di ricerca pubblici
- Creare un singolo punto di accesso alle informazioni sia interne che esterne
- Mantenersi aggiornati sulle nuove informazioni che arrivano nel sistema (monitoring/awareness)
- Dotarsi di soluzioni specifiche per l'analisi e quindi la valorizzazione del patrimonio informativo raccolto

# focusseek searchbox



# Riferimenti

Nicola Baldini

[nicola@focuseek.com](mailto:nicola@focuseek.com)

<http://www.focuseek.com>