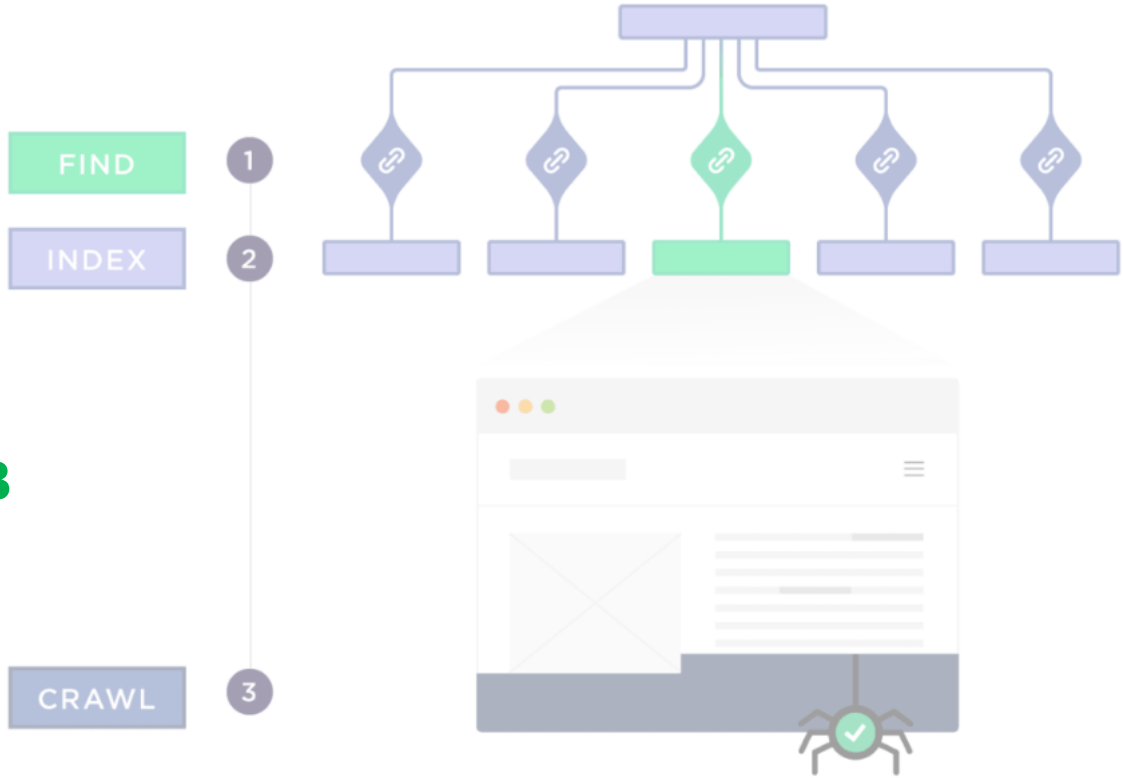
TECHNICAL SEO



SAHAR ANDALEEB

WHAT IS WEBSITE ARCHITECTURE?

“Website Architecture is how a website’s pages are structured and linked together.”

Why Is Website Architecture Important for SEO?

1. An optimized site architecture helps search engine spiders find and index all of the pages on your website.
2. The right website architecture makes it easy for visitors to find what they need on your site.

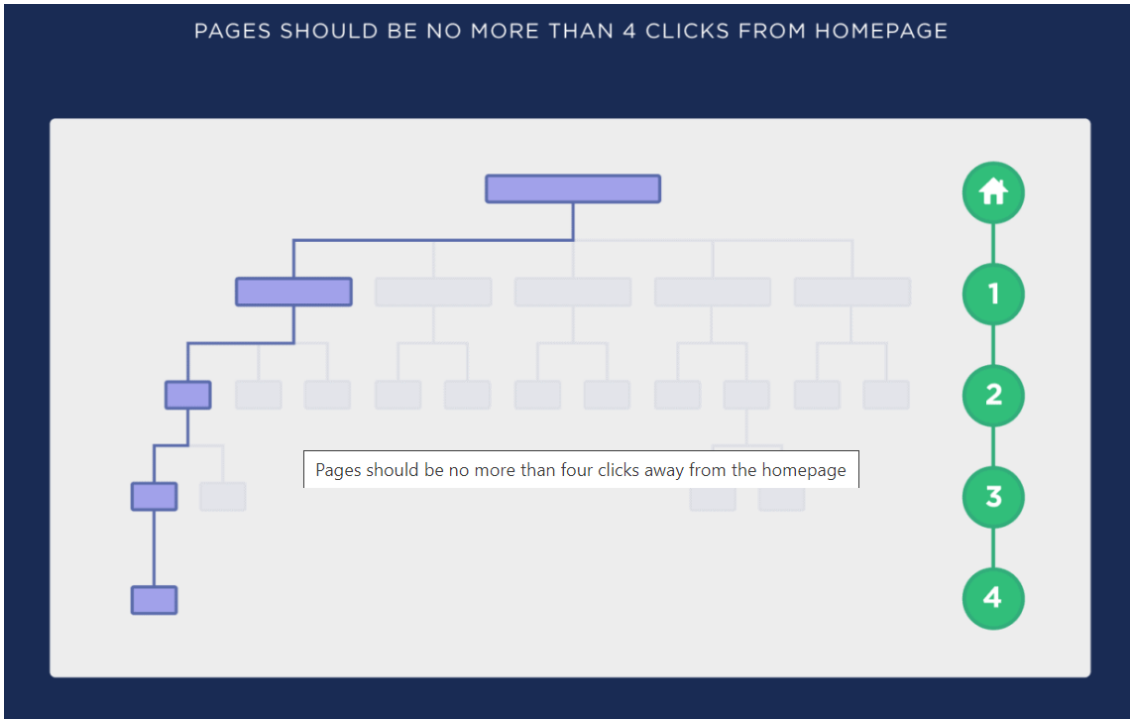
Homework (Important): Explore different website architecture and analyse which one or which combination suits you.

https://youtu.be/BXaXmDsZqZo?si=AHm5MxOFYCP29CKU

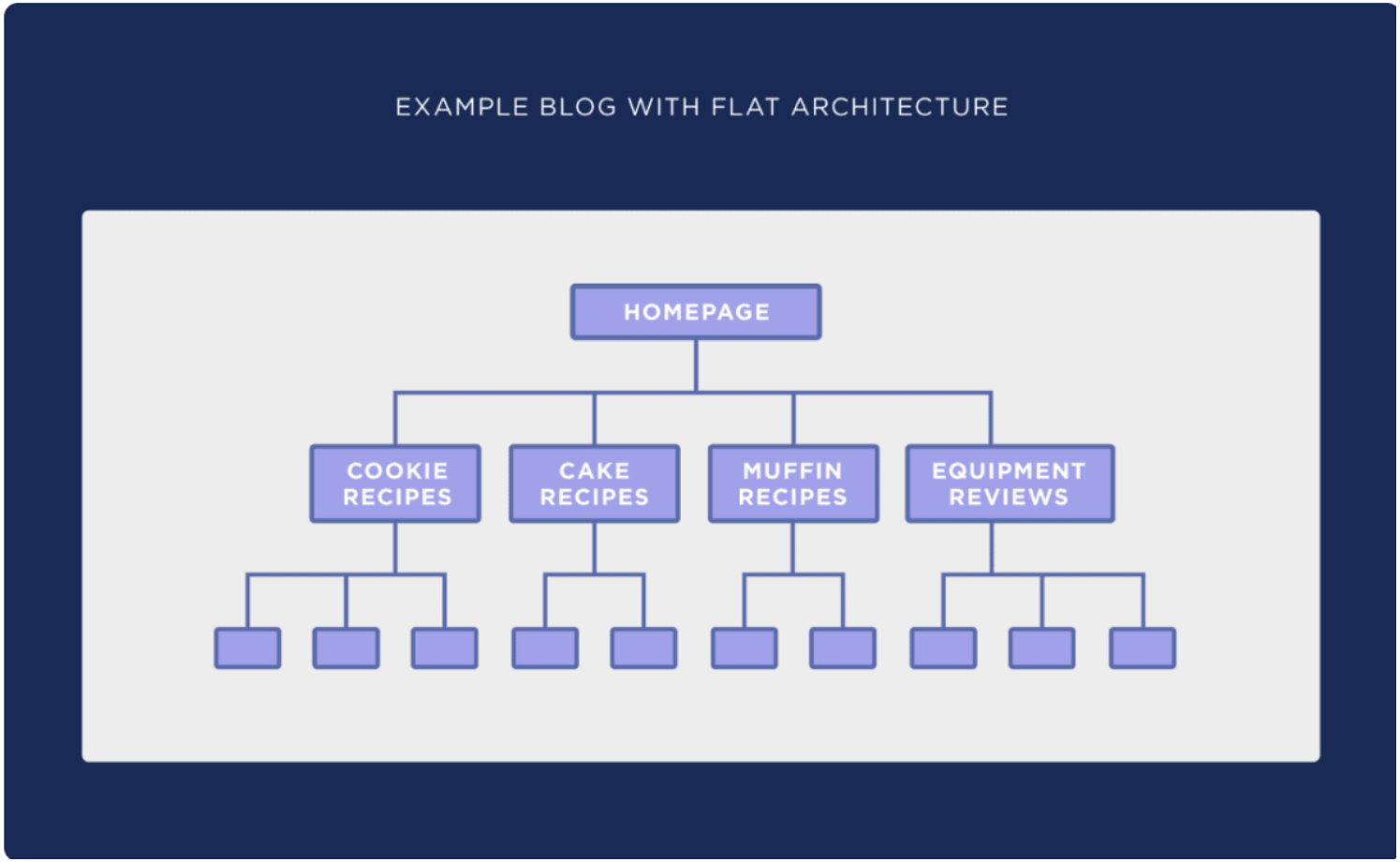
FLAT ARCHITECTURE LEVELS

- **Use a “Flat” Site Architecture**

- **Ideally 4 clicks or less.**



FLAT ARCHITECTURE



SEO-FRIENDLY WEBSITE ARCHITECTURE:

“A structure designed to enhance the visibility of a website on search engine results pages (SERPs) by optimizing its organization, navigation, and content.”

- It focuses on improving user experience, accessibility, and crawlability for search engine bots.

TYPES OF SEO-FRIENDLY WEBSITE ARCHITECTURE:

Flat Architecture: Shallow hierarchy with minimal levels of navigation. All pages are easily accessible from the homepage.

Hierarchical Architecture: Organized in a tree-like structure, with categories and subcategories, facilitating easy navigation.

Silo Architecture: Content grouped into thematic silos or categories, enhancing topical relevance and keyword focus.

IMPORTANT CONCEPTS

Silo Structure:

- Organizing content into distinct silos based on topics or themes.
- Helps search engines understand the topical relevance of content and boosts keyword targeting.

Topical Authority:

- Demonstrates expertise and authority on specific topics or niches.
- Achieved by consistently publishing high-quality, relevant content and earning backlinks from authoritative sources.

Content Hub:

- Centralized repository for comprehensive and authoritative content on a specific topic.
- Serves as a go-to resource for users and enhances topical relevance and authority.
- Often includes pillar pages, comprehensive guides, and related content to cover all aspects of the topic.

… CONT.

Link Equity: (a.k.a link juice)

- The value or authority passed from one webpage to another through *hyperlinks(backlinks).*
- Higher-quality and relevant links typically contribute more link equity to the receiving page, potentially improving its search engine ranking.

Page Authority:

A metric that predicts how well a specific webpage is likely to rank on search engine result pages (SERPs).

Influenced by factors like the quality and quantity of inbound links, content relevance, and overall website authority.

SITE MAP

“A **sitemap** is a file that *lists* all the pages of a website, providing search engines with structured information about the site's content.”

- Search engines read this file to crawl your site more efficiently
  - A sitemap can have a maximum of 50,000 links. Google will not crawl more than it.
  - If there are more than 50,000 URLs on a website, then multiple sitemaps will need to be created.

IMPORTANCE OF SITEMAPS IN SEO

- **Improved Crawling:** Sitemaps help search engine bots discover and index all pages of a website more efficiently.
- **Enhanced Visibility:** They ensure that search engines can access and understand the content hierarchy and importance.
- **Faster Indexing:** With a sitemap, new or updated content can be indexed quicker by search engines.
- **Two types:**
  - **Xml-site maps and html-site maps**

XML SITEMAP

“XML Sitemap is a structured file in XML format that lists the important URLs of a website along with additional metadata such as the last modified date, change frequency, and priority level of each URL.”

- XML sitemaps are **primarily designed for search engines** to *efficiently crawl and index* website content.
- **helps search engine bots** understand the website's structure, leading to better visibility and ranking in search engine results.
- **Tools:**
  - **https://www.xml-sitemaps.com/**
  - **Screaming frog**

XML SITEMAP SYNTAX

Xml Sitemap Syntax:

```
<url>
    <loc>https://www.example.com/page1.html</loc>
    <lastmod>2024-10-11</lastmod>
    <changefreq>monthly</changefreq> <priority>0.8</priority>
</url>
```

- <loc> specifies the URL of the page.
- <lastmod> indicates the last modification date of the page.
- <changefreq> suggests how frequently the page is likely to change.
- <priority> indicates the priority of this URL relative to other URLs on your site.

HTML SITEMAP

Html sitemap is a *webpage* on a website that serves as a navigational aid **for visitors** by providing a hierarchical list of links to all the pages within the site.

- **Unlike XML sitemaps, which are intended for search engines, HTML sitemaps are created for human users.**
- **provides a visual representation** of the website's *structure* and *interlinking* between pages

ROBOTS.TXT

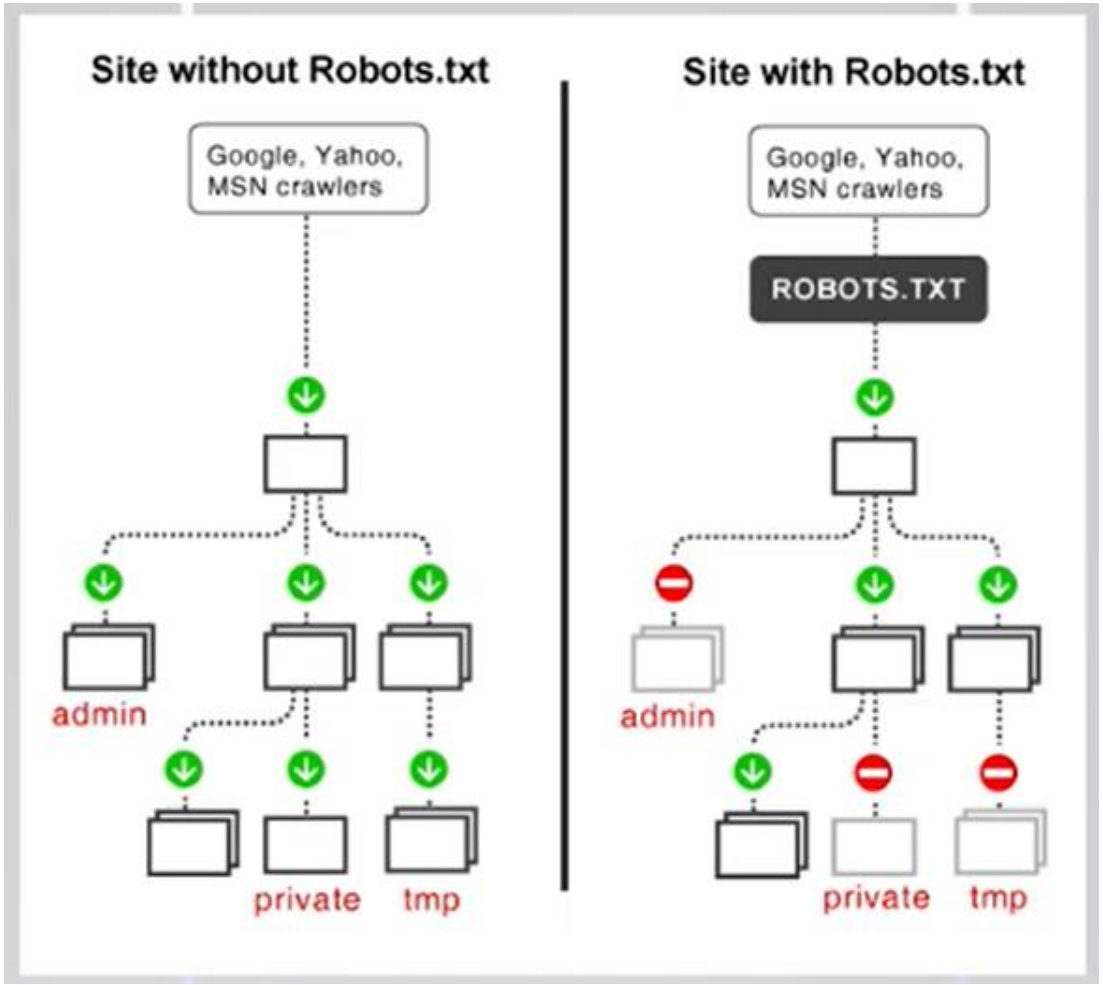
Robots.txt: "A text file located in the *root directory* of a website that instructs search engine crawlers on **how to** interact with the site's pages."

- Robots.txt can prevent certain pages from being indexed
  - useful for keeping sensitive information private.
  - User information, bank details etc.

IMPORTANCE:

- **Page Priority**: Prioritize crawling of important pages by allowing crawlers to focus on them rather than less important pages.
- **Resource Efficiency**: Prevents crawlers from wasting resources on non-essential pages, improving overall crawl efficiency.
- **Avoiding Penalties**: Properly configured robots.txt can prevent search engines from indexing content that could lead to SEO penalties, such as duplicate content or thin content.

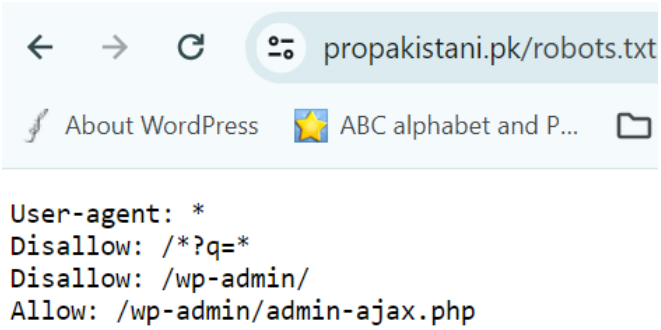
ROBOTS.TXT EXAMPLE



ROBOTS.TXT SYNTAX

Syntax:

- Use **"User-agent"** to specify which crawler the rule applies to.
  - *
- **"Disallow"** directive to specify which URLs should not be crawled/indexed.
- **"Allow"** directive to override a previous disallow rule.
- Also include **"Sitemap"** directive to indicate the location of the website's XML sitemap.



ROBOTS.TXT .. CONT.

Common Mistakes:

- **Blocking essential pages** inadvertently.
- Allowing crawlers to access sensitive or private information.

- **Keep your hosting safe… (it contains robots.txt..harmful manipulation in it can harm your SEO completely)**

CANONICAL ISSUE

Canonical issue: **“**A canonical issue in SEO refers to problems that arise from having multiple URLs that point to the same content within a website.”

- **This issue can dilute the site's search engine rankings**

- **http://propakistani.com**
- **https://propakistani.com**
- **https.://www. propakistani.com**
- **https://propakistani.com**
- **propakistani.com**
- **WWW vs. Non-WWW and HTTP vs. HTTPS, filters**

CANONICAL ISSUE EFFECT ON SEO

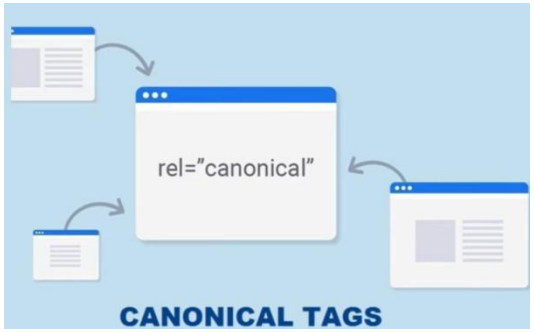
- **This issue can dilute the site's search engine rankings**

1. **Duplicate Content**: Search engines may crawl and index multiple versions of the same content, seeing them as separate and duplicate pages
2. **Split Page Authority**: Backlinks might point to multiple versions of the same content, splitting the authority (or link equity) among those versions instead of consolidating it to a single page.
3. **Wasted Crawl Budget**: Multiple versions of the same content can lead to inefficient use of crawl budget, as the bot spends time crawling duplicate content instead of discovering new or updated content.

CANONICAL TAG

Canonical Tag:

“A canonical tag is an HTML element that helps website administrators to prevent duplicate content issues by specifying the "canonical" or "preferred" version of a web page.



Purpose:

- Helps search engines understand which version of a page to index and present in search results.
- Prevents split ranking signals for pages with similar or duplicate content.

CANONICAL TAG SYNTAX

- Placed in the **<head>** section of an HTML document.

- **<link rel="canonical" href="URL-of-the-preferred-page"/>**

- **Example:**

  <link rel="canonical" href="https://propakistani.pk/" />

BROKEN LINKS

Reasons of broken links:

- Updated URL
- Change in URL without Redirections
- Deleted web pages
- Content/ post deleted
- Misspelled URL

The first step is to find the broken links, if any

- **Tools:**

→ https://brokenlinkcheck.com/

→ Google Search Console

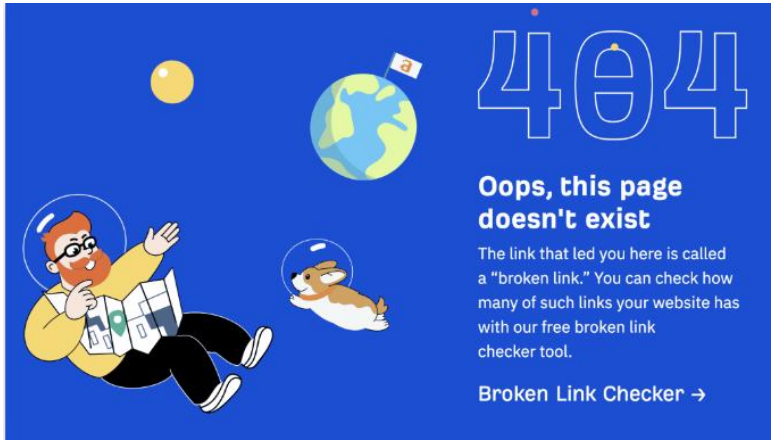
404 ERROR PAGE

“An HTTP error 404 occurs *when the web server cannot find a resource* (like a webpage) at a certain URL.”

404 error page is a response to missing links

Add it in your website

- A well-designed 404 page **can encourage users to explore other parts of your site instead of leaving** immediately after encountering an error.



WHAT IS A REDIRECT?

A redirect is a way to send **both** **users** and **search engines** to a different URL from the one they originally requested.

- Most commonly used redirects are **301** & **302**

301 redirect: It's a permanent redirect.

- It tells search engines that the original URL has been permanently moved to a new location.

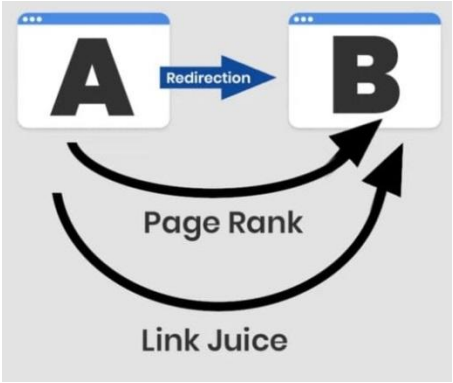
302 redirect: It's a temporary redirect.

- It indicates that the original URL has been temporarily moved to a new location.

301 & 302 REDIRECTS

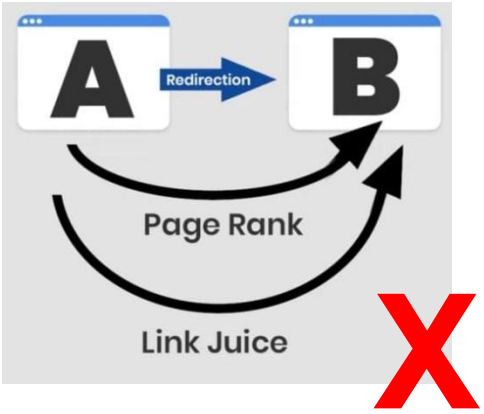
When to apply 301 redirection:

- Webpage/ website changed
- Domain name changed (TLD changed)
- Ownership changed
- http to https



When to apply 302 redirection:

- Page is updating / under maintenance
- Content is not ready

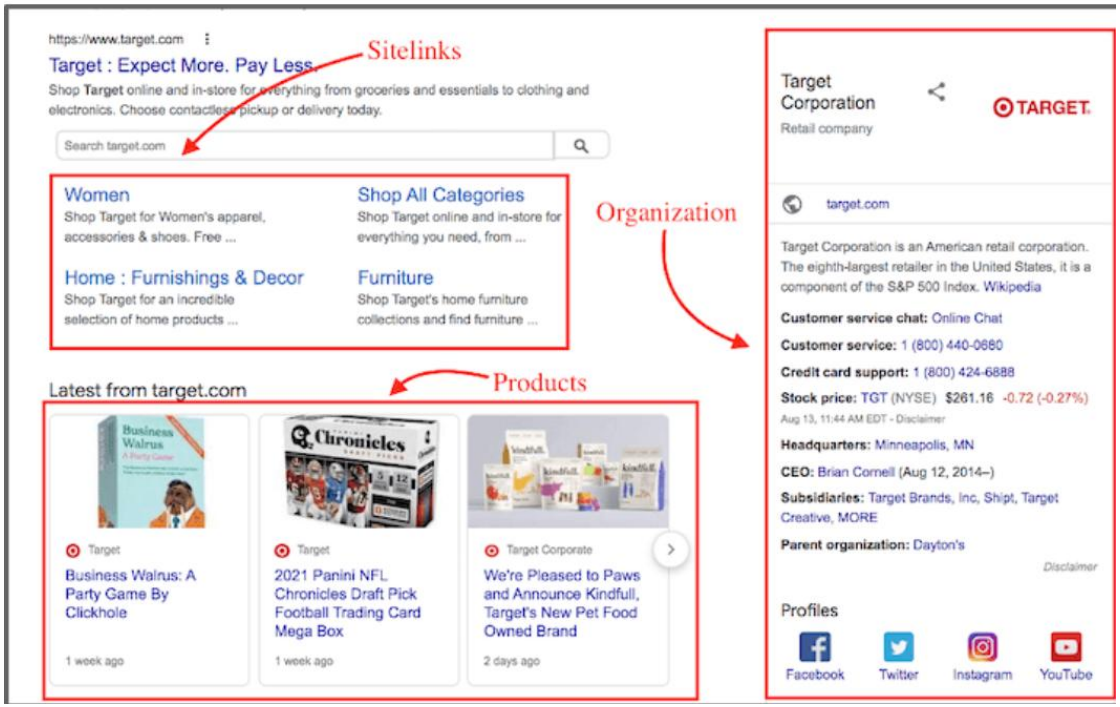


SCHEMA MARKUP

Schema markup (schema.org) is a *structured data vocabulary* that helps search engines to understand your website content *better*.

- **Enhances search results:** It can result in rich snippets or enhanced search results.
- **Structured data:** It provides a structured format for search engines to understand the content, improving relevance and visibility.
- **Types:** Includes various types like product, recipe, event, organization, bread crumbs etc.
- **SEO benefit:** Can lead to higher click-through rates and improved rankings by providing more relevant information to users.

ADDING SCHEMA



Schema markup generator Tools:

https://technicalseo.com/tools/schema-markup-generator/

DWELL TIME & BOUNCE RATE

Dwell time: Measures the duration visitors spend on a website or webpage before returning to search results or navigating away.

- Higher *dwell time* **indicates engaging** content and **user satisfaction**, which can positively impact SEO rankings.

Bounce rate: Indicates the percentage of visitors who navigate away from a website after viewing only one page.

- A lower *bounce rate* suggests that visitors **are finding relevant information** or **engaging content**, which can also positively affect SEO.

HOW TO DECREASE BOUNCE RATE

- Ensure that your meta title matches the content
- Improve content readability
- Enhance website UI
- Avoid too many ads
- Add images and media on the page
- Work on improving website speed
- Use internal linking in a strategic way
- Make sure there are no technical glitches

GOOD BOUNCE RATE

What bounce rate is good in SEO?

A good bounce rate varies on the niche of the website and the type of page.

On average, the bounce rate should be between 40% to 50%.

- **eCommerce sites:** 20% to 45%
- **Content-based sites and blogs:** 35% to 60%
- **Landing pages: 60% to 90%**
- **Web portals, news sites, event sites, and dictionaries:** 65% to 90%
- **B2B sites:** 25% to 55%

In general, you can categorize the bounce rate as follows:

- **Poor:** 70% or more
- **Average:** 55% to 70%
- **Good:** 41% to 70%
- **Excellent:** 25% to 40%

PAGE SPEED OPTIMIZATION

Page speed optimization is vital for SEO as it boosts **user experience**, **search rankings**, and **conversions**

User Experience: Faster pages enhance user satisfaction, reducing *bounce rates* and increasing *engagement.*

Search Rankings: Google considers page speed a **ranking factor**, impacting visibility.

Mobile Performance: Quick-loading sites are crucial for mobile users, who expect speed.

Conversions: Faster sites lead to *higher* conversion rates, as users prefer seamless experiences.

Crawl Efficiency: Slow pages can hinder search engine crawlers, affecting indexing.

Competitive Edge: A slight speed advantage can differentiate your site from competitors.

TOOLS TO ANALYZE PAGE SPEED

- These tools give a performance score from 0 to 100. The higher the number, the better.

1. Google's **PageSpeed Insights**
2. **Gtmetrix**
3. Chrome User Experience Report
4. Search Console (Core web vital report)
5. Web-vitals JavaScript Library
6. LightHouse
7. Chrome Dev Tools

WEB VITALS?

“**Web Vitals** are a set of standardized metrics from Google that help developers understand how users experience a web page.”

- Web Core Vitals optimization can improve a website's performance in search results, as Google prioritizes sites that offer a better user experience.
- **Important metrics to measure:**
  - Core Web Vitals (LCP, INP, CLS)

CORE WEB VITALS?

“Core Web Vitals are the subset of Web Vitals.”

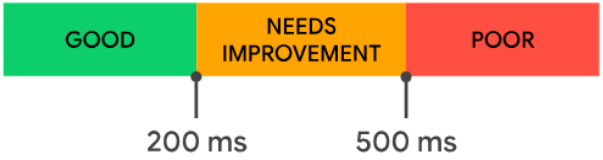
- Each of the Core Web Vitals represents a distinct facet of the user experience.
- These metrics are based on three primary areas of user experience, including:
  - Page ***loading*** performance
  - Ease of ***interaction***
  - ***Visual stability*** of a page from a user’s perspective
- Core Web Vitals **are measurable in the field**, and **reflects the real-world experience** of a critical user outcome

CORE WEB VITALS

1. **Visual Load → Largest Contentful Paint (LCP)**

2. **Visual Stability → Cumulative Layout Shift (CLS)**

3. **Interactivity → Interaction to Next paint (INP)**



- Focuses on *loading*, *interactivity*, and *visual stability*

LARGEST CONTENTFUL PAINT - LCP

- ***Visual Load***

"Largest Contentful Paint is a performance metric that measures the loading time of the main content of a web page."

  - how long it takes from the user initiating the page load until the largest text block, image, or video is rendered within the *viewport.*

- **LCP must occur within 2.5 seconds of when the page first starts loading.**

WHAT DOES LCP MEASURE?

- Images
- Video poster images
- Background images
- Block-level text

- The LCP element size usually reflects what's visible in the viewport.
  - For resized images, the reported size is the smaller of the **visible size** or the **intrinsic size**.

WHAT CAUSES LCP

Blocked Caching

Slow Hosting

Render-Blocking Javascript or CSS

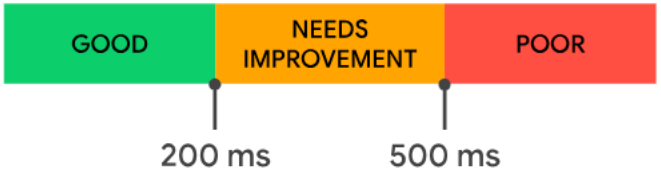
Un Optimized Images

INTERACTION TO NEXT PAIN

- *interactivity*

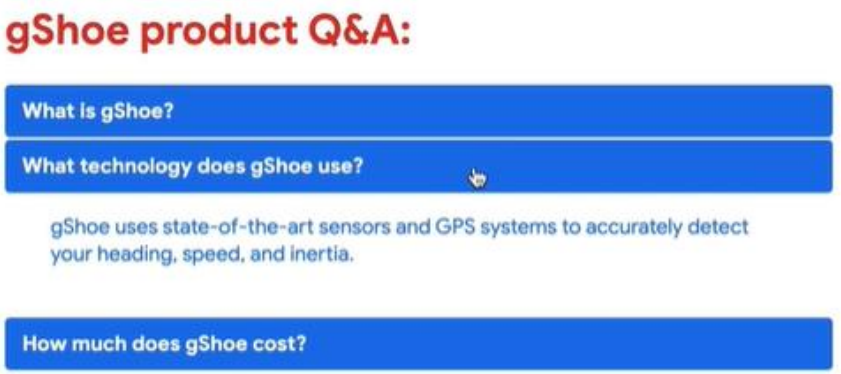
“INP (Interaction to Next Paint): Measures a page's overall responsiveness to user interactions by observing the latency of all click, tap, and keyboard interactions that occur throughout the lifespan of a user's visit to a page.”

- The final INP value is the longest interaction observed, ignoring outliers.
- The page's INP is calculated when the user leaves the page.
- **A low INP means that a page was reliably responsive to user input.**
- **25% of visitors will leave a website if it takes more than 4 seconds to respond.**
- pages must have a INP of **200 milliseconds** or less.



INP...CONT. EXAMPLE

Poor responsiveness



Good responsiveness

CUMULATIVE LAYOUT SHIFT (CLS)

- ***visual stability***

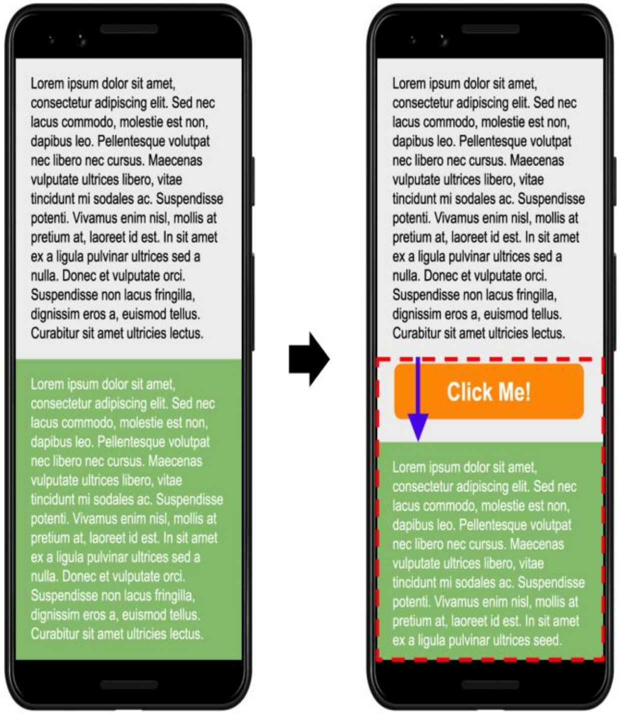
"Cumulative Layout Shift, is a metric that measures the stability of content on a webpage. It quantifies how often and how significantly elements on the page shift positions unexpectedly from the user's viewpoint during the loading phase."

- To provide a good user experience, must should maintain a CLS of **0.1.** or less.

CUMULATIVE LAYOUT SHIFT (CLS)…CONT.

- **<u>Not</u> all layout shifts are bad**. Dynamic web applications often change element positions. <u>A layout shift is problematic only if unexpected.</u>
- **Shifts due to user interactions** (like clicking, tapping, or typing) **are usually fine.**

Example: Elements in the visible viewport shift from their starting position while a user is reading content on a given page.



HOW TO IMPROVE PAGE SPEED

Here are few ideas for improving your website speed:

Compress your images:

Images are usually the biggest files on a webpage.

Use image compression tools to reduce their file sizes so they take as little time to load as possible.

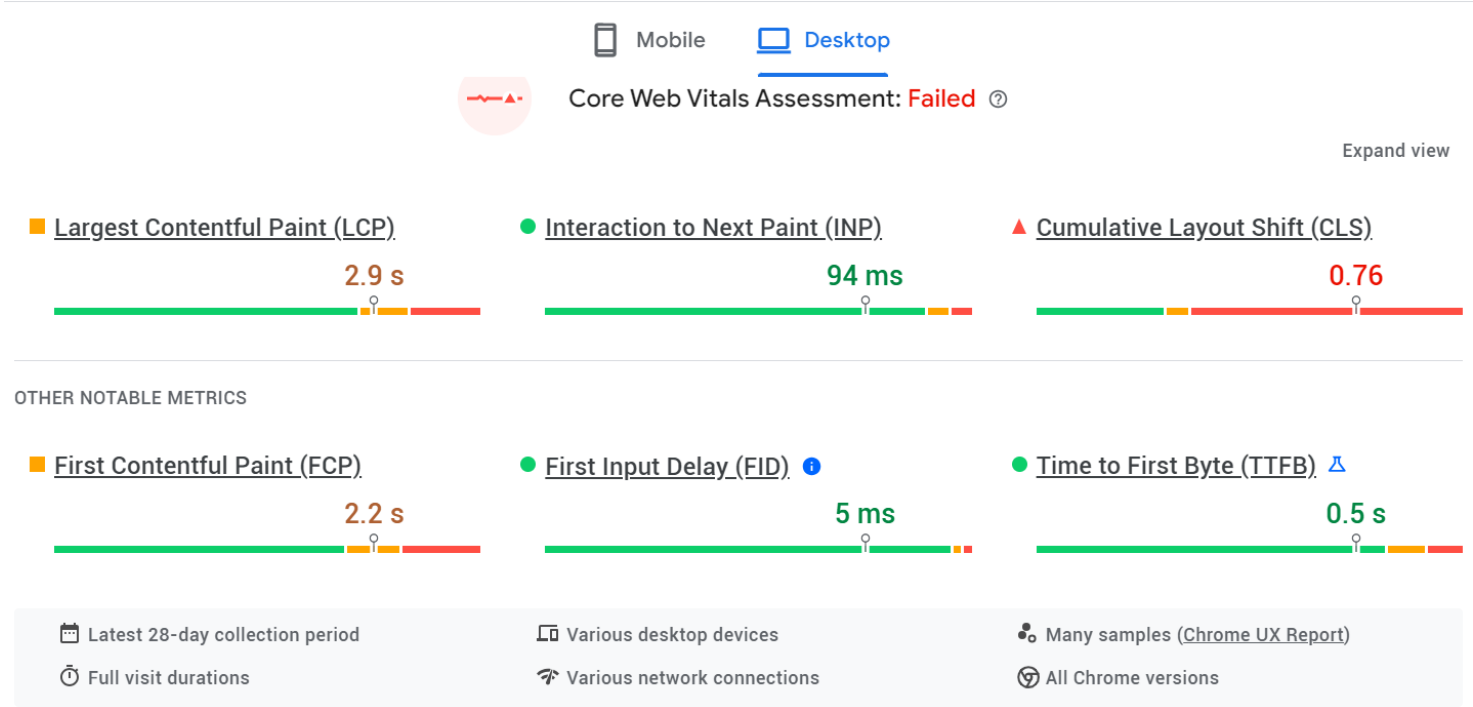
Use a content distribution network (CDN):

A CDN stores copies of your webpages on servers around the globe. It then connects visitors to the nearest server, so there's less distance for the requested files to travel.

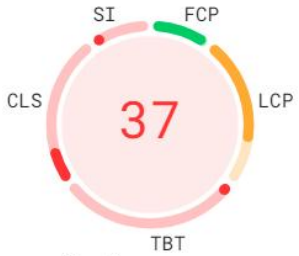
Minify HTML, CSS, and JavaScript files:

Minification removes unnecessary characters and whitespace from code to reduce file sizes. Which improves page load time.

SAMPLE LCP TEST (NUML.EDU.PK)



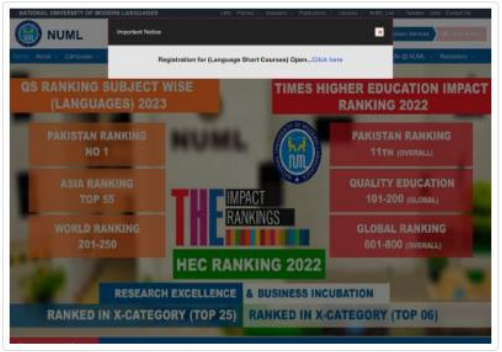
Mobile Desktop



Performance

Values are estimated and may vary. The performance score is calculated directly from these metrics. See calculator.

▲ 0–49 ■ 50–89 ● 90–100



METRICS Expand view

- ● First Contentful Paint **0.7 s**
- ■ Largest Contentful Paint **1.7 s**
- ▲ Total Blocking Time **840 ms**
- ▲ Cumulative Layout Shift **0.388**
- ▲ Speed Index **6.8 s**