**Anatomía de una gran escala de motores de búsqueda de hipertexto en la red**

Los motores de búsqueda han ido desarrollándose con el paso del tiempo, así mismo la cantidad de información guardada en la web ha ido creciendo rápidamente con el pasar de los años, y también se ha incrementado el número de usuarios que desconoce cómo usar motores de navegadores web. Sin embargo dichos motores de búsqueda suelen ser lentos a la hora de realizar concordancia de palabras claves, y también se ven afectados por la gente que emplea medidas que engañan a los motores de búsqueda automatizados. Por lo que Google nace con el ideal de lidiar con estos problemas de la actualidad. El nombre de Google se lo ideó en base al término googol.

Por el año 1994 y en secuencia de este cada vez más crecía el número de consultas que la gente realizaba a diario, lo que supondría que un motor de búsqueda soportaría un promedio de cientos de millones de consultas por año, esto fue previsto para el año 2000.

En el año 1994 se creyó que los índices de búsqueda harían excelente una consulta en un motor web para encontrar algún tipo de información , sin embargo, por el año 1997 este ideal cayó cuando una consulta arrojaba más “datos basura” que la información explicita del tema a investigar. Esto es muy importante ya que la gente no tiene la paciencia de buscar más allá de las 2 primeras páginas generalmente y tiende a rechazar toda la información restante.

La web tuvo un crecimiento comercial muy grande, ya que solo el 1.5% de las páginas finalizaban en .com por el año 1993, mientras que en el año 1997 este porcentaje creció al 60%, además la web empezó a abarcar el ámbito académico.

El fin de la creación de google era crear un motor de búsqueda que pudiera adaptarse a todas estas circunstancias de tal forma que no perdiera su eficacia a la hora de devolver información.

La creación de mapas que contienen hasta 518 millones de hipervínculos ha permitido el cálculo rápido de las páginas web “PageRank”. Lo cual ayuda con la importancia de la gente a la información que desea buscar en primera instancia.

La mayor parte de google está programada en C o C++, pudiéndose ejecutar en Solaris o Linux. La descarga de páginas web se realiza mediante varios rastreadores distribuidos. Existe un servidor URL el cual envía las direcciones URL que se obtengan de los rastreadores. Las páginas que se recuperan son enviadas al servidor de almacenamiento., este las comprime y las almacena en un repositorio. Cada página web tiene un número de identificación llamada docID que se asigna cuando una nueva dirección URL se analiza fuera de una página web. A continuación se lee el repositorio y se descomprimen los archivos, donde cada documento se lo denota como éxito. Todo esto desencadena en la lectura del archivo de anclas URL para convertir las búsquedas en URLs absolutas.

Es importante citar que la medida más importante de un motor de búsqueda es la calidad de sus resultados de búsqueda, sin embargo, aún la eficacia de un motor de búsqueda mediante este medio es difícil de lograr si evaluamos completamente la demanda de un usuario. Está se ha conseguido desarrollar debido a la experiencia que tiene Google con sus usuarios.