(TR-102) MASTERING THE SEMANTIC WEB –

Training Day 16 Report :

4 July 2024

Google Tag Manager:

Google Tag Manager (GTM) is a free tool that allows to manage and deploy marketing tags (snippets of code or tracking pixels) on the website or mobile app without modifying the code directly. GTM simplifies the process of adding, updating, and managing various tags, such as those used for analytics, conversion tracking, and remarketing.

Website Management Essentials: Robots.txt, Sitemap, and Error Pages

1. Robots.txt

The robots.txt file is used to manage and control how web crawlers (bots) interact with the website. This file is placed in the root directory of the website and provides instructions to

By: Jaspinder Kaur Walia   URN:2203843   CRN:2215199

web crawlers about which pages or sections of the site they are allowed to access and index. The below is the breakdown of its components and functionality:

- o User-agent: Specifies the web crawler to which the rules apply. For example, User-agent: * applies to all crawlers.
- o Disallow: Specifies the directories or files that should not be accessed by the crawler. For example, Disallow: /admin/ prevents crawlers from accessing the admin directory.
- o Allow: Specifies exceptions to the disallow rules, allowing access to specific files or directories.
- o Sitemap: Provides the location of the sitemap file to the crawlers, helping them better understand the site structure.

2.sitemap.html and sitemap.xml

Sitemaps are essential for helping search engines understand the structure and content of your website. There are two common types of sitemaps:

sitemap.html

Purpose: Primarily designed for human visitors, a sitemap.html file provides an overview of the website's structure, including links to important pages and sections.

Content: Typically includes a hierarchical list of links, helping users navigate the site more easily.

sitemap.xml

Purpose: Created for search engines, a sitemap.xml file lists the URLs of the site along with additional metadata such as the last modification date, change frequency, and priority.

By: Jaspinder Kaur Walia   URN:2203843   CRN:2215199

Content: Helps search engines crawl the site more efficiently and index the content

3. 404.html

The 404.html file is a custom error page displayed when a user tries to access a page that does not exist on your website. Customizing this page improves user experience by providing helpful navigation options and preventing users from leaving the site out of frustration.

How Priority Works in XML Sitemaps?

- The priority element in an XML sitemap is a way to provide search engines with information about the relative importance of pages on your website. This information can help search engines decide how to allocate their crawling resources.
- The priority element has a value range from 0.0 to 1.0.
- A value of 1.0 indicates the highest priority, while a value of 0.0 indicates the lowest priority.
- Typically, the homepage (/) has the highest priority, often set to 1.0.

<u>WordPress:</u>

WordPress is a widely used content management system (CMS) that allows to create and manage websites without needing to write code from scratch.

By: Jaspinder Kaur Walia   URN:2203843   CRN:2215199

Alternative to WordPress: Accelerated Mobile Pages (AMP)

Definition: AMP is an open-source framework designed by Google to create fast-loading mobile web pages.

Advantages:

- Speed: AMP pages load almost instantly on mobile devices.
- SEO Benefits: Google prioritizes AMP pages in search results, especially in mobile searches.
- User Experience: Improved user experience with faster loading times.

By: Jaspinder Kaur Walia   URN:2203843   CRN:2215199