

## YACY, UNA NUEVA FORMA DE BUSCAR POR INTERNET



### INTRODUCCION

La búsqueda y recuperación exitosa de información en la Web, se ha convertido en el principal problema que agobia a los usuarios, puesto que los resultados que se obtienen a sus demandas de información, en su mayoría, son irrelevantes, imprecisos e impertinentes.

Esta problemática es generada principalmente por la falta de un lenguaje de recuperación de información estandarizado que aplique para todos los sistemas de recuperación de información, es tratada de enmendar con la creación de buscadores o aplicaciones que permitan personalizar las búsquedas de acuerdo al comportamiento en línea y perfil de usuario, identificando con mayor precisión las necesidades del mismo.

Las redes P2P (punto a punto), hacen parte de las nuevas tendencias que buscan brindar mayor eficiencia por parte de los buscadores y sistemas de recuperación de información, permitiendo una conexión directa entre puntos<sup>1</sup>, con posibilidades de conformación de grupos o comunidades, intercambio rápido de información, comunicación constante, respuestas precisas, pertinentes y relevantes.

Yacy, es un buscador gratuito (P2P) diseñado por Michael Christen, que ofrece como principal características la descentralización de la información y el anonimato en cuanto al comportamiento de cada conjunto de documentos. El proyecto fue publicado a finales del año 2006, este se encuentra mejoramiento continuo, permitiendo que los usuarios expongan sus necesidades y sugerencias.

---

<sup>1</sup> PUNTOS: serie de nodos que se comportan simultáneamente como clientes y como servidores de los demás nodos de la red P2P. P2P. [termino de búsqueda: P2P]. [en línea-html]. Consulta [07-03-03]. P. 1. disponible en [www.wikipedia.com](http://www.wikipedia.com).

## RESUMEN

Con los avances tecnológicos, se ha modificado la manera como se maneja la información y el soporte donde se distribuye y pone en conocimiento la misma, concentrando la búsqueda y recuperación de información en los buscadores (motores de búsqueda) que ofrece Internet por la facilidad en su manipulación y el ahorro de tiempo en las consultas.

Esto exige que los creadores de estos motores de búsqueda, diseñen herramientas que satisfagan plenamente las necesidades de los usuarios, exponiendo ideas tan innovadoras como la que impone Michael Christen con Yacy (buscador que presenta como principal ventaja la aplicación P2P (punto a punto)), es decir, una red distribuida donde todos los integrantes son servidores, permitiendo la descentralización, la escalabilidad, el anonimato y el intercambio exitoso de información.

Esta herramienta desarrollada en Java esta disponible para los sistemas operativos GNU/Linux, Windows y Mac, es gratuito, esta diseñado bajo los términos de la licencia GPL y posee como principales características: la **Blacklist** (lista negra), los **Messages** (mensajes), el **Public file directory** (directorio de archivos públicos) y la **página Web**.

### PALABRAS CLAVES:

Buscadores, Motores de búsqueda, P2P, Descentralización, Yacy, Red, Descentralización.

### ABSTRACT

The way the information is handled has been modified with the technological advances. The searching and recovering of information using Google and other engines facilitate the manipulation and the time of consultations.

This technology demand that the creators of these searching engines satisfy fully users needs, having in mind new ideas as innovative as the ones mentioned by Michael Christen with Yacy, a searching engine that presents as its main advantage the application of P2P, i.e. a distributed network where all the members are servers, permitting the decentralization, the scalability, the anonymity and the successful inter exchange of information.

This technical tool developed in Java is available for the operating systems GNU/LINUX, Windows and Mac, is free, and is designed under the GPL's licensed terms having as its main characteristics: the **Blacklist**, the **Messages**, the **Public file directory** and the **web page**.

### KEY TERMS:

Information Retrieval Systems, Distributed Systems, Net, Metasearch, Crawlers.

## YACY, UNA NUEVA FORMA DE BUSCAR POR INTERNET

Los avances tecnológicos que se presentan a pasos agigantados año tras año, han modificado en muchas formas el estilo de vida de la humanidad, en especial la manera como se maneja la información, el soporte donde se distribuye y pone en conocimiento la misma.

Como resultado de estos avances, se ha concentrado la búsqueda y recuperación de información en la gran gama de buscadores que brinda Internet, los cuales son fabricados de manera más específica, es decir, su principal objetivo es lograr personalizar la búsqueda de acuerdo al perfil que se identifique para cada persona, realizando un seguimiento o análisis de las actividades que esta desarrolla en línea y a las especificaciones que el mismo demande en las consultas.

Estos buscadores pretenden satisfacer de manera rápida las necesidades de los usuarios, ofreciéndoles como resultados a sus consultas la información más relevante, precisa y, coherente conservada en sus servidores.

De esta forma, constantemente se generan nuevas ideas y estructuras para el diseño de buscadores, algunas muy innovadoras, como la que expone Michael Christen\*, que intenta: “ofrecer otro tipo de técnicas de selección de resultados, para que la búsqueda sea más atractiva<sup>3</sup>”

Su idea es un nuevo motor de búsqueda denominado Yacy, que posee como principal característica y ventaja la aplicación P2P (punto a punto), que permite que cada uno de los puntos<sup>4</sup> de la red de Yacy se convierta en un servidor, formando una base de datos distribuida, es decir, que no existe un servidor central, por el contrario todos los usuarios son un servidor.

Yacy, es una herramienta de software libre desarrollada en Java\*\*, que indexa la información a medida que se navega y se comparte información entre puntos, además facilita la creación de comunidades para el intercambio y administración de la información, que es de interés de determinado grupo; en la actualidad se encuentra en un periodo de prueba y mejoramiento, disponible para los sistemas

---

<sup>2</sup> \* Michael Christen es el creador del primer buscador descentralizado peer-2-peer YaCy.

<sup>3</sup> Entrevista al creador de Yacy. [termino de búsqueda: YACY]. [en línea-html]. 2006 consulta [07-03-03]. P. 1. disponible en <http://www.ojobuscador.com/2006/11/29/entrevista-el-creador-de-yacy/>

<sup>4</sup> Ibid P. 1

<sup>5</sup> \*\* **Java** es un lenguaje de programación orientado a objetos desarrollado por Sun Microsystems a principios de los 90, es compilado en un bytecode que es interpretado (usando normalmente un compilador JIT), por una máquina virtual Java. JAVA. [termino de búsqueda: JAVA]. [en línea-html]. Consulta [07-03-03]. P. 1. disponible en [www.wikipedia.com](http://www.wikipedia.com)

operativos GNU/Linux, Windows y Mac, es gratuito y esta diseñado bajo los términos de la licencia GPL<sup>6</sup>.

### ¿Pero qué es lo nuevo en Yacy?

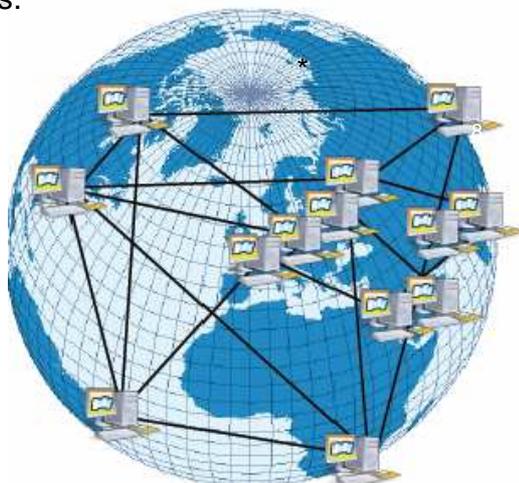
Las redes P2P poseen las siguientes características:

**Descentralización:** Este tipo de red (P2P) es por definición distribuida, cualquier miembro (punto<sup>7</sup>) se convierte en un servidor y un cliente, con las mismas características de todos los demás.

**Escalabilidad:** Una red P2P puede tener gran cantidad de usuarios, y funcionará mejor si más cantidad de los nodos se encuentran conectados, dado que entre más puntos se encuentren en línea, mayor será la cantidad de recursos que se pueden compartir.

**Anonimato:** Como no hay un servidor central, no se guarda el historial del comportamiento de los miembros, las peticiones, los autores, editores, lectores y el punto que conserva la información pueden conservarse en el anonimato.

**Intercambio:** porque se comparten, donan o intercambian unos recursos por otros recursos entre pares.



<sup>6</sup> **GPL** (General Public License o licencia pública general) es una [licencia](#) creada por la [Free Software Foundation](#) a mediados de los 80, orientada principalmente a proteger la libre distribución, modificación y uso de [software](#). Su propósito es declarar que el software cubierto por esta licencia es [software libre](#) y protegerlo de intentos de apropiación que restrinjan esas libertades a los usuarios. GPL. [termino de búsqueda: licencia GPL]. [en línea-html]. [07-03-31]. P. 1. disponible en [www.wikipedia.com](http://www.wikipedia.com).

<sup>7</sup> Ibid P. 1

<sup>8</sup> \* Imagen tomada de [www.yacy.net](http://www.yacy.net): ilustra como se establece una conexión P2P, directamente entre puntos.

En una entrevista realizada por Tomy Lorsch (vía correo electrónico) el 29 de noviembre de 2006 a Michael Christen<sup>9</sup>, el desarrollador de Yacy expone el principal enfoque ideológico que impulsó la creación de un sistema de esta magnitud, “Tenemos un enfoque ideológico, que tiene que ver con la libertad de la información, la imposibilidad de la censura y queremos que no se guarden los historiales de búsqueda de los usuarios. Eso lo podemos hacer y si logramos entregar buenos resultados a nuestros usuarios, también seremos más populares.<sup>10</sup>”

A diferencia de otros buscadores existentes en la Web, Yacy se instala en el PC, permitiendo personalizarlo de acuerdo al requerimiento de cada usuario, ajustando la configuración a las necesidades del mismo y facilitando el control de la gestión del contenido del buscador colectivo.

### **Características de Yacy**

Una vez configurado el motor de búsqueda y realizada la consulta, este ofrece diversas opciones, para facilitar la recuperación y conservación de la información consultada, por ejemplo:

El **Bookmarks** (registros) que permite guardar el listado de documentos, páginas o archivos, que son verdaderamente relevantes en la búsqueda.

La **Blacklist** (lista negra) que permite configurar las direcciones, páginas, y puntos de los cuales no deseamos que nos recupere ningún tipo de información.

Los **Messages** (mensajes) que facilitan la comunicación o el envío de mensajes entre puntos permitiendo visualizar constantemente su estado.

El **Public file directory** (directorio de archivos públicos) permite, subir o bajar documentos a la red P2P, los cuales, luego son indexados por el sistema.

Una **página Web**, que es creada automáticamente por el sistema de búsqueda Yacy, cuando se establece y configura un nuevo punto. Esta funciona como buscador para otros usuarios que deseen utilizarla, con la salvedad de no poder modificar la configuración, porque el sistema exige la clave de usuario.

### **¿Pero revolucionara Yacy la manera de buscar información en la Web?**

Esta respuesta, solo el tiempo nos la dará, porque aunque funcione muy bien, posiblemente se quede tan solo en simples grupos o comunidades con capacidad de compartir información, y aunque parezca que este nuevo mecanismo de búsqueda, intercambio y recuperación de información basado

---

<sup>9</sup> Ibid., p. 1

<sup>10</sup> Ibid., p. 1

en una red P2P, pueda revolucionar la forma como se buscan las cosas en Internet por las grandes ventajas que representa; también posee desventajas notorias, algunas de estas son:

- Cuando se busca entre una red prediseñada o una comunidad (grupo), la respuesta es rápida. Pero buscar en el índice global necesita más tiempo, ya que debe encontrar al punto alejado correcto, que contenga los mejores resultados de la búsqueda.
- Como todos los integrantes de la red, son servidores, se corre el riesgo de compartir virus e información inexacta.
- Solo funciona bien, o funcionaria mejor cuando tenga gran cantidad de usuarios, o parte de ellos, en línea. Y gran cantidad de documento e información indexada.
- La veracidad de la información depende de la integridad de los confortantes de la red.
- Por el momento Yacy es mucho más lento que los grandes buscadores comerciales y ofrece menos opciones de resultados frente a una demanda de información.
- Es importante entender y conocer la configuración que se aplica o modifica al sistema Yacy, ya que si por error se permite el acceso a todas las carpetas del PC, se corre el riesgo de pérdida, manipulación, entrada de virus y hackers a toda la información.

## BIBLIOGRAFIA

- Entrevista al creador de Yacy. [termino de búsqueda: YACY]. [en línea-html]. 2006 consulta [07-03-03]. P. 1. disponible en <http://www.ojobuscador.com/2006/11/29/entrevista-el-creador-de-yacy/>
- [www.yacy.net](http://www.yacy.net)
- YACY. [termino de búsqueda: YACY]. [en línea-html]. consulta [07-03-03]. P. 1. disponible en <http://meneame.net/story/yacy-aplicacion-p2p-podria-desplazar-google-search>.
- YACY. [termino de búsqueda: YACY]. [en línea-html]. consulta [07-03-17]. P. 1. disponible en <http://bitelia.com/2006/10/12/yacy-navegando-a-traves-de-p2p/> consulta realizada el 17 de marzo de 2007.
- YACY. [termino de búsqueda: YACY]. [en línea-html]. consulta [07-03-17]. P. 1. disponible en <http://es.wikipedia.org/wiki/YaCy>.

- YACY. [termino de búsqueda: YACY]. [en línea-html]. consulta [07-03-21]. P. 1. disponible en <http://software.hispavista.com/linux/utilidades/otros/11555/>.
- P2P. [termino de búsqueda: P2P]. [en línea-html]. Consulta [07-03-03]. P. 1. disponible en [www.wikipedia.com](http://www.wikipedia.com).

