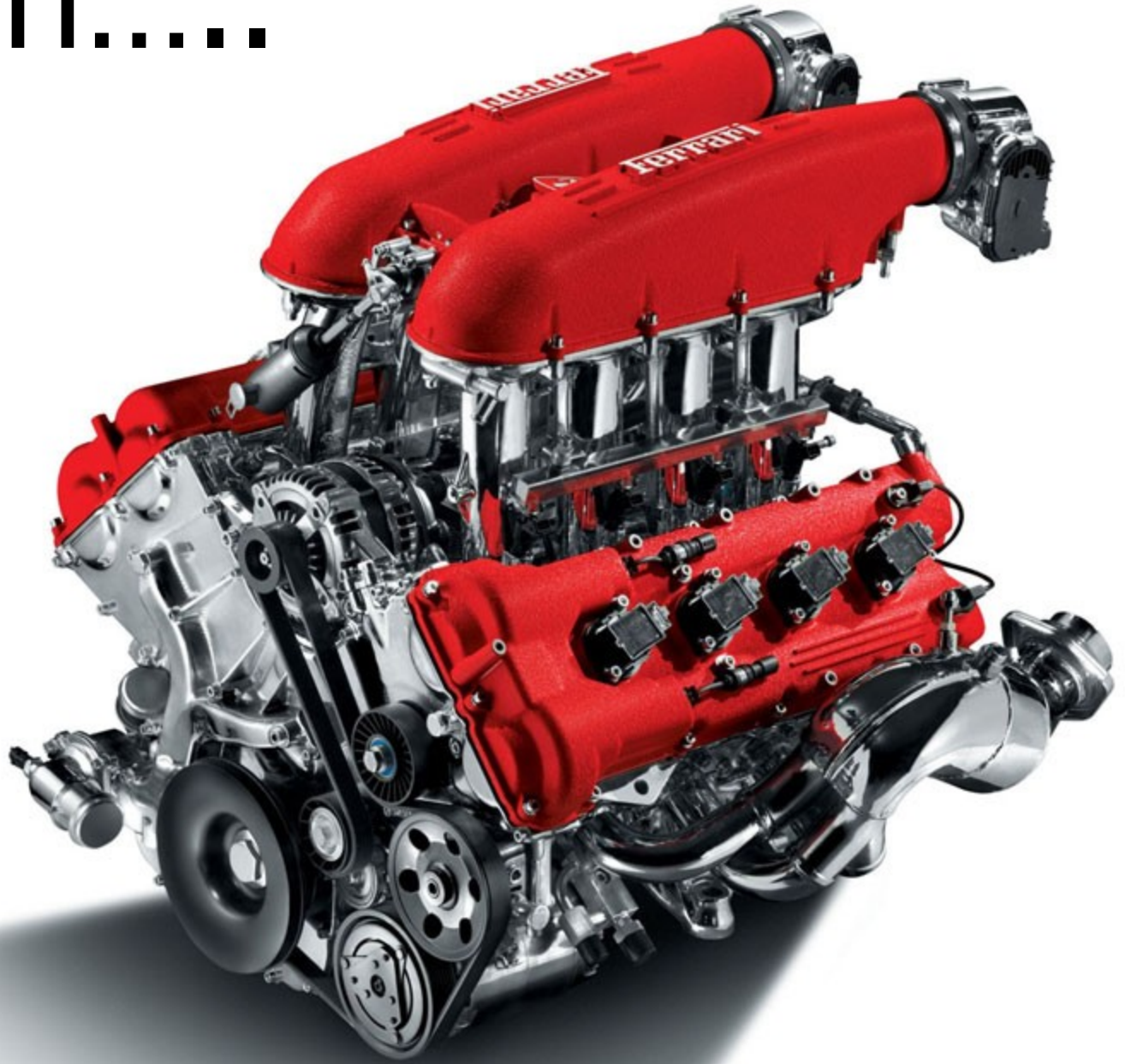


Motori.....



The image features the text "di Ricerca!" in a large, bold, blue font with a white outline and a slight drop shadow. The text is positioned on the right side of the frame. To the left of the text, there is a decorative graphic consisting of a horizontal bar with a gradient from blue to yellow, and a vertical black line intersecting it. The background is white.

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)



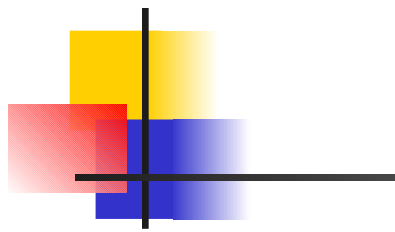
Perché i motori... (II)

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



bruno:

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





Motori ! v2.0 12/2007

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



cip:

Enorme!

Dimensioni del Web

Centinaia di miliardi di pagine/documenti

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



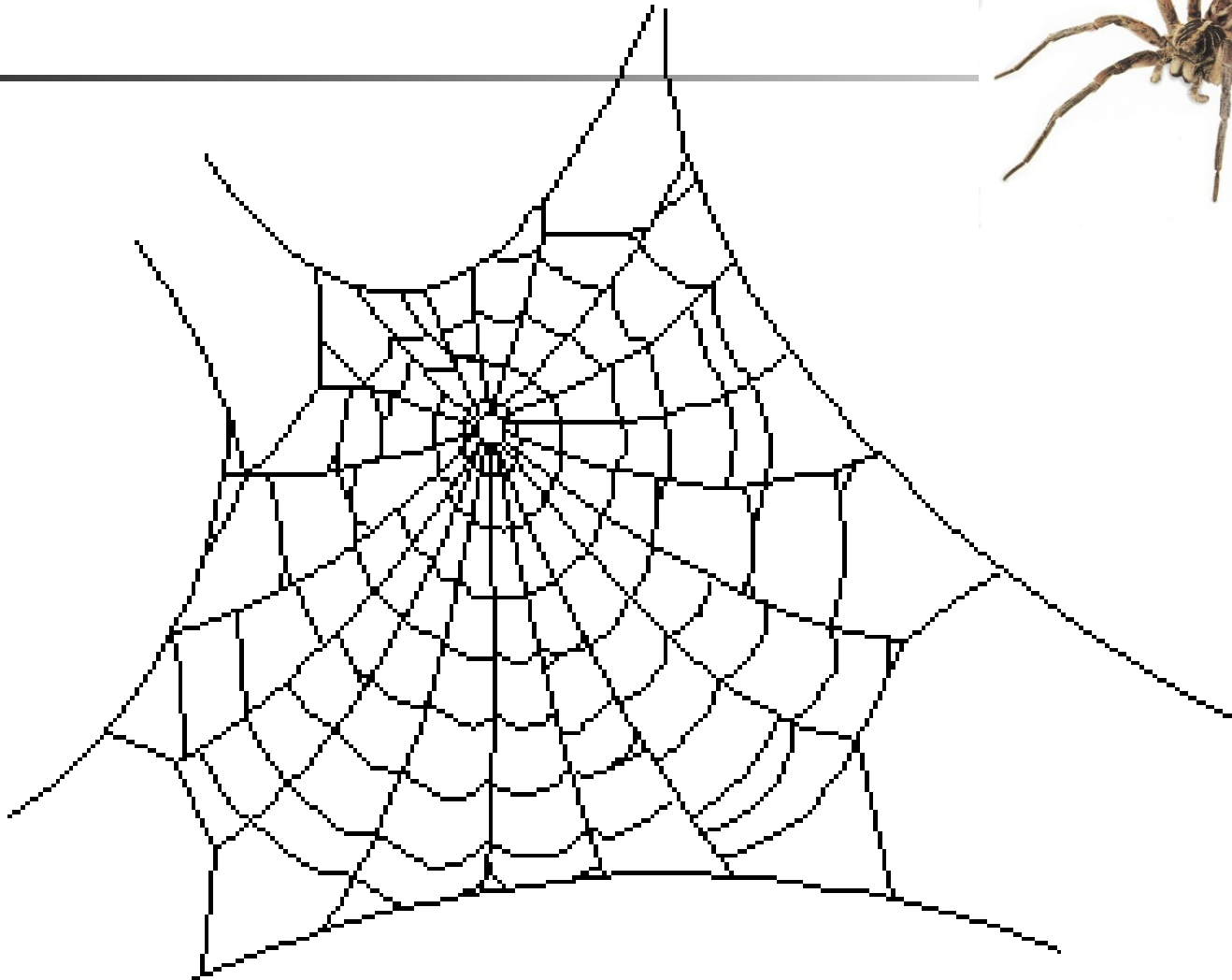
Esempio di Dominio e Sito

- Nel Dominio Tripod.com ...
- i Siti (sottodomini del tipo [nome.tripod.com](#)) sono moltissime migliaia
- Nel dominio blogspot.com...
- I blogs [nome.blogspot.com](#) sono **milioni**

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 (da aggiornare)

- 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 molti 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





- 150 anni di dati
- 224 gigabytes di nuovi dati al giorno
- 1.2 petabytes di informazioni digitali (+
++)



Tipi di documenti e oggetti sul web

- Pagine HTML
- .PDF (adobe Acrobat)
- .DOC (Microsoft Word)
- .PS Adobe Postscript)
- .PPT (Microsoft Powerpoint)
- .XLS (Microsoft 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