



Motori di Ricerca

Se non sai, sappi dove (e come)
cercare!



MOTORI DI RICERCA





Perché i motori... (I)

- Enorme quantità di Informazioni su Internet
- In crescita continua
- I Motori sono quasi **indispensabili** per trovare l'informazione che ti serve.
- Su Internet c'e' anche una impressionante **quantità** di ...

Spazzatura (rubbish)



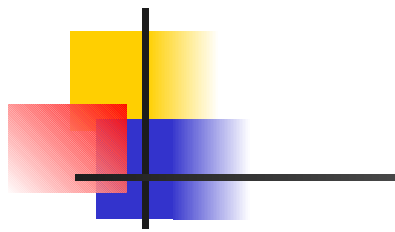
bruno:



Perché i motori... (II)

- Nonostante la sua potenza ...
Internet
- ...ed anche i Motori di Ricerca
- Sono **estremamente** primitivi

bruno:
http://it.wikipedia.org/wiki/World_Wide_Web#Dal_web_statico_al_web_semantico





Motori ! v1.8 05/2006

- Il web oggi: dimensioni, lingue etc.
- Glossario di ricerca
- Motori, metamotori, integratori
- Come funzionano,
- Come si utilizzano
- Directory
- Newsgroups





Dimensioni del Web

2005: ben oltre 100 miliardi di pagine presenti

- Decine e Decine di milioni di Domini
- Centinaia di milioni di Siti



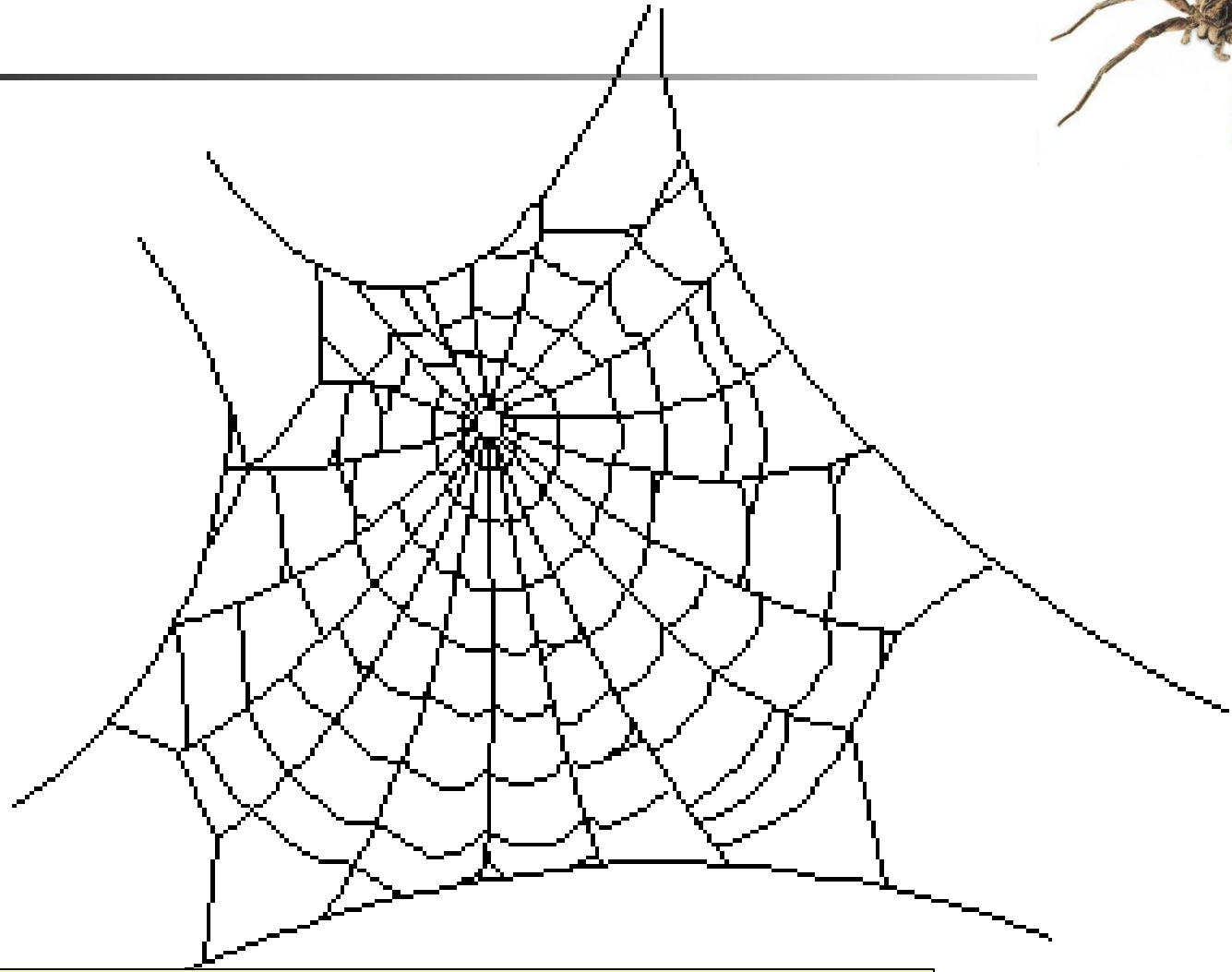
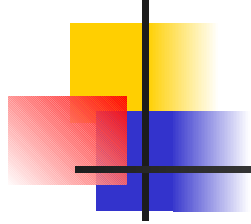
Esempio di Dominio e Sito

- Nel Dominio Tripod.com ...
- i Siti (nome.tripod.com) sono moltissime migliaia

cip:
Web: in Italiano "ragnatel" ;-)

bruno:
e lo spider

IL WEB!



bruno: http://en.wikipedia.org/wiki/World_Wide_Web



Il Web: alcune informazioni

- Siti pubblici
- Siti privati (accesso ristretto)
- provvisori o in stato transitorio (non terminati)
- Siti “fantasma” non più aggiornati da anni
- 10% di link “morti” (“dead links” o “broken links”)



Web: Info

- Lingue: Inglese 73% (50% madrelingua, tedesco 7%, giapponese 5%, francese e spagnolo 3% (e l'Italiano?))
- traduttori automatici (ohimè!)
- Il 47% dei siti **pubblici** sono negli USA

Quanto è grande Internet?

- “Surface” web (**SW**, Molti Miliardi di documenti)
- “Deep” web (DW) circa **500** volte più grande
- **SW** cresce di circa 7 milioni di pagine al giorno
- Media di **20KB** per pagina

Quanto è grande Internet?

- DW= SW più DB (Data Bases), pagine dinamiche, Intranets ...
- I due DB più grossi accessibili sono il "National climatic data center" ncdc.noaa.gov e la [NASA](http://NASA.gov)





- 150 anni di dati
- 224 gigabytes di nuovi dati al giorno
- 1.2 petabytes di info digitali



Tipi di documenti e oggetti sul web

- Pagine HTML
- .PDF (adobe Acrobat)
- .DOC (Microsoft Word)
- .PS Adobe Postscript)
- .PPT (Microsoft Powerpoint)
- .XLS (Microsft Excel)
- .RTF (Rich Text Format)
- ...



Tipi di documenti e oggetti sul web

- Immagini
- filmati
- musica
- Podcasts
- ...
- **codificati** in tanti formati diversi



Streaming su WEB

- Accesso diretto (senza attesa)
- a contenuti multimediali
- WEB radio
- WEB TV
- Audio/Video **on demand** (VOD)
- richiede (solitamente) banda larga
- ricerca contenuti multimediali