

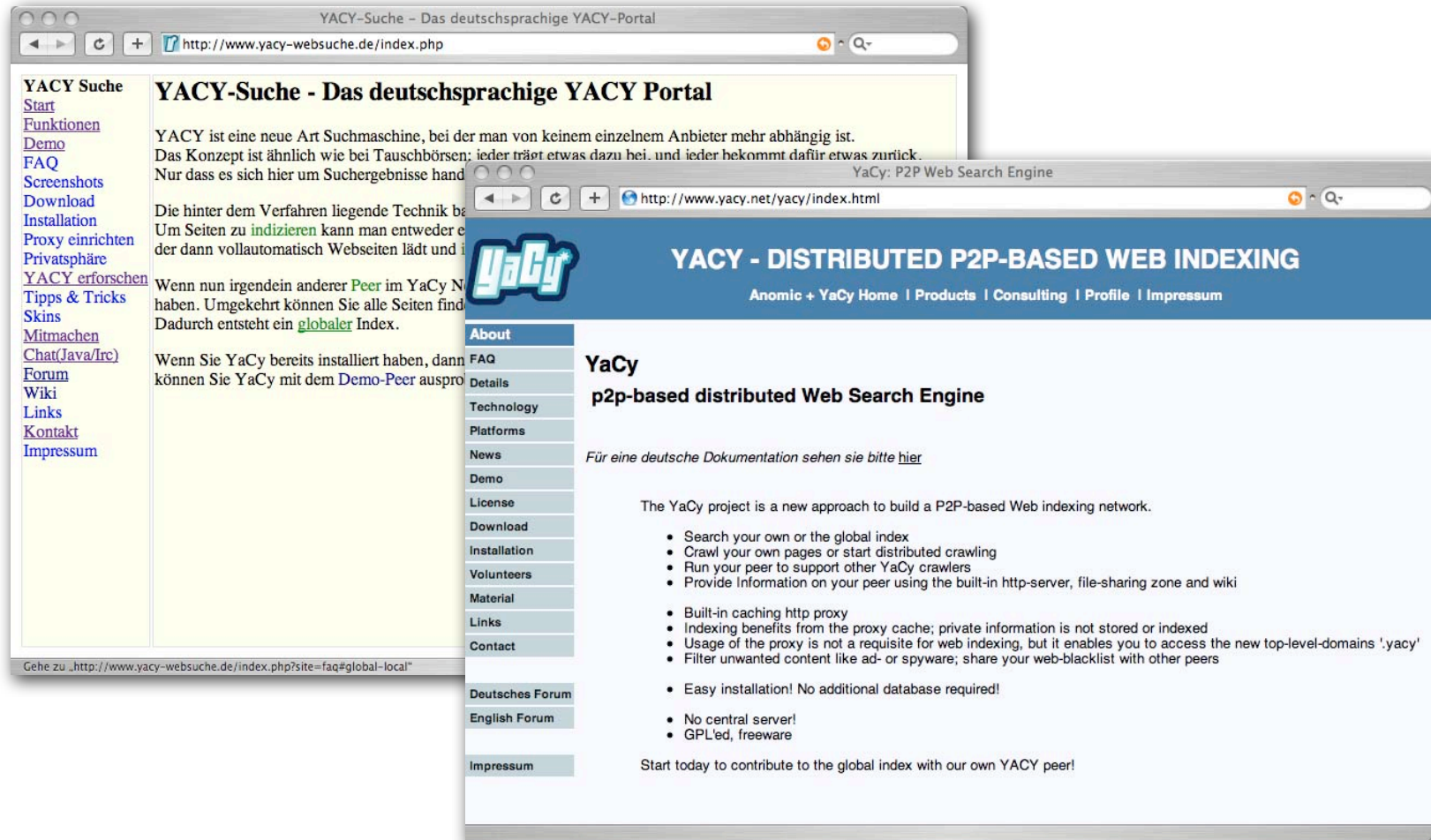


YaCy in der Presse & PR-Aktivitäten



YaCy in der Presse und PR-Aktivitäten

seit Juni 2004: Webauftritt in Englischer und Deutscher Sprache



Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



Dezember 2004:
Hannoversche Allgemeine
Zeitung

„Jedermanns Suchmaschine“

Michael Christen
+49 177 6424235

NR. 289 - DONNERSTAG, 9. DEZEMBER 2004

Internet & Computer

HANNOVERSCHE ALLGEMEINE ZEITUNG 23

Jedermanns Suchmaschine

Neuartige Software könnte in Zukunft Internet-Suchmaschinen wie Google und Yahoo Konkurrenz machen. Systeme wie „Yacy“ nutzen das Prinzip der unkontrollierbaren Tauschbörsen für Musik und Videos. Die Idee dahinter: Verteilte Datenbanken unter allen Teilnehmern enthalten einen stetig wachsenden Index über die Seiten des Internets – unabhängig von Konzernen.



Kopf hinter dem Yacy-Netzwerk: Michael Christen aus Frankfurt. Er hofft, dass sich neue Nutzer und Entwickler anschließen. Bisher entwickelte der Informatiker Yacy zum Großteil allein.

Die bisher nur von wenigen Computertweaks getestete kostenlose und von keinem Konzern gesteuerte Suchmaschine „Yacy“ wird soeben erst geboren. Die Idee basiert auf dem Open-Source-Gedanken und stammt aus Deutschland. Open Source heißt: Jeder kann daran mitarbeiten und hat Zugriff auf den Programmcode. Erste Tests und Beurteilungen von Experten lassen darauf schließen, dass das System möglicherweise das Zeug dazu hat, sich zu einem Kandidaten für eine Alternative zu Google & Co. zu entwickeln.

Das Prinzip ahmt dem der bekanntesten Internet-Suchmaschine Google. Google beherbergt seine Informationen auf mehreren hundert Servern. Ein Teil der Maschinen klappert eigenständig in Sekundenschnelle Hunderttausende von Internet-Seiten ab, folgt den darauf enthaltenen Links und verzeichnet die vorgefundenen Daten in einen riesigen Inhaltsverzeichnis. Der Anwender bekommt beim Aufruf von www.google.de von einem der Google-Server immer die gleiche, dankbar einfach zu bedienende Suchmaske ausgeliefert.

Die gerade entstehende neue Suchmaschine Yacy lagert ebenfalls auf mehreren Servern – allerdings nicht auf denen einer Firma, sondern auf den PCs der Anwender. Wer sich die kostenlose Software von <http://freemove.net/projects/yacproxy> installiert, macht seinen Computer für die Dauer seines Surfers im Internet zum Server. Die Software protokolliert auf Wunsch fast jede Seite mit, die der Nutzer aufruft – mit Ausnahme jener, für die ein Kennwort nötig ist. Indexiert werden also ausschließlich öffentlich zugängliche Daten. Wer möchte, kann auch einen eigenen „Crawler“ auf den Weg schicken. Der klappert dann von einer beliebig einstellbaren Internet-Seite aus alle Links ab, durchsucht die darin vorgefundenen Informationen in einer einstellbaren Tiefe und verschlagewertet sie. Das entstehende Inhaltsverzeichnis (der Index) ist am Privat-PC durchsuchbar – so wird nach der Eingabe der besonderen Adresse <http://localhost:8080> eine Suchmaschine auf dem eigenen PC aufgerufen.

Das wäre an sich noch nichts Besonderes, und die Suchmaske hat im derzeitigen Entwicklungsstadium auch noch keinen herausragenden Komfort. Eine weitweite unabhängige, von keinem Konzern kontrollierte Suchmaschine dürfte aber dadurch entstehen, dass die Nutzer ihre privaten Indizes anonym freigeben und unter-

einander austauschen können. Und das erlaubt Yacy schon in der jetzt verfügbaren Vorversion 0.3 filigran.

Mit den richtigen Einstellungen bekommt man an seinem privaten PC über die persönliche Suchmaschine auch Zugriff auf die Indizes anderer Internet-Surfer. Man kann auf Grund besonderer Mechanismen nicht erkennen, wer im einzelnen welche Seiten besucht hat – wohl aber, was die Seiten gezeigt haben. Umgekehrt erhalten andere Surfer Zugriff auf die Indizes selbst besuchter Seiten. Und mit Hilfe einer Sperre werden nur die Seiten verteilt, die keine persönlichen Informationen enthalten, wie etwa Nutzerkennungen und Passwörter.

Der Austausch dieser Informationen geschieht im Hintergrund direkt zwischen den Yacy-Teilnehmern. Es gibt keinen zentralen Server. Lediglich beim ersten Einstieg in das Yacy-Netz werden ein paar altbekannte zentrale Server kontaktiert, unter anderem die der Entwickler. Diese übermitteln zunächst nur die Adressen anderer gerade aktiver Yacy-Surfer. Und beim nächsten Aufruf werden nicht mehr die Server der Entwickler befragt, sondern die bekannten benachbarten PCs anderer Surfer von der letzten Sitzung.

Nach und nach entsteht so ein mächtiger dezentraler Datenverbund von Informationen über die besuchten Webseiten aller Yacy-Teilnehmer. Je mehr Leute daran teilnehmen, desto besser wird das Inhaltsverzeichnis.

„Noch gibt es ein Henne-Ei-Problem“, sagt Michael Christen, der Kopf hinter Yacy, ein Diplom-Informatiker aus Frankfurt/Main. Mit derzeit noch wenig Teilnehmern ist der gemeinsame Index klein, und bei einem so kleinen Index gebe es noch wenig passende Treffer bei der Stichwortsuche und daher auch zu geringe Akzeptanz. Ohnehin ist gerade erst die Version 0.3 der Software fertig geworden, eine Version für Auskenner. Die Fassung erlaubt bereits eine zuverlässige Suchen-Funktion über die selbst aufgerufenen Seiten, den Austausch mit den Indizes anderer Internet-Surfer und erstmals auch, einen eigenen „Crawler“ auf den Weg zu schicken.

Wie die Erfahrungen mit Tauschbörsen wie Kazaa eDonkey gezeigt haben, ist solch ein Tauschverbund kaum mehr zu stoppen, wenn er einmal eine bestimmte Schwelle von mehreren hundert Teilnehmern in aller Welt erreicht hat. Irgendwo sind immer gerade genügend Leute online, die einen Teil des Index bereitstellen. So haben die Musik- und die Filmindustrie ihre Mühe, das häufig illegale Treiben in dem dezentralen Netzwerk Kazaa zu stoppen.

Bei den Informationen, die das Yacy-Netzwerk sammelt, geht es dagegen nicht um kommerzielle Medien, sondern um Textinhalte von Internet-Seiten. Ähnlich wie Google & Co. ihre gesammelten Informationen in Datenbanken hinterlegen, agiert das verteilte Yacy-Netz.

Dabei hat Yacy durchaus hehre Motive: die Informationsfreiheit und die Unabhän-

gigkeit von herausragend bedeutenden Suchmaschinen. Es soll unter anderem zur alternativen Auffindstation für zensierte Inhalte werden. „Die Nutzer von Yacy dürften idealistisch orientiert sein“, sagt Michael Christen. „Dabei geht es um Meinungsfreiheit, Individualität und die Verfügbarkeit von Wissen und Information.“ Ein Beispiel: Eine zentrale Suchmaschine wie Google könnte derzeit durchaus bestimmte Internet-Seiten mit missliebigen Informationen herausfiltern – so wie nach einem Bericht von Heise.de kürzlich bei Google China gesehen. Im Yacy-Netz wäre das nicht möglich: Sobald einer der Teilnehmer die Seite besucht und freigegeben hat, wandern die Informationen in den gemeinsamen Index und werden umgehend auf andere Teilnehmer verteilt.

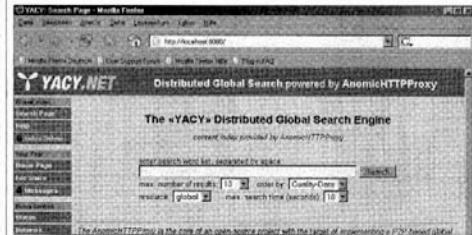
Experten halten solch ein Modell für Suchmaschinen für einen sinnvollen Ansatz. Bereits seit dem Jahr 2000 habe es solche Versuche gegeben, berichtet Wolfgang Sander-Beuermann, Geschäftsführer des SuMa-eV, einem Verein für die Entwicklung alternativer Suchmaschinen. „Das Yacy-Projekt ist am weitesten fortgeschritten“, urteilt der Professor. An der Universität Hannover betreibt er die Suchmaschine Metager.de; eine der letzten Suchmaschinen europaweit, die den US-Konkurrenten Google, Yahoo und MSN nennenswerte Zugriffszahlen abrotzt.

Dass ein solches verteiltes System für Informationen aus dem Internet konzernunabhängig betrieben werden kann, belegt nach seinen Worten das Usenet. Das war vor allem in den achtziger und neunziger Jahren ein an Universitäten betriebener, dezentraler Nachrichtenaustauschdienst. Allerdings hat es inzwischen an Bedeutung verloren. „Dem Otto-Normal-User erscheint Google immer noch gut genug. Das, was er mit Google nicht findet, sieht er ja auch nicht, und er kann sich dann auch nicht darüber ärgern, dass er es nicht findet“, sagt Sander-Beuermann.

Die wichtigste Eingabemaske zum Recherchieren im Usenet www.dejanews.com, hat 2001 ein amerikanischer Konzern gekauft. Seither können Nutzer zwar weiterhin alternative Programme installieren und so unabhängig von diesem Konzern im Usenet surfen. Komfortabler ist aber die Suchmaske des Konzerns.

Ironie der Geschichte: Dieser amerikanische Konzern war Google. Mit Yacy wäre das wohl nicht möglich.

MARCUS SCHWARZE



Die Suchmaschine Yacy liegt auf der eigenen Festplatte und ist im Internet-Browser unter der Adresse <http://localhost:8080> erreichbar. Ihre Daten bezieht sie aus den eigenen Surfsitzungen und denen anderer Nutzer.

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Januar 2005
c't 02/2005,
Seite 40,
Rubrik "aktuell"

"Suchmaschine
sucht Tauschpartner"

Suchmaschine sucht Tauschpartner

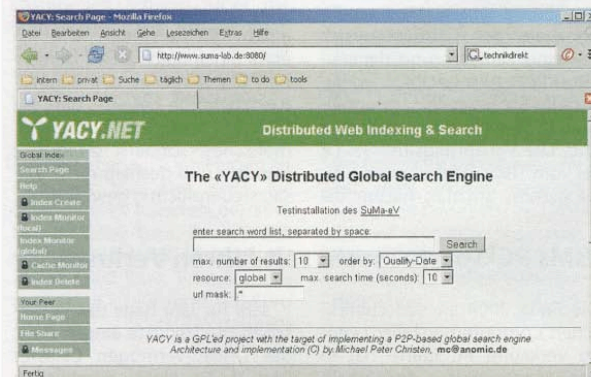
Inhalte von Suchmaschinen sollen nicht zensierbar sein. Darin sieht Entwickler Michael Christen die wertvollste Eigenschaft seines Projekts Yacy (<http://yacy.net>). Yacy ist ein P2P-artiger Verbund von Rechnern, die eine gemeinsame Suchmaschine bilden. Wie bei Dateitauschbörsen tragen alle Peers Suchergebnisse bei und nutzen die Ergebnisse der anderen.

Basis des Yacy-Netzes ist die als Open Source verfügbare Software, die sowohl als Client für die Suchmaschine als auch als Proxy-Server agiert. Der Proxy baut aus den Seiten, die seine Benutzer über ihn abrufen, einen Index auf. Stellt ein Benutzer eine Suchanfrage, so leitet der Proxy die Anfrage an alle Peers weiter, sammelt die Ergebnisse, bereitet sie auf und prä-

sentiert sie. Unter www.suma-lab.de:8080 lässt sich die Suchmaschine ausprobieren.

Die Qualität einer solchen Suchmaschine steht und fällt mit der Anzahl der Peers. Derzeit gehört nur ungefähr ein halbes Dutzend Rechner zum Verbund. Christen erhofft sich daher, dass sich noch mehr Mitstreiter an dem Projekt beteiligen. Die Proxy-Software ist in Java geschrieben und läuft unter Windows (NT, 2000, XP), Mac OS X und Linux. Die Projekt-Website enthält eine ausführliche Installationsanleitung. Auch für die Weiterentwicklung der Software wünscht sich Christen weitere Mitstreiter. Die Peers sollen zukünftig nicht nur Kontaktinformationen austauschen, sondern auch Indizes.

(Christiane Schulzki-Haddouti/jo)



Yacy lässt sich bereits online testen.



Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Februar 2005:
Eröffnung des
YaCy-Forums

innerhalb 8 Monate:
- über 10000 Artikel
- über 130 Teilnehmer

www.yacy-forum.de :: Index
http://www.yacy-forum.de/

www.yacy-forum.de
Das offizielle Forum zur Peer-to-Peer Suchmaschine YaCy.

Impressum | FAQ | Suchen | Mitgliederliste | Service
Profil | Du hast keine neuen Nachrichten | Logout [Orbiter]

Dein letzter Besuch war am: Do Okt 13, 2005 3:41 pm
Aktuelles Datum und Uhrzeit: Do Okt 13, 2005 6:41 pm
www.yacy-forum.de Foren-Übersicht

Neue und/oder editierte Beiträge anzeigen (129)
Keine permanent als gelesen markierte Themen
Beiträge seit dem letzten Besuch anzeigen
Eigene Beiträge anzeigen
Unbeantwortete Beiträge anzeigen

Forum	Themen	Beiträge	Letzter Beitrag
Allgemein			
Benutzung und Probleme Probleme mit der Benutzung von YaCy oder der Installation bitte hier posten... (Anonymer Post erlaubt) Moderator: Orbiter	283	2881	Do Okt 13, 2005 6:33 pm Huppi →
Anleitungen _Fertige_ Tipps und Anleitungen bitte hier posten. Moderator: Orbiter	27	162	Mo Okt 10, 2005 12:07 pm Allo →
Alles andere... ... bitte hier hinein. Alles was in keine andere Kategorie passt... (Anonymer Post erlaubt) Moderator: Orbiter	242	2414	Do Okt 13, 2005 6:19 pm roland →
Public Relations Wo findet sich ein Verweis auf YaCy in den Medien? Wie koordinieren wir uns möglichst effizient, um medienwirksam aufzutreten? Das klären wir hier... Moderator: Orbiter	52	339	Do Okt 13, 2005 4:35 am bischer →
Technisches			
Bugs Wenn sie meinen einen Bug gefunden zu haben, posten sie ihn hier, damit anderer Benutzer ihn nachvollziehen können... (Anonymer Post erlaubt) Moderator: Orbiter	331	2322	Do Okt 13, 2005 3:37 pm theli →
Wunschliste Wünsche oder Vorschläge für kommende YaCy-Versionen bitte hier posten, damit sie diskutiert werden können. (Anonymer Post erlaubt) Moderator: Orbiter	184	1612	Do Okt 13, 2005 6:10 pm Gast →
Coding Helfen Sie dem YaCy Projekt. Fertiger Quellcode oder Kommentare zum Quellcode sind hier erwünscht... Moderator: Orbiter	77	766	Fr Okt 07, 2005 5:19 pm Allo →
Drumherum...			
Off-Topic Alles, was nicht zu YaCy gehört, aber interessant sein könnte, kommt hier hin. Moderator: Orbiter	63	436	Do Okt 13, 2005 5:42 pm Allo →
Abstimmungen Wenn die Meinungen anderer User gefragt sind, ist hier der richtige Ort zum posten... Moderator: Orbiter	4	38	Fr Jul 01, 2005 5:21 pm Allo →
Mitteilungen Downtimes und andere Dinge dieses Board betreffend... Moderator: Orbiter	15	85	Di Sep 27, 2005 3:30 pm Allo →

Alle Foren als gelesen markieren

Alle Zeiten sind GMT + 1 Stunde

Wer ist online?

Unsere Benutzer haben insgesamt **11099** Beiträge geschrieben.
Wir haben **139** registrierte Benutzer.
Der neueste Benutzer ist [gnosis](#).

Insgesamt sind **6** Benutzer online: 4 registrierte, ein versteckter und ein Gast. [Administrator] [Moderator]
Der Rekord liegt bei **22** Benutzern am Do Sep 29, 2005 4:15 pm.
Registrierte Benutzer: [trans](#), [Huppi](#), [lulebad](#), [Orbiter](#)

Diese Daten zeigen an, wer in den letzten 5 Minuten online war.

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Februar 2005: News über YaCy auf golem.de

nur eine von vielen Berichten auf Web-Nachrichtenseiten und Blogs

YaCy - Suchmaschine auf Peer-to-Peer-Basis - Golem.de

<http://www.golem.de/0502/36260.html>

The screenshot shows the golem.de website interface. At the top, there are navigation links for FORUM, ARCHIV, NEWS-FEEDS, MARKT, SIDEREAL, WERBUNG, and IMPRESSUM. Below this is a search bar with the text 'Suchen' and a search button. The main content area features a headline: 'YaCy - Suchmaschine auf Peer-to-Peer-Basis' and a sub-headline: 'Prefetch-Funktion nimmt Umfeld von Webseiten in Suchindex auf'. The article text discusses the YaCy search engine's prefetch function, which indexes the surrounding environment of websites. It mentions that YaCy is a Java-based search engine that uses a local proxy server. The article also notes that YaCy is not dependent on individual providers like Google or Yahoo. Other elements visible in the screenshot include a sidebar with 'STELLENMARKT' and 'LETZTE MELDUNGEN', and a footer with copyright information and navigation links.

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Februar 2005:
Kommentar in der Technology-
Review:

Wolfgang Sander-Beuermann:
“YaCy ist meines Wissens das
weltweit einzige Proof-of-
Concept einer dezentralen Peer-
to-peer Suchmaschine”

PERSPEKTIVE 29

→ KOMMENTAR
Schürfrechte im Informationszeitalter Google hin,
Microsoft her – das Internet braucht eine freie Suchkultur

Wenn der Rohstoff des 21. Jahrhunderts wirklich Information ist, dann unterscheidet er sich grundlegend von seinen Vorgängern Eisenerz und Erdöl: Er verbraucht sich nicht, kann endlos wiederverwertet werden, ist einfach transportierbar und kann sich sogar durch Neuzusammensetzung vermehren.

Letztere Eigenschaft, so schön sie zunächst scheint, wird allmählich zur Plage. Die Menge der weltweit vorliegenden Information wächst seit Jahrhunderten stetig. Laut einem Bericht der University of California in Berkeley schafft die Menschheit derzeit ein bis zwei Exabyte (Milliarden Gigabyte) an Information pro Jahr. Dargestellt als Text entspricht das einem Bücherstapel, dessen Höhe etwa die 130fache Entfernung Erde-Mond erreichen würde. Die große Herausforderung ist, aus solch gigantischen Informationsmengen das subjektiv Wesentliche – also das Wissen – herauszusuchen.

Die Wissensextraktion wird im digitalen Zeitalter zunehmend von Internet-Suchmaschinen übernommen. Sie verarbeiten den Rohstoff Information zu Wissen, kontrollieren und verteilen ihn. Es kann keinem Nutzer ganz geheuer sein, dass diese Schlüsselfunktion der Informationsgesellschaft in die Hände weniger Konzerne gerät: Google hat mit einem Marktanteil von mehr als 80 Prozent in Deutschland ein De-facto-Monopol erreicht, das nun Microsoft mit seiner MSN Search angreifen will. Aber diese Alternative weckt schwerlich mehr Vertrauen.

Suchmaschinen-Monopolisten können bestimmen oder kontrollieren, welche Information wann und auf welchen Rechnern verfügbar ist und in welcher Reihenfolge die Ergebnisse angezeigt werden. Durch Beobachtung der Abrufe können sie genaue Profile ihrer Nutzer erstellen. Um die Vormacht der kommerziellen Wissenswächter zu brechen, bedarf es einer freien Suchkultur – so wie das offene Betriebssystem Linux die Welt vor einer Windows-Monokultur bewahrt hat.

Immerhin scheint man auch auf staatlicher Seite das Problem des Information Overkill erkannt zu haben. Die öffentliche Hand fördert zahlreiche Projekte, die Ordnung in den Datenwust bringen wollen. Doch die meisten davon sind mehr visionär als realistisch. Vom einst so gefeierten „Semantic Web“ etwa ist auch nach Jahren kaum Handfestes zu sehen. Kein Wunder: Solche Vorhaben setzen voraus, dass die Daten zunächst eingesammelt und suchgerecht indiziert werden. Mangels freier Software fehlt diese Voraussetzung.

Was also ist nötig, um im Informationszeitalter die freie Verfügbarkeit der Ressourcen sicherzustellen? Die Antwort ist die gleiche wie einst für Stahl und Öl: eine Vielfalt von Anbietern. Der beste Weg dorthin führt über freie Suchmaschinen-Software, auf welche die Betreiber solcher Maschinen zurückgreifen können. Dann entsteht ganz von selbst ein offener Wettbewerb.

Doch freie Suchmaschinen-Software ist rar. Es gibt Ansätze dazu in Russland und ein einziges Projekt in den USA (www.nutch.org). Auch Europa ist weitgehend Odnis – bis auf den Lichtblick Yacy, ein Vorhaben des Frankfurter Softwarespezialisten Michael Christen. Yacy ist meines Wissens das weltweit einzige Proof-of-Concept einer dezentralen Peer-to-peer-Suchmaschine.

Um die Suchmaschinen-Landschaft zu beleben, haben nun 13 Forscher, Politiker und Unternehmer den „Gemeinnützigen Verein zur Förderung der Suchmaschinen-Technologie und des freien Wissenszugangs“ (kurz: SuMa-eV, www.suma-ev.de) mit Sitz in Hannover gegründet. Zu den Gründungsmitgliedern gehören der MP3-Erfinder Karlheinz Brandenburg, der Vizepräsident für Forschung der Universität Hannover Wolfgang Ermer und ich selbst. Ziel des SuMa-eV ist die Etablierung einer auf möglichst viele autarke Systeme verteilten Suchmaschinen-Infrastruktur, die von ihrem Bauprinzip her kaum monopolisierbar ist. Der Kerngedanke dieser Struktur, die sich aus sehr vielen und sehr unterschiedlichen Bausteinen zusammensetzen kann, liegt in der Autarkie der Einzelsysteme: gesellschaftlicher Pluralismus wird netztopologisch abgebildet. Eigentlich wäre es im Interesse und in der Macht des Staats, die Meinungsvielfalt im Netz besser zu sichern. Während er – abgesehen von ein paar hellhörigen Parlamentariern – noch träumerische Visionen pflegt, müssen Initiativen wie SuMa-eV einspringen. ■

DER BESTE WEG ZU FREIEM ZUGANG ZU WISSEN IST EINE DEZENTRALE INFRASTRUKTUR FÜR DIE SUCHE

Das M.I.T.-MAGAZIN FÜR INNOVATION
TECHNOLOGY
REVIEW
MÄRKTE VON MORGEN
FOKUS: Werbefolle mit Verstand und Gefühl
EIN BISSCHEN BÖSE
Warum Börsenstür Google werden muss wie Microsoft, um zu überleben
KREBSMEDIZIN: DER STREIT UM DIE STRAHLENTHERAPIE
INTELLIGENZ: VON ROBOTERN DAS DENKEN LERNEN
NEUROLOGIE: DER UNHEIMLICHE DR. HUANG HONGYUN

DR. WOLFGANG SANDER-BEUERMANN
MANGELFÜRLEITER DER
SUCHMASCHINEN-ANWISSE AM
INSTITUT FÜR WISSENSINFORMATIK
DER UNIVERSITÄT HANNOVER

TECHNOLOGY REVIEW Februar 2005

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

März 2005: Artikel des YaCy-Autors in der Datenschleuder #86 über Architekturdetails und Zielsetzungen einer freien Suchmaschinentechnik



GOOGLE@HOME

YaCy – Peer-to-Peer Web-Suchmaschine

Michael Christen <mc@anomic.de>

Information ist im Web eine stark kontrollierte Ressource – Portale, Suchmaschinen und das DNS sind zentral in de-facto-Monopolen organisiert und bestimmen, welche Daten verfügbar sind. Mit YaCy wird die Kontrolle wieder an die Nutzer zurückgegeben.

Das YaCy-Projekt wurde Ende 2003 mit dem Ziel gestartet, eine freie, unabhängige und nicht zensurierbare PaP-basierte Web-Suchmaschine zu erstellen. Zu diesem Zeitpunkt existierten viele gut funktionierende PaP-Filesharingtechniken und mehrere freie Implementierungen von Crawlern/Indizierern, aber keine Technik, die PaP mit Suchmaschinentechnik verbindet. Wir stellen hier die über die Funktion einer Suchmaschine hinausgehenden Eigenschaften und die Architektur von YaCy vor.

PaP-Suchmaschine und caching http proxy: Für die Existenz des Proxy in YaCy gibt es drei Gründe:

Eine PaP-basierte Software ist auf lange Onlinelaufzeit angewiesen. YaCy soll nicht nur laufen, während der User eine Suche absendet. Synergie: Durch die Nutzung des Proxy als Mehrwert wird eine lange Onlinelaufzeit erreicht. Der Proxy ist ein 'Information Provider' für den Indexer.

Synergie: die Suchmaschine erhält als Option 'kostenlos' Web-Seiten zum Indizieren ohne Crawling.

Der Proxy bietet mit seinen eingebauten Filtern einen persönlichen Schutz vor ungewollten Webinhalten: das ist der notwendige 'Gegenpol' zur Zensurfreiheit durch die Suchmaschine. Die Idee ist, dass jeder sich (seine Familie, sein Unternehmen, etc.) wieder soweit selbst zensurieren kann, wie er es ggf. von einem Suchmaschinenbetreiber erwarten würde (wenn er denn diese Erwartung hätte).

Synergie: populäre Filter (bspw. Bannerblocker) können von Peer zu Peer importiert werden.

Crawling und Prefetching: während Crawling eine typische Aufgabe einer Indizierungssoftware ist, liefert ein proxy prefetch ggf. schnellere Zugriffszeiten für den Benutzer des Proxies.

Information Provider, Indexer, Datenbank, Suche, HTTP Proxy Cache, verteilbarer Crawler, Remote crawler, RWI Generierung, Daten-basierte AKA-Name, DHT, Peer-to-Peer - basierende DHT-Kontrolle, HTTP Server, User Interface, YaCy Protocol

Suchmaschine mit Mehrwert

YaCy ist nicht nur eine Suchmaschine mit Crawler und Suchfunktion, sondern auch ein Web-Server, ein caching http proxy mit optionalem prefetch, eine DNS-Erweiterung, ein Messagingsystem und ein Wiki. Warum das Ganze? Es gibt einen gemeinsamen Schlüssel für all diese Funktionen: Synergien zwischen Suchtechnik und der Lösung von zentral-gesteuerten Diensten im Internet. Im Detail:

54 die datenschleuder. #86 / 2005

lokal gestarteten Crawl aus. Der Peer erhält durch einen anderen Peer (der lokal einen Crawl ausführt) eine URL zum Indizieren zugewiesen und führt damit einen Remote Crawl aus. Der Peer bearbeitet eine lokal angestoßene Suchanfrage und fördert selektiv von anderen Proxies RWI-Fragmente ein, die anschließend permanent in der lokalen Datenbank bleiben. Der Peer erhält von anderen Proxies RWI-Fragmente zugewiesen, weil er für diese RWIs eine besse-

speichern sollen. Das funktioniert, weil RWIs ja im Vorfeld einer Suche bereits zu dem Peer, wo sie bei einer Suche erwartet werden, gewandert sind.

Mit steigender Peerzahl kann sich die DHT in dem wachsenden Netz weiter spezialisieren. Die Suchzeit für ein Wort steigt aber theoretisch nicht mit der Größe des Netzes, da der RWI ja immer nur an einer bestimmten Position

GOOGLE@HOME

Synergie: beide Funktionen benutzen prinzipiell den gleichen Algorithmus.

Eingebauter Webserver, Filesharing, Wiki & Messaging: wer durch die dezentrale Struktur der YaCy-Suche Informationsfreiheit sicherstellen möchte, will ggf. auch ein Publikationsmedium nutzen, das in gleichem Maße keiner Zensur unterliegt. Webinhalte können zwar durch YaCy ad-hoc erfasst werden, aber selbstgestellte Daten könnten weiterhin auf fremden Servern gesperrt oder entfernt werden.

Synergie: unzensurierte Suchergebnisse sind erst dann sinnvoll, wenn deren Ressource auch geladen werden kann. YaCy gibt dazu einige Basiswerkzeuge an die Hand. Da der dazugehörige Server dann dem Nutzer gehört, kann er nicht zensuriert werden.

Webserver, Suchinterface und Proxy: die natürlichste Umgebung für eine Webseite ist eine Webseite. Daher besteht YaCys GUI aus einem integrierten Webserver mit Servlet-Engine.

Synergie: der Proxy, das GUI und die eigenen Webinhalte (siehe oben) können den gleichen eingebauten httpd benutzen.

DNS-Umgebung und Erweiterung um die Top-Level-Domain „yacy“: Das DNS-System ist mit seiner zentral-hierarchischen Struktur ein einfacher Angriffspunkt für Zensuristen. YaCy bietet jedem Peerbetreiber seine eigene „peername-yacy“-Domain, die automatisch durch den Proxy zum YaCy-Webserver des Peerbetreibers aufgelöst wird.

Synergie: die Nutzung des Proxies macht den Eingriff in die DNS-Auflösung möglich und die YaCy PaP-Verwaltung stellt ganz selbstverständlich eine Peer-zu-IP-Datenbank dar, was auch mit dynamisch zugewiesenen IPs funktioniert. Es existiert außerdem ein schlüssiges Konzept, um ohne zentrale Datenbank einen Namensdiebstahl der yacy-Domains unterbinden zu können.

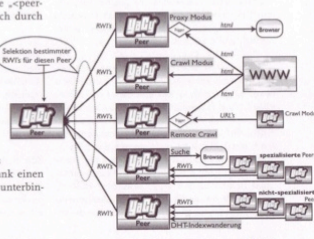
Zwar realisiert der eingebaute Proxy YaCys zentrale Konzepte, aber die Software kann auch betrieben werden, ohne den Proxy nutzen zu müssen. Dann dient YaCy „nur“ als Suchinterface, Crawler, Webserver etc.

Indizierung und Peearchitektur

YaCy besteht aus einem Crawler, einem Indexer, einer Datenbank, einem Suchinterface und der PaP-Organisation:

- Wir haben vom Crawler abstrahiert und sehen konzeptionell einen 'Information Provider' vor. Als solcher kann ein Crawler eingesetzt werden. Es ist aber auch möglich, Seiten aus dem integrierten Proxy Cache als Input für den Indexer zu benutzen.
- Der Indexer erzeugt konzeptionell einen Reverse Word Index (RWI), d.h. zu jedem Wort eine Liste der URLs plus Rankinginformationen. In unserer Implementation des RWI werden jedoch keine Wörter im Klartext gespeichert, sondern lediglich Wordhashes. Die Hashes können allerdings nicht wieder zurück in Wörter übersetzt werden. Das ist auch nicht notwendig und so können keine Klartextfragmente der ursprünglichen Webseiten auf den Rechnern der Peerbetreiber gefunden werden. Diese können daher auch nicht zur Verantwortung für die bei ihm/ihr gelagerten Wörter gezogen werden.

55 die datenschleuder. #86 / 2005



samenheit der YaCy-User auf bestimmte Webseiten. Dadurch wird die Indizierung auf populäre und für die Nutzer interessante Seiten fokussiert.

Bereits gefundene Webseiten werden vom suchenden Peer aus wieder per DHT wandern verteilt. Dadurch verstärkt sich die Präsenz interessanter Suchergebnisse. Dies kann als implizite Moderation des globalen Index

durch die Peersur angelesen werden.

Inhalte können nicht global zensuriert werden.

Es existieren Konzepte um die Peerszahl stark zu vergrößern

YaCy könnte für folgende Bereiche interessant werden

Als Betriebssoftware für Internetcafes: Der caching Proxy ist nutzbringend zur Bandbreitenbegrenzung, und eine geplante Abrechnungsfunktion mit Clientkontrolle bringt einen zusätzlichen Mehrwert. Einsatzgebiet wäre kland, eher weltweit. Die Interessiertheit in vielen Regionen fast bei Internetcafes und die Abreun auf Papier ohne Software

weiterung: YaCy könnte auch solche auslesen, die Proxyfunktionsfähig, aber alle anderen Vorzüge der Nutzung von YaCy verlieren. Open-Source-Konzepte und Firefox könnten in ihre Distribution aufnehmen, er Browser-Nutzer automatisch Kontrolle über Webseite und einsamen Suchindex.

YaCy ist Open-Source (GPL-Lizenz), kompakt (rund 1 MB Download), portabel (java), leicht zu installieren (nur unpacken, keine DB aufsetzen) und einfach zu betreiben. Wer mitmacht, der arbeitet aktiv an der Sicherstellung der Informationsfreiheit mit.

Links:
<http://www.yacy.net/> – YaCy-Homepage
<http://www.soma-lab.de/yacy> – deutsche Doku

Michael Christen
 +49 177 6424235

http://www.yacy.net/yacy
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

April 2005:
Zahlreiche Nachfragen zur
Mitarbeit an YaCy im YaCy-
Forum

Eröffnung des SVN-Archives (ein
Source Code Management
System) des YaCy-Sourcecodes
zur kooperativen
Softwareentwicklung auf
berlios.de

4 Monate später enthält das
SVN-Archiv bereits über 900
Codeänderungen von nur 5
Kernentwicklern.

BerliOS Developer: Project Filelist
http://developer.berlios.de/project/showfiles.php?group_id=3505

BerliOS : SourceBiz · SourceLines · DevCounter · DocsWell · SourceAgency · Partners · Contact Us

berlios Developer

Fostering Open Source Development

Project: YaCy P2P Web-Search - Files

Summary | Home Page | Forums | Bugs | Support | Features | Patches | Lists | Tasks | Docs | Surveys | News | SVN | **Files** | Screenshots | Wiki | Memberlist | Admin |

Below is a list of all files of the project. Before downloading, you may want to read Release Notes and Changelog (accessible by clicking on release version).

Package	Release & Notes	Filename	Size	D/L	Arch.	Date	Type
main release (generic)							
0.41						2005-10-04 00:00	
yacy_v0.41_20051004_848.exe			1785554	61	i386	.exe (32-bit Windows)	
yacy_v0.41_20051004_848.tar.gz			1996224	13	i386	.gz	
0.40						2005-08-16 19:11	
yacy_v0.40_20050816_547.exe			1612560	312	i386	.exe (32-bit Windows)	
yacy_v0.40_20050816_547.tar.gz			1782001	213	Any	.gz	
0.394						2005-08-03 21:50	
yacy_dev_v0.394_20050803_489.tar.gz			1762155	217	Any	.gz	
yacy_v0.394_20050803_489.exe			1600279	325	i386	.exe (32-bit Windows)	
0.391						2005-07-28 00:00	
yacy_v0.391_20050726_434.exe			1571275	208	i386	.exe (32-bit Windows)	
yacy_v0.391_20050726_434.tar.gz			1727855	90	Any	.gz	
0.39						2005-07-22 19:23	
yacy_v0.39_20050722_425.exe			1569751	154	i386	.exe (32-bit Windows)	
yacy_v0.39_20050722_425.tar.gz			1725500	124	Any	.gz	
0.38						2005-06-03 15:48	
yacy_v0.38_20050603_208.exe			1365865	1596	i386	.exe (32-bit Windows)	
yacy_v0.38_20050603_208.tar.gz			1476975	81	Any	.gz	
0.37						2005-05-02 00:00	
yacy_v0.37_20050502.exe			1389991	116	i386	.exe (32-bit Windows)	
yacy_v0.37_20050502.tar.gz			1490357	18	Any	.gz	
0.36						2005-04-09 00:00	
yacy_v0.36_20050326.exe			841370	0	i386	.exe (32-bit Windows)	
yacy_v0.36_20050326.tar.gz			949953	6	Any	.gz	
Project Totals:		8	16	24647665	3534		

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Mai 2005:
c't 11/2005,
Seite 60,
Rubrik "aktuell"

"Metager2 ist am Netz",
ein Bericht über die SuMa-eV -
Tagung bei der YaCy als Konzept
zur nicht-Zensierbarkeit
vorgestellt wurde

(siehe Präsentation: "YaCy -
Nicht-Monopolisierbarkeit von
Suchmaschinen")

aktuell | Suchmaschinen, Domains

Metager2 ist am Netz

Beim zweiten öffentlichen Forum des SuMa-eV (gemeinnütziger Verein zur Förderung der Suchmaschinen-Technologie und des freien Wissenszugangs e.V. [1]) wurden einige neue Entwicklungen der Suchmaschinenteknik vorgestellt. Als ein Highlight stellte der Schüler Markus Franz seine Metasuchmaschine Metager2 vor. Franz war auf eine Arbeit von Wolfgang Sander-Beuermann gestoßen, dem Vater der Metasuchmaschine MetaGer [2] und SuMA-eV-Geschäftsführer. Sander-Beuermann hatte bereits vor Jahren eine Weiterentwicklung von MetaGer skizziert, die nicht nur die in den Ergebnislisten der zugrunde liegenden Suchmaschinen enthaltene Informationen auswertet, sondern auch die Originaldokumente herunterlädt, analysiert und damit ein besseres Ranking erzeugen kann.

Seinerzeit ließ sich Metager2 nicht verwirklichen: Es dauerte einfach zu lange, alle Trefferdokumente herunterzuladen. Franz hat dieses Problem dank einer guten Netzwerkanbindung und der parallelen Anforderung der Dokumente im Griff [3]. Er will den Python-Quelltext seiner Suchmaschine unter einer offenen Lizenz zugänglich machen. Thorsten Blancke und Martin Gutschke haben sich mit Romso der Nachrichtenrecherche verschrieben. Ihr Suchdienst Romso bietet einige interessante neue Funktionen [4]. So gibt der Dienst zu jedem Treffer themenbezogene Schlagwörter aus, die dem Benutzer einen schnellen Überblick geben und die er zur Weitersuche nutzen kann. Zu jeder Abfra-


ge generiert Romso eine Zeitreihe, die den zeitlichen Verlauf der Medienpräsenz wiedergibt. Mit stündlich errechneten Mediencharts visualisiert Romso aktuelle Entwicklungen einzelner Themen, derzeit die Präsenz der Parteien und der DAX-30-Unternehmen in den Medien.

YaCy ist ein P2P-artiger Verbund von Rechnern, die eine gemeinsame Suchmaschine bilden. Wie bei Dateiaustauschbörsen tragen alle Peers Suchergebnisse bei und nutzen die Ergebnisse der anderen. Als ein wichtiger Vorteil im Vergleich zu gängigen Client/Server-Suchmaschinen sind die Inhalte von YaCy nicht zensierbar, so Entwickler Michael Christen.

Basis des YaCy-Netzes ist die als Open Source verfügbare Software, die sowohl als Client für die Suchmaschine als auch als Proxy-Server agiert. Der Proxy baut aus den Seiten, die seine Benutzer über ihn abrufen, einen Index auf. Außerdem durchsucht ein Crawler selbstständig nach neuen Seiten. Stellt ein Benutzer eine Suchanfrage, so leitet der Proxy die Anfrage an alle Peers weiter, sammelt die Ergebnisse, bereitet sie auf und präsentiert sie. Im Labor des SuMA-eV steht ein Peer, mit dem sich die Suchmaschine ausprobieren lässt [5]. (ja)

Literatur

- [1] Homepage SuMa-eV: <http://suma-ev.de>
- [2] MetaGer: <http://meta.rzn.uni-hannover.de>
- [3] Metager2: www.metager2.de
- [4] Romso: www.romso.de#8080
- [5] YaCy beim SuMa-eV: www.suma-lab.de#8080



Mit aktuellen Charts dokumentiert Romso die Entwicklung einzelner Themen, etwa die Medienpräsenz von Parteien.

60



<http://www.yacy.net/yacy>
mc@yacy.net

Michael Christen
+49 177 6424235



YaCy in der Presse und PR-Aktivitäten

Juni 2005:
c't 16/2005,
Seite 52,
Rubrik "aktuell"

YaCy wird ein ganzer Artikel über mehr als eine Seite gewidmet
"Wider die Monokultur"

Michael Christen
+49 177 6424235

aktuell | Web-Suchmaschinen

Richard Sietmann

Wider die Monokultur

P2P-Strategien gegen die Suchmaschinen-Monopolisierung

Überträgt man das Quasi-Monopol des Suchmaschinenbetreibers Google auf den Print-Bereich, entspräche die Situation im Web einer Welt, in der es neben ein paar lokalen Anzeigenblättern nur noch eine Zeitung gäbe. Nicht nur Meta-Suchmaschinen, auch ganz neue Ansätze versuchen sich an Abhilfen.

Steilt Google die ultimative Lösung dar oder inzwischen ein Problem? Für vier von fünf Websurfern in Deutschland ist die Antwort klar: Für sie ist Google die erste Anlaufstelle zum Aufspüren von Informationen im Internet und „google“ gleichbedeutend mit recherchieren. Die beiden anderen großen Suchmaschinenbetreiber, MSN Web-Suche und Yahoo, bringen es in der Online-Statistik von webstats.de [1] gerade mal auf 4,3 beziehungsweise 4,4 Prozent; AOL (2,3%) und T-Online (1,6%), reichen Suchanfragen an Google durch und erhöhen so den Marktanteil des Platzhirschen auf 85 Prozent. Alle anderen Recherche-Helfer hingegen, ob Web.de, MetaGer oder Fireball, bewegen sich in einer Nische.

Genau darin aber liegt auch ein Problem, denn Suchmaschinen lenken die Aufmerksamkeit auf Inhalte, „in keinem anderen Medium“, meint der Leiter des Suchmaschinenlabors am Regionalen Rechenzentrum Niedersachsen der Universität Hannover, Wolfgang Sander-Beuermann, „würde eine solche Machtkonzentration politisch geschildert“ in der Tat gilt bei den Fernsehsendern bereits eine Konzentration von 30 Prozent Zuschaueranteil als Gefahr für die Meinungsvielfalt. Und das Kartellrecht stuft Unternehmen mit einem Marktanteil von 33 Prozent als marktbeherrschend ein.

Im Fall der Suchmaschinen gibt es dagegen keine Regeln. Ein marktbeherrschender Anbieter kann kontrollieren, welche Informationen verfügbar sind und welche nicht, entscheiden in welcher Reihenfolge die Suchergebnisse auf dem Display des Nutzers erscheinen, und nachvollziehen, welche Informationen wann und von welchem Rechner aus abgerufen wurden.

chem Rechner aus abgerufen wurden.

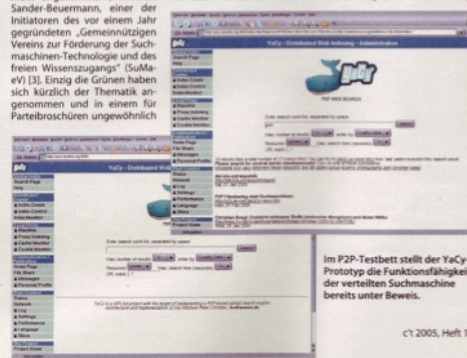
Hinter den Kulissen

Besonders kritisch wird der Umgang mit den Nutzerdaten, wenn die großen Suchmaschinenbetreiber verstärkt auf die personalisierte Suche setzen. Der kleine Startup Blinkx hat bereits die Richtung vorgegeben: Das System beobachtet fortlaufend die Arbeiten des Users am Desktop, liest die geöffneten Dokumente und Webseiten mit und bietet ihm individuell passende Informationen aus dem Web an, ohne dass es dazu einer expliziten Suchanfrage bedarf. Die Internet-Recherchen laufen dabei im Hintergrund als Web-Service über die Rechner der in San Francisco ansässigen Firma, die auf ihrer Homepage beteuert, keine Profile anzulegen und die Nutzer anonym zu behandeln [2].

„Die Risiken sind bei den staatlichen Stellen überhaupt noch nicht angekommen“, kritisiert Sander-Beuermann, einer der Initiatoren des vor einem Jahr gegründeten „Gemeinnützigen Vereins zur Förderung der Suchmaschinen-Technologie und des freien Wissenszugangs“ (SuMa-ev) [3]. Einzig die Grünen haben sich kürzlich der Thematik angenommen und in einem für Parteibroschüren ungewöhnlich

fundierten Papier unter dem Titel „Suchmaschinen: Das Tor zum Netz“ auf die Gefahren hingewiesen, die entstehen, wenn ein übermächtiger Pförtner den Zugang zu der Welt der Information im Internet kanalisiert. Darin fordern sie Transparenz im Page-Ranking sowie Datenschutz-Audits, setzen sich für die alternativen Anbieter ein und möchten die Medienkompetenz der Nutzer stärken [4].

Wie über Appelle an die Nutzer und den Gesetzgeber hinaus konkrete Alternativen aussähen könnten, erläuterte Sander-Beuermann auf einem Kolloquium im Konrad-Zuse-Institut in Berlin. Eine gewisse Abhilfe gegen den Tunnelblick aufs Web bieten bereits Meta-Engines, die Suchanfragen automatisch an mehrere Suchmaschinen weiterleiten und die gefundenen Treffer bündeln.



Im P2P-Testbett stellt der YaCy-Prototyp die Funktionsfähigkeit der verteilten Suchmaschine bereits unter Beweis.

ker Michael Christen entwickelter Ansatz, der erstmals die Suchmaschinen-Technik mit dem Filesharing verbindet: YaCy (Yet another Cyberspace) ist ein P2P-Verbund von Rechnern, die gemeinsam die Suchmaschine bilden [5]. Nach dem Prinzip der Tauschbörsen steuert jeder Nutzer einen lokalen Suchindex für die globalen Suchanfragen bei, und bei der Suche selbst werden immer auch die Indexdaten von anderen gerade aktiven YaCy-Nutzern abgefragt.

Die als Open Source unter der GNU General Public License verfügbare Java-Software des Prototypen – derzeit ist Version 0.38 aktuell – besteht im Wesentlichen aus der Search Engine PLASMA, einem http-Proxy-Server, dem YaCy-P2P-Protokoll und einer Datenbank. Wie eine herkömmliche Suchmaschine kann Plasma über eine http-Schnittstelle als Crawler das Internet nach Webseiten abgraben und zu einem Reverse Word Index (RWI) verschlagworten, der zu jedem erfassten Suchbegriff die zugehörigen Links enthält. Der lokale Proxy-Server kann aber auch die vom Nutzer besuchten Seiten in einer Cache abspeichern, verschlagworten und in den Index integrieren; mit einer Prefetch-Funktion lassen sich mit einstellbarer Tiefe zudem die verlinkten Seiten für andere unsichtbar bleiben, setzt YaCy auf den P2P-Anony-

misierungsmechanismus der Distributed Hash Tables. Genügend Peers vorausgesetzt, lässt die verteilte Erschließung eine weitaus gründlichere und aktuellere Erfassung erwarten, als zentral verwaltete Search Engines sie leisten können. Denn selbst dem Platzhirsch Google mit den nach eigenen Angaben mehr als acht Milliarden Webseiten in seinem Index bleiben Untersuchungen zufolge rund 40 Prozent des Web verborgen [6], weil längst nicht jede Webseite verlinkt ist, sodass sie von Crawlern gefunden werden kann. Dagegen würden auch solche Seiten aus dem Invisibilib Web ans Licht geholt und in YaCy-Index auftauchen, sobald ein Peer sie einmal besucht hat. Aber kann ein Open-Source-Projekt mit P2P-Netzwerkern ernsthaft gegen den Nasdaq-Liebling Google antreten – ist eine Linux und Firefox mit steigendem Erfolg dem Microsoft-Monopol bei Desktop-Betriebssystemen und Internet-Browsern Paroli bieten? Wenn sich dieses P2P-Suchmaschinen-Netz ausbreiten würde, dann könnte es zu einer Alternative zu kommerziellen Suchmaschinen werden, glaubt Sander-Beuermann.

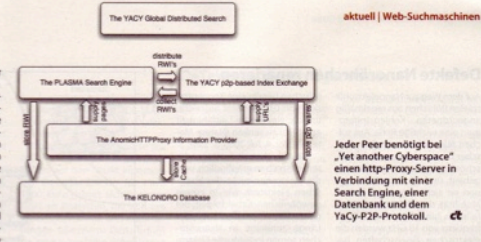
Noch ist das Rennen nicht einmal gestartet, bisher haben an dem Testprojekt insgesamt 695 Peers teilgenommen. Doch die Suchmaschinen-Experte ist zuversichtlich: „Kalkuliert man einmal über den Daumen, wie viele Peers es geben müsste, um mit Google zu konkurrieren, kommt man in die Größenordnung von einigen zehntausend – das ist also nicht absurd.“ [6]

Literatur

- [1] Online-Statistik von webstats.de: www.webstats.de/deutsch/index.shtml?deutsch/webstats.html
- [2] Blinkx, www.blinkx.com
- [3] Gemeinnütziger Verein zur Förderung der Suchmaschinen-Technologie und des freien Wissenszugangs, www.suma-ev.de
- [4] Bündnis 90/Die Grünen, Suchmaschinen: Das Tor zum Netz, www.gruene-fraktion.de/cms/publikationen/dokbin/63/63265.pdf
- [5] Yet another Cyberspace, www.yacy.net
- [6] Andreas Pothe, Nachgeziht, Wie groß ist das WWW? c't 26/04, S. 164
- [7] Jo Rager, Wiettsuchen, Suchmaschinen-Test, c't 26/04, S. 156

c't 2005, Heft 16

c't 2005, Heft 16



aktuell | Web-Suchmaschinen

Jeder Peer benötigt bei „Yet another Cyberspace“ einen http-Proxy-Server in Verbindung mit einer Search Engine, einer Datenbank und dem YaCy-P2P-Protokoll. ct

ct magazin für computer technik

Fotos, Musik, Dokumente ...

Datenschätze retten

Kratzer und Fehlbrände reparieren
Tools, Laufwerke, Tricks

A3-Tintendrucker
LCD-TVs mit PC-Anschluss
Business English auf CD

Online-Shopper auf schwarzen Listen

Für Workstation und Server
Dual-CPU-Boards

.NET für Linux und Mac OS
Power-Sauger kaufen mehr CDs
Die Technik der Tauschbörsen

Besser, billiger, vielseitiger
Smarte Handys
Winzige PCs für die Hosentasche

http://www.yacy.net/yacy
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

August 2005:
News-Ticker auf giga.de:

“Peer-to-Peer Suchmaschine wächst stetig”

The screenshot shows a web browser window displaying an article on the GIGA.DE website. The browser's address bar shows the URL <http://www.giga.de/index.php?storyid=124054>. The article is titled "Peer-to-Peer Suchmaschine YaCy wächst stetig" and is categorized under "Software". The text discusses the growth of the YaCy search engine, its P2P architecture, and its advantages over traditional search engines like Google and Yahoo. It mentions that YaCy is a free, independent search engine that uses a distributed index of local search engines. The article also provides information on how to use YaCy, including the need for a special Java software and a local proxy server. At the bottom of the article, there are links for "SMS", "CHAT", "MAIL", "FORUM", and "COMMENTS". The footer of the page indicates the copyright is © 1998 - 2005 GIGA Television GmbH.

GIGA.DE – Peer-to-Peer Suchmaschine YaCy wächst stetig!

http://www.giga.de/index.php?storyid=124054

NEWS Ticker +++ Vermal GeForce-Power auf einem Board +++

Software
Peer-to-Peer Suchmaschine YaCy wächst stetig!

(Alex.B) Das Peer-to-Peer Prinzip dürfte den meisten Leuten mittlerweile bekannt sein. Jeder User trägt sein kleines Stück zu einem Gesamtsystem bei - "gemeinsam zum Ziel", könnte das Motto lauten. Jetzt scheint diese Vorgehensweise auch im Suchmaschinen-Bereich Einzug zu halten - und zwar recht erfolgreich!

Die Rede ist von "YaCy", eine unabhängige Suchmaschine, bei der sich der Suchindex aus den lokalen Suchindizes der einzelnen User zusammensetzt. Die Projektentwickler wollen sich nicht länger auf einzelne Anbieter wie etwa Google oder Yahoo verlassen, und "YaCy" scheint tatsächlich auf dem besten Weg zu sein, eine gescheite Alternative zu werden.

Wie funktioniert denn nun "YaCy" im Detail? Ganz einfach: Damit die Suchmaschine genutzt werden kann, muss eine spezielle Java-Software auf dem heimischen Rechner installiert sein, die ähnlich wie ein lokaler Proxy-Server funktioniert. Dort landen dann alle besuchten Webseiten, die dann mit Schlagworten und versehen werden. Diese landen schließlich in einem Suchindex. Mithilfe eine Blacklist können auch besondere Webseiten gezielt aus dem Index herausgehalten werden. Bei "YaCy" übernimmt der Benutzer die Arbeit der Software bei herkömmlichen Suchmaschinen, die ständig die Tiefen des Internets absuchen.

Soviel zu den Grundlagen, nun zur P2P-Funktion: Sobald Suchanfragen für die YaCy Software erfolgen, gesellen sich neben dem eigenen Suchindex auch Indexdaten der anderen, aktiven YaCy-Benutzer in die eigene Anfrage mit dazu. Das Ergebnis: ein ständig wachsender Suchindex – und zwar Weltweit!

Die Yacy-Software gibt es für Windows, Linux und Mac OS X. Wer mehr über die alternative Suchmaschine erfahren möchte, findet unter yacy-forum.de ein Forum, das alle weiteren Fragen beantwortet.

published: 08.08.05

SMS CHAT MAIL FORUM COMMENTS

© 1998 - 2005 GIGA Television GmbH

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net

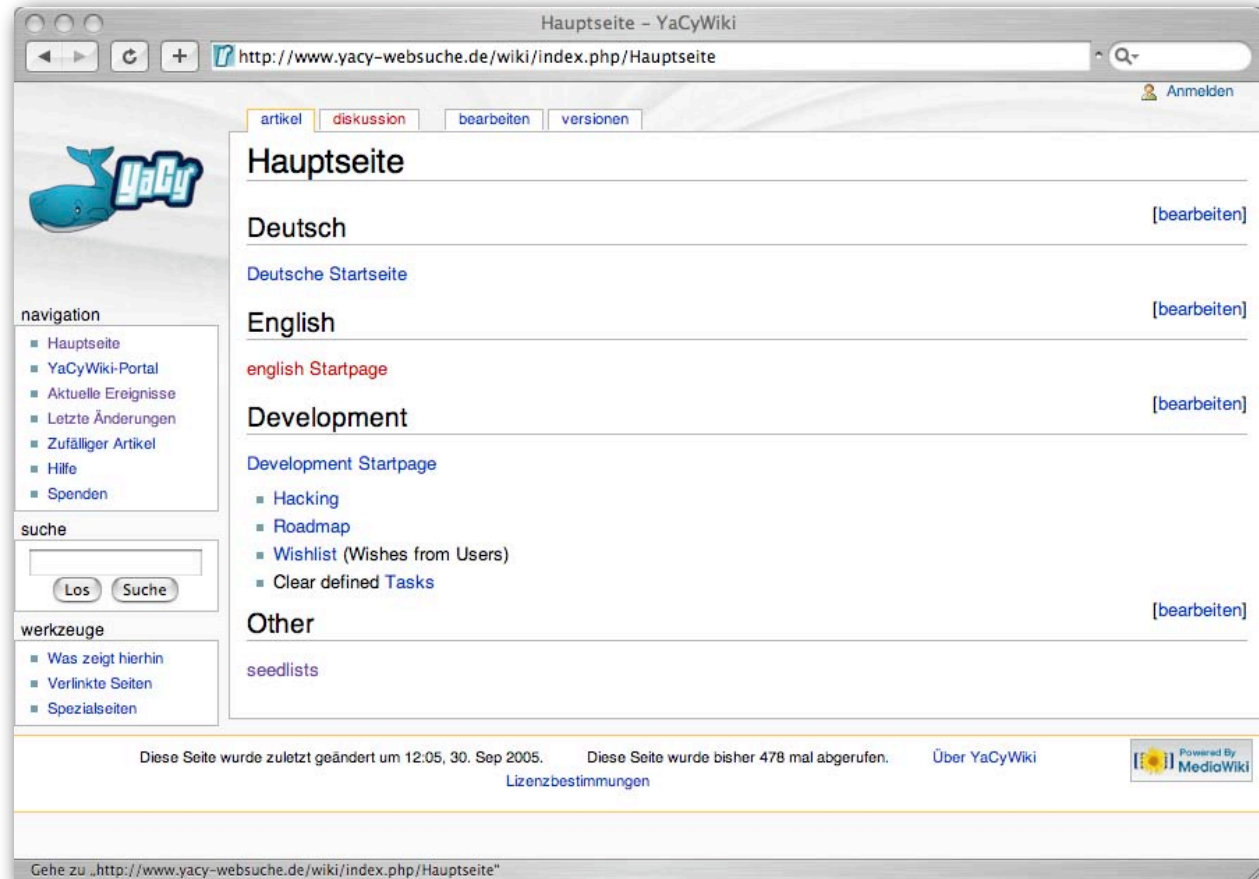


YaCy in der Presse und PR-Aktivitäten

September 2005:

Eröffnung des
YaCy-Wiki

Dort entsteht
gegenwärtig ein YaCy-
Handbuch, wird die
zukünftige
Entwickungs-Roadmap
entwickelt und
Entwickler-Treffen
koordiniert.



Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Oktober 2005:
Ausgabe 41 der "Die Zeit" macht
das Suchmaschinen-Thema
zum Leitartikel "Unser Schlüssel
zur Allwissenheit":
"David gegen Google"

Bei einer Aufzählung von
Alternativen zu Google wird
YaCy erwähnt. Der Online-
Artikel enthält einen Link zu
yacy.net:
"Der Suchmaschinenkrieg tobt.
[..] An der deutschstämmigen
Suchmaschine YaCy arbeiten
Entwickler nach dem Open-
Source-Prinzip"



Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net

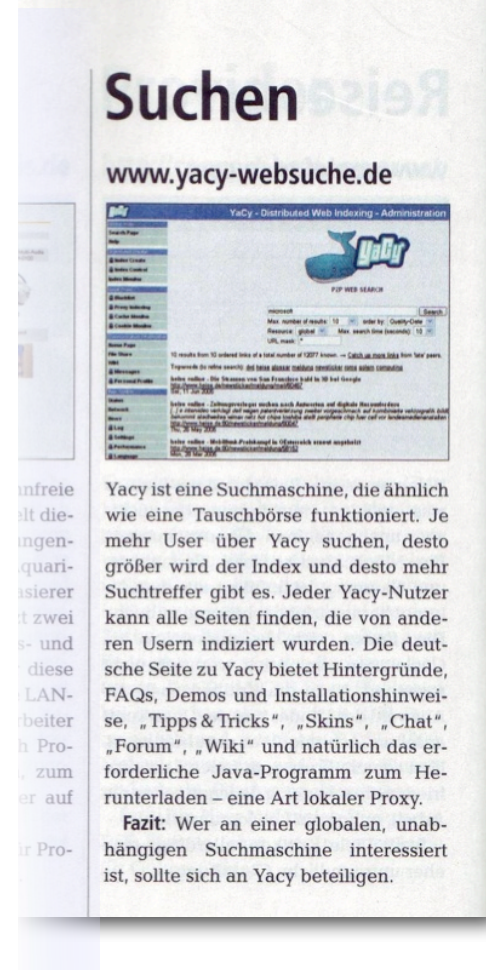


YaCy in der Presse und PR-Aktivitäten

INTERNET Surf-Tipps

Oktober 2005:

Surf-Tipp in der com! 11/2005



anfreie
elt die-
ngen-
quari-
asierer
t zwei
s- und
diese
LAN-
rbeiter
h Pro-
zum
er auf
r Pro-

YaCy ist eine Suchmaschine, die ähnlich wie eine Tauschbörse funktioniert. Je mehr User über YaCy suchen, desto größer wird der Index und desto mehr Suchtreffer gibt es. Jeder YaCy-Nutzer kann alle Seiten finden, die von anderen Usern indiziert wurden. Die deutsche Seite zu YaCy bietet Hintergründe, FAQs, Demos und Installationshinweise, „Tipps & Tricks“, „Skins“, „Chat“, „Forum“, „Wiki“ und natürlich das erforderliche Java-Programm zum Herunterladen – eine Art lokaler Proxy.

Fazit: Wer an einer globalen, unabhängigen Suchmaschine interessiert ist, sollte sich an YaCy beteiligen.

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



Das YaCy-Netz im Oktober 2005



Michael Christen
+49 177 6424235

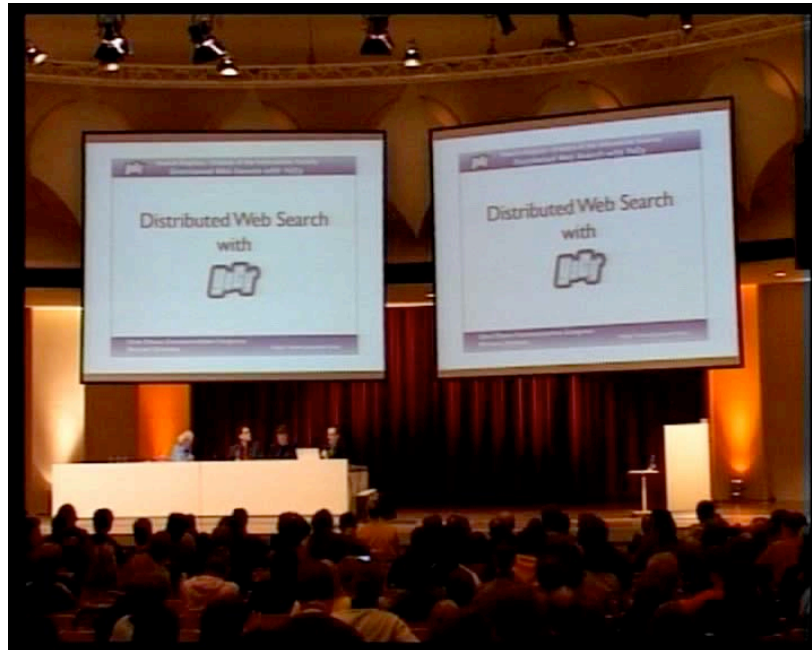
<http://www.yacy.net/yacy>
mc@yacy.net



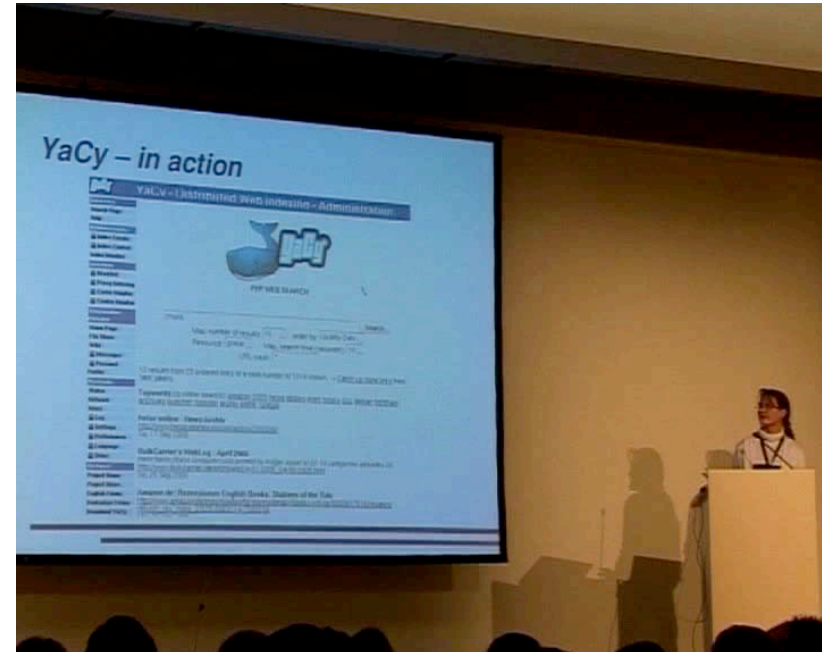
YaCy in der Presse und PR-Aktivitäten



Dezember 2005:
YaCy ist Thema in 2 Vorträgen
beim 22. Chaos Communication Congress (22C3) in Berlin.



„Search Engines - Oracles Of The Information Society“
von
Frédéric Philipp Thiele, Wolfgang Sander-Beuermann,
Michael Christen und Hendrik Speck



„Developing Intelligent Search Engines“
von
Isabel Drost

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net

Januar 2006:
Kritischer Artikel über Google
und Ausblick für YaCy als
Alternative

PC Magazin 1/2006
PRAXIS • TECHNIK • TESTS
Nur € 1,99
XP Starter Kit KOMPLETT AUF CD
All-in-One: Einmal klicken, alles installiert
Open Office 2.0, Adobe Reader 7, Real Player, DVD-Player, AntiVir, Bildbearbeitung, Zip-Tool, u.v.m.

Blaze Video Magic V1.0
Der Film-Konverter
• Rechnet Filme aus gängigen Formaten in MPEG 4 um
• Ideal für Portable Video-Player wie PSP & iPod Video

WinOptimizer 2005
Der Windows-Tuner
• Räumt die Registry und Festplatte auf
• Ideal für Online- und Desktop-PC
• Beschleunigt Windows nachhaltig

Die besten DVD-Brenner • Festplatten • Beamer • Lüfter • Notebooks unter 800 €

Besser als Google
Trotz Chaos alles finden: Die besten Suchtools für den eigenen PC & fürs Internet

Premiere geknackt
So schauen Hacker Play-TV & ORF 1, ohne dafür zu zahlen

DER PURE WAHNSINN MIT XBOX 360
Diese 2009-€-Konsolle ist besser als jeder PC &...

TECHNIK PRAXIS
• HD-DVD kontra Blu-ray & DVD
• Test Songs PSP & PS
• Cool: Generiere ab 1 € 10 € in Funktionen &...

Gewinnen Sie Preise im Wert von 180.000 Euro!

SPECIAL ALLES FINDEN

SUCHE: NADEL

WIR SIND DIE GUTEN!

Wird Google zum Microsoft der Suchmaschinen? Google muss sich gegen Monopolvorwürfe zur Wehr setzen. von CATHRIN GÜNZEL

Mit 30 Milliarden Seiten im Index sortieren Google inzwischen das, was die Welt wissen will. Die Firma will drahtlose Internet-Zugänge gratis anbieten, treibt die Entwicklung der kostenfreien Bürosoftware Open Office voran und speichert in Zukunft vielleicht Texte und Dokumente im Internet. Die Kalligrafen kaufen sich in den letzten Jahren mehr als ein Dutzend Technologiefirmen zusammen...

GoogleWatch
Die Webseite google-watch.org beobachtet den Suchverkehr ganz genau: News und Gerüchte...

Local for mobile
Take the power of Google Local with you on your mobile phone.

„Google ist kein Betriebssystem“

DAS INTERVIEW FÜHRTE CATHRIN GÜNZEL

Medienrechtler Wolfgang Schulz, Direktor des Hamburger Hans-Bredow-Instituts für Medienforschung, zu Googles Vernachlässigung des Suchmaschinenmarktes

„Google ist kein Betriebssystem“

Das Interview führte Cathrin Günzel

„Haben andere Suchmaschinen überhaupt eine Chance, Googles Marktstellung anzufordern?“
Googles Monopol ist zurecht. Die Suchmaschine ist kein Betriebssystem, das ist auch der grundlegende Unterschied zu Microsoft. Nichts hindert daran, jederzeit zu einer anderen Suchhilfe zu wechseln. Schon morgen kann ein anderer Anbieter eine bessere, schnellere Lösung auf den Markt bringen.

„Worin liegt dann das Problem?“
Kombiniert, reguliert den Markt. Funktionierendes Wettbewerb ist ein Schutzmittel gegen jede Art des Missbrauchs von Marktmacht. Google limitiert den Informationsfluss im Internet zurecht oder nahezu unbemerkbar.

„Muss Google nun staatliche Regulierung oder kartellrechtliche Maßnahmen befrachten?“
So wie ich zurzeit keine Sinnhaftigkeit sehe, wenn Google eine andere Suchmaschine kauft, müsste das Unternehmen mit einem Kartellverfahren rechnen.

„Ist Googles Weltgeltungsmacht medienrechtlich unbedenklich?“
Es besteht Bindungsbedarf. Bisher ist die Kontrolle auf das klassische Rundfunkrecht beschränkt. Ich habe Zweifel, ob das noch zeitgemäß ist. Das Ausdehnen der Kontrolle auf neue Dienste wie Suchmaschinen würde mir einleuchten. So dürfen zum Beispiel Rundfunkanbieter nicht mehr als 30 Prozent des Zuschauermarktes beherrschen. Googles Marktanteil liegt bei über 80 Prozent.

„Genügt nicht die Selbstregulierung, der sich Suchmaschinen zum Beispiel mit dem Ausfiltern inaktiver Inhalte unterwerfen haben?“
Das ist ein guter Anfang. Da ist Deutschland den USA voraus. Doch es muss für die Nutzer noch transparenter werden, nach welchen Regeln die Ergebnisse der Suchmaschinen zustande kommen. Im Übrigen sind deutsche Online-Anbieter zur Sperrung illegaler Inhalte verpflichtet. Über ein künftiges Selbstregulierungsverfahren für solche Entscheidungen sollte man jedoch diskutieren.

PC Magazin 1/2006 www.pc-magazin.de

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacymc@yacy.net>



YaCy in der Presse und PR-Aktivitäten

März 2006:
YaCyBar - eine Firefox Toolbar
wird offiziell von mozilla.org
gehostet

The screenshot shows the Firefox Add-ons page for the YaCyBar extension. The page header includes the Mozilla logo and a search bar. The main content area features a sidebar with navigation links for YaCyBar, including Overview, Previews & Screenshots, Comments, Add a Comment, Version History, and About the Author. The main content area displays the extension's overview, including its name, version (0.1), author (Alexander Schier), and release date (Mar 6, 2006). A description states it is a searchtoolbar for the Peer to Peer searchengine YaCy, with a link to the website. The page also shows compatibility information (Works with: Firefox 1.5 - 1.5.0.* ALL) and an 'Install Now for MacOSX (38 KB)' button. A 'Rating:' section is visible on the right, along with a 'More Previews >' link.

Michael Christen
+49 177 6424235

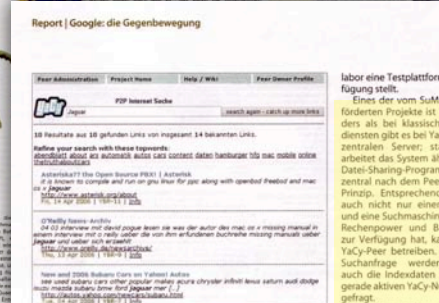
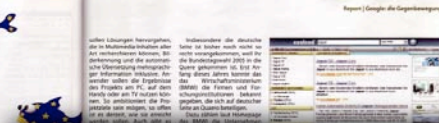
<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

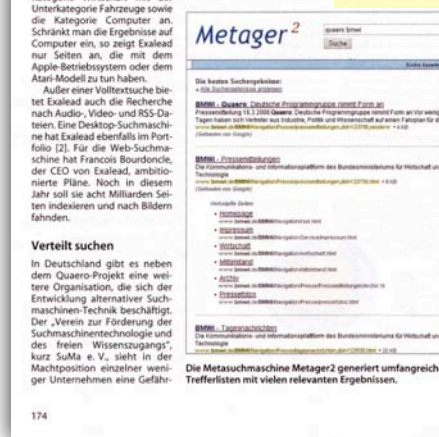
April 2006:
große Artikelsammlung über
Google und Alternativen in der
c't 10/2006

In einem Artikel zu europäischen
Suchmaschinenalternativen wird
YaCy unter nur insgesamt 4
Alternativen ausführlich
besprochen



labor eine Testplattform zur Verfügung stellt.
Eines der vom SuMa e.V. geförderten Projekte ist YaCy. Anders als bei klassischen Suchdiensten gibt es bei YaCy keinen zentralen Server; stattdessen arbeitet das System ähnlich wie Datei-Sharing-Programme dezentral nach dem Peer-to-Peer-Prinzip. Entsprechend gibt es auch nicht nur einen Crawler und eine Suchmaschine, wer die Rechenpower- und Bandbreite zur Verfügung hat, kann einen YaCy-Peer betreiben. Bei einer Suchanfrage werden immer gerade aktive YaCy-Nutzer abgefragt.
YaCy lässt sich über mehrere Peers ausprobieren, etwa auf der Projekt-Homepage [4] oder im Labor des SuMa e.V. Die Ergebnisse sind allerdings noch recht dürftig. Zu obiger Abfrage „quero brw“ zum Beispiel liefert YaCy keine Treffer. Die YaCy-Betreiber suchen noch Mitstreiter, die weitere Peers betreiben – je mehr Teilnehmer das YaCy-Netz hat, desto reichhaltiger werden die Ergebnisse. Die YaCy-Software wurde unter Java entwickelt und läuft auf diversen Linux-Varianten, Mac OS X und

Windows. Die Installation der Software wird auf der Homepage detailliert beschrieben.
Wer nicht die Möglichkeit hat, eine vollständige Server-Software zu installieren, der kann das Projekt auch mit der YaCy-Toolbar für Firefox unterstützen. Mit der Werkzeugleiste lassen sich Anfragen direkt an YaCy senden oder Seiten zum Indexieren an einen der Peers übermitteln. Der Benutzer kann Spam-Seiten melden; YaCy setzt sie auf eine schwarze Liste.
Metasucher
Unterstützung vom SuMa e.V. erhält auch die Metasuchmaschine Metager2, eine Weiterentwicklung des Klassikers Metager (www.metager.de), wie im Namen schon anknüpft. Anders als der Vorgänger und viele andere Metasuchmaschinen sammelt Metager aber nicht nur die Treffer mehrerer Suchdienste ein und bereitet sie auf. Stattdessen lädt Metager2 die Trefferseiten, analysiert sie und generiert daraus eine Übersicht. Dies soll es ermöglichen, Spam auszufiltern und das Ranking wesentlich zu verbessern.
Der Abdruck der einzelnen Trefferseiten dauert natürlich ein wenig länger als die Recherche mit Google oder Yahoo. Metager2 benutzt derzeit die Ergebnisse von Google, Yahoo, MSN, Ask und Exalead. Etwa fünf bis zehn Sekunden muss der Benutzer warten, bis Metager2 die Ergebnisse eingesammelt und zu einer langen Liste zusammengestellt hat. Die Resultate können sich aber sehen lassen: In Stichproben lieferte die Suchmaschine gute, relevante Treffer; Spam-Sites hatten Selbstenwert. Außerdem der Volltextsuche hat Metager2 derzeit keine weiteren Funktionen.
Literatur
[1] Homepage der „deutsch-französischen Arbeitsgruppe“ mit einer Beschreibung des Quero-Projekts: www.dflap-2005.com
[2] Volltexte, Queresktop 4.0 free, c't 9/06, S. 84
[3] Homepage des SuMa e.V.: www.suma-ev.de
[4] Homepage von YaCy: www.yacy-technology.com
[5] Metager 2, www.metager2.de



Michael Christen
+49 177 6424235

http://www.yacy.net/yacy
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

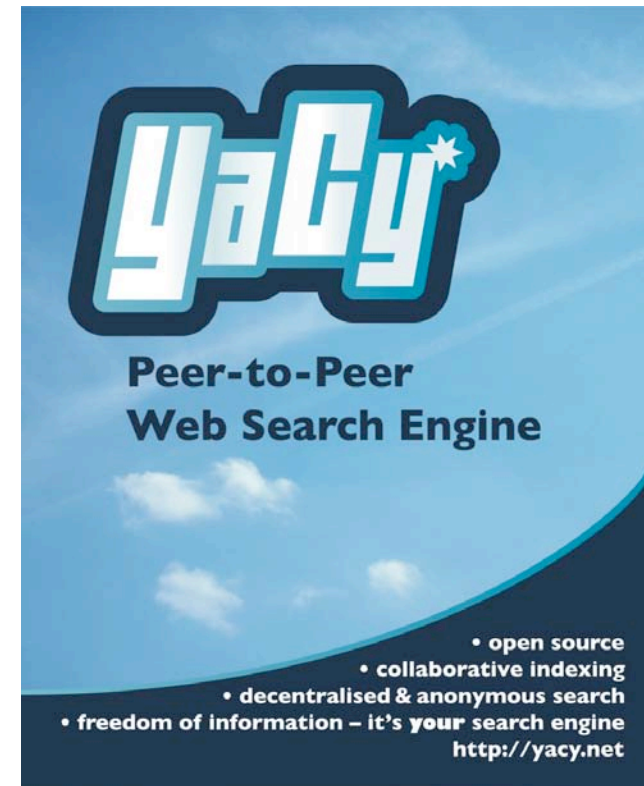
Mai 2006:
YaCy hat einen Messestand
beim LinuxTag 2006 in
Wiesbaden



YaCy ist auch die Portal-
Suchmaschine für
<http://www.linuxtag.org>



Vortrag:
„Informationsfreiheit und freie
Information Retrieval Software“
im Rahmen des ‚Practical Linux
Forum‘



Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



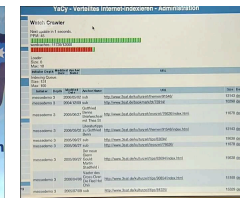
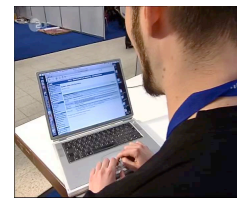
YaCy in der Presse und PR-Aktivitäten

Mai 2006:
das 3sat-Magazin
,neues‘ berichtet
über YaCy vom
Linuxtag

In einem fast 1-
Minütigen Beitrag
inklusive Interview
wird YaCy als
Alternative zu
Google dargestellt.

Die Sendung wurde
mehrfach wiederholt,
u.a. auf 3sat

The screenshot shows a website with a red header. On the left is a navigation menu with categories like 'Programm', 'Service', 'Film', 'Gespräch', etc. The main content area features the article title and a detailed text about the YaCy project. A 'Links' box on the right contains the URL 'www.yacy.net'. At the bottom, there are navigation links like 'zurück', 'Seitenanfang', and 'Druckversion'.



<http://www.3sat.de/neues/sendungen/magazin/91715/index.html>

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Juli 2006:
Das Linuxmagazin 07/2006
berichtet über aktuelle
Geschehnisse bzgl Open-Source
Projekte. YaCy wird als alternative
zu Google genannt und dessen
Monopolstellung kritisiert.



106
Linux Magazin 006
Forum

Abbildung 3: Ein Netzwerk von Yacy-Clients liefert die Ergebnisse für eine Websuche.

So verfügt die Software über einen eigenen Webproxy. Aus ihm speichert das System die Zugriffe abzüglich aller personalisierten Informationen, um die Datenbank für die Suchmaschine aufzubauen. Parallel arbeitet sich ein Crawler-Programm durch Webseiten, um diese in den Index aufzunehmen. Yacy erlaubt es, den eigenen Index mit den Indizes anderer zu verknüpfen. Dabei übermüht die Software nur die Indizes, nicht aber die Ergebnisse, da die Datenmenge sonst zu schnell wächst.

Setzt ein Benutzer eine Anfrage ab, werden auf Basis des globalen Index die Ergebnisse verschiedener Yacy-Installationen gesammelt, um sie dem Nutzer als eine konsolidierte Ausgabe zu präsentieren. Der dezentrale Aufbau schützt Yacy nicht nur gegen Manipulation, er ermöglicht es auch, Informationen mit hoher Aktualität zu erfassen und zu suchen. Zudem geschieht das Erstellen des Index um Größenordnungen schneller als bei vergleichbaren Projekten.

Yacy ist in Java geschrieben. Es läuft auf Linux ebenso wie auf Mac OS X und Windows, hängt aber von der Java-Implementierung ab. Im Laufe des Linxtags 2006 hat Michael aber bereits mit dem Maintainer von GNU Classpath erste Erfolge erzielt, um Yacy auf eine freie Java-Variante zu portieren. Bis zur ersten nativ übersetzten Version ist es also wohl nicht mehr so weit. Daher ist es höchste Zeit, Yacy mal dem Präzisiest zu unterziehen. Vielleicht wäre dies ein guter Kandidat für den nächsten Preis-Software-Preis. (agr)

Nummer. Daher haben Duane, der Gründer von Cacerot, und sein Freund Matthew damit begonnen, auf Basis des Standards eine eigene, nicht kommerzielle Plattform aufzubauen, die diese Technologie allgemein verfügbar macht. Diese Plattform findet sich unter E164.org [1] und erlaubt es, dort eigene Telefonnummern zu registrieren.

Der Nutzen eines solchen Dienstes hängt natürlich von der Qualität der Einträge ab. Um Mißbrauch zu verhindern, findet automatisches Validieren mit PIN-Eingabe statt. Um konventionelle Nummern einzutragen, setzt dies aber einen Anbieter voraus, der die Anrufe ins PSTN weiterleitet. Genau hier liegt einer der aktuellen Engpässe: Als nicht kommerzieller Service sind diese Anrufe ein wesentlicher Kostfaktor. E164.org bittet daher die Provider um eine Sachspende. Da sie immer öfter E164.org nutzen, hilft die Spende den eigenen Interessen.

Im Prinzip erlaubt der Standard das Zuordnen verschiedener Dienste, beispielsweise auch der so genannten Instant Messenger (IM). Ein sehr beliebtes Programm in diesem Gebiet ist Gaim [2], das ab der nächsten Version auch VoIP unterstützt (Abbildung 2). Daher reicht mit E164.org künftig eine einzige Nummer aus, um verschiedenen Dienste zu identifizieren. E164.org will dies fördern und setzt daher immer wieder Prämien aus, um Enun-Support in verschiedenen Programmen einzubauen.

Die Prämie für Firefox hat sich bereits jemand gesichert und damit das Zuweisen von Webseiten über Enun ermöglicht. Für Gaim und Thunderbird stehen noch 100 beziehungsweise 20 US-Dollar bereit. Vielleicht reicht es bald schon aus, auf Visitenkarten nur noch eine Nummer anzugeben. Darunter wären Webseite, Mail, VoIP, Fax und Jabber als Dienste zu finden. Eine Kombination dieser Ansätze mit Peer-to-Peer-Methoden ergäbe einige sehr interessante Möglichkeiten.

Suchmaschine Yacy

Auf der Peer-to-Peer-Methodik basiert auch das zweite Projekt, eine freie Suchmaschine, die sowohl Anonymität als auch Informationsfreiheit bewahren will: die Yacy-Suchmaschine [3]. Es ist in letzter Zeit immer häufiger Unbehagen über die Vormacht von Google zu spüren. Viele Leute fragen sich, ob die Abhängigkeit von dem Suchdienst nicht zu groß sei. Michael Christen gehörte zu denen, die sich diese Frage gestellt haben. Schließlich begann er damit, eine Suchmaschine auf Basis des Peer-to-Peer-Prinzips zu entwerfen. Yacy besteht aus mehreren Komponenten, die gemeinsam beeindruckende Funktionen bieten (Abbildung 3).

Der Autor
Dipl.-Phys. Georg C. F. Freve beschäftigt sich seit etlichen Jahren mit freier Software und kam früh zu GNU/Linux. Nach Mitarbeit im GNU-Projekt und seiner Aktivität als dessen europäischer Sprecher hat er die Free Software Foundation Europe initiiert, deren Präsident er ist. Mehr Informationen finden sich unter <http://gnuth.org>

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

September 2006:
Die c't 22/2006 berichtet über das jährliche Suchmaschinenforum des SuMa-eV. Der Artikel berichtet über den Vortrag „Entwicklungen bei der Peer-to-Peer Suchmaschine YaCy“ und den neuen Features, u.a. dezentrales social bookmarking.



Suchmaschinenforum: über Google, Quaero und vielversprechende Neulinge

Suchmaschinen sind die wichtigsten Orientierungshilfen im Datenmeer des Web. Allerdings dominieren immer weniger Unternehmen den Suchmaschinenmarkt. Einer der schärfsten Kritiker dieser „monopolartigen Strukturen“ ist der Verein zur Förderung der Suchmaschinen-Technologie und des freien Wissenszugangs, kurz SuMa-eV, der in Berlin zu seinem dritten Forum einlud.

Ursprünglich versprach das Treffen Neuigkeiten über Quaero, das europäische Suchmaschinen-Forschungsvorhaben, das in der Vergangenheit bereits zum „Google-Killer“ hochstilisiert wurde [1]. Doch der Beitrag von Stefan Wess, Geschäftsführer des deutschen Quaero-Konsortialführers Empolis, wurde kurzfristig abgesagt. Obwohl Quaero also offiziell gar nicht mehr auf der Agenda stand, war es dennoch allgegenwärtig.

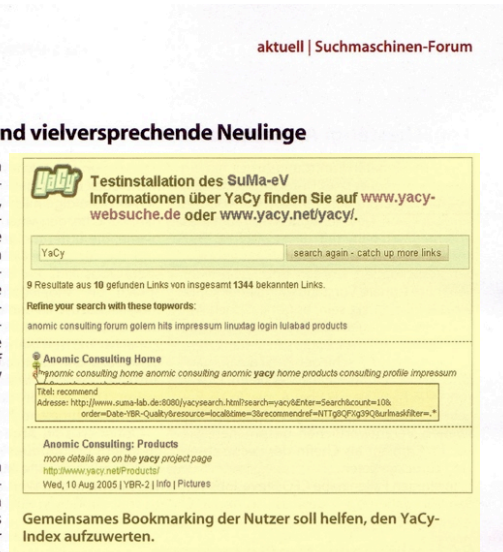
So stellte Dagmar Wöhrl, Parlamentarische Staatssekretärin im Wirtschaftsministerium, in ihrem Grußwort die Wichtigkeit von Quaero heraus. Als sogenanntes „Leuchtturmprojekt“ soll das Forschungsvorhaben mit 400 Millionen Euro gefördert werden. Wolfgang Sander-Beuermann, Vorsitzender des SuMa-eV, mahnte in seiner Begrüßungsrede, Quaero sei „für Europa die letzte Chance, den Zug in die Informationsgesellschaft nicht zu verpassen“. Wenn Quaero kein Erfolg werde, würde der technologische Rückstand Europas so groß werden, dass er „endgültig und unaufholbar“ sei. Wie groß der Vorsprung von Google ist, zeigte der Vortrag

von Hendrik Speck, Professor an der Fachhochschule Kaiserslautern. Egal ob in Deutschland, Frankreich oder den USA: In vielen wichtigen Märkten habe Google bereits die Nase vorn. In anderen Ländern, etwa Russland, sei es nur noch eine Frage der Zeit, bis Google eine ähnlich beherrschende Stellung einnehme. Der Beitrag ist, wie die meisten anderen auch, auf der Homepage des SuMa-eV hinterlegt [2].

GYM-Vorherrschaft

An der Vorherrschaft der großen drei – Google, Yahoo und Microsoft, kurz GYM – werden auch die Start-ups des Web 2.0 nichts ändern, so der Marktforscher Alexander Linden. Von den Netzwerkeffekten, der „Währung“ im Web 2.0, werden am ehesten die Unternehmen profitieren, die bereits eine große Benutzerbasis besitzen – allen voran das GYM-Dreigespann. Für sie sei es am einfachsten, einen neuen Dienst zu etablieren. Während andere viel Werbeaufwand betreiben müssen, reiche es ihnen häufig, einen Link auf ein neues Angebot zu setzen. Google selbst unterstrich diese These in persona des Pressesprechers Stefan Keuchel, der etliche kürzlich neu eingeführte Dienste präsentierte.

Dass auch ohne Quaero-Förderung interessante neue Suchmaschinen entstehen können, ist engagierten Entwicklern und der Unterstützung des SuMa-eV zu verdanken, der mit seinen Labs eine Entwicklungsplattform bereitstellt [3]. Michael



Christen stellte Neuerungen der dezentralen Suchmaschine YaCy vor [4]. YaCy arbeitet verteilt auf privaten Rechnern ohne zentrale Kontrolle. Auf diese Weise soll zum einen Zensur verhindert, zum anderen die Anonymität der Suchenden gewährleistet werden. Mittlerweile 65 Knotenrechner des Projekts indexieren 300 Millionen Seiten. Interessierte sind nach wie vor willkommen, sich an YaCy zu beteiligen. Die Server-Software soll sich jetzt so einfach einrichten wie ein Desktop-Programm. Über eine Firefox-Suchleiste können Surfer bequem mit YaCy suchen, Seiten zum Indexieren anmelden oder auf die Blacklist setzen.

Mit OpenCrawl bietet der SuMa-eV Interessierten die Möglichkeit, eine individuelle Suchmaschine zu hosten. Der Web-

master kann ihre Ergebnisse in seine Webseiten einbetten [5]. Um OpenCrawl nutzen zu können, muss man sich formlos mit einer E-Mail bewerben. Auch Markus Franz, der Entwickler von Metager2, stellte ein neues Projekt vor: Die Multimediasuchmaschine M4 wirft gleichzeitig Text-, Audio-, Bild- und Videoeergebnisse aus. (jo)

Literatur

- [1] Jo Bager, Europäische Gegenströmung, Bestrebungen zur Entwicklung europäischer Google-Konkurrenten, c't 10/06, S. 172
- [2] Programm des Forums mit vielen Beiträgen, <http://suma-ev.de/forum06>
- [3] Labs des SuMa-eV, <http://suma-lab.de>
- [4] YaCy, www.yacy-websuche.de
- [5] OpenCrawl, www.opencrawl.de

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten

Mit DVD Top-Programme für Freizeit & Beruf € 9,90 November-Januar 1/07

MAGAZIN FÜR FREIE SOFTWARE

OpenSource

1.800 aktuelle Programme für Windows und Linux

4,7 GB gratis auf DVD

Projekte zum Mitmachen: Linux im Eigenbau Schritt für Schritt zur ganz persönlichen Distribution

Google-Alternative So helfen Sie beim Aufbau einer vernetzten Suchmaschine

Software auf Heft-DVD

Ubuntu Linux 6.10 RC

Dezember 2006 - Januar 2007: Das OpenSource Magazin 01/2007 bringt einen 4-Seitigen Artikel über YaCy mit Hinweis auf den Artikel auf der Titelseite.

Internet-Projekte

Gemeinsam gegen Google

Als Kombination aus P2P-Technologien und Distributed Computing möchte die freie Suchmaschine YaCy Google und Co. die Stirn bieten. Abwegig ist dieser Plan nicht - stehen doch einige zehntausend Rechner in Googles Cluster Millionen potenzieller YaCy-Peers gegenüber.

Die Philosophie von YaCy ähnelt der von OpenSearch, möchte auch hier aber ein bisschen anders sein. In OpenSearch werden die Suchmaschinen von den Servern der Suchmaschinenbetreiber übernommen, indem sie als Server oder Client über YaCy verbunden werden.

YaCy ist eine Suchmaschine, die sich selbst organisiert. Sie besteht aus einem zentralen Server und mehreren Clients. Die Clients sind über ein Peer-to-Peer-Netzwerk verbunden und können sich gegenseitig Suchergebnisse liefern. Dies ermöglicht es, die Suchleistung zu skalieren und die Suchergebnisse zu verbessern.

Das Projekt ist ein Beispiel für die Macht von Open Source und Distributed Computing. Es zeigt, wie eine Gruppe von Freiwilligen zusammenarbeiten kann, um eine Lösung für ein Problem zu finden.

Internet-Projekte

So geht's Installation von YaCy

Man können denken, dass eine umfangreiche, detaillierte Server-Software eine erhebliche Investition in Geld ist. Das Gegenteil ist der Fall: Eine Basisinstallation ist mit wenigen Klicks erledigt.

1. Auf dem Server muss eine aktuelle Version von YaCy installiert sein. Dies kann über die YaCy-Website oder über ein Paketmanager wie apt-get oder yum erreicht werden.
2. Die YaCy-Server-Software muss über ein Netzwerk mit den Clients verbunden sein. Dies kann über die YaCy-Website oder über ein Netzwerk-Tool wie nmap erreicht werden.
3. Die YaCy-Server-Software muss über ein Netzwerk mit den Clients verbunden sein. Dies kann über die YaCy-Website oder über ein Netzwerk-Tool wie nmap erreicht werden.

Die Installation ist ein einfacher Prozess, der nur wenige Minuten dauert. Die YaCy-Website bietet eine detaillierte Anleitung für die Installation und die Konfiguration der Software.

Internet-Projekte

YaCy als Suchmaschine

Die YaCy-Website ist eine Suchmaschine, die sich selbst organisiert. Sie besteht aus einem zentralen Server und mehreren Clients. Die Clients sind über ein Peer-to-Peer-Netzwerk verbunden und können sich gegenseitig Suchergebnisse liefern. Dies ermöglicht es, die Suchleistung zu skalieren und die Suchergebnisse zu verbessern.

Die YaCy-Website ist eine Suchmaschine, die sich selbst organisiert. Sie besteht aus einem zentralen Server und mehreren Clients. Die Clients sind über ein Peer-to-Peer-Netzwerk verbunden und können sich gegenseitig Suchergebnisse liefern. Dies ermöglicht es, die Suchleistung zu skalieren und die Suchergebnisse zu verbessern.

Internet-Projekte

YaCy als Suchmaschine

Die YaCy-Website ist eine Suchmaschine, die sich selbst organisiert. Sie besteht aus einem zentralen Server und mehreren Clients. Die Clients sind über ein Peer-to-Peer-Netzwerk verbunden und können sich gegenseitig Suchergebnisse liefern. Dies ermöglicht es, die Suchleistung zu skalieren und die Suchergebnisse zu verbessern.

Die YaCy-Website ist eine Suchmaschine, die sich selbst organisiert. Sie besteht aus einem zentralen Server und mehreren Clients. Die Clients sind über ein Peer-to-Peer-Netzwerk verbunden und können sich gegenseitig Suchergebnisse liefern. Dies ermöglicht es, die Suchleistung zu skalieren und die Suchergebnisse zu verbessern.

Michael Christen
+49 177 6424235

<http://www.yacy.net/yacy>
mc@yacy.net



YaCy in der Presse und PR-Aktivitäten



Januar 2007: Die c't 01/2007 berichtet über individuelle Suchmaschinen-Software

YaCy ist eine der Suchmaschinen-Optionen, und erhält als einzige der genannten Software einen Beschreibungskasten zur Installation und Inbetriebnahme

Report | Individuelle Suchmaschinen

Sites hinzufügen und löschen. Der Inhalt behält aber immer alle Fäden in der Hand und kann die Änderungen von Dritten oftmals rückgängig machen. Beliebig benennbare Etiketten, sogenannte Refinements,

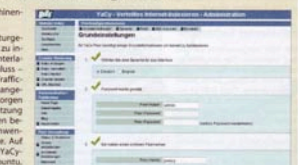
soßen Besuchern helfen, die Ergebnisse einer Suche geschärft durchzuführen. Der Inhaber versteht von der Suchmaschine erfasste Websites mit solchen Refinements. Die Suchmaschine zeigt die Etiketten oberhalb der

Ergebnisse an. Klickt der Nutzer auf ein Refinement, so bevorzugt Google die betreffenden Sites in den Ergebnissen. Außer der Web-Administrationsoberfläche stellt Google ein Bookmarklet namens Google

Marker bereit, mit dem der Inhaber beim Surfen per Klick Sites zum Suchmaschinenindex hinzufügt. Daneben unterhält Google eine Schnittstelle für den im- und Export kompletter Suchmaschinenbeschreibungen

Privatdetektiv: YaCy – Suchmaschine selbst aufgesetzt

So leistungsfähig die persönlichen Suchangebote auch sein mögen – Yahoo und MSN – per Suchmaschinenindex haben sie den Nachteil, dass sie nur die Suchergebnisse der Suchmaschinenbetreiber zeigen. Alternativ setzt YaCy auf einen cachenden Proxy-Server auf. Benutzern Surfen den Proxy-Server, so indiziert YaCy die angeschurften Seiten sowie, je nach Einstellung, darin verlinkte Seiten. Auf diese Weise kann YaCy Seiten aus dem sogenannten Deep Web, also zum Beispiel aus Online-Datenbanken, indizieren, die anderen Suchmaschinen verschlossen bleiben.



ambitioniertes Suchmaschinenprojekt. Ein Crawler erzeugt regelmäßig viel Traffic, da er alle zu indizierenden Seiten heruntergeladen muss. Ein DSL-Anschluss mit Flatrate, damit die Traffic-Kosten nicht für eine unangenehme Überraschung sorgen – ist daher eine Voraussetzung für den Betrieb. Ansonsten benötigt YaCy als Java-Anwendung eine Virtual-Machine. Auf der Projektseite stehen YaCy-Pakete für Debian, Ubuntu, Gentoo, iPCop und Windows zum Herunterladen bereit. Zur Installation muss man die Linux-Pakete nur in ein Verzeichnis der Wahl entpacken, die Windows-Version kommt in Form eines Installers. Anschließend startet man YaCy mit der Start-Batch-Datei, die sich im Installationsverzeichnis findet. Für Linux heißt sie startACY.sh, für Windows startACY.bat. Der Windows-Installer richtet einen Autostart-Eintrag ein, der darauf sorgt, dass YaCy bei jedem Systemstart hochgefahren wird.

Nach der Installation lässt sich die Verwaltungsoberfläche von YaCy unter <http://localhost:8080> aufrufen. Jetzt muss der angehende Suchmaschinenbetreiber noch unter <http://localhost:8080/manager.html> ein Administratorpasswort eingeben und seinen YaCy-Peer benennen – fertig ist der individuelle Rechner.

YaCy kann einen Proxy benutzen – praktisch zum Beispiel, wenn man die Suchmaschine in einem Firmennetz einsetzen will. Dazu ruft man die Seite www.yacy.net/faq.html auf und wählt den Punkt 'Remote Proxy'. Mit dem Punkt 'Lokaler Proxy/Proxy-Indizierung' gibt es nicht mehr möglich eine individuelle Suchmaschine, die sich genau auf seine persönlichen Interessen festlegen lässt. Unter 'Admin' können die YaCy als Proxy weitergeleitete Seiten

tracht – oder man muss das Crawlern in die Stunden verlegen, in denen man selbst nicht am PC sitzt. Unter <http://localhost:8080/index.html> ist die Suche erreichbar. In den erweiterten Optionen kann der Benutzer YaCy veranlassen, nur die lokalen Ergebnisse, nicht die Gesamtresultate des YaCy-Verbinds zu benutzen. A propos: Ein frisch installierter YaCy-Peer mit Internet-Zugriff hat den Status 'Junior'. Damit kann er die Indizes der anderen YaCy-Peers nutzen, aber selbst nicht von außen erreicht werden. Um auch selbst Ergebnisse für die Allgemeinen bereit zu stellen, muss der Administrator den Port 8080 seiner Firewall für eingehende Verbindungen öffnen. Details hierzu wie zu den umfangreicheren weiteren Einstellungsöglichkeiten von YaCy, die über das Rahmen dieses Artikels sprengen, liefert die ausführliche Online-Dokumentation.

YaCy wird permanent weiterentwickelt. In Kürze wird eine Version erscheinen, mit der man gezielt nach MP3, Videos und anderen Medientypen suchen kann. Auch ist ein Robbing-Mechanismus für eine völlig private Suche geplant.

Report | Individuelle Suchmaschinen

Der persönliche Fahnder

Gezielt suchen mit individuellen Suchmaschinen

Manchmal will man nicht die ganze Web-Welt durchsuchen, sondern nur diejenigen Sites, die man für seine Recherche als relevant ansieht. Bei einer Hand voll Domänen kann man zu diesem Zweck individuelle Suchmaschinen aufsetzen und die Suche auf der eigenen Homepage anbinden. Auf diese Weise sind schon Tausende Spezialsuchmaschinen entstanden, die jahrelang als Phänomen...

Wie ich diese mal selbst gemacht habe... (Text continues with details about setting up a search engine)

118 | c't 2007, Heft 1

Report | Individuelle Suchmaschinen

YaCy lässt sich in vielerlei Hinsicht konfigurieren.

bei der Indexierung berücksichtigt. Als Standard ist eine Crawler-Tiefe von 1 eingestellt, das heißt, YaCy indiziert die gecachten Seiten sowie diejenigen, die darin verlinkt sind. Auf keinen Fall indiziert YaCy persönliche Seiten, also etwa solche, die eine Passwort-Authentifizierung erfordern.

Unter 'Crawler Steuerung/Indizes erzeugen' lassen sich eigene Crawler-Aufträge festlegen. Dazu gibt der Administrator unter Startpunkt die URL ein, auf der der Crawler starten soll. Die 'Crawling-Tiefe' legt fest, wie viele Ebenen tief er den Links auf dem Startpunkt folgen soll. Die Optionssseite lässt den Administrator noch viele weitere Parameter festlegen, etwa, ob YaCy dynamische URLs mitwiderlegen soll. Jeder Parameter wird auf der Einstellungsseite detailliert beschrieben.

Unter 'Crawler Monitor' zeigt YaCy die zur Indizierung anstehenden Seiten an. Während des Crawlens erzeugt YaCy bei uns eine derzeit hohe CPU-Last, dass ein flüssiges Arbeiten mit dem PC nicht mehr möglich war. Daher sollte man für den YaCy-Betrieb einen eigenen Rechner einsetzen – hierfür kommen auch betagtere Modelle in Betracht – oder man muss das Crawlern in die Stunden verlegen, in denen man selbst nicht am PC sitzt.

Unter <http://localhost:8080/index.html> ist die Suche erreichbar. In den erweiterten Optionen kann der Benutzer YaCy veranlassen, nur die lokalen Ergebnisse, nicht die Gesamtresultate des YaCy-Verbinds zu benutzen. A propos: Ein frisch installierter YaCy-Peer mit Internet-Zugriff hat den Status 'Junior'. Damit kann er die Indizes der anderen YaCy-Peers nutzen, aber selbst nicht von außen erreicht werden. Um auch selbst Ergebnisse für die Allgemeinen bereit zu stellen, muss der Administrator den Port 8080 seiner Firewall für eingehende Verbindungen öffnen. Details hierzu wie zu den umfangreicheren weiteren Einstellungsöglichkeiten von YaCy, die über das Rahmen dieses Artikels sprengen, liefert die ausführliche Online-Dokumentation.

YaCy wird permanent weiterentwickelt. In Kürze wird eine Version erscheinen, mit der man gezielt nach MP3, Videos und anderen Medientypen suchen kann. Auch ist ein Robbing-Mechanismus für eine völlig private Suche geplant.

180 | c't 2007, Heft 1

Report | Individuelle Suchmaschinen

181 | c't 2007, Heft 1

Report | Individuelle Suchmaschinen

182 | c't 2007, Heft 1

Michael Christen +49 177 6424235

<http://www.yacy.net/yacymc@yacy.net>



YaCy in der Presse und PR-Aktivitäten

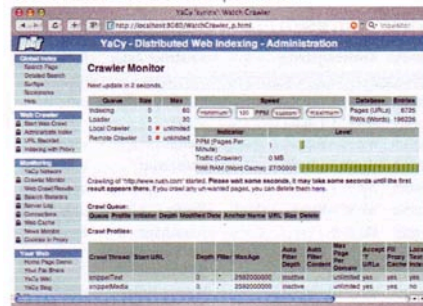


Mai 2007:

Die c't 11/2007 bietet YaCy in ihrer Software-Kollektion auf DVD an. Die Software wird im Heft besprochen.

Gleichzeitig wird YaCy im Heise Software-Archiv in den Top-10 der besten bewerteten Software für MacOS und Linux aufgeführt

c't-Software-Kollektion | Internet



Mit vereinten Kräften: Die P2P-Suchmaschine YaCy indiziert alle geladenen Seiten lokal und teilt ihr Wissen auf Wunsch mit anderen Nutzern.

Die P2P-Suchmaschine YaCy soll eine unabhängige Alternative zu Google und Co. werden – indiziert wird, was die Peer-Betreiber aus dem Netz laden. Den persönlichen Index kann man der Allgemeinheit zur Verfügung stellen. Er wird in Form einer verteilten Hash-Tabelle vorgehalten, aus der man nicht ableiten kann, wer

welche Seiten ursprünglich geladen hatte. Die Installation des eigenen Peers ist einfach, administriert wird über ein Web-Interface. Damit YaCy wirklich floriert, müssen aber noch mehr Peers teilnehmen – die Suchergebnisse sind bislang noch nicht umfassend. Die Suchfunktion kann man auf yacy.net (unter „Demo“) testen. (Tobias Engler/ad)



Windows Mac OS Linux

- Am besten bewertet
1. PuTTY
 2. SelfHTML
 3. Audacity
 4. SciTE
 5. Etherpad
 6. YaCy
 7. Video Disk Recorder...
 8. Python
 9. TestDisk
 10. amaroK
- alle ...

Windows Mac OS Linux

- Am besten bewertet
1. SelfHTML
 2. Audacity
 3. YaCy
 4. Python
 5. TestDisk
 6. POV-Ray
 7. JavE - Java Ascii V...
 8. Devwx
 9. jEdit
 10. FreeMind
- alle ...