



پایان نامه برای اخذ درجه کارشناسی ارشد

دانشکده علوم اجتماعی و اقتصادی

رشته مدیریت فناوری اطلاعات

عنوان:

شناسایی الگوهای رفتار جوییش اطلاعات کاربران دانشگاهی ایران در اینترنت

استاد راهنما:

جناب آقای دکتر غلامعلی منتظر

استاد مشاور:

سرکار خانم دکتر معصومه حسین زاده شهری

تحقیق و نگارش:

منا محبی عدل

mona.mohebi@gmail.com

تیرماه ۱۳۸۷

بسمه تعالی

دانشگاه: دانشگاه الزهراء،

دانشکده: دانشکده علوم اجتماعی و اقتصادی

رشته تحصیلی: مدیریت فناوری اطلاعات

عنوان: شناسایی الگوهای رفتار جویس اطلاعات کاربران دانشگاهی در اینترنت

استاد راهنما: جناب آقای دکتر غلامعلی منتظر

دانشجو: منا محبی عدل

این پروژه تحت قرارداد پژوهشی شماره ۵۰۰/۶۲۶۰/ت مورخ ۸۷/۴/۲۴ از پشتیبانی معنوی و مادی

مرکز تحقیقات مخابرات ایران بهره‌مند شده است.

تقدیم به

پدر بزرگوار و مهربانم به پاس حمایت‌ها، رهنمودها و دوستی‌های همیشگی‌اش

و مادرم که حضورش جاودان و روشنی‌بخش است.

سپاس از

خداوند جان و خرد که به من جان داد تا خرد ورزم.

پدر و مادر بزرگوام به پاس حمایت‌ها و دلگرمی‌هایشان.

استاد ارجمندم جناب آقای دکتر غلامعلی منتظر که راه صحیح علم‌آموزی و پژوهش را به من آموختند و راهنمایی

این پژوهش را صبورانه پذیرا شدند.

استاد عزیزم سرکار خانم دکتر معصومه حسین‌زاده که رهنمودها و همراهی‌های ایشان همواره روشنی‌بخش بود.

و استاد گرانقدرم سرکار خانم دکتر فریبا لطیفی که بسیار نکته‌ها از ایشان آموختم و شاگردی ایشان از افتخارات

من است.

چکیده

هدف از این پژوهش شناسایی الگوهای رفتاری کاربران اینترنت در دانشگاه‌های ایران است. از این‌رو رفتار کاربران از هنگام انتخاب ابزار کاوش، فرموله‌سازی پرسمان جستجو، بررسی و بازدید از نتایج ارائه شده توسط موتور کاوش، فرموله‌بندی مجدد پرسمان جستجو، انواع حرکت‌ها، راهبردها و الگوهای رفتاری کاربران در هنگام کاوش در وب و نیز عوامل مؤثر در میزان موفقیت جستجو و رضایت کاربران از نتایج حاصل، مورد بررسی قرار گرفته است. بدین منظور پس از مطالعه و بررسی تحقیقات مشابه در سایر نقاط جهان، مهمترین متغیرها و عوامل تأثیرگذار در موفقیت جستجو و حصول رضایت کاربران شناسایی شده و بر اساس این متغیرها، پرسشنامه‌ای شامل هفت بخش و ۹۰ پرسش تخصصی در زمینه رفتار جستجوی اطلاعات کاربران در وب طراحی و توزیع شده است. جامعه نمونه بر اساس روش نمونه‌گیری خوشه‌ای طبقه‌بندی شده از میان دانشجویان و استادان هشت دانشگاه معتبر ایران (تهران و شهرستانها) در مقاطع و رشته‌های مختلف تحصیلی برگزیده شده است. در نهایت ۷۹۳ پاسخنامه مورد بررسی قرار گرفت. متغیرهای مورد بررسی در این تحقیق عبارتند از: میزان آشنایی با ابزارها و خدمات مختلف کاوش در وب، انتخاب ابزارهای مناسب کاوش، آشنایی کافی با روش‌های فرمولبندی پرسش جستجو، طرح پرسش‌های مناسب و واضح، صرف وقت و توجه کافی و معیارهای صحیح در انتخاب پیوندهای مرتبط با موضوع جستجو. نتایج نشان می‌دهد بین متغیرهای یادشده با متغیر میزان تجربه کاری و تجربه جستجو در اینترنت، و موفقیت و رضایت کاربران از جستجو در اینترنت رابطه معناداری وجود دارد.

کلیدواژه‌ها: رفتار کاربران وب، جستجو و بازیابی اطلاعات، فرمولبندی پرسش، رضایت جستجوگران وب، راهبردهای جستجوی وب، وب‌کاوی

Abstract

The purpose of this research is the recognition of collegiate information searcher behavior on Internet in Iran. Therefore, we surveyed on user's behaviors while the searching on the Internet step by step, such as selection of search tools and using them, query formulation, surveying and visiting of search's result pages, query reformulation method and all the types of moves, strategies and patterns behavior which they use while searching. Also we surveyed on the critical factors and variables which they affect on amount of successfulness of the search and user satisfactory. So, after studying and surveying on the same researches on the other points of world, we recognized the main factors and variables which are affective on successful search and user satisfactory. Then we designed and distributed a questionnaire based on the mentioned variables. The questioner consists of seven parts which they cover 90 special questions about users' behavior while they search on the web. Sample community has selected based on categorized cluster sampling method, and the questionnaire has distributed between some of the students and faculty members of 8 main universities of Iran in different grade and disciplines. Finally, we've done our research based on 793 filled questionnaires. Our variables in this research are: amount of user's knowledge about search tools and their different services, selection of suitable search tools, enough knowledge about query formulation methods, designing the suitable and clear queries, spending enough time and attention while selecting the relative links on the result pages and having the correct criteria about that.

Key words: web searcher behavior, information retrievals, query formulation, searcher satisfactions, information search tools, web search strategies, web mining.

فصل اول: کلیات پژوهش.....

- ۱.۱: مقدمه..... ۱
- ۲.۱: مساله پژوهش..... ۳
- ۳.۱: ضرورت و اهمیت پژوهش..... ۴
- ۴.۱: اهداف پژوهش..... ۵
- ۵.۱: فرضیات پژوهش..... ۵
- ۱.۵.۱: پرسش‌های پژوهش..... ۵
- ۲.۵.۱: فرضیه‌های پژوهش..... ۶
- ۶.۱: متغیرهای پژوهش..... ۶
- ۱.۶.۱: متغیرهای وابسته..... ۶
- ۲.۶.۱: متغیرهای مستقل..... ۶
- ۷.۱: سازمان‌هایی که از نتایج کاربردی پژوهش استفاده خواهند کرد..... ۷
- ۸.۱: اطلاعات مربوط به روش‌شناسی پژوهش..... ۷
- ۱.۸.۱: تعریف جامعه آماری و ویژگی‌های آن..... ۷
- ۲.۸.۱: روش نمونه‌گیری و حجم نمونه..... ۷
- ۳.۸.۱: ابزار سنجش، مقیاس‌های سنجش (پایایی و اعتبار)..... ۷
- ۴.۸.۱: روش تجزیه و تحلیل داده‌ها..... ۷
- ۹.۱: تعریف واژگان کلیدی..... ۷

فصل دوم: منابع موضوع پژوهش.....

- ۱.۲: پیشگفتار..... ۱۰
- ۲.۲: وب و اینترنت..... ۱۱
- ۱.۲.۲: نحوه کارکرد وب..... ۱۳
- ۲.۲.۲: انواع صفحات و محتوای وب..... ۱۴
- ۳.۲.۲: وب نامرئی..... ۱۵
- ۴.۲.۲: وب ۲..... ۱۷
- ۵.۲.۲: آمار رشد وب و اندازه اطلاعات در اینترنت..... ۱۸
- ۶.۲.۲: استفاده از اینترنت در ایران..... ۲۱
- ۳.۲: وب‌کاوی و بازیابی اطلاعات..... ۲۲
- ۱.۳.۲: مقدمه..... ۲۲
- ۲.۳.۲: محتوا کاوی وب..... ۲۳
- ۳.۳.۲: ساختار کاوی وب..... ۲۴
- ۴.۳.۲: کاربرد کاوی وب..... ۲۵
- ۵.۳.۲: تمایز بین سه دسته محتوا، ساختار و کاربرد کاوی وب..... ۲۵

۲۶	۶.۳.۲: بازیابی اطلاعات
۲۷	۷.۳.۲: ارتباط پژوهش حاضر با وب‌کاوی
۲۸	۴.۲: ابزارهای کاوش اطلاعات در وب
۲۸	۱.۴.۲: شیوه‌های بازیابی اطلاعات در وب
۳۰	۲.۴.۲: انواع ابزارهای کاوش در وب
۳۵	۳.۴.۲: طرز کار موتورهای جستجو
۳۷	۴.۴.۲: معیارهای ارزیابی موتورهای جستجو
۴۰	۵.۲: رفتارهای بازیابی اطلاعات کاربران
۴۰	۱.۵.۲: مقدمه
۴۲	۲.۵.۲: عوامل موثر بر رفتارهای بازیابی اطلاعات کاربران
۴۵	۳.۵.۲: تحقیقات انجام شده در زمینه شناسایی رفتار جستجوگران وب
۴۵	الف: مدل‌سازی رفتار کاربران و روشهای جستجو
۴۷	ب: بررسی رفتار جوییش اطلاعات کاربران یک پایگاه اطلاعاتی خاص
۵۲	ج: فرمول‌بندی پرسمان، تاکتیک‌ها و راهبردهای جستجو
۵۶	د: میزان موفقیت جستجو
۵۷	و: تحقیقات کلی
۵۹	ط: نتایج آماری
۶۰	۴.۵.۲: بررسی یافته‌های پژوهش‌های انجام شده در زمینه جستجوی وب
۶۰	الف: راهبردهای بازیابی اطلاعات کاربران
۶۸	ب: فرموله‌سازی پرسمان جستجو در وب
۷۳	ج: تاکتیک‌ها و حرکت‌های کاربران
۷۷	۵.۵.۲: نتیجه‌گیری

فصل سوم: روش پژوهش

۷۸	۱.۳: مقدمه
۷۸	۲.۳: روش پژوهش
۸۰	۳.۳: جامعه آماری و ویژگی‌های آن
۸۱	۴.۳: ابزار اندازه‌گیری، طراحی پرسشنامه و نحوه گردآوری داده‌ها
۸۵	۵.۳: تعیین روایی و پایایی ابزار اندازه‌گیری
۸۶	۶.۳: روش آماری تجزیه و تحلیل داده‌ها

فصل چهارم: یافته‌های پژوهش

۸۷	۱.۴: مقدمه
۸۷	۲.۴: توصیف نمونه آماری
۹۰	۳.۴: بررسی فراوانی هر یک از سوالات پرسشنامه
۱۰۶	۴.۴: آزمون فرضیه‌های پژوهش
۱۰۶	۱.۴.۴: مقدمه
۱۰۶	۲.۴.۴: روش تجزیه و تحلیل و تفسیر داده‌ها
۱۰۷	۳.۴.۴: آزمون فرضیات تحقیق

۴.۴.۴: آزمون برخی از سوالات و جداول پرسشنامه ۱۱۱

۵.۴.۴: بررسی و آزمون برخی روابط بین متغیرها ۱۱۷

فصل پنجم: نتیجه‌گیری و پیشنهادها

۱.۵: مقدمه ۱۱۸

۲.۵: نتایج پژوهش ۱۱۸

۳.۵: پیشنهاد برای تحقیقات آینده ۱۲۴

۴.۵: محدودیت‌های تحقیق ۱۲۴

۵.۵: قوت‌های تحقیق ۱۲۵

۶.۵: پیشنهادات کاربردی ۱۲۵

منابع و مأخذ انگلیسی ۱۲۶

منابع و مأخذ فارسی ۱۳۰

پیوست‌ها ۱۳۱

فهرست جداول

- جدول ۱-۲: مقایسه صفحات وب ایستا و پویا ۱۴
- جدول ۲-۲: توزیع موضوعی اطلاعات روی وب ۱۵
- جدول ۳-۲: تعداد صفحات وب نمایه شده توسط هر یک از موتورهای جستجو تا نوامبر سال ۲۰۰۴ ۲۰
- جدول ۴-۲: تمایز بین دسته‌بندیهای وب‌کاوی ۲۶
- جدول ۵-۲: بررسی رفتارجوییش اطلاعات کاربران یک پایگاه اطلاعاتی خاص ۴۷
- جدول ۶-۲: بررسی رفتارجوییش اطلاعات کاربران از دیدگاه فرمول‌بندی پرسمان، تاکتیک‌ها و راهبردهای جستجو ۵۲
- جدول ۷-۲: بررسی رفتارجوییش اطلاعات کاربران از دیدگاه کلی ۵۷
- جدول ۸-۲: راهبردهای جستجو ۶۷
- جدول ۱-۳: برخی یافته‌های جدول مورگان در مورد حجم نمونه آماری ۸۱
- جدول ۲-۳: متغیرهای مورد بررسی در سوالات پرسشنامه و پژوهش‌های انجام شده مرتبط با آنها ۸۲
- جدول ۱-۴: توزیع فراوانی میزان تحصیلات پاسخ دهندگان ۸۷
- جدول ۲-۴: تنوع رشته‌های تحصیلی پاسخ دهندگان ۸۹
- جدول ۳-۴: تنوع دانشگاه‌های مورد بررسی ۹۰
- جدول ۴-۴: شرح مقادیر فراوانی‌ها و نتایج به دست آمده از هر یک از سوالات پرسشنامه ۹۱
- جدول ۱-۵: یافته‌های پژوهش بر اساس مقادیر فراوانی‌های به دست آمده از پرسشنامه ۱۲۰

فهرست نمودارها و شکل‌ها

- شکل ۱-۲: آمار وبگاه‌ها از سال ۱۹۹۵ تا ۲۰۰۶ ۱۹
- شکل ۲-۲: درصد استفاده از موتورهای کاوش مشهور در سال ۲۰۰۶ ۱۹
- شکل ۳-۲: میزان همپوشانی نتایج قابل ارائه توسط موتورهای جستجو ۲۰
- شکل ۴-۲: دسته بندی مجموعه روش‌های وب کاوی ۲۳
- شکل ۵-۲: دسته‌بندی‌های مرتبط با زیرشاخه بازبایی اطلاعات ۲۷
- شکل ۶-۲: مدل فرموله سازی مجدد پرسش جستجو ۷۰
- شکل ۷-۲: نمونه‌ای از حرکت چشم کاربران در بازدید از صفحه نتایج موتور جستجو ۷۵
- نمودار ۱-۴: تعیین جنسیت پاسخ دهندگان ۸۷
- نمودار ۲-۴: درصد فراوانی موقعیت دانشگاهی پاسخ دهندگان (دانشجو، استاد، دانش آموخته) ۸۸
- نمودار ۳-۴: گستره سنی پاسخ دهندگان ۸۸
- نمودار ۴-۴: گروه رشته‌های تحصیلی پاسخ‌دهندگان ۸۸
- نمودار ۵-۴: میزان استفاده از انواع ابزارهای کاوش وب ۹۹
- نمودار ۶-۴: میزان استفاده از پایگاه‌های اطلاعاتی وب ۹۹
- نمودار ۷-۴: میزان استفاده از سرویس‌های ارائه شده توسط موتورهای کاوش ۱۰۰
- نمودار ۸-۴: میزان استفاده از سرویس‌های خبری روی وب ۱۰۰
- نمودار ۹-۴: میزان استفاده از تاکتیک‌های مختلف در طرح پرسمان جستجو ۱۰۰
- نمودار ۱۰-۴: انواع تغییرات اعمال شده در پرسمان پس از بازبایی ناموفق اطلاعات ۱۰۱
- نمودار ۱۱-۴: انواع تاکتیک‌ها در طرح پرسش‌های چندمفهومی با کلیدواژه‌های مختلف ۱۰۱
- نمودار ۱۲-۴: محتوای موضوعی اطلاعات مورد جستجو در وب ۱۰۱
- نمودار ۱۳-۴: دلایل رجوع به وب ۱۰۲
- نمودار ۱۴-۴: میزان بازدید از صفحات نتایج ارائه شده توسط موتورهای کاوش وب ۱۰۲
- نمودار ۱۵-۴: معیارهای انتخاب پیوندهای مرتبط در صفحه نتایج موتور جستجو ۱۰۲
- نمودار ۱۶-۴: متوسط زمان گذرانده شده برای انتخاب پیوند مرتبط در صفحه نتایج ۱۰۳
- نمودار ۱۷-۴: زمان گذرانده شده برای جستجوی اطلاعات مطلوب در یک بار اتصال به اینترنت ۱۰۳
- نمودار ۱۸-۴: استراتژی‌های محدود کردن حیطه جستجو ۱۰۳
- نمودار ۱۹-۴: استراتژی‌های جستجو ۱۰۴
- نمودار ۲۰-۴: برخی تاکتیک‌های جستجو ۱۰۴
- نمودار ۲۱-۴: میزان تجربه پاسخ دهندگان در جستجو در اینترنت ۱۰۴
- نمودار ۲۲-۴: میزان استفاده پاسخ دهندگان از اینترنت در هفته ۱۰۵
- نمودار ۲۳-۴: فرآیند جستجوی اطلاعات مطلوب در اینترنت آسان است یا پیچیده ۱۰۵
- نمودار ۲۴-۴: میزان رضایت پاسخ دهندگان از روشهای جستجوی اطلاعات خود در وب ۱۰۵
- نمودار ۲۵-۴: معیارهای رضایت کاربران از موتورهای کاوش ۱۰۶

فصل اول

کلیات پژوهش

فصل دوم

مرور ادبیات پژوهش

فصل سوم

روش پژوهش

فصل چهارم

یافته‌های پژوهش

فصل پنجم

نتیجه گیری و پیشنهادها

پیوست‌ها

۱.۱: مقدمه

در حال حاضر اینترنت یکی از منابع مهم اطلاعات توزیع شده است و موتورهای جستجو ابزار کاوش و بازیابی اطلاعات^۱ در این شبکه جهانی هستند. با توجه به حجم وسیع اطلاعات روی اینترنت، این شبکه اطلاعاتی جهانی به عنوان منبع مهمی برای جستجوی اطلاعات، کسب و کار و رفع نیازهای اطلاعاتی افراد مورد توجه قرار گرفته است. خصوصاً در مجامع علمی و دانشگاهی استفاده از اینترنت به عنوان یک منبع عظیم اطلاعات علمی مورد توجه و استفاده قرار دارد. از این رو بررسی این مساله مهم است که آیا اصولاً اینترنت می‌تواند به عنوان یک منبع قابل اعتماد در جایگاه اول منابع مورد استفاده افراد متخصص در جستجوی اطلاعات قرار گیرد؟ محتوای اطلاعات موجود و در دسترس در اینترنت چیست؟ و این محتوای اطلاعاتی تا چه میزان پاسخگوی نیازهای متفاوت کاربران هست؟

از سوی دیگر موتورهای کاوش ابزار جستجو در این سیستم اطلاعات جهانی هستند. در این مرحله بررسی کارایی و موثر بودن عملکرد موتورهای کاوش در جستجوی اطلاعات مورد نیاز، مورد بررسی و اهمیت قرار می‌گیرد. موتورهای کاوش وب تا چه حد در شناسایی و دسترسی به محتوای اینترنت، در شناسایی نیازهای اطلاعاتی کاربران بر اساس کلیدواژه‌ها که تنها ورودی‌های این ابزارها و معرف نیاز اطلاعاتی کاربران است، و در ارائه نتایج مفید و مرتبط با پرسش کاربر در کوتاهترین زمان ممکن، توانا هستند؟

مطلب مهم بعدی این است که کاربران و جستجوگران اطلاعات در اینترنت تا چه حد با محتوای اطلاعاتی این شبکه آشنا هستند؟ تصور می‌کنند که کدام نیازهای اطلاعاتی خود را می‌توانند با جستجو در وب مرتفع سازند؟ تا چه حد با ابزارهای مختلف کاوش در وب، حیطه موضوعی مورد شناسایی آنها و خدمات ارائه شده توسط این ابزارها آشنا هستند؟ و تا چه حد و از کدام یک از این ابزارها و خدمات استفاده می‌کنند؟ برای معرفی نیاز اطلاعاتی خود در وب (وارد کردن کلیدواژه‌ها در موتورهای کاوش، یا وارد کردن آدرس وبگاه‌های شناخته شده در وب) چگونه عمل می‌کنند؟ تا چه حد با روش‌ها و ابزارهای فرموله‌سازی پرسش جستجو آشنا هستند و از آنها بهره می‌گیرند؟ برای شناسایی پیوندهای مرتبط با نیاز اطلاعاتی خود در صفحه نتایج ارائه شده توسط موتورهای کاوش غالباً چه معیارهایی را در نظر می‌گیرند، چقدر وقت صرف می‌کنند و تا چه حد به بازدید پیوندهای ارائه شده توسط موتورهای کاوش می‌پردازند؟

در پایان آگاهی به این نکته اهمیت دارد که در نهایت کاربران از جستجوی اطلاعات در اینترنت جهت رفع نیازهای اطلاعاتی خود تا چه حد احساس رضایت می‌کنند و اینترنت را تا چه حد پاسخگوی نیازهای اطلاعاتی خود می‌دانند و نیز تا چه حد از روش‌های جستجوی خود در وب رضایت دارند و خود را تا چه حد در جستجوی اطلاعات مفید و مرتبط با نیازشان در وب در کوتاهترین زمان و با موفقیت، متبحر می‌دانند؟

^۱ Information Retrieval

برای پاسخگویی به این سوالات، به ویژه بخش سوم سوالات (روش‌های جستجوی اطلاعات کاربران در اینترنت) که موضوع اصلی این پایان‌نامه است، تحقیقات بسیاری در سالهای اخیر انجام شده است. در این تحقیقات جوامع آماری مختلفی مورد بررسی قرار گرفته‌اند. به عنوان مثال در بسیاری از این تحقیقات رفتار جستجوی اطلاعات کاربران در یکی از اصلی‌ترین موتورهای کاوش وب در یک بازه زمانی معین بر اساس ثبت ورود کاربر^۲ روی کارساز^۳ آن موتور کاوش خاص مورد بررسی و مطالعه قرار گرفته است. در برخی از این پژوهش‌ها یک نمونه آماری خاص مورد مطالعه قرار گرفته است و رفتار جستجوی اطلاعات آنها در اینترنت با استفاده از ابزارهای مختلف جستجو در وب بر اساس ثبت ورود کاربر روی کارساز و یا به صورت بازدید حضوری بررسی شده است. ابزار پرسش‌نامه در این تحقیقات به ندرت مورد استفاده قرار گرفته است. اما در خصوص این پایان‌نامه خاص با توجه به دلایلی که در ذیل به آنها اشاره خواهد شد، پرسش‌نامه به عنوان ابزار مطالعه در این پژوهش مورد استفاده قرار گرفت:

۱. استفاده از ثبت ورود کاربر روی کارساز، اطلاعات کاملی را در مورد روش‌های جستجوی اطلاعات کاربران به دست نمی‌دهد و بیشتر برای بررسی مسیریابی کاربران در وب و آگاهی از مدت زمان جستجوی آنها در وب و اینکه از چه صفحاتی برای چه مدتی بازدید می‌کنند، اطلاعات مفیدی ارائه می‌کند.

۲. طبق پیگیری‌های به عمل آمده دسترسی به اطلاعات ثبت ورود کاربر روی کارسازهای دانشگاه‌های مختلف به دلیل محدودیت‌های امنیتی که مسئولان شبکه‌های کامپیوتری دانشگاه‌ها ایجاد می‌نمودند، امکان‌پذیر نبود.

۳. بررسی حضوری روش‌های جستجوی کاربران نیز برای این تحقیق روش مناسبی به نظر نمی‌رسید و انتخاب نشد. زیرا اولاً اطلاعات بدست‌آمده چون صرفاً بر اساس مشاهده تصادفی رفتار کاربران است، کاملاً مستند نیست. ثانیاً تمامی جوانب فرایند جستجو مورد بررسی قرار نمی‌گیرد. ثالثاً در مورد جامعه آماری وسیع این پژوهش به هیچ عنوان امکان‌پذیر نمی‌نماید. رابعاً ممکن است فردی با توجه به موضوع خاص مورد جستجو، روش خاصی را در جستجو پیش گیرد که به صورت معمول مورد استفاده او نیست. خامساً با توجه به محدودیت‌های زمانی و امکانات موجود، استفاده از این روش برای این تحقیق ممکن نبود.

قابل ذکر است کلیه سوالات مطرح شده در پرسش‌نامه با استناد به موارد مورد بررسی در تحقیقات مشابه که در سایر نقاط جهان انجام پذیرفته، طراحی گردیده است. شرح این مطلب در فصل دوم این پایان‌نامه به تفصیل ذکر شده است.

² Log File

³ Server

۲.۱: مساله پژوهش

همچنانکه اشاره شد، بازیابی اطلاعات^۴ از روی اینترنت و وب بسیار حائز اهمیت است. موتورهای جستجو، مهمترین ابزارکاوش و بازیابی اطلاعات در شبکه اینترنت هستند. از دیدگاه مدیریت سیستم‌های اطلاعاتی^۵، چهار ویژگی دقت^۶، پیوستگی زمانی^۷، روزآمد بودن^۸ و مرتبط بودن^۹ درمورد اطلاعات بازیابی شده باید مورد توجه قرار گیرد (منتظر، ۱۳۸۲). ویژگی‌های ذکر شده می‌توانند معیاری برای ارزیابی عملکرد و نتایج موتورهای جستجو محسوب گردند. از آنجا که جستجوی اطلاعات از اینترنت و توسط موتورهای جستجو هنوز با محدودیت‌های بسیاری مواجه است و بازیابی اطلاعات مفید و مطلوب از وب هنوز برای بسیاری از کاربران میسر نیست، در این پژوهش سه مساله مهم تاثیرگذار در کاهش عملکرد کاربران وب در جستجوی اطلاعات، مطرح و شناسایی الگوهای رفتاری کاربران وب در جستجوی اطلاعات به عنوان یک راه‌حل پیش‌نیاز برای رفع این مسائل ارائه می‌گردد. این سه مساله عبارتند از:

الف: اینترنت یک سیستم اطلاعات توزیع‌شده و مهمترین منبع اطلاعاتی به شمار می‌آید. مهم است اسنادی که بیشترین ارتباط را با پرسش جستجو دارند از کارسازهای مختلف بازیابی و در اختیار کاربر قرار گیرند. اما معمولاً برخی اسناد مرتبط با موضوع جستجو در وب که حاوی کلیدواژه مطرح شده در پرسش کاربر نیستند، بازیابی نمی‌شوند و برخی اسناد غیرمرتبط که حاوی آن کلیدواژه هستند کشف می‌شوند. بنابراین جستجوی موضوعات در وب به صورت مفهومی نیست (Perlovsky, 2006. Pedrycz, 2003. Olivas, 2003).

ب: از آنجا که اسناد وب در پایگاه‌های اطلاعاتی موتورهای جستجو بر اساس کلیدواژه نمایه می‌شوند، نتایج حاصل از موتورهای جستجو به صورت سیاهه‌ای از اسناد مرتبط یا غیرمرتبط که غالباً بر اساس دو معیار محل قرارگرفتن کلیدواژه در سند^{۱۰} و تعداد تکرار کلیدواژه در سند، رتبه‌بندی^{۱۱} شده‌اند، برای کاربر به نمایش گذاشته می‌شود. بنابراین انتخاب کلیدواژه مناسب در حصول نتیجه مطلوب بسیار حائز اهمیت می‌باشد. بدین منظور عموماً کاربران برای جستجوی اطلاعات ناگزیر به استفاده از عبارت‌ها یا کلیدواژه‌های ترکیبی یا استفاده از عملگرهای بولی به همراه کلیدواژه‌های مختلف و در موتورهای جستجوی گوناگون هستند (Kacprezyk, 2005, Gils, 2007).

⁴ Information Retrieval
⁵ Information System Management
⁶ Accuracy
⁷ Online
⁸ Updating
⁹ Relevancy
¹⁰ Document or Web Page
¹¹ Ranking