
بسم الله الرحمن الرحيم

دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

دانشکده ریاضی و کامپیوتر

پایان‌نامه

برای دریافت درجهٔ کارشناسی ارشد علوم کامپیوتر

گرایش سیستم‌های هوشمند

بررسی استفاده از تکنیک‌هایِ هوش‌مصنوعی برای بهبود نتایج موتورهای جستجو

نگارش

حسن زارعی ماژین

استاد راهنمای اول

جناب آقای دکتر محمد ابراهیم شیری

استاد راهنمای دوم

جناب آقای دکتر احمد عبدالله زادهٔ بارفروش

استاد مشاور

جناب آقای دکتر مجتبی مظفری

زمستان ۱۳۸۴

بسمه تعالی

فرم اطلاعات پایان‌نامه

کارشناسی ارشد و دکترا

دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

معاونت پژوهشی

تاریخ:.....

پیوست:.....

نام و نام‌خانوادگی:حسن زارعی ماژین **دانشجوی: آزاد** (√) **بورسیه** (...) **معادل** (...)

شماره دانشجویی:82113159 **دانشکده**:ریاضی و علوم کامپیوتر **رشته تحصیلی**:علوم کامپیوتر

نام و نام‌خانوادگی استاد راهنما:دکتر محمد ابراهیم شیری

دکتر احمد عبدالله زادهٔ بارفروش

عنوان پایان‌نامه به فارسی:بررسی استفاده از تکنیک های هوش مصنوعی برای بهبود نتایج جستجو

عنوان پایان‌نامه به انگلیسی: Improving Search Results with Data Mining in a Thematic Search Engine applying AI Techniques

نوع پروژه: کارشناسی ارشد(√) / دکترا **کاربردی** (√) **بنیادی** (...) **توسعه‌ای** (...) **نظری** (√)

تاریخ شروع:83/7/1 **تاریخ خاتمه**:84/12/20 **تعداد واحد**: 6

سازمان تأمین کننده اعتبار:دانشگاه صنعتی امیرکبیر(معاونت پژوهشی)

واژه‌های کلیدی به فارسی: خوشه بندی ، رتبه بندی ، پروفایل کاربر ، پرس وجو ،بردار بندی.

واژه‌های کلیدی به انگلیسی: Clustering ,Ranking, User Profile , Query, Vectorization

نظرها و پیشنهادها به منظور بهبود فعالیت‌های پژوهشی دانشگاه:

استاد راهنما:

دانشجو:

امضاء استاد راهنما: **تاریخ**:

نسخه 1: معاونت پژوهشی

نسخه 2: کتابخانه و به انضمام دو جلد پایان‌نامه به منظور تسویه حساب با کتابخانه و مرکز اسناد و مدارک علمی

تقدیم و تشکر

از زحمات استادان گرانقدرم بخصوص دکتر محمد ابراهیم شیری و دکتر احمد عبدالله زادهٔ بارفروش کمال تشکر را دارم که با صبر و حوصله‌ای مثال زدنی در تهیهٔ این پایان نامه یاری‌گر من بودند و از خانوادهٔ رویایی‌ام که همواره در همه حال فرزند خویش را یاریگر بودند، همچون گذشته یاری‌ها و صبوری ها دیدم که در حوصلهٔ انسان‌های خاکی نمی‌گنجد.

تقدیم به آیدا ، بهار ، آرش ، آریا ، معین ، مستانه ، احسان و هوشنگ و مادر و پدرم.

چکیده

بهبود نتایج موتورهای جستجو و رتبه‌بندی آن‌ها با توجه به پروفایل کاربر موضوع اصلی این پایان‌نامه می‌باشد که بدین‌منظور از تکنیک‌های هوش‌مصنوعی مانند الگوریتم‌های ژنتیک و تکنیک‌های فازی برای شاخص‌بندی و رتبه‌بندی در یک موتور جستجو استفاده گشته است.نگارش مستندات به فضای برداری و خوشه بندی مستندات نیز همچون راهکاری برای نزدیک شدن به علایق کاربر و رتبه بندی نتایج پیاده‌سازی گشته است. الگوریتم‌های کاهش ابعاد ماتریس‌های عبارت – مستند و روش LSI به عنوان راهکار ارائه شده برای کاهش پیچیدگی و کشف روابط پنهان مفهومی پیاده سازی گشته و نتایج الگوریتم‌های خوشه‌بندی نیز با معیارهای متعددی ارائه شده‌اند.

فصل اول

موتورهای جستجو

۱موتورهای جستجو

۱,۱ *مقدمه*

موتورهای جستجو امروز همچون قلب تپندهٔ اینترنت عمل می‌کنند و عملاً با افزایش سرسام‌آور تعداد صفحات موجود در وِب، تنها راه کاوش و یافتنِ اطلاعات در این شبکه ، استفاده از موتورهای جستجو می‌باشد. امروزه کمپانی‌هایی که موتورهای جستجوی خود را عرضه کرده‌اند رقابتی سنگین و فشرده را در عرصهٔ عمومی آغاز کرده‌اند که نمونهٔ آن رقابت شرکت‌های گوگل و یاهو می‌باشد که بتازگی مایکروسافت نیز موجودیت خود را در این عرصه شروع کرده‌است و قصد سرمایه‌گذاری وسیع در این حوزه را دارد. شرکت گوگل با الگوریتم انقلابی خود که البته پیشینه‌اش در کارهای موتورهای قبلی به چشم می‌خورد در حال حاضر محبوب‌ترین موتور جستجو می‌باشد و سهام آن ارزشی افسانه‌ای یافته است و این موفقیت مدیون نگرش جدید مؤسسین این موتور به ماهیت مسئله می‌باشد. موتورهای سنتی با الگوریتم‌هایی که مبتنی بر شمارش کلمات کلیدی است پیش می‌روند و این نوع نگرش بتنهایی جوابگویِ نیازهای ذهنی مراجعه‌کنندگان به موتورهای جستجو نیست. در ارائهٔ پاسخها الگوریتم‌های رتبه‌بندی[۱] مطابق خواست کاربر عمل نمی‌کنند و نمونه‌های کم و غیر تجاری‌ای از موتورهای جستجو وجود دارد که به مسئلهٔ شخصی‌سازی[۲] جستجو اهمیت داده‌اند. لذا با توجه به اهمیت مسئله ، ضرورت مطرح گشتن روشهایی نوین در این حوزه آشکارتر می‌گردد.

تاریخچهٔ موتورهای جستجو به اوایل دههٔ ۹۰ می‌رسد که این موتورها بصورت جدی مطرح گشتند البته پیش از آن مبانی این مسئله در مبحث بازیابی اطلاعات[۳] دارای پیشینهٔ کافی بود و در آنجا شاخصی بر روی اطلاعات و مجموعهٔ داده‌ها تعریف می‌گشت و پاسخ‌ها در جواب یک پرس‌وجو بصورت یک مجموعهٔ رتبه‌بندی‌شده بازگردانده می‌شد.

در آغاز تولد این تکنولوژی موتورهای جستجوئی مانند WWWW [۴] بودند که در سال ۱۹۹۴ تنها ۱۱۰،۰۰۰ صفحه را شاخص‌بندی می‌کرد و بنا به‌گزارش سایت‌رسمی Search Engine Watch در سال ۱۹۹۷ موتورهای جستجویِ پیش‌رو مانند Alta Vista تنها تعداد صفحاتی بین ۲ تا ۱۰ میلیون را

---

[۱] Ranking

[۲] personalization

[۳] Information Retrieval

[۴] World Wide Web Worm

شاخص‌بندی می‌کردند. این تعداد در سال ۲۰۰۳ بنا به‌گزارش[۱۰] به حدود ۴ بیلیون رسیده‌است و همچنین تعداد پرس‌وجوها نیز از ۱۵۰۰ عدد در روز در سال ۱۹۹۴ به نزدیک ۲۰ میلیون مورد در روز در سال ۹۷ توسط Alta Vista رسیده‌است که همین تعداد در سال ۲۰۰۳ به نزدیک ۲۰۰ میلیون مورد در روز برای گوگل رسیده است.

اکنون دیگر موتورهای جستجو بعنوان یکی از پیشروترین تکنولوژی‌ها شناخته می‌شوند و شکی نیست که با پشتوانهٔ اقتصادی‌ای که این موتورها خواهند داشت بزودی باید منتظر انقلاب‌های دیگری در این عرصه بود. این فصل به معرفی چند موتور جستجوی مطرح می‌پردازد و وجه تمایز این موتورها را از همدیگر با سهم آنها در عرصهٔ جستجوی اینترنتی بررسی می‌کند.

۱,۲ *گوگل*[۵]

موتورجستجوئی است که استفادهٔ قابل توجهی از ساختار موجود در ابرمتن[۶]ها را می‌کند و رویکردی متفاوت و کاملاً خودکار نسبت به گونه‌نگاری[۷]های دستی را پیش‌رو گرفته‌است که این گونه‌نگاری‌ها از نظر ساخت و نگهداری پرهزینه و مستعد لغزیدن در دام قضاوت شخصی می‌باشند و همچنین توسیع آنها کند و پرهزینه می‌باشد.

گونه‌نگاری‌های دستی تمام موضوعات حاشیه‌ای و تخصصی را نمی پوشانند. در این راستا گوگل معتقد است که تطابق کلمه‌ای[۸] بین پرس‌وجو و صفحات نمی‌تواند جواب صحیح را به کاربر بدهد و دارای کیفیت پایینی می‌باشد[۳۲]. موتورهای جستجوی خودکار نیز در معرض خطر توسط کسانی هستند که به‌منظور تبلیغ ، معیارهای رتبه‌بندیِ[۹] این موتورها را در خود می گنجانند بدون آنکه واقعاً رتبهٔ بالاتری از نظر محتوا داشته باشند. لذا گوگل می‌کوشد که بیشترین استفاده را از ساختارهای اضافی ابرمتن‌ها بکند تا به نتایجی با کیفیت بالاتر برسد.

در آغاز تولد موتورهای جستجو گمان برآن بود که با شاخص‌بندی[۱۰]تمامیِ صفحات وب به یک موتور ایده‌آل می‌رسیم ، اما رشد سریع وب نشان داد که دستیابی به کیفیت بالا در یک موتور جستجو

---

[۵] Google : www.google.com/
[۶] hypertext
[۷] Taxonomy
[۸] Keyword Matching
[۹] ranking
[۱۰] Indexing

تنها وابسته به شاخص‌بندی نیست و نمونهٔ گوگل ثابت کرد که پارامترهای دیگری بسیار نقش تعیین کننده‌تری دارند.

گوگل دارای دو خصوصیت عمده است که آن را در قادر به دستیابی به نتایجی با کیفیت بالاتر می‌کند :

استفاده از ساختار اتصال[۱۱] : این ساختار رتبهٔ کیفیت را برای هر صفحه مشخص می‌کند که این رتبه را PageRank می‌نامند. از این ساختار می‌توان برای بهبود نتایج نیز سود جست [۳۲].

متن لنگری[۱۲]: در موتورهای قبل از گوگل متن یک اتصال[۱۳]را جزو صفحه‌ای که در آن گنجانده‌شده بود به حساب می‌آوردند ، اما بدلیل اینکه این متن کوتاه توصیف دقیق‌تری از صفحه‌ای که به‌آن اشاره می‌کند را می‌دهد و در بسیاری از موارد آن صفحه ممکن است فاقد اطلاعات متنی باشد، گوگل تصمیم برآن گرفت که این متن‌های لنگری را جزو صفحهٔ مقصد در نظر بگیرد[۳۲]. این امر ممکن است منجر به آن شود که صفحاتی که دیگر در وب نیستند و بنا بدلایلی حذف گشته‌اند و تنها یک اتصال خطابِ به آنها وجود دارند بعنوان نتایج جستجو بازگردانیده شوند اما با مرتب کردن نتایج گوگل تا حد زیادی این مشکل را مرتفع کرده است.

۱,۲,۱ صفحه رتبه[۱۴]

این استراتژی که با تصوری که انسان از اهمیت[۱۵]دارد همخوان است بعنوان یک برگ برندهٔ گوگل بسیار کارآ در عمل نشان داده‌است[۳۲]. در این استراتژی رتبهٔ یک صفحه در دو صورت بالا می‌رود :

صفحات زیادی با رتبهٔ کم به آن اشاره کنند.

صفحاتی معدود اما با رتبهٔ بالا به این صفحه اشاره کنند. مثلاً اشارهٔ سایت یاهو به یک صفحه دلیل بر اهمیت فراوانِ آن صفحه می‌باشد و رتبهٔ آن را به شدت بالا می‌برد.

در فصل دوم جزئیات این الگوریتم ارائه می‌گردد.

---

[۱۱] Link Structure
[۱۲] Anchor Text
[۱۳] Link
[۱۴] Page Rank
[۱۵] Importance

۱,۲,۲ جنبه‌های دیگر گوگل

گوگل علاوه بر استفاده از متن‌های لنگری و PageRank خصوصیات ویژهٔ دیگری نیز دارد که بعنوان مثال در تحلیل متن یک صفحه اطلاعات در مورد مکان هر کلمه و نوع فونت آن را نیز همچون پارامترهایی مهم نگاه می‌دارد و از آن استفادهٔ وسیعی در جستجوهایش می‌کند. در ضمن گوگل ابرمتن خامِ تمامی‌صفحات را برای رجوع به آنها با الگوریتم‌های خاصی فشرده‌می‌کند و همواره نگهداری می‌کند[۳۱].گوگل درآمدهای خود را از تبلیغات سایت‌ها بدست می‌آورد که بصورتی تقریباً نامزاحم در سمت راست نتایج و مرتبط با زمینهٔ جستجو می‌آیند.

۱,۳ *یاهو*[۱۶]

این موتور جستجو که اشتهار خود را علاوه بر قابلیت جستجویش مدیون سرویس پست الکترونیکی‌اش نیز می‌باشد همواره یکی از بازیگران اصلی در عرصهٔ جستجویِ اینترنتی بوده‌است و اگر بتوان یک رقیب تجاریِ عمده برای گوگل برشمرد قطعاً غیر از یاهو نخواهد بود.

تاریخچهٔ این موتور محبوب را می‌توان به سه قسمت مجزا تقسیم کرد :

یاهو قبل از گوگل که در سال ۱۹۹۴ تأسیس شدو مبتنی بر گونه‌نگاریِ دستی بود که در آن ویراستاران مجموعهٔ صفحات را در مقولات می‌گنجاندند.

یاهو باگوگل که در سال ۲۰۰۲ بر آن شدند که سیستم خود را مبتنی بر خزشگر[۱۷] بنمایند و این نتایج تا سال ۲۰۰۴ از گوگل گرفته می‌شدند و در یاهو به نمایش در می‌آمدند.

یاهو بدون گوگل که از سال ۲۰۰۴ به بعد یاهو از تکنولوژی خاص خود برای خزش[۱۸] در وب استفاده می‌کند.

علاوه بر نتایج جستجویِ با کیفیت بالا می‌توان از انواع جستجوهایِ دیگر یاهو نیز استفاده کرد که شامل جستجو در تصاویر و تکه‌های فیلم و صفحاتِ زرد[۱۹] می‌باشد. با وجود اینکه یاهو به تکنولوژی مبتنی بر خزش روی‌آورده است اما هنوز گونه‌نگاری یاهو به حیات خود ادامه می‌دهد و از طریق صفحهٔ

---

[۱۶] Yahoo :www.yahoo.com/

[۱۷] Crawler based

[۱۸] Crawling

[۱۹] Yellow Pages

خانگی یاهو[۲۰] قابل دستیابی است. سایتهای تجاری برای اینکه در لیست تجاری و مقولاتِ تجاری یاهو گنجانده شوند مبلغی را پرداخت می‌کنند ، اما در هر صورت باید قبلاً صلاحیت آنها نیز توسط مسئولین این امر در یاهو تأیید گردد. برنامهٔ CAP [۲۱] یاهو با دریافت مبلغی سایت‌ها را در نتایج خزشی نیز قرار می‌دهد اما رتبهٔ آنها را تضمین نمی‌کند. یاهو با استفاده از برنامهٔ Overture که از سال ۲۰۰۳ با خرید شرکتی به همین نام در خود گنجاند مکان‌های تبلیغاتی را در صفحهٔ اصلیِ خود به فروش می‌رساند که در واقع شبیه استراتژی گوگل در این زمینه می‌باشد. شرکت اوِرتور قبل از آنکه جزئی از یاهو گردد خود چند شرکت مهم دیگر مانند Altavista و AllWeb را خریداری کرده بود که یاهو با این خرید عملاً تکنولوژی‌های این شرکت‌ها را نیز مالک شد.

خزشگر یاهو مبتنی بر Inktomi می‌باشد که در دانشگاه برکلی توسعه داده شد و بنیانی شد برای خزشگر مستقل یاهو نسبت به گوگل.

۱,۳,۱,۱ قابلیت‌های دیگر یاهو

در کنار جستجویِ کلمات کلیدی یاهو خیز بلندی بسوی شخصی‌سازی[۲۲] برداشته است که در پورتال My Yahoo می‌توان پیشرفت‌هایی محسوس با گزینه‌هایی فراوان مشاهده کرد. با استفاده از تکنولوژی جستجو یاهو به محدودهٔ خرید[۲۳] نظر دوخته است که با استفاده از Smart Sort بسیاری از کالاهای الکترونیک اکنون با استفاده از این سرویس موتور یاهو براحتی با واردکردن مشخصات محدوده و سایر اطلاعات قابل خرید می‌باشند و این سرویس در عمل بسیار موفق نشان داده است.

۱,۳,۲ موتور Ask Jeeves[۲۴]

این موتور ابتدا در سال ۱۹۹۸ و ۱۹۹۹ به شهرت رسید که عمدهٔ شهرتش بواسطه قابلیت پردازش زبان طبیعیِ آن بود. در واقع در پسِ پردهٔ این موتور تکنولوژیِ پیشرفته‌ای دست‌اندرکار نبود بلکه حدود ۱۰۰ ویرایشگر بصورت دستی همگانی‌ترین پرسش‌ها را انتخاب کرده و جواب‌های مناسب را برای آنها فراهم می‌کردند. اما امروز این موتور به سیستم مبتنی بر خزشگر روی آورده‌است که آن نیز نتایجش را از Teoma می‌گیرد که اکنون قسمتی از این شرکت گشته است.

---

[۲۰] Dir.yahoo.com/

[۲۱] Content acquisition program

[۲۲] Personalization

[۲۳] Shopping

[۲۴] http://www.askjeeves.com/

AskJeeves دو خصوصیت عمدهٔ Tab های نامرئی و SmartSearch را در خود دارد که به‌نظر می‌رسد که خصوصیاتی متعلق به موتورهای نسل آینده باشند[۳۳]. این موتور نتایج اصلی خود را از Teoma می‌گیرد و نتایج پولی را از گوگل دریافت می‌کند.

۱,۳,۳ موتور AOL [۲۵]

این موتور جستجو دارای دو نسخهٔ داخلی و خارجی است که هر دو نسخه نتایج خود را از گوگل دریافت می‌کنند ولی نسخهٔ داخلی دسترسی کامل به اطلاعات داخل AOL دارد و این نسخه تنها برای اعضایِ آن قابل دسترسی می‌باشد. در واقع با وجود گوگل نیازی به نسخهٔ خارجیِ این موتور احساس نمی‌شود[۲۸] و تنها به سلیقهٔ کاربر بستگی دارد و علاوه بر این بسیاری از قابلیت‌های نتایج گوگل در نتایج AOL دیده نمی‌شود که بعنوان مثال گزینهٔ cached در نتایج AOL قابل مشاهده و دسترسی نیست.

۱,۳,۴ موتور AllTheWeb [۲۶]

این موتور که توسط یاهو تغذیه می‌گردد نسخه‌ای سبکتر و قابل‌انتخاب[۲۷]تر در مقایسه با یاهو می‌باشد که به جستجوی وِب در قیاس با جستجویِ فایلهای صوتی و تصویری یا FTP ، تمایل بیشتری دارد.

این موتور قبلاً بعنوان دارائی شرکتFAST وجود داشت و همچون نمونهٔ نمایشی برای جستجوی وِب ارائه شده‌بود. بنابراین ممکن است بعنوان FAST نیز به آن اشاره شود . سرانجام در آوریل ۲۰۰۳ توسط Overture خریداری شد که بعدها نیز جزئی از یاهو شد[۳۶].

۱,۳,۵ موتور Hot Bot [۲۸]

با وجود اینکه این موتور یک اَبرموتور[۲۹] نمی‌باشد اما نتایج سه موتور خزشگر عمده را مانند یاهو، گوگل و Teoma را قابل دسترسی آسان می‌نماید. این موتور نتایج را باهم نمی‌آمیزد ولی با این‌وجود راهی آسان برای گرفتن نتایج از موتورهایی که به آنها اشاره شد می‌باشد. این موتو که از سال

---

[۲۵] American Online : www.aol.com/ also http://aolsearch.aol.com/ (internal) http://search.aol.com/ (external)

[۲۶] www.allheweb.com

[۲۷] customizable

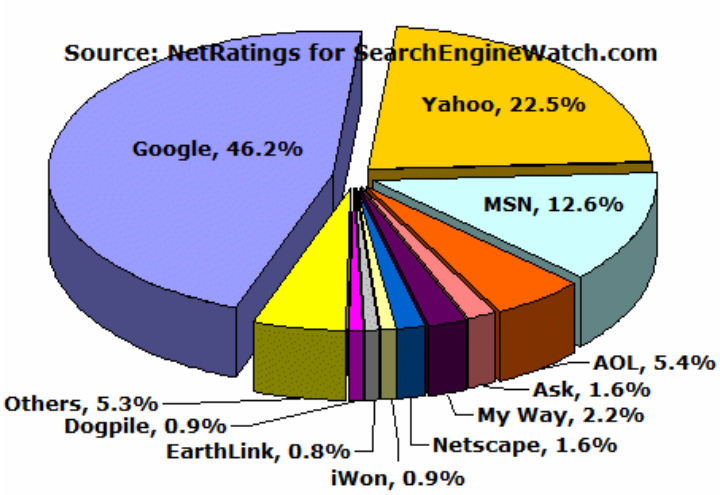
[۲۸] www.hotbot.com

[۲۹] Meta Search Engine

۱۹۹۶ پا به عرصه نهاد در آغاز بین افراد حرفه‌ایِ این حوزه به مقبولیت خوبی دست یافت که به خاطر کیفیت بالا و جامعیت نتایجی بود که توسط Inktomi فراهم می‌گشتند و در ضمن ظاهر و رنگهای غیرعادیِ آن نیز باعث جلب توجه بسیاری از کاربران می‌شد[۳۳].

۱,۴ *سهم موتورهای جستجو*

بنابر شکل ۱-۱,۴ که از سایت دیده‌بانِ موتورهای جستجو[۳۰] برگرفته شده‌است مشاهده می‌گردد که در حدود نصف جستجوها توسط موتور گوگل انجام می‌گردد و پس از آن یاهو و MSN دارای سهمی تقریباً مشابهند که با توجه به اینکه نتایج موتورهای سهیم باقیمانده اکثراً طبق شکل ۱-۱,۷ توسط گوگل تغذیه می‌شود می‌توان جستجویِ اینترنتی را در انحصار گوگل دانست که البته با برنامه‌های بلندپروازانه‌ای که یاهو و مایکروسافت در این حوزه ارائه کرده‌اند انتظار می‌رود که بزودی بتوان شاهد تقسیم‌بندیِ متفاوتی نسبت به شکل ۱-۱,۴ بود.



شکل ۱-۱,۴ : بررسی سهم موتورهای جستجو در اینترنت [۲۸]

۱,۵ *تغذیهٔ موتورهای جستجو از همدیگر*

موتورهایی که اعمال شاخص‌بندی و جستجو را با تکنولوژی و امکانات ویژهٔ خود انجام می‌دهند معدودند. درشکل ۱-۱,۷ و شکل ۲-۱,۷ می‌توان مشاهده کرد که یاهو و گوگل دو منبع اصلی تغذیهٔ موتورهای دیگر هستند که در این میان نقش گوگل به‌مراتب پررنگ‌تر می‌باشد. علاوه بر آن ODP [۳۱]

---

[۳۰] Search Engine Watch

[۳۱] Open Directory Project

بعنوان یک منبع اصلی گونه‌نگاری برای گوگل و وابستگانش مانند Netscape و AOL مورد استفاده قرار می‌گیرد. در دو شکل مذکور گوگل و یاهو بصورت مجزا به نمایش در آمده‌اند هر چند که در دورهٔ دوم یاهو همانگونه که ذکر آن رفت یاهو از تکنولوژی خزشگر گوگل استفاده می‌کرد ولی اکنون یاهو با ادغام در چند شرکت دیگر که بیان شد ، از تکنولوژی و تکنیک‌های مستقل و خاص خود سود می‌جوید.

چون موتورهای جستجو نتایج را از منابع متعددی نمایش می‌دهند معمولاً نتایج یک منبع خاص بر منابع دیگر تسلط دارد. بعنوان مثال در گوگل نتایج اصلی عمدتاً نتایجی هستند که از خزشگر خاص گوگل بدست آمده‌اند.

نتایجی که یک موتور جستجو ارائه می‌کند از انواع زیر می‌توانند باشند :

- *نتایج حاصل از خزشگر* : که با خزش در کل وِب بدست می‌آیند که نمونهٔ آن گوگل و Overture می‌باشند.
- *تلاش انسان* : که نتایج اصلی با استفاده از فهرست‌ها و دایرکتوریهایی که بدست انسان ساخته می‌شوند بدست می‌آیند که نمونهٔ آن ODP یا نمونهٔ قبل از گوگلِ یاهو می‌باشند.
- *نتایج پولی :* که بسته به مبلغی که بجهت مکان‌دهیِ آنها پرداخت می‌شود با ترتیب به کاربر ارائه می‌شون که اکثریت قریب به اتفاق موتورهای عمده چنین نتایجی را نیز به موازات نتایج اصلیِ خود ارائه می‌کنند.

۱,۶ *نتایج انسانی*

بسیاری از موتورهای جستجو علاوه بر نتایج بدست آمده از خزشِ وِب نتایجی که از گونه‌نگاری‌های دستی و دسته‌بندی‌هایی که توسط انسان انجام می‌گردد، را نیز فراهم می‌سازند که در جدولِ شکل ۱۱,۸- موتوری که این نتایج را فراهم می‌سازد در یک ستون ظاهر گشته‌اند.

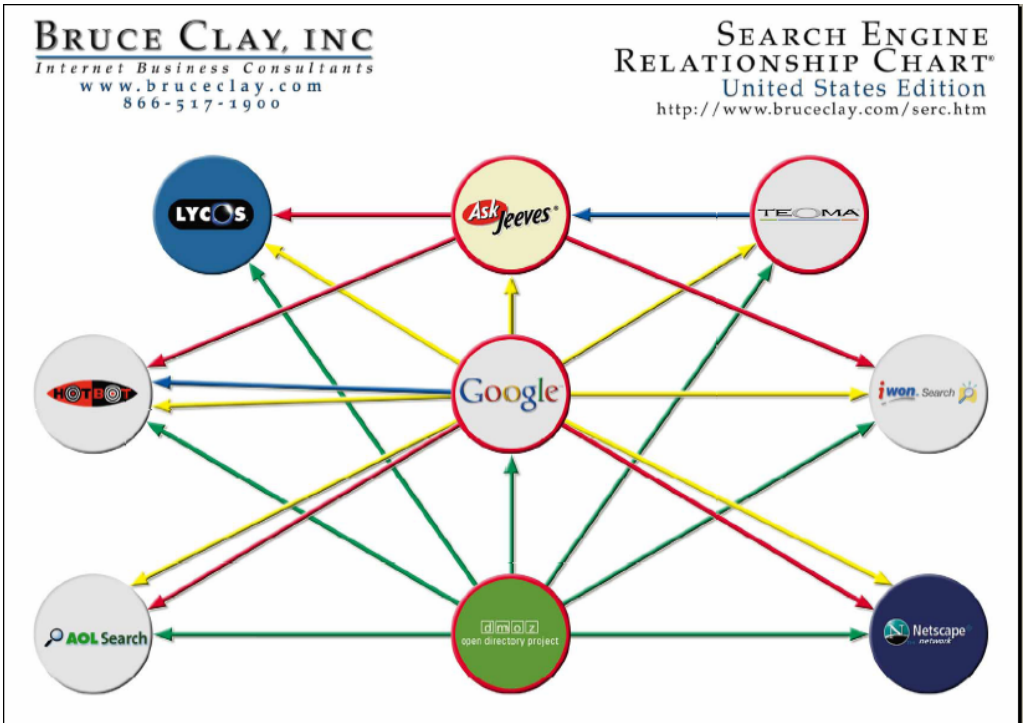
برای هر موتوری که از نتایج دسته‌بندیِ انسانی استفاده می‌کند لازم است که یک نسخهٔ فراهمگر[۳۲] مبتنی بر خزش نیز بعنوان پشتیبان در دسترس باشد تا در مواقعی که دسته‌بندی انسانی جواب نمی‌دهد به این نسخه، ارجاع شود. بعنوان مثال اگر یک جستجو در Lycos از یافتن نتیجهٔ

[۳۲] Provider

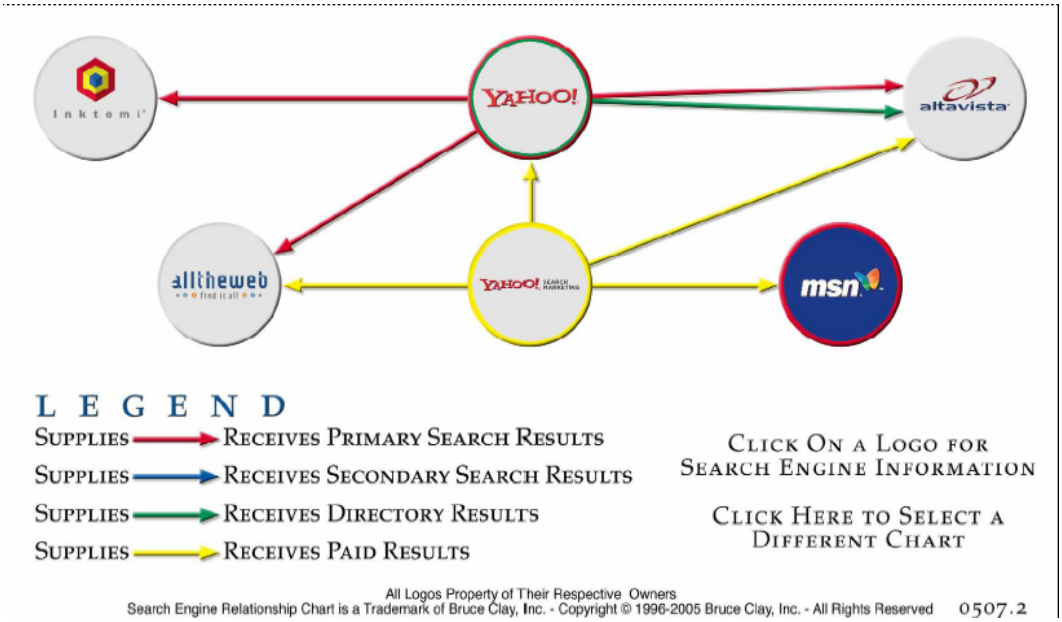
مناسب عاجز بماند آنگاه یاهو نتایج مورد نظر را با استفاده از خزشگرِ خود به Lycos ارائه می‌کند. بعضی نیز مانندOverture هر دونوع نتیجه را خود فراهم می‌سازند.

۱,۷ *نتایج پولی*

تمامیِ موتورهای جستجویِ عمده به‌موازات نتایج اصلی‌شان یک فهرست پولی نیز ارائه می‌کنند که در جدولشکل ۳۱,۷- فراهمگر این نتایج نیز در یک ستون معرفی گشته‌اند. بعنوان مثال Overture برای بسیاری از همکارهای خویش چنین نتایجی را فراهم می‌سازد.



شکل ۱,۷-۱ بررسی تغذیهٔ موتورها از یکدیگر مربوط به گوگل[۳۴]



شکل ۱,۷-۲ : تغذیهٔ موتورها مربوط به یاهو [۳۴]

درشکل ۱۱,۸- می‌توان جدولی که نمایش دهندهٔ موتورهای جستجوی عمده و منابع تغذیه‌شان می‌باشد را مشاهده کرد. در این جدول ده موتور عمده به نمایش درآمده‌اند که در یک ستون نوع تکنولوژی جستجوی آنها و در ستون‌های دیگر منابع تغذیهٔ آنها در مورد نتایج پولی یا نتایج دسته‌بندی و منبع فراهم کننده برای نتایج اصلی آنها آمده است[۳۶].