

INTRODUZIONE AL WEB: IL WORLD WIDE WEB

**CORSO DI LAUREA IN ECONOMIA AZIENDALE
ING. PAOLA LAPADULA - UNIVERSITÀ DEGLI STUDI
DELLA BASILICATA**

A.A. 2023/2024

SOMMARIO

- World Wide Web
 - Web
 - Server Web, HTTP e URL
 - Browser
- Motori di ricerca

INTRODUZIONE

- La crescita di internet ha aumentato notevolmente, l'uso di numerosi servizi che sono oggi molto diffusi:
 - Posta elettronica
 - Commercio on-line
 - Scambio messaggi (Forum/ chat)
 - Telelavoro, Teleconferenze
 - E-learning
- Un servizio fondamentale... WWW

INTRODUZIONE

- Tim Berners Lee propose il Web al CERN (Consiglio Europeo per la Ricerca Nucleare) nel 1989
 - per consentire agli studiosi di avere accesso a vari database di lavori scientifici attraverso i propri computer
 - Studiosi da tutto il mondo potevano ricercare e recuperare documenti da un gran numero di server, ovunque fossero dislocati

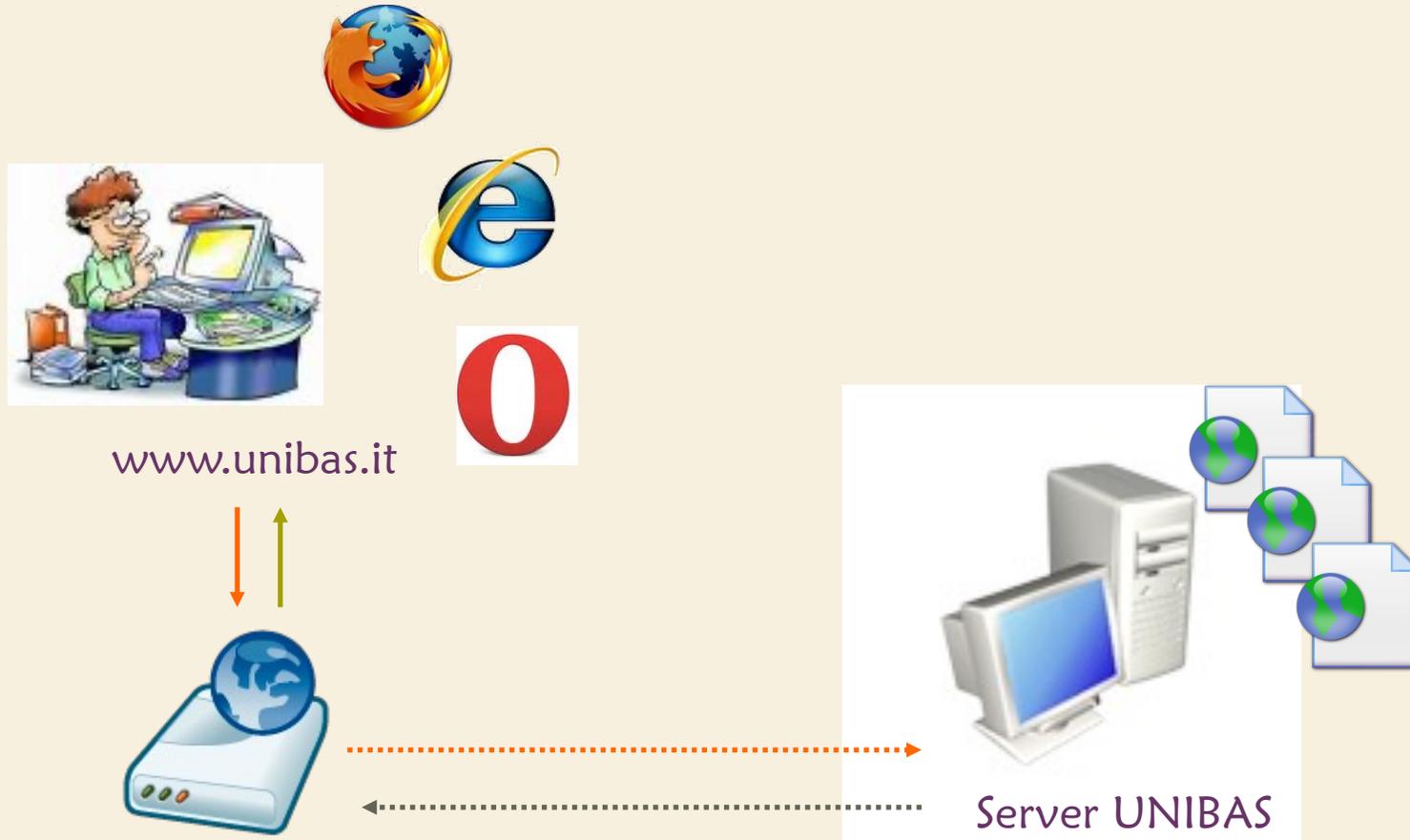
INTRODUZIONE

- Nel 1990, Tim Berners Lee conia il nome di *World Wide Web* (W.W.W) per la pubblicazione in rete di ipertesti
- Un **ipertesto** è formato da un insieme di documenti, collegati tra loro tramite riferimenti ipertestuali, denominati **link**
 - A differenza di un documento standard, letto in modo sequenziale, l'ipertesto viene *navigato*

INTRODUZIONE

- Il World Wide Web è un insieme vastissimo di contenuti multimediali e di servizi che possono essere resi disponibili dagli stessi utenti di Internet
 - I contenuti devono essere presenti su un server e vengono fruiti tramite siti web
 - **Sito web**: insieme delle risorse presenti su uno o più server che contengono informazioni (risorse)
 - È possibile accedere a tutte le risorse in rete tramite un unico protocollo (**HTTP**) e un unico client (**browser**)

IL WEB



IL WEB

- L'utente richiede una risorsa (solitamente un documento) identificata da un **URL**
- Lo **user agent** (browser) interpreta l'URL e inoltra una richiesta al server opportuno (ftp, http, ...)
- Il **server** fornisce la risorsa richiesta o un messaggio di errore
- (cont.)

IL WEB



HTML
HyperText
Markup Language

- Lo user agent interpreta i contenuti del messaggio di risposta
- Se il messaggio è un documento **HTML** il browser si occupa automaticamente di reperire le sottoparti referenziate nel documento (immagini, applet java...)
- La risorsa viene presentata all'utente, eventualmente con l'ausilio di programmi esterni (mpeg player, audio player)

IL SERVER WEB

- Risponde alle richieste del browser
 - Trasmettendo i dati richiesti (file html, immagini, applet, script flash, ...)
- Spesso include:
 - un interprete per i principali linguaggi di scripting lato server ...
- Esempi: MS Internet Information Server, Netscape Server, Apache

IL PROTOCOLLO HTTP

- Il protocollo per la comunicazione tra browser e server web
- Sfrutta una connessione TCP sulla porta 80
- Usa gli **URL** (>>) come meccanismo di indirizzamento
- Caratteristiche:
 - ogni connessione permette una singola transazione
 - il trasferimento dati è binario

UNIFORM RESOURCE LOCATOR

- Gli URL definiscono una sintassi e una semantica per l'identificazione e l'accesso a risorse su Internet
- Un URL definisce:
 - quale protocollo (schema) usare per accedere alla risorsa (cioè quale server contattare)
 - il nome del server e la porta presso cui è disponibile il servizio
 - il path della risorsa
 - il nome della risorsa
 - eventuali parametri da passare alla risorsa

UNIFORM RESOURCE LOCATOR



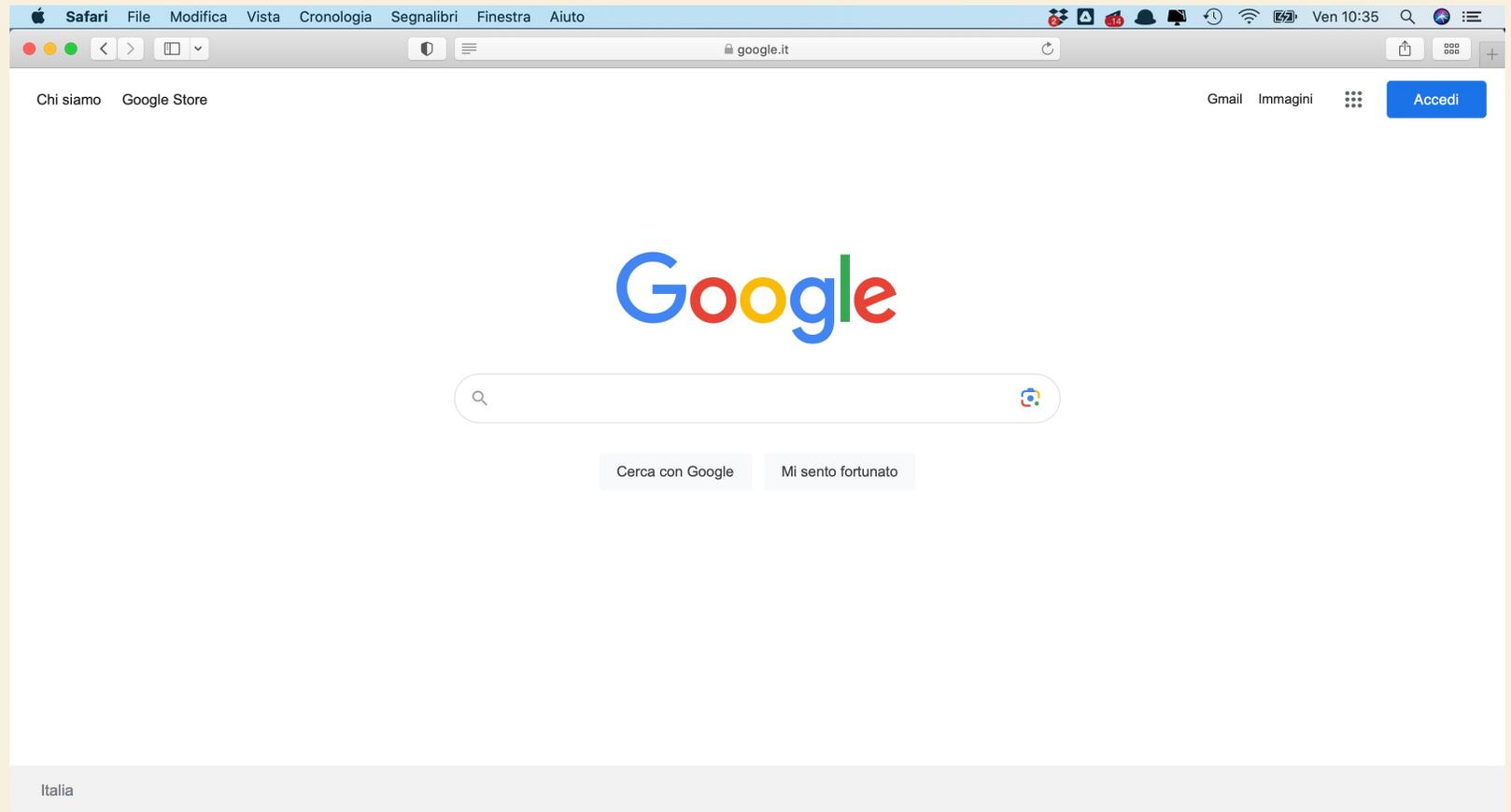
Indirizzo simbolico
DNS
Domain Name
System

- Per esempio:
 - `http://www.google.it/index.html`
 - `http` (è il protocollo)
 - `www.google.it` (indirizzo simbolico del server google)
 - `index.html` (risorsa web ovvero il file .html)
 - Il carattere `'/'` (delimitatore)
- Esempi con altri protocolli:
 - `ftp` (es: <ftp://...../...../risorsa.txt>)
 - `telnet://ipese3.elet.polimi.it`
 - `mailto:lapadula.paola@gmail.com`

IL BROWSER

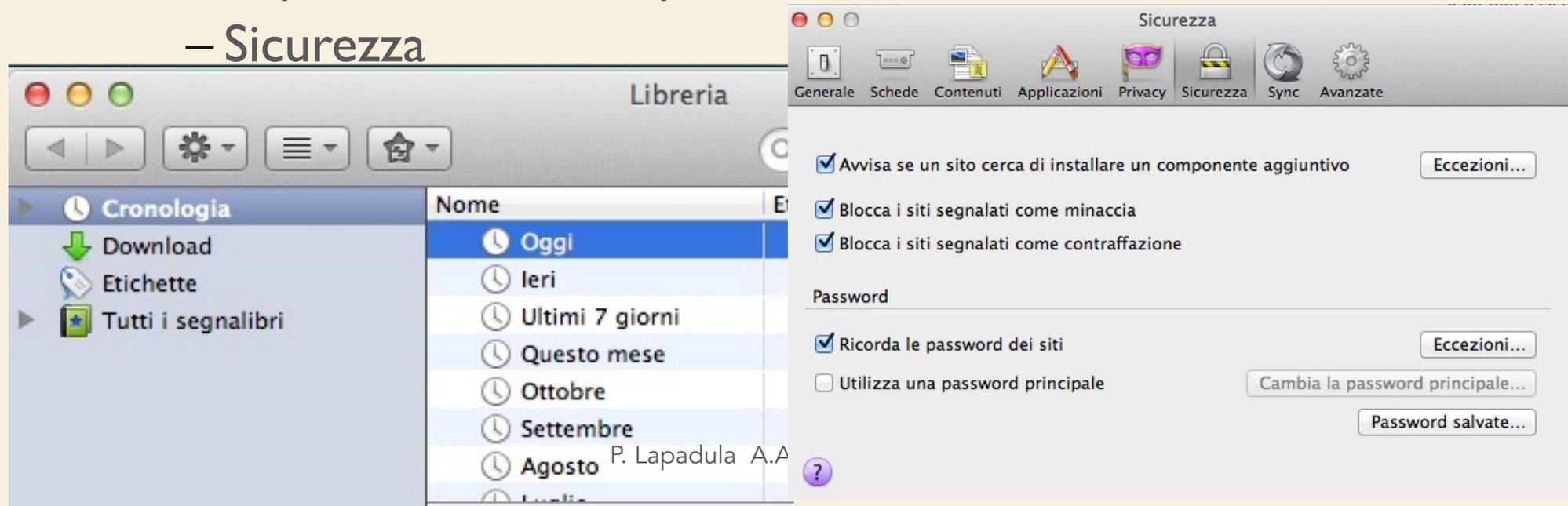
- Gioca il ruolo del client
 - Gestisce l'interazione con l'utente
- Interpreta il linguaggio **HTML** e presenta i dati a video
- Generalmente include:
 - un interprete per i principali linguaggi di scripting lato client...
 - ... e la possibilità di invocare plug-in per visualizzare i formati non direttamente supportati
- Esempi: MS Internet Explorer, Mozilla, Opera, etc.

IL BROWSER



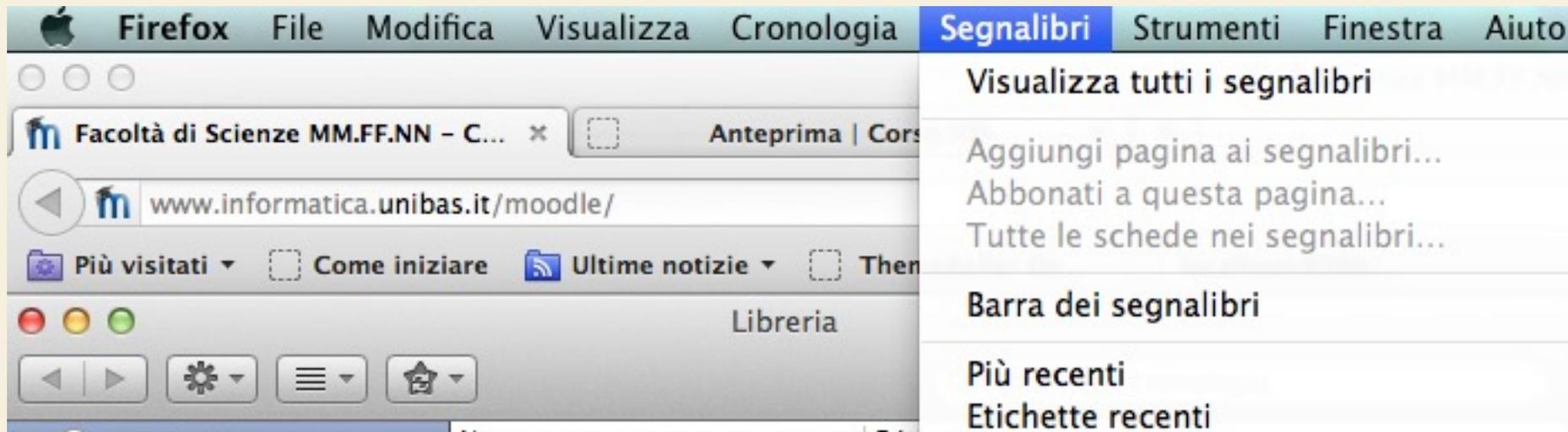
IL BROWSER

- Configurazione:
 - Home Page
 - Cronologia
 - Impostare browser predefinito
 - Sicurezza



IL BROWSER

- Bookmarks o Preferiti
 - Aggiunta/Rimozione
- Visualizzazione sorgente di una pagina html
- Analizza pagina html (Componenti Aggiuntivi Firebug)

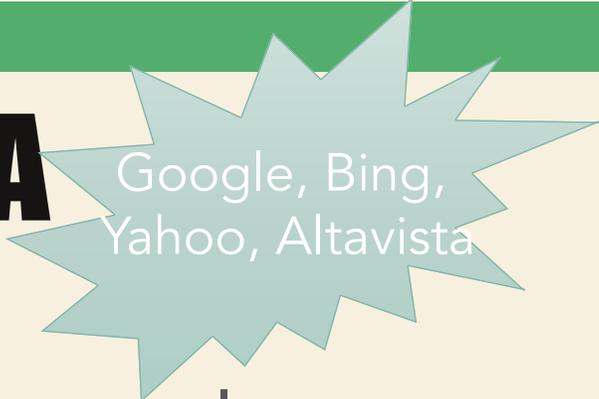


PORTALE WEB



- Contenitore di risorse web ben organizzato, dal quale è possibile reperire molte informazioni e offre numerosi servizi
- Sono detti anche portali internet o semplicemente portali
- Esempi di portali:
 - www.basilicatanet.it
 - <http://www.informatica.unibas.it/moodle/>

I MOTORI DI RICERCA

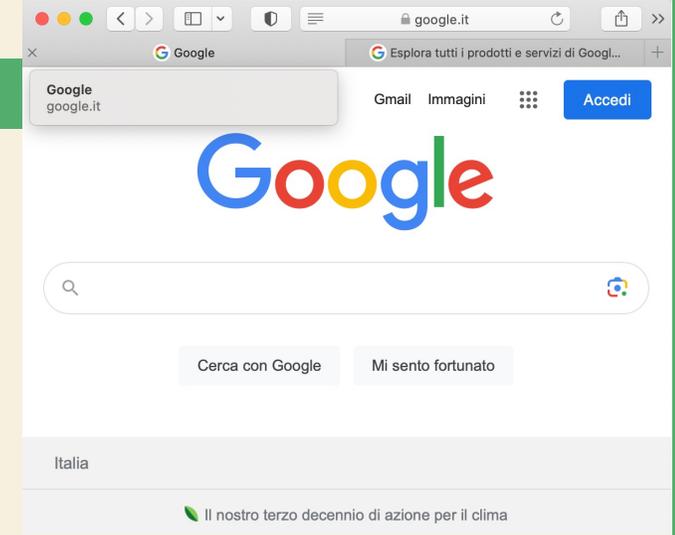


Google, Bing,
Yahoo, Altavista

- Sono software in grado di aiutare a trovare le informazioni che si stanno cercando nel web
- Permettono di effettuare un monitoraggio automatico delle risorse presenti in rete e selezionare quelle potenzialmente interessanti in base ai criteri di ricerca impostati dall'utente
- Strumento dinamico per la costruzione di una mappa di contenuti

COME FUNZIONANO

- Motori per parole chiave
- Motori che utilizzano indici di contenuti (detti più propriamente *cataloghi*)
- I motori di ricerca sono accessibili attraverso appositi siti web
 - (es.: <http://www.google.it>)
 - Ogni motore di ricerca possiede una sorta di archivio, costantemente aggiornato, che cataloga circa il 10-15% di tutti i siti disponibili in Internet



MOTORI DI RICERCA PER PAROLE CHIAVE

- Permettono di effettuare una selezione all'interno delle risorse web visualizzando gli indirizzi dei siti che contengono, nel titolo o all'interno delle pagine, le **parole chiave** digitate nell'apposito modulo di ricerca
- Questi motori funzionano attraverso agenti software chiamati **spider** che si occupano di monitorare continuamente la rete classificando i siti in base al tipo di parole contenute al loro interno

MOTORI DI RICERCA PER PAROLE CHIAVE

- Il motore fornisce una lista delle risorse trovate a partire dai siti che contengono una **maggiore occorrenza della parola chiave**,
- ma cerca anche di effettuare **valutazioni semantiche** sui contenuti del sito analizzando il ruolo che tali parole hanno all'interno del testo

MOTORI DI RICERCA PER PAROLE CHIAVE

- Se digitiamo *palio di siena* verranno cercati tutti i siti che contengono la parola *palio* quelli che contengono la parola *siena*, quelli che contengono sia *palio* che *siena*;
- Ci ritroviamo un'infinità di siti



The screenshot shows a Google search interface. At the top left is the Google logo. To its right are navigation links: Web, Immagini, Gruppi, Directory, News, and altro ». Below these is a search bar containing the text 'palio di siena' and a 'Cerca' button. To the right of the search bar are links for 'Ricerca avanzata' and 'Preferenze'. Below the search bar are radio buttons for search filters: 'il Web' (selected), 'pagine in Italiano', and 'pagine provenienti da: Italia'. The search results are under the heading 'Web'. The first result is 'Risultati da Google News su palio di siena - Mostra le ultime notizie'. Below this are two news snippets: 'Perché Berlinguer è di nuovo attuale - L'Espresso - 31 dic 2005' and 'QUATTRO PASSI A SIENA - Il Sole 24 Ore - 29 dic 2005'. At the bottom, there is a link for 'Il Palio di Siena' with a brief description: 'Il Palio di Siena visto dai contradaiooli, notizie, statistiche, gioco e momenti di confronto nel sito NON UFFICIALE La Veste della Piazza'.

MOTORI DI RICERCA PER PAROLE CHIAVE

- **Criteri di ricerca:**
 - Operatori booleani: OR, AND, NOT
- **OR:** informatica OR telecomunicazioni
 - Tutti i documenti indicizzati dal motore di ricerca che contengono o la parola informatica o la parola telecomunicazioni
- **AND:** informatica AND telecomunicazioni
 - Tutti i documenti indicizzati dal motore di ricerca che contengono sia la parola informatica sia la parola telecomunicazioni

MOTORI DI RICERCA PER PAROLE CHIAVE

- **NOT**: NOT informatica
 - Tutti i documenti indicizzati dal motore di ricerca che non contengono la parola informatica
- **NOTE**:
 - l'operatore AND può essere omissso
 - Gli operatori vanno scritti con le lettere maiuscole

MOTORI DI RICERCA PER PAROLE CHIAVE

- Operatore site
 - effettua la ricerca nei documenti di uno specifico sito
 - Per esempio >> potenza site:www.basilicatanet.it



The screenshot shows a Google search interface. The search bar contains the query "potenza site:www.basilicatanet.it". Below the search bar, the word "Search" is displayed in red, followed by the text "About 93,300 results (0.34 seconds)". The search results are organized into categories: "Web", "Images", "Maps", "Videos", "News", and "Shopping". The "Web" category is expanded, showing two results. The first result is titled "Potenza, oggi e domani due seminari su reati ambientali" and includes a link to "www.basilicatanet.it/basilicatanet/.../detail.jsp?... - Translate this page". The second result is titled "Provincia di Potenza. Lavori di manutenzione sulla Sp 83" and includes a link to "www.basilicatanet.it/basilicatanet/.../detail.jsp?... - Translate this page".

MOTORI DI RICERCA PER PAROLE CHIAVE

- **Con le virgolette**
 - Se si scrive un gruppo di parole fra virgolette, verrà cercata la frase, ossia le parole in quell'ordine preciso
 - si tratta di una funzione molto potente, resa possibile dal fatto che il motore indicizza ciascuna parola trovata nei siti e non solo: ne ricorda il contenuto ed anche l'ordine esatto delle parole

MOTORI DI RICERCA PER PAROLE CHIAVE



Senza virgolette?

- **Con le virgolette**
 - Infatti, nell'indicizzazione, il motore cerca semplicemente una corrispondenza uno a uno di simboli senza conoscerne il significato ed è perciò in grado di trovare non solo parole in italiano, ma in qualsiasi lingua
 - Se si cerca “Roberto Benigni”, il servizio trova tutti i documenti che contengono entrambe le parole, con la maiuscola, disposte l'una a fianco all'altra

MOTORI DI RICERCA PER PAROLE CHIAVE

- Segni + o -
 - Perché una parola sia sempre inclusa nel risultato della ricerca, la si deve fare precedere da un segno (+) senza spazi
 - perché invece quella parola ne sia esclusa, la si deve far precedere da un segno (-) senza spazi
 - Per esempio:
 - +“*Roberto Benigni*” –“*la vita è bella*”: trova tutti i documenti che contengono la frase “*Roberto Benigni*” ed esclude quelli che contengono “*la vita è bella*” quindi troverà tutti i siti che parlano dell’attore italiano ma non relativamente al film *La Vita è Bella*

MOTORI DI RICERCA PER PAROLE CHIAVE

- Operatore *
- Esempio:
 - se questo è un *
 - Operatore che restituisce tutti i documenti indicizzati dal browser che contengono le parole: se questo è un e poi qualsiasi altra parola
- L'operatore * indica qualsiasi parola
- ATTENZIONE tutto questo vale anche per la ricerca di immagini

MOTORI DI RICERCA PER PAROLE CHIAVE

- Operatore *
- Altri esempi:
 - Chi cerca la forma inglese e americana di una parola come color (colour) può inserire l'asterisco al posto dell'eventuale "u" : colo*r
 - Al posto dell'ultima lettera di una parola per cercare forme maschili e femminili, plurali o singolari: can* cattiv*
 - Quando ci si vuole assicurare dell'ortografia di una parola, ad esempio in caso di dubbio sulla presenza di una doppia: soprat*utto

GOOGLE – RICERCA AVANZATA

Ricerca avanzata

Trova pagine web che contengono...

tutte queste parole:

questa esatta parola o frase:

una qualunque di queste parole:

nessuna di queste parole:

numeri da:

a

Per fare questo nella casella di ricerca.

Digita le parole importanti: labrador retriever nero

Racchiudi le parole esatte tra virgolette: "labrador retriever"

Digita OR tra tutte le parole che vuoi: miniatura OR standard

Anteponi il segno - (meno) alle parole da escludere: -roditore, - "Jack Russell"

Inserisci due punti (..) tra i numeri e aggiungi un'unità di misura: 10..35 kg, € 300..€ 500, 2010..2011

Poi limita i risultati per...

lingua:

tutte le lingue

Trova le pagine nella lingua selezionata.

area geografica:

tutti i Paesi

Trova le pagine pubblicate in un'area geografica specifica.



Android



Android Auto

androidtv

Android TV



Assistente Google



Google Play Giochi



Google Shopping



Google Store



Google Strumenti di immissione



Google TV



Google Voli



Google Wallet



Google Wifi



Keep



Lens



Messaggi



Moduli



Nest



News



Pixel



Play Protect



Podcasts



Presentazioni



Ricerca



Scholar



Sites



Tilt Brush



Traduttore



Travel



Voice



Waze



Wear OS by Google



YouTube



YouTube Kids



YouTube Music



YouTube TV

SOMMARIO

- World Wide Web
 - Web
 - Server Web, HTTP e URL
 - Il Browser
- Motori di ricerca

TERMINI DELLA LICENZA

- This work is licensed under the Creative Commons Attribution-ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/1.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.
- Questo lavoro viene concesso in uso secondo i termini della licenza “Attribution-ShareAlike” di Creative Commons. Per ottenere una copia della licenza, è possibile visitare <http://creativecommons.org/licenses/by-sa/1.0/> oppure inviare una lettera all’indirizzo Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.