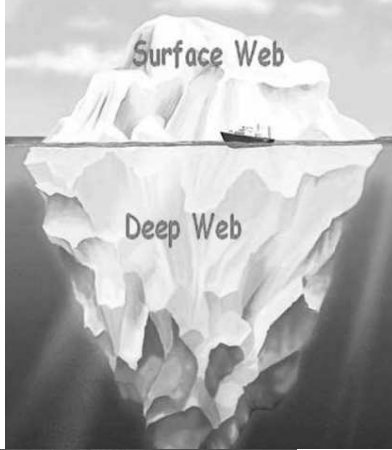


এবার আসুন শুনি অদৃশ্য ওয়েবের কথা

মঈনুল বারী শুভ

আমাদের মত দেশেও, বিশেষ করে শহর এলাকায় বারো-চোদ্দ বছরের কোনো বালককেও যদি জিজ্ঞেস করি: ওয়ার্ল্ড ওয়াইড ওয়েব কী? তাহলেও হয়ত ঠিকই তার কাছ থেকে উত্তর পাওয়া যাবে 'থ্রিডব্লু' সম্বন্ধে। ইন্টারনেট নামের বিশ্ববিস্তৃত নেটওয়ার্কের একটি গুরুত্বপূর্ণ সেবা এই ওয়ার্ল্ড ওয়াইড ওয়েবকে চেনে না এমন কম মানুষই আছেন। কিন্তু সেই একই মানুষগুলোকে যদি জিজ্ঞেস করা হয়: তাহলে বলুন তো, অদৃশ্য ওয়েব কী? অথবা ডিপ ওয়েব? আমার ধারণা ওয়েব সম্বন্ধে এতক্ষণ ধরে কথার ফুলঝুরি ছড়ানো মানুষগুলোর শতকরা আশি ভাগই চুপ হয়ে যাবেন। অদৃশ্য ওয়েব? ডিপ ওয়েব? এ আবার কোন চিজ? এ ধরনের ভাবনাই হয়ত খেলা করবে তাদের মনে।

প্রিয় পাঠক, মনে হচ্ছে আপনিও তাদের মতই ভয়াবাচ্যাকা খেয়ে গেছেন? ওয়েব আবার অদৃশ্য হয় কীভাবে? গভীরই বা হয় কীভাবে? এটাই হয়ত ভাবছেন মনে মনে। কিন্তু আসলেই আমাদের অতি পরিচিত ওয়ার্ল্ড ওয়াইড ওয়েবের পাশাপাশি অদৃশ্য বা ইনভিজিবল ওয়েব বলেও একটা ওয়েব



আছে। এবং জেনে আরো অবাক হবেন, আমাদের পরিচিত ওয়ার্ল্ড ওয়াইড ওয়েবের চাইতে অদৃশ্য বা ডিপ ওয়েবের আয়তন আরো কয়েক গুণ বেশি। আসুন তাহলে, আর ভূমিকা না করে জেনে নেয়া যাক অদৃশ্য ওয়েব আসলে কী!

তার আগে বলি ভিজিবল বা দৃশ্যমান ওয়েব কী। আমরা ইন্টারনেট সার্চ ইঞ্জিন এবং সাবজেক্ট ডিরেক্টরির মত অনুসন্ধান হাতিয়ার ব্যবহার করে ওয়ার্ল্ড ওয়াইড ওয়েবের যে অংশটিকে পাই সেটি হল ভিজিবল ওয়েব। আর ইনভিজিবল বা অদৃশ্য ওয়েব হচ্ছে সেটি, কোনো সার্চ ইঞ্জিন বা সার্চ ডিরেক্টরি ব্যবহার করে যার নাগাল পাওয়া সম্ভব নয়।

সব কিছু কেন দৃশ্যমান হয় না?

দিন দিন ইন্টারনেট সার্চ ইঞ্জিনগুলোর শক্তি বাড়ছে। আজ থেকে বছর তিনেক আগেও 'অদৃশ্য' বলা হত এমন অনেক ওয়েব সাইট বা

ডকুমেন্টও এখন সার্চ ইঞ্জিনগুলোর মাধ্যমে সার্চ করে বের করা যাচ্ছে। তারপরও কিছু কিছু তথ্যসম্পদ আছে যেগুলো সার্চ ইঞ্জিনগুলোর আওতার বাইরেই থেকে গেছে, যেমন:

◆ অনুসন্ধানযোগ্য ডাটাবেসের কনটেন্ট: অদৃশ্য ওয়েবের একটি বড় অংশ জুড়েই আছে হাজার

দিয়ে বোঝাই না হয়ে যায় সেটি নিশ্চিত করার জন্য। এর মধ্যে আছে ডায়নামিকালি জেনারেট করা এমন সব পেজ যেগুলো দু'চারজন ইউজার ছাড়া বাকি আর কারো কাছে কোনো মূল্য বহন করে না। লাইব্রেরি ক্যাটালগ, পাবলিক রেকর্ড ডাটাবেস ইত্যাদিতে শত শত মানুষ সার্চ করে

নানারকম তথ্য সংগ্রহ করে, এসব তথ্য শুধু এসব মানুষের কাছেই গুরুত্বপূর্ণ। সার্চ ইঞ্জিনগুলো বৃহত্তর জনগোষ্ঠীর কাছে গুরুত্বহীন এসব ডাটা দিয়ে তাদের ডাটাবেস ভর্তি করতে চায় না। আবার কোনো

কোনো ওয়েব পেজ নির্মাতাই চায় না তার পেজের সব তথ্য সবার কাছে দৃশ্যমান হোক। এজন্য তারা ওয়েব সাইটে বিশেষ মেটাট্যাগ দিয়ে রাখে যাতে সার্চ ইঞ্জিনগুলো সেই সাইট বা এর কোনো কোনো অংশ সার্চ করতে না পারে। এক্ষেত্রে সার্চ ইঞ্জিনের পক্ষে এসব সাইটের তথ্য সার্চ করে বের করা সম্ভব হয় না। তবে আশার কথা হচ্ছে, সার্চ ইঞ্জিনগুলোর শক্তি বৃদ্ধির সঙ্গে

সঙ্গে অদৃশ্য ওয়েবের কিছু কিছু অংশ এখন সার্চ ইঞ্জিনের মাধ্যমে অনুসন্ধান করা যাচ্ছে। যেমন কিছু কিছু ওয়েব পেজের কথা বলা যায়, এগুলো আগে ইনভিজিবল হলেও এখন ভিজিবল। যেমন: ◆ এইচটিএমএল নয় এমন ওয়েব পেজ (pdf, Word, Excel, PowerPoint), যেগুলো এখন এইচটিএমএল ফরম্যাটে কনভার্ট করা হয়েছে বা হচ্ছে।

◆ স্ক্রিপ্টভিত্তিক পেজ, যেগুলোর ওয়েব অ্যাড্রেসে '?' চিহ্ন বা কোনো স্ক্রিপ্ট কোডিং আছে।

◆ ভিন্ন ধরনের ডাটাবেস সফটওয়্যার (যেমন, Active Server Pages, Cold Fusion) ব্যবহার করে ডায়নামিক উপায়ে তৈরি পেজ। এগুলোকেও ইনডেক্স করা সম্ভব যদি কোথাও এগুলোর নির্দিষ্ট একটি ওয়েব অ্যাড্রেস থাকে যেটি সার্চ ইঞ্জিনের ক্রলার প্রোগ্রামগুলো খুঁজে বের করতে পারে।

◆ স্ক্রিপ্টভিত্তিক পেজ, যেগুলোর ওয়েব অ্যাড্রেসে '?' চিহ্ন বা কোনো স্ক্রিপ্ট কোডিং আছে।

◆ ভিন্ন ধরনের ডাটাবেস সফটওয়্যার (যেমন, Active Server Pages, Cold Fusion) ব্যবহার করে ডায়নামিক উপায়ে তৈরি পেজ। এগুলোকেও ইনডেক্স করা সম্ভব যদি কোথাও এগুলোর নির্দিষ্ট একটি ওয়েব অ্যাড্রেস থাকে যেটি সার্চ ইঞ্জিনের ক্রলার প্রোগ্রামগুলো খুঁজে বের করতে পারে।