

# تحلیل رفتار اطلاع‌یابی پژوهشگران در موتور جستجوی سامانه‌ی ملی اطلاعات پایان‌نامه‌ها / رساله‌های دانش‌آموختگان داخل کشور (گنج)

فصلنامه علمی - پژوهشی



دوره ۲، شماره ۵  
زمستان ۹۵

سمیه فاتحی

استادیار مهندسی کامپیوتر، پژوهشگاه علوم و فناوری اطلاعات ایران (ایراندک)<sup>۱</sup>

علی نعیمی صدیق

استادیار مهندسی صنایع، پژوهشگاه علوم و فناوری اطلاعات ایران (ایراندک)

**چکیده:** یکی از رسالت‌های مهم موتورهای جستجوی پایگاه‌های اطلاعاتی، ارائه بهترین خدمات متناسب با نیازهای کاربران است. این امر باعث افزایش رضایت کاربران و علاقه‌مندی آنان به استفاده از پایگاه‌های اطلاعاتی می‌شود. دانشجویان، اساتید دانشگاه و پژوهشگران به‌طور روزانه به جستجو در موتورهای جستجو پرداخته و نیازهای اطلاعاتی خود را برآورده می‌سازند. پایگاه اطلاعاتی گنج (گنجینه علمی ایرانیان) پژوهشگاه علوم و فناوری اطلاعات ایران، با برخورداری بیش از هزاران رکورد علمی، امکان جستجو برای پژوهشگران را در پایان‌نامه‌ها، نشریات علمی داخلی، مقالات، همایش‌ها، طرح‌های پژوهشی و گزارش‌های دولتی را فراهم می‌کند. واضح است که کاربران تحت تأثیر عوامل مختلف از شیوه‌ها و روش‌های گوناگون برای کسب اطلاعات استفاده می‌نمایند و رفتارهای اطلاع‌یابی متفاوتی را از خود بروز می‌دهند. هدف اصلی پژوهش حاضر، تحلیل رفتار اطلاع‌یابی کاربران در موتور جستجوی سامانه‌ی گنج است. تحلیل رفتار اطلاع‌یابی کاربران می‌تواند به درک بهتر نیازهای آنان کمک کند. نتایج تحلیل، بیانگر آن است که میزان مراجعه حضوری در مقایسه با غیرحضوری بسیار اندک است و ساعات مراجعه حضوری در بین پژوهشگران حوزه علوم انسانی و فنی-مهندسی متفاوت است. همچنین نتایج این تحلیل می‌تواند جهت ارزیابی و رفع نواقص احتمالی موجود در زمینه بازیابی اطلاعات توسط موتور جستجو کمک نماید.

**کلیدواژه‌ها:** پایگاه‌های اطلاعاتی، تحلیل رفتار اطلاع‌یابی، سامانه گنج، موتور جستجو.

## ۱- مقدمه

امروزه، اطلاعات نقش مهمی را در نوع و کیفیت زندگی بشر بازی می‌کند و به‌عنوان منبع اساسی و مهم در فعالیت‌های روزمره مطرح است و تمام تصمیم‌گیری‌ها تحت تأثیر اطلاعات است. در پژوهش‌های حوزه علوم اطلاعات، قسمت مهم و قابل توجه، رفتار مرتبط با اطلاعات است که شامل جستجوی اطلاعات مورد نیاز، رفتار اطلاع‌یابی و مدل‌های جستجوی اطلاعات است (Pareek and Rana 2013). رفتار اطلاع‌یابی، شامل راه‌هایی است که افراد نیازهای اطلاعاتی‌شان را بیان می‌کنند (Pareek and Rana 2013; Ahiauzu and Ani 2015). از نظر ویلسون، جستجوی اطلاعات توصیف روش‌هایی است که افراد اطلاعات را جستجو، بررسی، ارزیابی و استفاده می‌کنند. در مسیر جستجوی اطلاعات جدید، فرد ممکن است با افراد مختلف، کتابخانه‌ها، مراکز اطلاعاتی و سیستم‌های اطلاعاتی از جمله کامپیوتر تعامل داشته باشد (Wilson 2000). واضح است که رفتارهای اطلاعاتی در محیط‌های علمی و دانشگاهی، به دلیل نیاز پژوهشگران و دانشجویان به اطلاعات درست و به‌روز برای پژوهش‌هایشان، نقش بسیار مهمی دارد (Pareek and Rana 2013). نتایج پژوهش‌هایی که درباره الگوهای جستجوی اطلاعاتی از پژوهشگران در کتابخانه‌های دانشگاهی انجام شده است، نشان می‌دهد حتی بین الگوهای مورد استفاده از اطلاعات میان پژوهشگران رشته‌های مختلف تفاوت وجود دارد (Shivalingaiah 2010) و عوامل گوناگونی از جمله سهولت و سرعت دسترسی، آشنایی با منابع، سطح و موضوع پژوهش بر رفتار جستجوی کاربران تأثیرگذارند (Ahiauzu and Ani 2015).

از طرف دیگر، امروزه شبکه جهانی وب، منبع عظیمی از اطلاعات است که این اطلاعات با محتوای تحت وب توسط میلیون‌ها صفحه در دسترس عموم مردم قرار دارد (Facca and Lanzi 2005; Jansen and Spink 2006). در طول دهه گذشته، موتورهای جستجو به ابزار اصلی پیدا کردن اطلاعات برای بسیاری از مردم تبدیل شده‌اند (Gan et al. 2008; Silvestri 2009). انتشار محتوای علمی و دستاوردهای پژوهشی به‌صورت الکترونیکی به‌ویژه بر روی اینترنت، روش توزیع اطلاعات علمی در دنیا را تغییر داده است (Ahiauzu and Ani 2001; Gleeson 2015). با توسعه انتشارات الکترونیکی و دیجیتالی شدن نه‌تنها توسعه کتابخانه‌ها بهبود یافته است بلکه آن‌ها را از منابع نگه‌دارندگان اطلاعات به منابع دسترسی اطلاعات تغییر داده است (Ellis and Oldman 2005). نتایج پژوهشی که به‌تازگی در سال ۲۰۱۵، توسط آهیاز و آنی صورت گرفته است (Ahiauzu and Ani 2015) نشان داد

1. Wilson
2. Ahiauzu
3. Ani

اینترنت به‌عنوان ابزار اطلاع‌یابی، بالاترین رتبه‌بندی را با ۳۱,۶ درصد داشته است و تماس‌های اشخاص، کارگاه‌های کنفرانس و کتابخانه‌ی دانشگاه، در رتبه‌های بعدی قرار می‌گیرند. حال با توجه به اینکه اینترنت یک منبع اطلاعاتی غنی برای تمامی کاربران و به‌ویژه کاربران دانشگاهی است، توجه به کیفیت موتورهای جستجوی پایگاه‌های علمی از اهمیت ویژه‌ای برخوردار است. دو شاخص کلیدی که کیفیت این موتورهای جستجو را به‌طور کلی مشخص می‌کند (Kurth 1993; Silvestri 2009): کیفیت نتایج برگشت داده‌شده و سرعت نتایجی که برگشت داده‌شده است. از این‌رو برای ارتقای شاخص اول، فهم کاربران و نیازهای آنان راه‌حل اصلی است (White and Morris 2007; Jansen and Spink 2006). به همین دلیل، پژوهشگران به بررسی لاگ‌های پرس‌وجو<sup>۱</sup> جستجو می‌پردازند تا درک بهتری از آنچه کاربران به دنبالش هستند، داشته باشند.

در ایران، پایگاه اطلاعاتی گنج (گنجینه علمی ایرانیان) پژوهشگاه علوم و فناوری اطلاعات ایران، با برخورداری بیش از هزاران پایان‌نامه و رساله دانشگاه‌های کشور، امکان جستجوی حضوری و غیرحضوری برای پژوهشگران را فراهم می‌کند. همچنین، اینکه پژوهشگران چگونه جستجو می‌کنند و چگونه این فرآیند می‌تواند بهبود یابد، به ارائه خدمات مناسب‌تر و مفیدتر به پژوهشگران کمک شایانی می‌نماید.

در این مقاله، به بررسی رفتار اطلاع‌یابی کاربران و اولویت نیازهای جستجوی آنان از طریق بررسی لاگ‌های جستجوی سامانه‌ی گنج پرداخته می‌شود. اهداف این مقاله، بررسی رفتارهای جستجوی کاربران موتور جستجوی سامانه‌ی گنج، تعیین ویژگی‌های مفید اطلاعاتی جهت تحلیل رفتار بهتر پژوهشگران و دسته‌بندی پژوهشگران از لحاظ عبارات جستجو، رشته‌های تحصیلی، تعداد نتایج مشاهده‌شده و زمان‌های استفاده (روزهای هفته و ساعات کاری) است. از این‌رو در بخش دوم این مقاله پیش‌زمینه‌ای از تحلیل لاگ و مزایا و معایب آن ارائه خواهد شد. در بخش سوم، پژوهش‌هایی که تاکنون در حوزه تحلیل لاگ موتورهای جستجو در پایگاه داده‌های عمومی و علمی صورت گرفته، شرح داده خواهند شد. در بخش چهارم، روش پژوهش و در بخش پنجم، لاگ‌های موتور جستجوی سامانه‌ی گنج، تجزیه و تحلیل یافته‌ها و نتایج به‌دست‌آمده ارائه خواهد شد. در نهایت جمع‌بندی، نتیجه‌گیری و پیشنهادهایی برای مطالعات آتی در بخش ششم ارائه خواهد شد.

---

<sup>1</sup> Query

## ۲- پیش‌زمینه

### ۲-۱- تحلیل لاگ<sup>۱</sup>

لاگ یک رکورد الکترونیکی از تعاملاتی است که بین سیستم و کاربران سیستم اتفاق می‌افتد (Jansen, Murray and Teevan 2007). داده‌های لاگ، داده‌هایی غنی همراه با جزئیات فردی هستند (Taksa, and Spink 2009). تحلیل لاگ، فرآیند تبدیل داده‌های خام به اطلاعاتی برای حل مسائل است (Alspaugh et al. 2014). پیترز<sup>۲</sup> تحلیل لاگ را مطالعه‌ی تعاملات الکترونیکی ثبت‌شده بین سیستم‌های بازیابی اطلاعات برخط<sup>۳</sup> و افرادی که در این سیستم‌ها جستجو می‌کنند، تعریف کرد (Peters 1993).

از طرفی پرس‌وجو<sup>۴</sup>، پرسشی است که کاربران در سیستم جستجو می‌کنند. لاگ‌کاوی پرس‌وجو<sup>۵</sup> شاخه‌ی از لاگ‌کاوی<sup>۶</sup> و تحلیل لاگ است (Jansen 2009)، هدفش کشف الگوهای موردعلاقه از لاگ‌های پرس‌وجوهای موتورهای جستجوی وب در جهت افزایش کارایی یک سرویس آنلاین است (Silvestri 2009). در این پژوهش تمرکز ما بر روی لاگ‌های سامانه‌ی گنج و هدف، لاگ‌کاوی پرس‌وجوهای تولیدشده توسط کاربران این سامانه است.

### ۲-۱-۱- مزایا و معایب جمع‌آوری و تحلیل داده‌های لاگ پرس‌وجو

جمع‌آوری و تحلیل لاگ‌های پرس‌وجو شامل مزایایی است از جمله: بهبود جستجوی عمومی و جستجوی شخصی، امکان ثبت تاریخچه‌ی جستجوی افراد در سیستم و ارزان‌تر بودن هزینه جمع‌آوری لاگ‌ها نسبت به جمع‌آوری داده‌های اصلی (Jansen, Taksa, and Spink 2009; Murray and Teevan 2007). در کنار مزایای ذکرشده برای تحلیل و جمع‌آوری لاگ‌های پرس‌وجو، خطراتی نیز وجود دارد از جمله اینکه داده‌ها ممکن است توسط مؤسسه جمع‌آوری‌کننده مورد سو استفاده قرار بگیرد و یا توسط افراد با نیت مخرب به بیرون درز پیدا کنند. همچنین ممکن است توسط دولت‌ها استفاده شوند. عیب بزرگ داده‌های لاگ این است که معمولاً داده‌های جامع و کاملی نیستند، زیرا با توجه به سؤالات یک پژوهش خاص جمع‌آوری نشده‌اند. معمولاً برای بررسی محدوده خاصی در حوزه لاگ‌کاوی که برخی از پژوهشگران تمایل به مطالعه آن دارند، داده‌های لاگ داده‌های مناسبی نیستند.

نکته دیگر این است که در زمان جمع‌آوری این داده‌ها اگر کاربران سیستم از جمع‌آوری لاگ‌هایشان آگاه باشند ممکن است رفتار خود را تغییر دهند و رفتارهای آن‌ها کاملاً طبیعی نیست (Murray and

- 
1. Log analysis
  2. Peters
  3. Online
  4. Query
  5. Query Log Mining
  6. Log mining

Teevan 2007). قابل ذکر است که برخی از مزایا و معایب ذکر شده فوق، فرصت‌ها و چالش‌هایی را در این پژوهش به همراه داشته است.

### ۳- پیشینه پژوهش

تحلیل لاگ سامانه‌های اطلاعاتی روشی است که از ۱۹۶۰ تاکنون جهت ارزیابی سامانه‌ها و نیز مطالعه‌ی رفتار کاربران مورد استفاده بوده است (Alspaugh et al. 2014). با توجه به مطالعات انجام شده در این حوزه، پژوهش‌های انجام شده تاکنون را به دو قسمت می‌توان تقسیم کرد: پژوهش‌های بخش اول بر انواع موتورهای جستجوی مبتنی بر وب و بررسی لاگ‌ها، پرس‌وجوهای مطرح بر روی آن‌ها و بررسی روش‌های خوشه‌بندی پرس‌وجوها تأکید دارند. بخش دوم، پژوهش‌هایی هستند که صرفاً بر روی موتورهای جستجوی پایگاه‌های علمی مانند ساینس دایرکت تمرکز دارند.

### ۳-۱- تحلیل لاگ جستجوی کاربران در پایگاه‌های غیرعلمی

مارکیونینی<sup>۲</sup> (۱۹۸۹) معتقد است که فرآیند جستجوی اطلاعات با چندین عامل از قبیل ویژگی‌های فردی، نوع جستجو، ویژگی سامانه‌های اطلاعاتی، حوزه‌های دانش، محیط اطلاعاتی، تجربه و دانش پیشین در جستجو و نتایج جستجو در ارتباط است. یانسن<sup>۳</sup> و همکارانش (۲۰۰۰) گزارش کردند که ۶۷ درصد از کاربران فقط یک جستجو را در موتور جستجو انجام داده‌اند و ۱۹ درصد آن‌ها تنها یک تغییر در جستجوی خود ایجاد کرده‌اند. در همین سال، مورزی<sup>۴</sup> و همکارانش (۲۰۰۰) یک الگوریتم جدید برای تقسیم‌بندی خودکار کاربران وب بر اساس تحلیل الگوهای دسترسی‌شان ارائه دادند که منجر به بهبود سازمان‌دهی اسناد وب برای راحتی هدایت کاربران به سمت نیازهایشان شد. در واقع، کاربرانی که در دسترسی به منابع اطلاعاتی مسیرهای مشابهی را دنبال کرده‌اند در یک خوشه قرار می‌گیرند. ژئی<sup>۵</sup> و همکارانش (۲۰۰۱) لاگ پرس‌وجوهای سرویس جستجوی شرکت آنلاین آمریکایی<sup>۶</sup> را تحلیل کردند. نتایج نشان داد اگر درصد اندکی از پرس‌وجوها قابل پیش‌بینی باشند و نتایج جستجو برای آن‌ها بهینه شود، کل سیستم می‌تواند برای ۸۰ درصد ارائه‌دهنده‌های خدمات بهینه شود. همچنین نتایج نشان داد که پرس‌وجوها در طول زمان پایدارند و فقط تنها ۲۰ درصد در طول زمان تغییر می‌کنند. این نتایج مشخص کرد که هرچه یک پرس‌وجو در طول زمان بیشتر دوام بیاورد احتمال دوامش در آینده نیز بیشتر است.

1. ScienceDirect
2. Mchionini
3. Jansen
4. Morzy
5. Xi
6. AOL (www.aol.com)

ون<sup>۱</sup> و همکارانش (۲۰۰۲) یک روش خوشه‌بندی جدید برای پرس‌وجوها را ارائه کردند که در آن از لاگ‌های کاربر بر روی داده‌های وب‌سایت Encarta استفاده کردند. در این پژوهش، شباهت بین دو پرس‌وجو بر اساس اسنادی که کاربران برای آن انتخاب کرده‌اند، مشخص می‌شود. خوشه‌بندی پرس‌وجوها فرآیندی جهت کشف پرسش‌های متداول<sup>۲</sup> یا موضوعات معروف روی یک موتور جستجو است. یافته‌های این پژوهش، فرضیه پشت دیدگاه پرسش‌های متداول را تأیید می‌کند که تعداد زیادی از کاربران به موضوعات مشابه و پرسش‌های متداول علاقه‌مندند. همچنین نتایج نشان داد که اغلب کاربران از مجموعه کوچکی از کلمات استفاده می‌کنند و اسناد را برای مطالعه انتخاب می‌کنند، زیرا ۴۹ درصد از پرس‌وجوها شامل تنها یک کلمه کلیدی و ۳۳ درصد از پرس‌وجوها شامل دو کلمه کلیدی بودند.

بیتزل<sup>۳</sup> و همکارانش (۲۰۰۴)، با بررسی صدها میلیون لاگ پرس‌وجو دریافتند که تغییرات در محبوبیت و یگانگی پرس‌وجوها به ساعتی که در طول روز جستجو می‌شوند، بستگی دارد. برای مثال، پرس‌وجوهای مربوط به امور مالی افراد در ساعت هفت تا ده صبح محبوب‌تر هستند. درحالی‌که پرس‌وجوهای مربوط به موسیقی کمتر در این ساعات مورد تقاضا هستند. نتایج این پژوهش، دیدگاه ارزشمندی را جهت ارتقای اثربخشی بازبایی مسیریابی و الگوریتم‌های ذخیره (caching) فراهم نمود.

رز و لوینسون<sup>۴</sup> (۲۰۰۴)، چارچوبی را برای فهم اهداف جستجوهای کاربران ارائه دادند. جهت توسعه و آزمایش چهارچوب پیشنهادی، نمونه‌ای از پرس‌وجوهای موتور جستجوی AltaVista مورد بررسی قرار گرفت. نتایج این پژوهش نشان داد که چارچوب پیشنهادی می‌تواند جهت ارتباط اهداف با پرس‌وجوها استفاده شود و درواقع «حساسیت به هدف» عامل مهمی در جستجوی کاربران است. از نتایج این پژوهش، یعنی درک اهداف کاربر از جستجو می‌توان جهت بهبود الگوریتم‌های موتورهای جستجو استفاده کرد. همچنین می‌توان شخصی‌سازی صفحات جستجو را متناسب با نیاز کاربر فراهم کرد.

در پژوهشی که در سال ۲۰۰۶ توسط اسپینک و جانسن<sup>۵</sup> (۲۰۰۶) انجام شد، آن‌ها دریافتند که کاربران دو یا سه واژه را جستجو کرده و همواره فقط اولین صفحه‌ی نتایج را مشاهده می‌کنند. در پژوهش دیگری که توسط آن‌ها انجام گرفت (Jansen and Spink 2006)، تعاملات بین کاربران و موتورهای جستجوی وب از دیدگاه طول نشست، طول پرس‌وجو، پیچیدگی پرس‌وجو و محتوای مشاهده‌شده در میان موتورهای جستجوی وب مقایسه شد. یافته‌ها نشان داد که میزان استفاده از اپراتورهای پیشرفته پرس‌وجو در موتورهای جستجوی ایالات‌متحده آمریکا بسیار بالاتر از اروپاست. همچنین درصد پرس‌وجوهای تک‌واژه‌ای، برای موتورهای جستجوی وب ایالات‌متحده آمریکا در محدوده‌ی بین ۲۰ تا ۲۹ درصد و برای موتورهای جستجوی وب اروپایی، بین حدود ۲۵ تا ۳۵ درصد است. همچنین یافته‌های این پژوهش نشان می‌دهد که موضوعات درباره‌ی تجارت، مسافرت، اقتصاد،

1. Wen
2. Frequently asked questions (FAQ)
3. Beitzel
4. Rose and Levinson
5. Spink and Jansen

کامپیوتر، اینترنت و فناوری تقریباً ۲۵ درصد از پرس‌وجوها را تشکیل می‌دهند. در سال ۲۰۰۷، وایت و موریس<sup>۱</sup> (۲۰۰۷) لاگ تعامل کاربران حرفه‌ای و غیرحرفه‌ای موتور جستجو را مورد مطالعه قرار دادند. در واقع کاربران حرفه‌ای کاربران فعالی هستند که زمان کمتری را جهت جستجو و پیمایش دنباله جستجو صرف می‌کنند و انحراف کمتری در دنباله‌های جستجویشان وجود دارد و در جستجویشان موفق‌ترند. نتایج این پژوهش نشان داد که تفاوت‌های معنی‌داری در پرس‌وجوها، کلیک بر روی نتایج، مرور کردن صفحات وب پس از اجرای پرس‌وجو و جستجوی موفق کاربرانی حرفه‌ای نسبت غیر حرفه‌ای‌ها وجود دارد.

گان<sup>۲</sup> و همکارانش (۲۰۰۸) پژوهشی را جهت بهبود جستجوهای جغرافیایی انجام دادند. یافته‌های پژوهش نشان می‌دهد که تفاوت‌های مهمی بین پرس‌وجوهای جغرافیایی و غیر جغرافیایی وجود دارد. نتایج نشان می‌دهد که طبقه‌بندی‌کننده طراحی‌شده توسط آنان، با دقت حدود ۹۶٫۶۹ درصد همه پرس‌وجوهای را به درستی طبقه‌بندی می‌کند. همچنین پرس‌وجوهای جغرافیایی اغلب هدفشان در مکان‌یابی و دریافت خدمات است. در حالی که پرس‌وجوهای غیر جغرافیایی بیشتر هدفشان سرگرمی، دانلود یا مشاهده صفحات اطلاعاتی است. همچنین پرس‌وجوهای جغرافیایی به استفاده از تعداد واژه‌های بیشتری نسبت به پرس‌وجوهای غیر جغرافیایی تمایل دارند. از دیگر یافته‌های این پژوهش این است کاربرانی که مکرراً جستجوهای را روی نواحی جغرافیایی مشابه انجام می‌دهند، به‌سادگی می‌توان شهر آن‌ها را با توجه به اصطلاحات جغرافیایی که جستجو می‌کنند، استخراج کرد.

بندرسکی و کرافت<sup>۴</sup> (۲۰۰۹) بر روی تحلیل پرس‌وجوهای بلند (طولانی) در لاگ‌های جستجوی کاربران باهدف شناسایی ویژگی‌ها پرتکرار در این نوع پرس‌وجوها تمرکز کردند. جهت تجزیه و تحلیل پرس‌وجوها سه ویژگی طول، نوع پرس‌وجوها و فرکانس آن‌ها در نظر گرفته شد. پرس‌وجوهایی با طول کمتر از چهار حدود ۹۰٫۳ درصد کل پرس‌وجوها را تشکیل می‌دهند و برای ۹۹ درصد پرس‌وجوها، طول پرس‌وجوها کمتر از ۱۲ است. نتایج نشان می‌دهد که کاربران در زمان مشاهده لیست نتایج پرس‌وجو تمایل دارند نتایج پایین لیست را برای پرس‌وجوهای طولانی‌تر نسبت به پرس‌وجوهای کوتاه‌تر کلیک کنند. همچنین در رفتار کلیک کاربر بین انواع مختلف پرس‌وجوهای بلند تفاوت وجود دارد. براساس بررسی‌های پژوهش‌های پیش‌ازاین، پژوهشی توسط داوئی<sup>۵</sup> و همکارانش (۲۰۰۸)، مشخص شد که پرس‌وجوهای ناقص (کمیاب و کمتر درخواست شده) کمتر کلیک می‌شوند و بازدید می‌شوند و پرس‌وجو-های دوباره بازنویسی شده پرتکرار هستند آن‌ها معتقدند که فرکانس پرس‌وجو بسیار مهم‌تر از طول پرس‌وجو است، اما در نتایج این پژوهش مشاهده شد که فرکانس و طول پرس‌وجو با یکدیگر برای بازیابی نتایج بهینه مؤثر هستند.

1. Morris
2. Gan
3. Geographic
4. Bendersky and Croft
5. Downe

تیلر و تیوان<sup>۱</sup> (۲۰۱۰) پژوهشی را جهت بازیابی مجدد نتایج موتورهای جستجو انجام دادند. نتایج نشان داد که پرس‌وجوهای مجدد بازیابی شده، وقتی تغییرات اساسی هستند، نسبت به پرس‌وجوهایی که قبلاً استفاده شده‌اند، تمایل بیشتری به بهتر شدن دارند و به‌طور معمول کوتاه‌تر هستند. همچنین یافته‌های این پژوهش تأکید دارند که زمانی که کاربری روی نتیجه پرس‌وجو کلیک می‌کند، احتمال دارد که مسیر مشابهی را مانند سایر کاربرانی که بر همان URL کلیک کرده‌اند دنبال کند. تمامی یافته‌های این پژوهش می‌توانند جهت بهبود عملکرد موتورهای جستجو مورد استفاده قرار گیرند.

ستیندیل<sup>۲</sup> و همکارانش (۲۰۱۲) یک شمای<sup>۴</sup> طبقه‌بندی کننده برای رفتار تایپ کردن کاربران پیشنهاد دادند. هدف آن‌ها بررسی چگونگی تحلیل لاگ پرس‌وجوهای یک سیستم جستجوی فوری و یافتن رفتارهای کاربر در جستجوی فوری است. جستجوی فوری روی کاراکتر به کاراکتر پرس‌وجو که کاربر تایپ کرده، پرس‌وجو را به‌عنوان نتیجه پیشنهاد می‌دهد. روش پیشنهادی این پژوهش بر روی لاگ‌های پرس‌وجو سیستم جستجوی PSearch، انجام شد. نتایج نشان داد که جستجوی فوری می‌تواند به‌طور معمول دو ثانیه زمان را در هر جستجو صرفه‌جویی کند و همچنین تلاش کاربر را برای تایپ کردن را کاهش و همچنین باعث افزایش نرخ موفقیت در بازگرداندن نتایج مورد نظر کاربر شود.

در سال ۲۰۱۲، پژوهش دیگری جهت تحلیل ۷۵۰ میلیون پرس‌وجو به‌دست‌آمده از یک موتور جستجوی تجاری واقعی انجام شد. طول پرس‌وجو، نرخ پالایش پرس‌وجو، نرخ کلیک، نرخ پرس‌وجو از طریق جستجوی پیشرفته از فاکتورهای مورد بررسی در این پژوهش بودند. نتایج نشان داد که برخی از رفتارهای کاربران همیشه ثابت است. به‌عنوان مثال بیش از ۹۰ درصد کاربران تنها به پنج نتیجه اولی که موتور جستجو برمی‌گرداند توجه دارند (Xie, Yu, and Cen 2012).

در سال ۲۰۱۳، توجه به کاربران و ویژگی‌های منحصر به فرد هر کاربر افزایش یافت. به‌طوری که در پژوهشی که انجام گرفت (جریری، اسدی و نوشین فرد ۱۳۹۲)، نتایج نشان داد که بین سبک‌شناختی و رفتار اطلاع‌یابی کاربران رابطه وجود دارد و افراد با سبک‌شناختی کلامی از تعداد واژه‌های بیشتری نسبت به افراد با سبک‌شناختی تصویری استفاده می‌کنند.

### ۳-۲- تحلیل لاگ جستجوی کاربران در پایگاه‌های علمی

پژوهش‌های بالا، اطلاعات زیادی در مورد رفتار جستجو در وب در اختیار پژوهشگران می‌گذارد؛ اما بشک، رفتار جستجو در موتورهای جستجو عمومی با رفتار جستجو در یک پایگاه تخصصی علمی متفاوت است. از این رو برخی از پژوهش‌ها تنها به بررسی رفتارهای جستجو در پایگاه‌های علمی پرداختند.

- 
1. Tyler and Teevan
  2. Re-finding
  3. Cetindil
  4. Scheme



در سال ۲۰۰۰، یو و اپس<sup>۱</sup> از داده‌های لاگ ۲۳ ماهه برای بررسی رفتار کاربر در پروژه SuperJournal استفاده کردند. داده‌های لاگ شامل ۱۰۲,۹۶۶ فعالیت ثبت‌شده کاربران در ارتباط با سه موتور جستجوی وب بود. متغیرهایی مانند تعداد نشست‌ها برای هر مجله، تعداد نشست‌ها در هرماه، فراوانی استفاده و زمان استفاده بررسی شدند. نتایج این پژوهش نشان می‌دهد که لاگ‌ها می‌توانند منابع مقرون‌به‌صرفه و خوبی برای استخراج اطلاعات از نحوه‌ی تعامل کاربران با سیستم باشند اما آن‌ها نمی‌توانند دلیل رفتار کاربران را شرح دهند، لذا منابع اطلاعاتی دیگر برای بررسی رفتار کاربران موردنیاز است.

در پژوهش دیگری در سال ۲۰۰۲، کیا<sup>۲</sup> و همکارانش لاگ فایل رفتاری کاربران ScienceDirect در Elsevier را بررسی کردند. گزارش‌های آماری از جمله تعداد بازدیدها در یک دوره زمانی، موضوعات مورد جستجو، تعداد پرس‌وجو برای هر مکان دسترسی کاربر یافته‌های این پژوهش است.

در پژوهشی، ۵۴۱۹۰۲۰ پرس‌وجو کاربر بر روی یک موتور جستجوی وبسایت دانشگاهی مورد بررسی قرار گرفت (Wang, Berry and Yang 2003). تمرکز آن‌ها بر روی شناسایی مشکلاتی از جمله تایپ غلط پرس‌وجوها توسط کاربران و پیشنهادهایی جهت رفع آن‌ها، فهم رفتارهای جستجوی کاربر با توجه به فرکانس‌های تکرار جستجو و توسعه تکنیک‌هایی جهت تحلیل پرس‌وجوها بود. یافته‌های این پژوهش نشان داد که الگوهای مشابهی در طول سال‌ها در بین رفتار جستجوی کاربران وجود دارد. پرس‌وجوهای انجام‌شده توسط کاربران، بر اساس موضوع و تعداد جستجو، دارای تغییرات فصلی هستند. از دیگر یافته‌های این پژوهش، تحلیل‌های آماری است که نشان می‌دهد ۳۸ درصد از تمام پرس‌وجوها تنها شامل یک واژه هستند و اکثر پرس‌وجوها منحصربه‌فرد هستند.

یی<sup>۳</sup> و همکارانش (۲۰۰۶) در بررسی رفتار جستجوی کاربران، بر روی داده‌های پایگاه روانشناسی PsycINFO و دو پایگاه حوزه‌ی تاریخ با نام ABC-Clio و America تمرکز کردند. پژوهشگران این پژوهش دریافتند که ترکیبی از واژه‌ها، شناخت بهتری از رفتار نگارش پرس‌وجو کاربران را در اختیار قرار می‌دهد. همچنین نتایج این پژوهش نشان داد که میانگین طول پرس‌وجو در پایگاه داده‌های روانشناسی PsycINFO، ۳,۱۶ واژه و در دو پایگاه دیگر ۳,۴۲ واژه بوده است. نتایج نشان داد علیرغم اینکه طول پرس‌وجوها کوتاه است، اصطلاحات چند واژه‌ای که شامل دو یا سه واژه هستند اغلب در پرس‌وجوها استفاده می‌شوند.

در پژوهشی، به بررسی لاگ پایگاه علمی ScienceDirect پرداخته شد (Nicholas, Huntington, and Jamali 2008). هدف آن‌ها انجام تحلیل عمیقی از لاگ کاربران و فهم دقیقی از رفتار اطلاع‌یابی آنان بود. نتایج تحلیل نشان داد که میان رفتار جستجوی افراد در رشته‌های مختلف علمی، تفاوت‌های زیادی وجود دارد که این رفتار شامل تعداد دفعات بازدید و تعداد جستجو در یک نشست بود. برای مثال

1. Yu and Apps
2. Kea
3. Y

محققان علوم اجتماعی، بیشتر از سایرین تنها چکیده مقاله یا مقالات در حال انتشار را مشاهده می‌کردند. در حالی که محققان علم مواد در یک نشست به‌طور متوسط اکثر مجلات منحصربه‌فرد را مشاهده کردند.

ژیاندونگ<sup>۱</sup> و همکارانش (۲۰۱۱) ویژگی‌هایی از جمله توزیع زمانی دسترسی کاربران، توزیع طول عبارات جستجو و نسبت بازبایی اطلاعات پایگاه داده‌ی علمی-تخصصی مجلات چینی را تحلیل کردند. نتایج نشان داد که کاربران با رشته‌های مختلف و سطوح علمی متفاوت، رفتارهای متفاوتی را در جستجو از خود بروز می‌دهند. در نهایت، بر اساس نتایج این پژوهش، یک مدل استراتژی جستجوی آکادمیک کاربران دانشگاهی به دست آمد.

پژوهشی بر روی لاگ یک وبگاه علمی تحلیلی انجام شد که هدف کسب دانشی در خصوص رفتار کاربر جهت بهبود سیستم‌های بازبایی اطلاعات بود. یافته‌های این پژوهش نشان می‌دهد که اغلب پرس‌وجوها در علم و فناوری کوتاه و ساده هستند که نشان از این است که کاربران تمایل دارند حداقل تلاش را در جهت پرس‌وجوها انجام دهند (Park and Lee 2013).

ادانتن<sup>۲</sup> (۲۰۱۴)، لاگ فایل سامانه‌ی بازبایی اطلاعات کتابخانه دانشگاه Strathclyde را در بازه‌ی زمانی دوماهه مورد بررسی قرارداد. وی در این پژوهش، پرس‌وجوهای کاربران و الگوهای جستجو را مورد تحلیل قرارداد. نتایج به‌دست‌آمده از این پژوهش در جهت بهبود طراحی واسط کاربر و سازمان‌دهی محتوا به کار گرفته شد.

اخیراً پژوهشی بر روی استفاده از لاگ‌ها و تحلیل آن جهت یافتن رابطه بین کاربران، مسیرهای دسترسی و مقالات دستیابی شده توسط آن‌ها ارائه شد (Sato et al., 2015). آن‌ها از پنج پایگاه داده ژاپنی جهت جمع‌آوری لاگ‌ها استفاده کردند. نتایج این پژوهش نشان می‌دهد که تنها پژوهشگران نیستند که از منابع علمی استفاده می‌کنند بلکه تعداد افراد غیر محقق بسیار بیشتر است و بازه‌ی کاربرانی که در موتورهای جستجو به دنبال اطلاعات علمی هستند، بسیار وسیع است. دیگر یافته این پژوهش نشان می‌دهد از آنجاکه دسترسی به منابع علوم کاربردی از علوم پایه بیشتر است، لذا یادگیری افراد در این حوزه‌ها بیشتر است.

پژوهشی بر روی پرس‌وجوهای کاربرانی که توسط کاربران در سال‌های ۲۰۰۸ تا ۲۰۱۱ ارسال شده بود، صورت گرفت (Park et al., 2016). هدف آن‌ها فهم طولی تعاملات کاربر و سیستم در حوزه‌ی علم و فناوری بود. نتایج نشان داد که رفتار جستجوی کاربران و الگوهای رفتاری آنان ثابت است و آن‌ها تمایل به استفاده از پرس‌وجوهای ساده و کوتاه دارند. زمان تصمیم‌گیری برای کاربران پس از صدور پرس‌وجو به‌طور میانگین ۱۰ ثانیه گزارش شد. به‌طور متوسط زمان ۱۰ ثانیه ثبت‌شده است که آیا کاربران می‌خواهند روی نتایج به‌دست‌آمده از پرس‌وجو تمرکز کنند یا خیر. اگر کاربران قرار است نتایج به‌دست‌آمده را مشاهده کنند به‌طور میانگین زمان حدود ۱۰ دقیقه را برای این کار صرف می‌کنند.

1. Jandong
2. Quntan
3. Longitudinal

همان‌طور که پیشینه‌ی پژوهش نشان می‌دهد با وجود اینکه تحلیل لاگ به‌عنوان یکی از شیوه‌های ارزیابی سامانه‌های اطلاعاتی و همچنین مطالعه‌ی رفتار اطلاع‌یابی کاربران در خارج از ایران جایگاه خود را پیدا کرده است، اما در ایران با توجه به دانش نویسندگان، کار پژوهشی چندانی در این زمینه انجام نشده است. تعدادی پژوهش نظری در زمینه تحلیل لاگ انجام شده است (جمالی مهموئی ۱۳۸۳؛ رجبعلی بگلو ۱۳۸۶؛ زوارقی ۱۳۸۴؛ نستوده ۱۳۸۲). در این میان، تنها دو پژوهش عملی بر روی تحلیل لاگ‌های پایگاه اطلاعات و مدارک علمی ایران (ایراندک) و رفتار جستجوی کاربران آن در انجام گرفته است (خسروی و مهموئی ۱۳۹۲). در پژوهشی که در سال ۱۳۹۲ انجام گرفت یک سری تحلیل‌های آماری بر روی یک دوره ۲۵ روزه لاگ جستجوی کاربرانی که به جستجوی اطلاعات در این پایگاه اطلاعاتی پرداخته‌اند، انجام شد (خسروی و مهموئی ۱۳۹۲). نتایج این پژوهش روزهای پر استفاده و پربسامدترین طول عبارت جستجو و موضوع جستجو را مشخص می‌کند. در این پژوهش تنها به ارائه یکسری گزارش‌های آماری اکتفا شده است که می‌تواند شروع‌کننده خوبی در مسیر این دست پژوهش‌ها باشد اما تحلیلی بر روی رفتارهای کاربران صورت نگرفته است.

در پژوهشی که اخیراً انجام گرفت (Rabiei and Haeri 2017)، محققان لاگ جستجوی کاربران را برای یک بازه طولانی‌تر (فوریه ۲۰۱۳ تا جولای ۲۰۱۵) را که شامل ۱۸۹۴۳۸ رکورد است، بررسی نمودند. تمرکز آن‌ها تنها بر روی جستجوهای انجام شده در حوزه علوم محیطی بود و از تکنیک متن‌کاوی و پردازش زبان طبیعی استفاده کردند.

با توجه به آمار بالای تعداد پژوهشگران با مراجعه حضوری و غیرحضوری به سامانه‌ی گنج و تنوع رشته‌های تحصیلی پژوهشگران، مطالعه رفتار اطلاع‌یابی پژوهشگران از طریق بررسی لاگ‌های جستجو بسیار حائز اهمیت است؛ بنابراین در این مقاله، تعیین ویژگی‌های مفید اطلاعاتی جهت تحلیل رفتار بهتر پژوهشگران و دسته‌بندی پژوهشگران از لحاظ عبارات جستجو، رشته‌های تحصیلی، تعداد نتایج مشاهده‌شده و زمان‌های استفاده (روزهای هفته و ساعات کاری) مورد بررسی قرار گرفت.

#### ۴- روش‌شناسی پژوهش

سامانه‌ی ملی اطلاعات پایان‌نامه‌ها/ رساله‌های دانش‌آموختگان داخل کشور تحت عنوان سامانه‌ی گنج در پژوهشگاه علوم و فناوری اطلاعات ایران، پیاده‌سازی شده است. بخش پایگاه داده‌ی آن با استفاده از نرم‌افزار MySQL پیاده‌سازی شده است. کاربران از طریق این سامانه قادر به جستجوی اطلاعات موردنیاز خود هستند. نمایی از سامانه در شکل ۱ آمده است.

خوش آمدید: مهمان(ورود)

جستجو جستجو پیشرفته جستجوی حرفه ای میز کار تماس با ما

عبارت جستجو:  ضمن:  همه جستجو

تعداد نتایج: 9408

مرتب سازی بر اساس:  ربط:  به صورت:  زوئی:

ارتزایی محلات شهری با تاکید بر احساس تعلق و هویت محله ای (نمونه موردی: محلات شهر قزوین)

**مقاله های همایش های ایران** ، اولین همایش ملی جغرافیا، شهرسازی و توسعه پایدار ، انجمن محیط زیست کوشش ، 1392

نویسنده: زهره داوودز ، رفیه بابایی

دسترسی به فایل نامرئض پیشینه های (رکوردفایل) که نشانه «پی.دی.اف.» ندارند، شدنی نیست.

محلات شهری سلول های حیوانی و سازنده پیکره شهرها می باشند که در حیات و زوال شهرها تأثیر بسیاری دارا هستند. با مطرح شدن رویکرد توسعه پایدار محله ای و توجه به توسعه پایدار در این مفیاس، صاحب نظران برای سعش پایدار محلات قری جغرافیا و شاخص های بسیاری را مطرح کردند. از جمله اشتراکات م وجود در میان دیدگاه های گوناگون طرح شده، تدکید و مکن آنها بر هویت محله ای و احساس تعلق ساکنین محلات نسبت به محله می باشد. در این تحقیق به ارتزایی محلات شهر قزوین بر اساس احساس تعلق و هویت محله ای پردا ...

**نمایه ها:**  
توسعه پایدار  ارتزایی محله ای  محله  هویت  احساس تعلق

**مختص کن**

**شماره ک انتشار:**  
 همه  
 گزارش دولتی (22)  
 پایانه (7254)  
 طرح پژوهشی (584)  
 مقاله های همایش های ایران (951)  
 مقاله های مجله های علمی (587)

**سال انتشار:**  
 همه  
 1384  
 1392  
 1156  
 1391  
 1037  
 1390

**نمایه:**  
 همه  
 تهران (127)  
 ایران (106)  
 ایران (103)  
 کارکنان (103)

شکل ۱- نمایی از سامانه گنج

در زمانی که کاربران از موتور جستجوی این سامانه استفاده می کنند، اطلاعات جستجوی آنان نیز توسط نرم افزار MySQL ثبت و ذخیره می شود. نمونه ای از لاگ های ذخیره شده در شکل ۲ آمده است:

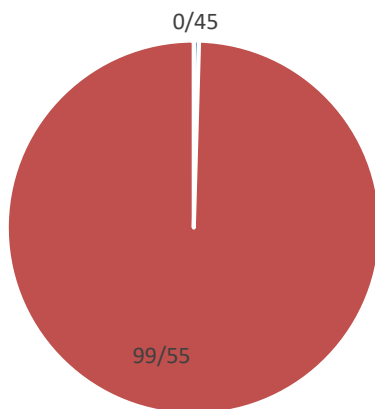
query	user_id	created_at	number	ip
1 انگوی مصرف	375827	2016-09-22 00:00:03	107	66.249.73.129
2 فن آوری اطلاعات	375827	2016-09-22 00:00:03	37	66.249.73.129
3 صور بیانی	375827	2016-09-22 00:00:04	99	5.160.134.165
4 برناجه توسعه	375827	2016-09-22 00:00:06	16552	66.249.73.129
5 معیار آزمون	375827	2016-09-22 00:00:08	3115	157.55.39.81
6 حمل و نقل دریایی	375827	2016-09-22 00:00:08	280	207.46.13.49
7 ساختار جنگل	375827	2016-09-22 00:00:10	11	157.55.39.7
8 استخراج با حلال	375827	2016-09-22 00:00:12	47	157.55.39.7
9 همی پلزی	375827	2016-09-22 00:00:16	55	66.249.73.129
10 (ان+۲)	375827	2016-09-22 00:00:17	2583	157.55.39.197
11 سوخت فسیلی	375827	2016-09-22 00:00:18	1870	66.249.73.129
12 تاب آوری	375827	2016-09-22 00:00:20	465	66.249.73.129
13 دست رفتن فرصت	375827	2016-09-22 00:00:23	77	188.34.160.35
14 ... اشتراک دانش و عوامل فنا	375827	2016-09-22 00:00:25	41	66.249.73.129
15 مسیربایی	375827	2016-09-22 00:00:26	206	66.249.73.129
16 ارزش افزوده اقتصادی	375827	2016-09-22 00:00:28	49	207.46.13.49
17 کارگران کارخانه	375827	2016-09-22 00:00:29	14	207.46.13.49
18 مدیریت پرستاری	375827	2016-09-22 00:00:33	250	66.249.73.129
19 دست رفتن فرصت	375827	2016-09-22 00:00:35	77	188.34.160.35
20 تخصیص ارز	375827	2016-09-22 00:00:36	90	66.249.73.129
21 سامانی	375827	2016-09-22 00:00:38	1274	66.249.73.129
22 آهن	375827	2016-09-22 00:00:38	302	68.180.228.226

شکل ۲- نمایی از لاگ موتور جستجوی گنج

فایل لاگ موتور جست‌وجوی سامانه گنج شامل اطلاعاتی از قبیل شناسه کاربر، عبارت جستجو، تعداد دفعات جستجو، مکان دسترسی کاربر، زمان و تاریخ جستجو است. در این پژوهش بررسی‌ها بر روی لاگ‌های سامانه‌ی گنج در سه ماه سوم سال ۱۳۹۵ انجام گرفت. حجم رکوردهای این لاگ فایل ۱۱۲۵۲۹۷۴ عدد است. با استفاده از زبان SQL، گزارش‌هایی بر روی این لاگ فایل اجرا شد و اطلاعات موردنیاز استخراج شد که نتایج آن در بخش بعدی ارائه می‌شود.

## ۵- تجزیه و تحلیل یافته‌ها

لاگ فایل موردبررسی در این پژوهش، در قالب سه ماه مهر، آبان و آذر تفکیک شده است. جهت تحلیل بهتر و منطقی‌تر لاگ فایل جستجوی کاربران را در دو حالت بررسی کرده‌ایم. در حالت اول کل جستجوهای که بر روی سامانه‌ی گنج و در حالت دوم تنها جستجوهایی که از طریق تالار جستجوی پژوهشگاه علوم و فناوری اطلاعات ایران صورت گرفته است بررسی شدند، یعنی جستجوی کاربرانی که به‌صورت حضوری و از داخل پژوهشگاه اقدام به جستجو کرده‌اند. با توجه به هدف مقاله که تحلیل رفتار اطلاع‌یابی پژوهشگران است، شایان ذکر است که جستجوهای داخل پژوهشگاه از جستجوهای خارج از آن، به دلایل حجم بالای جستجوها توسط خزش گرها و غیر هدفمند بودن جستجوها تفکیک شده است. قابل ذکر است که ۹۹،۵۵ درصد جستجوها از خارج از پژوهشگاه و تنها ۰،۴۵ درصد از جستجوها به‌صورت حضوری انجام گرفته است (شکل ۳).



شکل ۳- نسبت جستجوهای داخل و خارج از پژوهشگاه

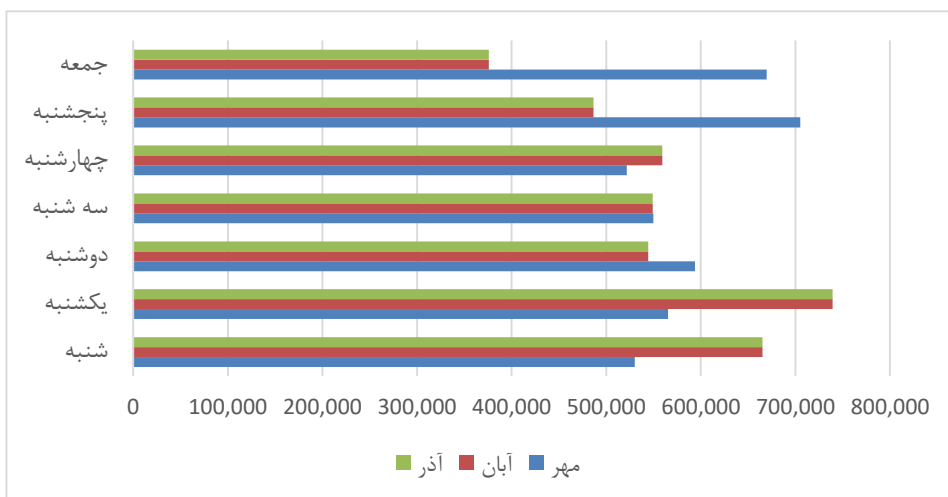
نتایج به دست آمده بر روی تحلیل داده‌های لاگ فایل سه‌ماهه‌ی سوم سال ۱۳۹۵ سامانه گنج به شرح زیر است:

جدول ۱- آمار تعداد جستجوها به تفکیک ماه

بازه جستجو	تعداد کل جستجوها	تعداد جستجوها از داخل پژوهشگاه
مهرماه ۱۳۹۵	۴,۱۳۵,۵۳۳	۵,۳۱۶
آبان ماه ۱۳۹۵	۳,۹۱۹,۹۰۱	۱۷,۱۹۲
آذرماه ۱۳۹۵	۳,۱۹۷,۵۴۰	۲۶,۳۸۱
تعداد کل	۱۱,۲۵۲,۹۷۴	۴۸,۸۸۹

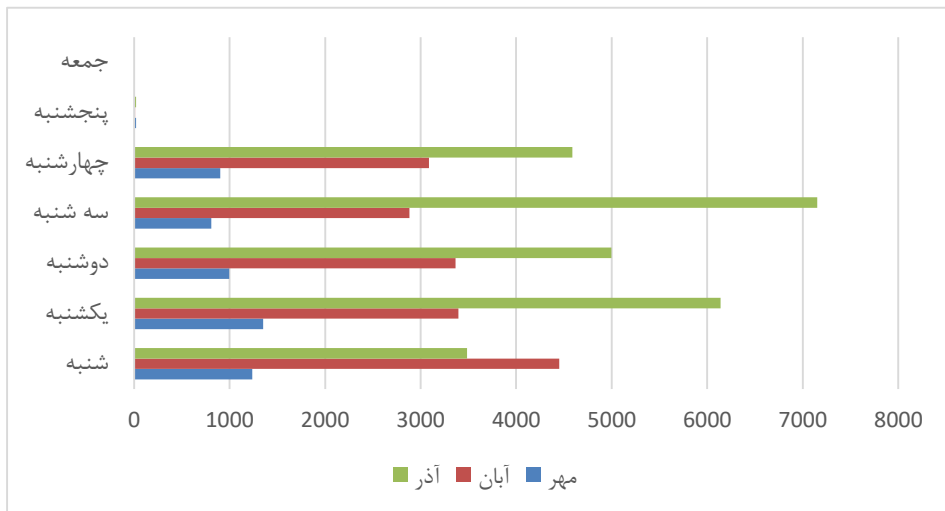
همان‌طور که مقادیر جدول ۱ نشان می‌دهد بیشترین آمار جستجو مربوط به مهرماه ۱۳۹۵ است که می‌توان دلیل آن را شروع سال تحصیلی و وجود شور و نشاط در ابتدای سال تحصیلی بین دانشجویان جهت انجام فعالیت‌های پژوهشی دانست؛ اما این تعداد برای جستجوهای حضوری در آذرماه به بیشینه مقدار خود رسیده است که دلیل آن را می‌توان نزدیکی به انتهای ترم تحصیلی، ارائه پروپوزال و دفاع از رساله در نظر گرفت.

با نگارش و اجرای یک پرس‌وجو بر روی فایل لاگ در نرم‌افزار Sql server حجم جستجوهای انجام شده بر روی سامانه‌ی گنج برای هر یک از روزهای هفته در هر یک از ماه‌های مهر، آبان و آذر به تفکیک به دست آمد که در شکل ۴ و ۵ به‌طور کامل نمایش داده شده است.



شکل ۴- میزان کل جستجوها به تفکیک روزهای هفته

شکل ۵- میزان جستجوها از داخل پژوهشگاه به تفکیک روزهای هفته



روزهای پنجشنبه در مهرماه بالاترین آمار جستجوی کاربران را گزارش می‌دهد و این حالت در روزهای یکشنبه در آبان و آذرماه مشاهده می‌شود و روزهای انتهای هفته در آبان و آذرماه از نرخ جستجوی بسیار کمی نسبت به سایر روزها برخوردار هستند (شکل ۴). درحالی‌که در جستجوهای حضوری با توجه به تعطیلی پژوهشگاه در روزهای پنجشنبه و جمعه، بالاترین حجم جستجو مربوط به سه‌شنبه در آذرماه و روزهای شنبه در آبان ماه و روز یکشنبه در مهرماه است.

حجم جستجوی انجام‌شده بر روی سامانه‌ی گنج به‌طور کلی بر اساس روزهای هفته در جدول ۲ آمده

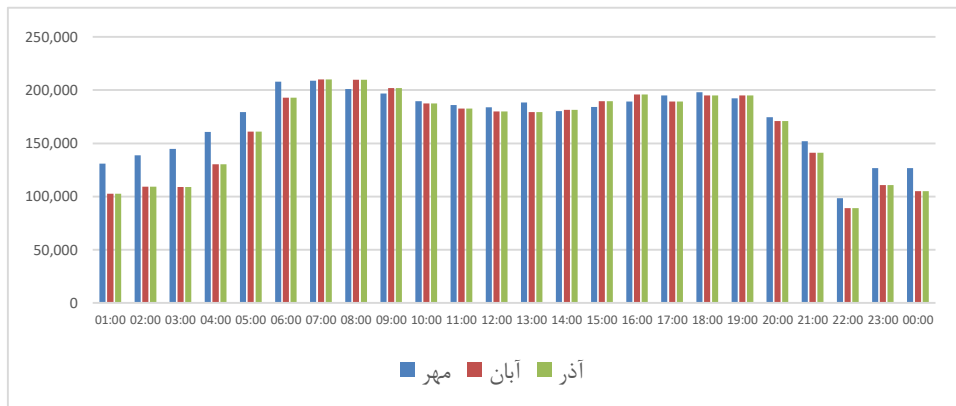
است:

جدول ۲- حجم جستجو بر اساس روزهای هفته سامانه گنج سه ماه سوم سال ۱۳۹۵

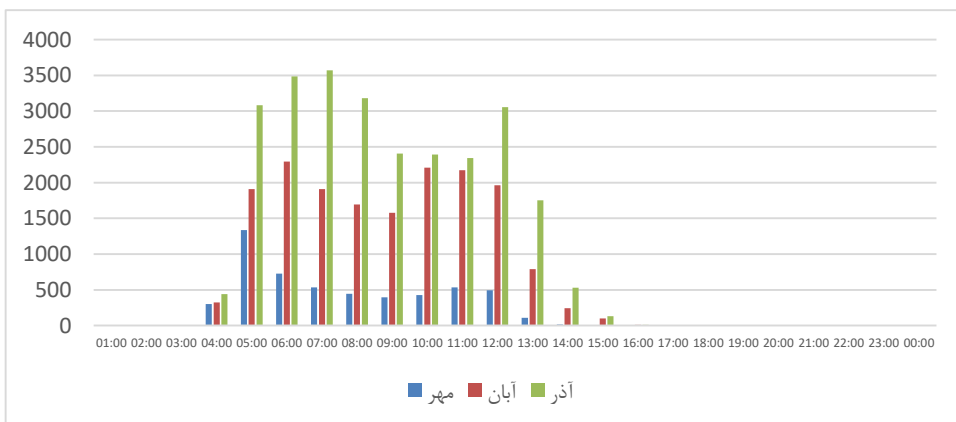
روزهای هفته	حجم جستجو بر اساس روزهای هفته	کل جستجوها	داخل پژوهشگاه
شنبه	۱۵۹۸۶۲۷	۹۱۷۴	
یکشنبه	۱۷۲۹۴۵۸	۱۰۸۸۶	
دوشنبه	۱۶۸۳۵۰۰	۹۳۵۷	
سه‌شنبه	۱۶۲۵۷۶۱	۱۰۸۴۵	
چهارشنبه	۱۶۳۴۹۸۱	۱۰۲۴۴	
پنجشنبه	۱۶۱۱۰۸۳	۵۱	
جمعه	۱۴۸۵۰۵۳	۰	

نتایج نشان می‌دهد که روز یکشنبه در بین سه‌ماهه‌ی سوم سال ۱۳۹۵ دارای بالاترین حجم جستجو هم به‌طور کلی و هم در داخل پژوهشگاه است.

در تحلیل دیگری که بر روی داده‌های لاگ کاربران سامانه‌ی گنج صورت گرفت، حجم جستجوها بر اساس ساعات شبانه‌روز به دست آمد.



شکل ۶- حجم کل جستجوها بر اساس ساعات شبانه‌روز به تفکیک ماه

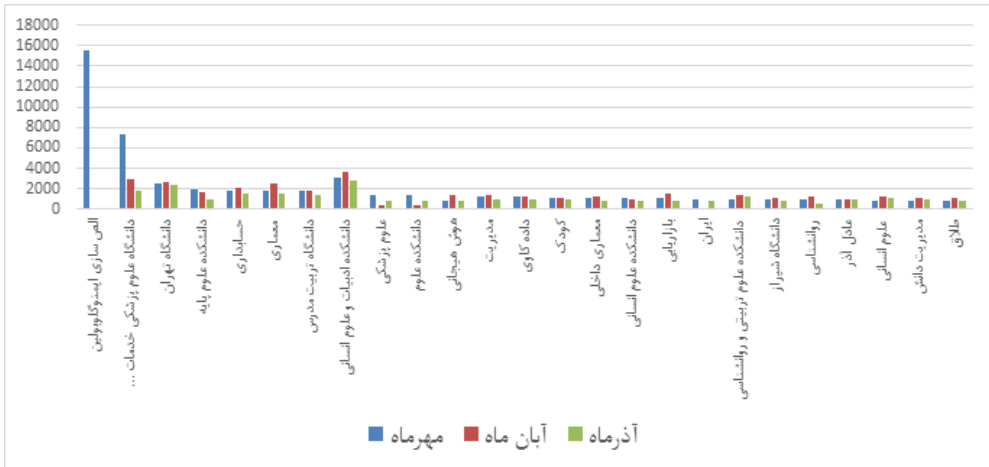


شکل ۷- حجم جستجوهای داخل پژوهشگاه بر اساس ساعات شبانه‌روز



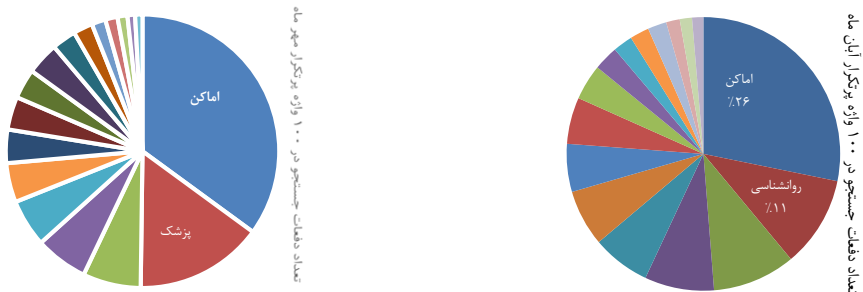
نتایج نمودار نشان می‌دهد که بیشترین حجم جستجوها در بین ساعات ۷ الی ۹ صبح و ۵ الی ۷ بعدازظهر است. در زمانی که افراد از داخل پژوهشگاه جستجو را انجام می‌دهند، زمان پرتراфик جستجو در ساعت ۶ صبح و ۱۲ ظهر است.

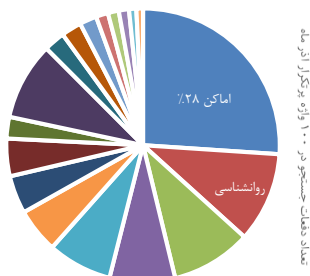
در تحلیل دیگری، بیشترین واژه‌های مورد جستجو در سامانه‌ی گنج نیز به تفکیک ماه به دست آمد (شکل ۸).



شکل ۸- بیشترین واژه‌های مورد جستجو در سامانه‌ی گنج در سه ماه سوم سال ۱۳۹۵

با توجه به اینکه در نگاه اولیه بر روی نمودارها، اکثر عبارات مورد جستجو مربوط به مکان‌ها هستند؛ تصمیم گرفته شد ۱۰۰ واژه‌ی پرتکرار در هرماه به‌صورت جداگانه بررسی شود و برای هرکدام از آن‌ها یک برچسب متناسب با موضوع در نظر گرفته شود. این برچسب‌ها که شامل اماکن، روانشناسی، علوم اجتماعی، تجارت، مدیریت، فنی مهندسی، فناوری اطلاعات، فرهنگ، هنر، پزشکی، علوم تربیتی، حسابداری، علوم انسانی، ادبیات، افراد، محیط‌زیست، حقوق، سیاست و غیره است، به عبارت جستجو نسبت داده شدند.





شکل ۹- سهم جستجوی لغات در بین ۱۰۰ واژه پرتکرار در جستجوها

همان‌طور که نتایج نشان می‌دهند اکثر جستجوهای انجام‌شده خارج از پژوهشگاه توسط خزش‌گرها صورت گرفته است. به‌طور مثال واژه‌ی «خالص‌سازی ایمنوگلوبولین» توسط یک IP به تعداد بیش از ۱۵۰۰۰ مرتبه در یک بازه زمانی کوتاه مورد جستجو واقع شده است. همچنین مشاهده می‌شود تعداد زیادی از جستجوها عبارت‌هایی هستند که به مکان‌ها مربوط می‌شوند. به نظر می‌رسد برای کاربران مکان انجام پژوهش از اهمیت بسزایی برخوردار است. به‌طور مثال واژه دانشگاه تهران یکی از عبارت‌های پرتکرار است. حال، اگر اماکن را در نظر نگرفته و کلیدواژه‌های پرتکرار را در نظر گرفته شود، لیست آن‌ها در جدول ۳ قابل مشاهده است.

جدول ۳- لیست کلیدواژه‌های پرتکرار در رشته‌های مختلف در کل جستجوها به همراه تعداد دفعات جستجو

<ul style="list-style-type: none"> <li>• هوش هیجانی (۳۱۸۵)</li> <li>• روانشناسی (۲۸۳۶)</li> <li>• رضایت زناشویی (۲۵۳۷)</li> <li>• رضایت شغلی (۲۵۳۳)</li> <li>• افسردگی (۲۳۰۸)</li> </ul>	روانشناسی (۲۸۰۱۷)
<ul style="list-style-type: none"> <li>• خالص‌سازی ایمنوگلوبولین (۱۵۶۱۶)</li> <li>• علوم پزشکی (۲۳۸۳)</li> <li>• اوتیسم (۱۷۸۰)</li> </ul>	پزشکی (۱۹۷۷۹)
<ul style="list-style-type: none"> <li>• معماری (۵۹۷۴)</li> <li>• معماری داخلی (۳۲۲۸)</li> <li>• زنجیره تأمین (۱۹۳۱)</li> <li>• طراحی معماری (۱۷۰۸)</li> <li>• مهندسی صنایع (۱۷۰۱)</li> </ul>	فنی مهندسی (۱۶۱۲۷)
<ul style="list-style-type: none"> <li>• طلاق (۲۸۰۹)</li> <li>• سرمایه اجتماعی (۲۰۸۰)</li> <li>• اعتیاد (۲۰۱۶)</li> <li>• اخلاقی حرفه‌ای (۱۹۷۸)</li> </ul>	علوم اجتماعی (۱۵۹۸۹)

<ul style="list-style-type: none"> <li>• سرمایه فکری (۱۹۶۳)</li> </ul>	
<ul style="list-style-type: none"> <li>• مدیریت (۳۷۴۸)</li> <li>• مدیریت دانش (۳۰۳۱)</li> <li>• مدیریت مالی (۱۷۷۰)</li> <li>• تعهد سازمانی (۱۷۶۹)</li> <li>• کارآفرینی (۱۷۶۹)</li> </ul>	مدیریت (۱۵۳۰۸)
<ul style="list-style-type: none"> <li>• داده‌کاوی (۳۴۹۵)</li> <li>• فناوری اطلاعات (۲۴۳۹)</li> <li>• تحلیل محتوا (۲۲۴۰)</li> <li>• فضای مجازی (۱۸۴۷)</li> <li>• پردازش تصویر (۱۸۰۴)</li> </ul>	فناوری اطلاعات (۱۵۱۳۱)
<ul style="list-style-type: none"> <li>• بازاریابی (۳۳۵۵)</li> <li>• تولید (۲۰۱۰)</li> <li>• توسعه پایدار (۱۹۹۸)</li> <li>• اقتصاد (۱۹۸۲)</li> <li>• اقتصاد مقاومتی (۱۷۹۵)</li> </ul>	تجارت (۱۲۷۷۹)
<ul style="list-style-type: none"> <li>• تصویرسازی (۲۳۷۱)</li> <li>• هنر (۲۳۶۴)</li> <li>• سینما (۱۷۸۱)</li> <li>• گرافیک (۱۷۷۸)</li> </ul>	هنر (۹۹۰۷)
<ul style="list-style-type: none"> <li>• علوم انسانی (۳۲۶۸)</li> <li>• مسئولیت مدنی (۲۳۳۹)</li> <li>• برنامه‌ریزی درسی (۲۲۰۷)</li> </ul>	علوم انسانی (۷۸۱۴)

در حالت کلی لیست کلیدواژه‌های پرتکرار ارائه شد، اگر تنها IP های داخل پژوهشگاه مدنظر قرار گیرد، لیست کلیدواژه‌های مورد جستجوی کاربران تا حدی متفاوت است (جدول ۴).

جدول ۴- لیست کلیدواژه‌های پرتکرار در رشته‌های مختلف در جستجوهای داخل پژوهشگاه

<ul style="list-style-type: none"> <li>• هوش چندگانه (۱۳)</li> <li>• سلامت روان تاب‌آوری (۱۱)</li> </ul>	روانشناسی
<ul style="list-style-type: none"> <li>• بیماری‌های کودکان (۱۲)</li> <li>• حسی-حرکتی (۱۲)</li> </ul>	پزشکی
<ul style="list-style-type: none"> <li>• معماری (۱۲۳)</li> <li>• برق (۳۰)</li> <li>• فرآیند طراحی معماری (۱۴)</li> <li>• طراحی رویکرد مصرف انرژی (۱۳)</li> </ul>	فنی مهندسی

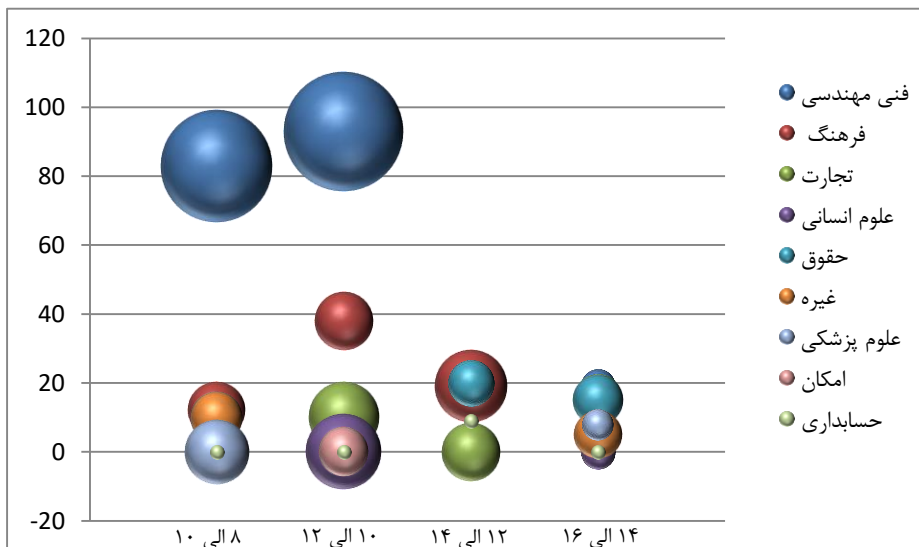
علوم اجتماعی	<ul style="list-style-type: none"> <li>• تغییر جنسیت حقوق   حقوقی (۱۲)</li> <li>• فسخ نکاح (۱۱)</li> </ul>
مدیریت	<ul style="list-style-type: none"> <li>• مدیریت دانش (۹۱)</li> <li>• نقشه دانش (۲۶)</li> <li>• برنامه استراتژیک &amp; مدیریت دانش (۱۷)</li> </ul>
فناوری اطلاعات	<ul style="list-style-type: none"> <li>• آموزش مجازی (۶۳)</li> <li>• آموزش الکترونیکی (۲۵)</li> <li>• امکان‌سنجی آموزش مجازی (۱۹)</li> <li>• دانشگاه مجازی (۱۷)</li> <li>• سایبر (۱۷)</li> </ul>
تجارت	<ul style="list-style-type: none"> <li>• اقتصاد مقاومتی (۳۴)</li> </ul>
هنر	<ul style="list-style-type: none"> <li>• فرهنگ &amp; رسانه (۱۹)</li> </ul>
علوم انسانی	<ul style="list-style-type: none"> <li>• علم اطلاعات و دانش شناسی (۳۶)</li> <li>• فلسفه (۲۳)</li> <li>• مثنوی (کتاب) (۲۲)</li> <li>• رفتار اطلاعاتیابی (۱۴)</li> </ul>

همچنین کلمات پرتکرار برحسب زمینه‌های مختلف جستجو در ساعات مختلف روز مورد بررسی قرار گرفت و نتایج آن در جدول ۵ و شکل ۱۰ قابل مشاهده است. به‌عنوان مثال افراد علاقه‌مند به موضوعات فنی-مهندسی و فناوری اطلاعات بیشتر در صبح و ساعات ۸ الی ۱۲ جستجوهای خود را انجام می‌دهند. این در حالی است که موضوعات مربوط به علوم انسانی بیشتر در ساعات ۱۲ الی ۱۶ مورد جستجو است. شایان ذکر است که توجه به نتایج این رفتارها می‌تواند در ارائه خدمات بهتر آموزشی به تفکیک گروه‌های مختلف کمک قابل توجهی نماید. افزون‌براین، بهبود سرعت بازیابی نتایج جستجو به تفکیک گروه‌های مختلف توسط الگوریتم‌های موتور جستجو می‌تواند یکی دیگر از دستاوردهای این نتایج باشد.

جدول ۵- آمار مربوط به جستجوی کلمات برحسب ساعات مختلف جستجو

موضوعات مورد جستجو	۸ الی ۱۰ صبح	۱۰ تا ۱۲ صبح	۱۲ الی ۱۴ بعدازظهر	۱۴ الی ۱۶ بعدازظهر
فنی مهندسی	۸۳	۹۳	۱۹	۱۹
فناوری اطلاعات	۶۷	۷۶	۰	۶
فرهنگ	۱۲	۳۸	۱۹	۶
مدیریت	۱۷	۱۸	۲۸	۰
تجارت	۳۱	۱۰	۰	۱۶
روانشناسی	۰	۲۶	۱۸	۱۱
علوم انسانی	۰	۰	۴۹	۰

۶	۰	۳۰	۱۰	علوم اجتماعی
۱۵	۲۰	۰	۱۰	حقوق
۱۳	۱۱	۰	۱۴	افراد
۵	۲۱	۰	۱۰	غیره
۱۲	۰	۹	۱۴	مذهب
۸	۱۰	۱۲	۰	علوم پزشکی
۵	۰	۰	۲۲	هنر
۰	۱۴	۰	۰	امکان
۰	۰	۱۳	۰	علوم پایه
۰	۹	۰	۰	حسابداری



شکل ۱۰- آمار مربوط به جستجوی کلمات برحسب ساعات مختلف جستجو

با توجه به اینکه در سه‌ماهه پاییز ۹۵، ۵۶ روز کاری وجود داشته است و تعداد کل کلیدواژه‌های جستجو ۴۸،۸۸۹ عدد بوده است، بنابراین به‌طور متوسط در هر روز ۸۷۳ کلیدواژه جستجو شده است و در هر ساعت تنها ۱۰۹ واژه در سامانه گنج به‌عنوان کلیدواژه مورد جستجو قرار گرفته است. همچنین با توجه به آمار تعداد حضور کاربران که روزانه به‌طور متوسط ۱۸۰ نفر ثبت شده است، کاربران با متوسط ۴،۸۵ واژه به نتیجه رسیده‌اند و یا جستجوی خود را متوقف نموده‌اند. بدیهی است برای جستجوی حوزه پژوهشی در نگارش پیشنهاد، پایان‌نامه و یا رساله اکتفا نمودن به حداکثر پنج واژه نمی‌تواند پژوهش‌مدنظر را جامع نماید.

همان‌طور که در جدول ۵ و شکل ۱۰ ملاحظه می‌گردد کاربران علاقه‌مند به حوزه فنی-مهندسی بیشتر در ابتدای روز به پژوهشگاه مراجعه می‌نمایند، درحالی‌که علاقه‌مندان به حقوق در ساعات بعدازظهر

مراجعه می‌نمایند. کلیدواژه «تجارت» در کل روز تقریباً به‌طور مساوی مراجعه‌کننده خود را دارد و علاقه‌مندان کلیدواژه «فرهنگ» بیشتر ساعات میانی را ترجیح می‌دهند. به‌طور کلی کاربران علوم انسانی ساعات بعدازظهر را برای جستجوی خود مطلوب‌تر می‌دانند؛ بنابراین به نظر می‌رسد که می‌توان تبلیغات برنامه‌های آموزشی را به تفکیک حوزه‌های کاربران در طول روز بخش‌بندی نمود به‌طوری‌که این تبلیغات نیمه اول روز بیشتر معطوف به حوزه فنی مهندسی باشد و حوزه علوم انسانی در بعدازظهر انجام گیرد. همچنین با توجه به اینکه اماکن بیشترین جستجو را به خود اختصاص داده است، می‌توان با آموزش کاربران جستجوی آن‌ها را طوری هدایت نمود که بتوانند به نحو مطلوب‌تری به تسلط نسبی در حوزه پژوهش موردعلاقه خود دست یابند و استادان و پژوهشگران این حوزه را با سرعت بیشتری شناسایی نمایند.

## ۶- نتیجه‌گیری

در این پژوهش به تحلیل لاگ‌های کاربران سامانه‌ی گنج پرداخته شد و نتایج این تحلیل نشان می‌دهد که بیش از ۹۹ درصد کاربران تمایل دارند از بیرون پژوهشگاه به جستجو بپردازند و کمتر از یک درصد از آن‌ها به دلیل دسترسی به متن کل پایان‌نامه و رساله به تالار جستجوی پژوهشگاه مراجعه حضوری دارند. این امر حاکی از آن است که خدمات بهتر و مطلوب‌تری به پژوهشگران حضوری می‌بایست ارائه گردد تا تالار جستجو بتواند تعداد پژوهشگران بیشتری را جذب نماید. همچنین با توجه به اینکه هر کاربر در مراجعه حضوری به‌طور متوسط پس از جستجوی پنج کلیدواژه فرآیند جستجوی خود را متوقف می‌نماید نیاز است که پایگاه گنج به نحو جذاب‌تری طراحی گردد تا پژوهشگر بتواند در حوزه پژوهش خود دانش بیشتری را کسب نماید.

نتایج پژوهش بیانگر آن است که پژوهشگران حوزه علوم انسانی بیشتر علاقه‌مند به حضور در ساعات بعدازظهر هستند و پژوهشگران فنی-مهندسی بیشتر ساعات اولیه روز را ترجیح می‌دهند. بر اساس این یافته، می‌توان خدمات و دوره‌های آموزشی تخصصی را بهتر برنامه‌ریزی نمود. افزون بر این، تعیین کلیدواژه‌های پرتکرار رشته‌های تحصیلی به مکان پژوهش اختصاص دارد که خود می‌تواند به‌عنوان مؤلفه‌ای در بهبود الگوریتم‌های موتور جستجو جهت افزایش سرعت بازایی اطلاعات کمک بسزایی نماید. قابل‌ذکر است این پژوهش مانند هر پژوهشی، محدودیت‌هایی داشته است. محدودیت نخست این بود که تحلیل لاگ‌ها تنها بر روی ۳ ماهه سوم سال ۹۵ صورت گرفت و حجم محدودی از کل لاگ‌ها را شامل می‌شد. محدودیت دوم این بود که لاگ‌های سامانه‌ی گنج، اطلاعات کمی را ذخیره می‌کنند که خود این مسئله یافته‌های کمتری را در اختیار پژوهش قرار می‌دهد. ویژگی‌های ذخیره‌شده در لاگ‌ها، تنها شامل عبارت مورد جستجو، زمان جستجو، تعداد نتایج جستجو و محل جستجو است برای پژوهش‌های آتی می‌توان به ارتباط کلیدواژه‌ها، خوشه‌بندی آن‌ها و تشکیل گراف برای تحلیل بیشتر و دقیق‌تر رفتار کاربران اشاره نمود. همچنین می‌توان ویژگی‌های بیشتری را در لاگ‌ها ذخیره نمود از جمله اینکه کلیک کاربر بر روی کدام نتایج بوده است، زمان کلیک کاربر بر روی هر نتیجه، مدت‌زمان حضور کاربر بر روی هر یک از نتایج، جنسیت کاربران، مقطع تحصیلی و سن آن‌ها. واضح است با ذخیره کردن

ویژگی‌های بیشتر از کاربران و رفتار آن‌ها در زمان تعامل با سامانه گنج، تحلیل‌های بیشتری می‌توان ارائه کرد.

## فهرست منابع

چریری، نجلا؛ اسدی، مریم؛ نوشین فرد، فاطمه. ۱۳۹۲. تحلیل رفتاری جستجوی اطلاعات پژوهشگران حوزه‌های مختلف علوم از وب بر اساس سبک‌های شناختی کلامی و تصویری. پژوهشنامه‌ی پردازش و مدیریت اطلاعات

جمالی مهموئی، حمیدرضا. ۱۳۸۳. یادداستی بر روش تحلیل گزارش‌های وب (وبلاگ). اطلاع‌شناسی ۱. (۳): ۱۸۶-۱۷۵

خسروی، مریم؛ مهموئی، حمیدرضا. ۱۳۹۲. تحلیل لاگ پایگاه اطلاعات و مدارک علمی ایران (ایراندک) و رفتار جستجوی کاربران آن. پژوهشنامه‌ی پردازش و مدیریت اطلاعات. پژوهشگاه علوم و فناوری اطلاعات ایران.

رجبعلی بگلو، رضا. ۱۳۸۶. آشنایی با شیوه‌های تحلیل گزارش: فرایند تعامل در نظام‌های الکترونیکی. فصلنامه کتابداری و اطلاع‌رسانی. ۱۰ (۳۹).

زوارقی، رسول. ۱۳۸۴. تحلیل گزارش‌های وب، روشی نوین برای ارزیابی عملکرد وبسایت‌ها: مطالعه‌ی موردی: وب ایران سایت مرکز اطلاعات و مدارک علمی. اطلاع‌شناسی ۲ (۳-۴): ۱۱۶-۸۸.

ستوده، هاجر. ۱۳۸۲. «روش تحلیل گزارش‌های وب (وبلاگ). اطلاع‌شناسی. ۱ (۱): ۸۴-۶۹»

Alspaugh, S., Di Chen, B., Lin, J., Ganapathi, A., Hearst, M. A., & Katz, R. H. (2014, November). Analyzing Log Analysis: An Empirical Study of User Log Mining. In *LISA* (pp. 53-68).

Ahiau, B. E., & Ani, O. E. (2015). A Survey of Information Seeking-Behavior of Academic Staff in a Nigerian University in Digital Age. *Science*, 3(4), 89-94.

Beitzel, S. M., Jensen, E. C., Chowdhury, A., Grossman, D., & Frieder, O. (2004, July). Hourly analysis of a very large topically categorized web query log. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 321-328). ACM.

Bendersky, M., & Croft, W. B. (2009, February). Analysis of long queries in a large scale search log. In *Proceedings of the 2009 workshop on Web Search Click Data* (pp. 8-14). ACM.



- Cetindil, I., Esmaelnezhad, J., Li, C., & Newman, D. (2012). Analysis of Instant Search Query Logs. In *WebDB* (pp. 7-12).
- Downey, D., Dumais, S., Liebling, D., & Horvitz, E. (2008, October). Understanding the relationship between searchers' queries and information goals. In *Proceedings of the 17th ACM conference on Information and knowledge management* (pp. 449-458). ACM.
- Ellis, D., & Oldman, H. (2005). The English literature researcher in the age of the Internet. *Journal of Information Science*, 31(1), 29-36.
- Facca, F. M., & Lanzi, P. L. (2005). Mining interesting knowledge from weblogs: a survey. *Data & Knowledge Engineering*, 53(3), 225-241.
- Gan, Q., Attenberg, J., Markowetz, A., & Suel, T. (2008, April). Analysis of geographic queries in a search engine log. In *Proceedings of the first international workshop on Location and the web* (pp. 49-56). ACM.
- Gleeson, A. C. (2001). Information-seeking behavior of scientists and their adaptation to electronic journals. *University of North Carolina dissertation*.
- Jansen, B. J., Taksa, I., & Spink, A. (2009). Handbook of Research on Web Log Analysis, chapter Research and Methodological Foundations of Transaction Log Analysis. *IGI Global*, 226-253.
- Jansen, B. J. (2009). Understanding user-web interactions via web analytics. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1), 1-102.
- Jansen, B. J., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the web. *Information processing & management*, 36(2), 207-227.
- Jansen, B. J., & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information processing & management*, 42(1), 248-263.
- Ke, H. R., Kwakkelaar, R., Tai, Y. M., & Chen, L. C. (2002). Exploring behavior of E-journal users in science and technology: Transaction log analysis of Elsevier's ScienceDirect OnSite in Taiwan. *Library & Information Science Research*, 24(3), 265-291.

- Kumar, A., Salmani, N., & Baweja, S. (2014). Information Seeking Behaviour by the Research Scholars & Faculty Members: A Survey Study of Kurukshetra University Kurukshetra in the Disciplines of Life Science. *IOSR Journal Of Humanities And Social Science (IOSR-JHSS)*, 19(6), 119-138.
- Kurth, M. (1993). The limits and limitations of transaction log analysis. *Library Hi Tech*, 11(2), 98-104.
- Marchionini, G. (1989). Information-seeking strategies of novices using a full-text electronic encyclopedia. *Journal of the American Society for Information Science*, 40(1), 54.
- Morzy, T., Wojciechowski, M., & Zakrzewicz, M. (2000). Web users clustering. In *Proc. of the 15th International Symposium on Computer and Information Sciences* (pp. 374-382).
- Murray, G. C., & Teevan, J. (2007, December). Query log analysis: social and technological challenges. In *ACM SIGIR Forum* (Vol. 41, No. 2, pp. 112-120). ACM.
- Nicholas, D., Huntington, P., & Jamali, H. R. (2008). User diversity: as demonstrated by deep log analysis. *The Electronic Library*, 26(1), 21-38.
- Oduntan, O. (2014). Information seeking behavior on an academic library search engine.
- Park, M., & Lee, T. S. (2013). Understanding science and technology information users through transaction log analysis. *Library Hi Tech*, 31(1), 123-140.
- Park, M., Park, M., Lee, T. S., & Lee, T. S. (2016). A longitudinal study of information needs and search behaviors in science and technology: A query analysis. *The Electronic Library*, 34(1), 83-98.
- Pareek, A. K., & Rana, M. S. (2013). Study of information seeking behavior and library use pattern of researchers in the Banasthali University. *Library Philosophy and Practice* (e-journal), Paper 887, <http://digitalcommons.unl.edu/libphilprac/887>
- Peters, T. A. (1993). The history and development of transaction log analysis. *Library hi tech*, 11(2), 41-66.

- Rose, D. E., & Levinson, D. (2004, May). Understanding user goals in web search. In *Proceedings of the 13th international conference on World Wide Web* (pp. 13-19). ACM.
- Rabiei, M., Hosseini-Motlagh, S. M., & Haeri, A. (2017). Using text mining techniques for identifying research gaps and priorities: a case study of the environmental science in Iran. *Scientometrics*, 110(2), 815-842.
- Sato, S., Nishiura, M., Nagai, Y., Itsumura, H., ساتو شيو، نيشورا ميناكو، ... & 永井裕子. (2015). Usage Log Analysis of Articles in Five Japanese Institutional Repositories: the Relationships between Users, Access Paths, and Accessed Articles. *同志社図書館情報学= Doshisha Journal of Library and Infomation Science*, (25), 20-37.
- Shivalingaiah, D. (2010). Information seeking patterns of researchers in the university libraries in Karnataka state. *SRELS Journal of Information Management*, 47(1), 83-101.
- Silvestri, F. (2009). Mining query logs: Turning search usage data into knowledge. *Foundations and Trends® in Information Retrieval*, 4(1-2), 1-174.
- Spink, A., & Jansen, B. J. (2006). Searching multimedia federated content web collections. *Online Information Review*, 30(5), 485-495.
- Tyler, S. K., & Teevan, J. (2010, February). Large scale query log analysis of re-finding. In *Proceedings of the third ACM international conference on Web search and data mining* (pp. 191-200). ACM.
- Wang, P., Berry, M. W., & Yang, Y. (2003). Mining longitudinal Web queries: Trends and patterns. *Journal of the Association for Information Science and Technology*, 54(8), 743-758.
- Wen, J. R., Nie, J. Y., & Zhang, H. J. (2002). Query clustering using user logs. *ACM Transactions on Information Systems*, 20(1), 59-81.
- White, R. W., & Morris, D. (2007, July). Investigating the querying and browsing behavior of advanced search engine users. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 255-262). ACM.
- Wilson, T. D. (2000). Human information behavior. *Informing science*, 3(2), 49-56.

Xie, K., Yu, H., & Cen, R. (2012). Using log mining to analyze user behavior on search engine. *Frontiers of Electrical and Electronic Engineering*, 7(2), 254-260.

Xi, W., Chowdhury, A., Sidhu, K., & Pass, G. (2001). Query Log Analysis. *Laporan teknis, American Online, Inc.*

Yi, K., Beheshti, J., Cole, C., Leide, J. E., & Large, A. (2006). User search behavior of domain-specific information retrieval systems: An analysis of the query logs from PsycINFO and ABC-Clio's Historical Abstracts/America: History and Life. *Journal of the Association for Information Science and Technology*, 57(9), 1208-1220.

Yu, L., & Apps, A. (2000). Studying e-journal user behavior using log files: the experience of superjournal. *Library & Information Science Research*, 22(3), 311-338.