

Saat ini, keberadaan mesin pencari (search engine) di internet seperti halnya Yahoo! atau Google sudah tidak dapat lagi dipisahkan dari kehidupan kita. Pengguna internet menggunakan mesin pencari untuk mendapatkan berbagai macam informasi, mulai dari mencari lowongan, mencari klien, mencari data, menemukan teman, hingga mencari cara membuat atau melakukan sesuatu. Bagaimana cara mesin pencari tersebut menemukan halaman web yang kita cari?

Menurut pengelola Google, cara kerja mesin pencari sama seperti mencari kata yang tepat di dalam indeks sebuah buku yang tebal. Saat kita mencari di Google, program mereka akan memeriksa indeks guna menentukan hasil yang paling relevan untuk disampaikan kepada Anda. Proses pencarian ini memiliki tiga komponen penting, yakni crawling atau menjelajahi, menyusun index, dan menyajikan hasilnya. Crawling adalah sebuah proses dimana Googlebot menemukan halaman baru atau yang telah diperbaharui untuk ditambahkan pada indeks Google.

Google menggunakan sekelompok komputer untuk mengambil miliaran halaman web. Program yang melakukan pengambilan ini dinamai Googlebot (juga dikenal sebagai robot, bot, atau spider). Googlebot menggunakan proses algoritmis: program komputer untuk menentukan situs mana yang akan diambil, seberapa sering, dan seberapa banyak halaman yang akan diambil dari masing-masing situs.

Proses pencarian Google dimulai dengan satu daftar URL halaman web, yang dihasilkan dari proses pencarian sebelumnya, dan ditambahkan data peta situs (sitemap) yang disediakan oleh pengelola web (webmaster). Saat Googlebot mengunjungi masing-masing website ini, ia mendeteksi semua link pada masing-masing halaman dan menambahkannya pada daftar halaman yang akan dijelajahi. Situs baru, perubahan pada situs yang ada, dan link yang mati, akan dicatat untuk digunakan mengupdate Google Index.

Googlebot memeroses setiap halaman yang dijelajahnya guna mengumpulkan indeks semua kata yang dilihatnya berikut lokasinya di masing-masing halaman. Selain itu, Google juga memeroses informasi yang disertakan di label (tag) dan ALT attribute konten kunci. Googleboot mampu memeroses banyak jenis konten, namun tidak semuanya. Sebagai contoh, Google tidak bisa memeroses konten beberapa file rich media atau halaman dinamis.

Saat pengguna memasukkan permintaan pencarian mereka, mesin Google akan mencari indeks untuk menemukan halaman yang sesuai dan kembali dengan hasil yang menurutnya paling relevan bagi pengguna. Relevansi ditentukan oleh lebih dari 200 faktor. Salah satunya adalah pagerank (peringkat halaman) tersebut. Pagerank adalah ukuran nilai penting halaman

berdasarkan link yang datang dari halaman lain. Dalam bahasa sederhana, setiap link menuju halaman di situs Anda dari situs lain akan menambah pagerank situs Anda.

Kedudukan masing-masing link tidak sama. Google berusaha keras meningkatkan kenyamanan pengguna dengan mengidentifikasi link spam dan produk lain yang memengaruhi hasil pencarian secara negatif. Tipe link terbaik adalah yang diberikan berdasarkan kualitas konten Anda. Agar situs Anda memiliki peringkat pencarian yang baik di halaman hasil pencarian, sangat penting untuk memastikan bahwa Google bisa menjelajahi dan mengindeks situs Anda dengan benar. Panduan Webmaster Google menjabarkan beberapa hal terbaik yang akan membantu Anda menghindari hal-hal negatif dan memperbaiki peringkat Situs Anda.

*Sumber: Google*