

1.3 EXCHANGING DATA · 1.3.4

Search indexing, PageRank & server/client-side

How search engines find and rank pages, and where processing happens. Spec 1.3.4(b)(c)(d).

01 Indexing (b)

Crawler Bot/spider follows **links**; records URL, content, relevance.

Index Database of words → pages, **built in advance**.

Why Query searches the index → **fast** results, not the live web.

Metatags Hidden keywords in HTML; help indexing.

02 PageRank (c)

Idea Importance = **quality + quantity** of incoming links.

Better A link from a **high-ranked** page counts more.

Diluted A page's vote is shared across its outbound links.

d **Damping factor** (~0.85): prob. user keeps clicking links.

03 Client-side vs server-side (d)

● Client-side (browser)

Validation, menus, animations. Fast; less server load.
Can be bypassed → **untrusted**.

● Server-side

Database, payments, logins, recalculating prices. **Secure**; needs a round-trip.

Rule Money, stored data or security → **server-side** (even if also checked client-side).

FINAL PASS BEFORE THE EXAM

Rapid exam tips

Six slips on search & processing questions.

01

Crawlers build the **index in advance**; your query searches that index, not the live web.

02

Metatags = **hidden** keywords in the HTML.

03

PageRank: a link from an **important** page counts more; votes are **diluted** by many outbound links.

04

Damping factor **d** = probability a user keeps clicking links.

05

Client-side advantage = **speed/no round-trip**, never security.

06

Payments, logins, final prices → **server-side** (can't be bypassed).