

ORIGINAL ARTICLE

Detecting emerging Trends in Articles on Iranian Medical librarianship and Information Using the TFIDF Algorithm

Meisam Dastani¹, Soraya Ziaei², Faeze Delghandi^{3*}

1. Ph.D, Determinants of Knowledge and Information Science, Social Health Research Center, Gonabad University of Medical Sciences, Gonabad, Iran.
2. Associate Professor, Department of Knowledge & Information Science, Payame Noor University, Tehran, Iran.
3. Assistant Professor, Department of Knowledge & Information Science, Payame Noor University, Tehran, Iran.

Correspondence
Faeze Delghandi
Email: delghandi@pnu.ac.ir

How to cite
Dastani, M., & Ziaei, S., & Delghandi, F (2024). Detecting emerging Trends in Articles on Iranian Medical librarianship and Information Using the TFIDF Algorithm. *Digital and Smart Libraries Researches*, 11(3), 61-74.

ABSTRACT

The present study identifies emerging trends in medical library and information science articles published in Iranian scientific-research journals. This exploratory research analyzes medical library and information science articles published in this field's journals in Iran from 1997 to 2020 using text mining techniques. The TF-IDF algorithm was employed to identify the most important terms used in the articles. Python programming language was utilized to implement the text mining algorithms.

The examination of emerging terms in articles published in medical library and information science journals indicates that terms such as LibQUAL, practical, and bibliotherapy have recently entered the articles and studies of this field in domestic journals during the period from 2005 to 2015. Similarly, terms like invention, altmetrics, and repository have recently appeared in the articles and studies of this field in domestic journals during the period from 2015 to 2020.

The results indicate that the terms used in medical library and information science articles have not remained constant over time and have undergone changes during different periods. This reflects that, in line with the emergence and growth of technology, this scientific field has also evolved.

Keywords: Library and Information Science, Medicine, Analysis, Keywords, Text Mining, Iran.

KEYWORDS

Library and Information Science, Medicine, Analysis, Keywords, Text Mining, Iran.



شناسایی روندهای نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی منتشر شده در ایران با استفاده از الگوریتم TFIDF

میثم داستانی^۱، ثریا ضیائی^۲، فائزه دلقدی^{۳*}

چکیده

پژوهش حاضر به شناسایی روندهای نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی منتشر شده در مجلات علمی-پژوهشی ایرانی پرداخته است. مطالعه اکتشافی حاضر به تجزیه و تحلیل مقالات کتابداری و اطلاع‌رسانی پزشکی در مجلات این حوزه در ایران از سال ۱۳۷۶ تا ۱۳۹۸ با استفاده از فنون متن کاوی پرداخته است. جهت شناسایی مهم‌ترین واژگان به کار رفته در مقالات از الگوریتم TF-IDF استفاده شده است. زبان برنامه‌نویسی پایتون نیز جهت اجرای الگوریتم‌های متن کاوی به کار گرفته شده است. بررسی واژگان نوظهور در مقالات منتشر شده در مجلات کتابداری و اطلاع‌رسانی پزشکی نشان می‌دهد که واژگان لیب کوال، عملی، بایلیوترایی در بازه زمانی ۱۳۸۴ تا ۱۳۹۴ به‌تازگی وارد مقالات و مطالعات حوزه کتابداری و اطلاع‌رسانی پزشکی در مجلات داخلی شده است. همچنین واژگان اختراع، آلت‌متریک، مخزن، بازه زمانی ۱۳۹۴ تا ۱۳۹۹ به‌تازگی وارد مقالات و مطالعات حوزه کتابداری و اطلاع‌رسانی پزشکی در مجلات داخلی شده است. نتایج نشان‌دهنده آن است که واژگان مقالات کتابداری و اطلاع‌رسانی پزشکی در طول زمان ثابت نبوده و در بازه‌های زمانی مختلف، دچار تغییراتی شده است. این امر نشان‌دهنده آن است که همگام با ظهور و رشد فناوری، این رشته علمی نیز تغییر یافته است.

واژه‌های کلیدی

کتابداری و اطلاع‌رسانی، پزشکی، تجزیه و تحلیل، کلیدواژه، متن کاوی، ایران.

۱. دکتری، گروه علم اطلاعات و دانش‌شناسی، مرکز تحقیقات عوامل اجتماعی مؤثر سلامت، دانشگاه علوم پزشکی گناباد، گناباد، ایران.
۲. دانشیار، گروه علم اطلاعات و دانش‌شناسی، دانشگاه پیام نور، تهران، ایران.
۳. استادیار، گروه علم اطلاعات و دانش‌شناسی، دانشگاه پیام نور، تهران، ایران.

نویسنده مسئول: فائزه دلقدی
رایانامه: delghandi@pnu.ac.ir

استناد به این مقاله:

داستانی، میثم؛ ضیائی، ثریا و دلقدی، فائزه (۱۴۰۳). شناسایی روندهای نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی منتشر شده در ایران با استفاده از الگوریتم TFIDF. پژوهش‌های کتابخانه‌های دیجیتالی و هوشمند، ۱۱(۳)، ۶۱-۷۴.

مقدمه

سال ۱۸۹۸ خود دلیلی بر این ادعا است (دیمیتروف، ۱۹۹۲). تحقیقات اطلاعات سلامت برای کتابداران پزشکی در جهت ایجاد دانش جدید و پیشبرد عمل مبتنی بر شواهد ضروری است. انجمن کتابخانه پزشکی که شامل متخصصان حرفه‌ای اطلاعات علوم پزشکی می‌شود، بر اهمیت و مزایای پژوهش در علوم اطلاع‌رسانی پزشکی تأکید دارند (لسیک و همکاران، ۲۰۱۶). بر همین اساس متخصصان و پژوهشگران کتابداری و اطلاع‌رسانی پزشکی در راستای اهداف و رسالت این حوزه علمی، پژوهش‌های مختلفی را انجام و در مجلات تخصصی رشته منتشر می‌نمایند. بنابراین با توجه به موارد مذکور و اهمیت موضوع، این پژوهش به شناسایی و تجزیه و تحلیل مهم‌ترین واژگان به کار رفته در مقالات کتابداری و اطلاع‌رسانی پزشکی منتشر شده در مجلات مرتبط با این موضوع در ایران پرداخته است.

مبانی نظری

امروزه مقالات علمی می‌توانند در پیشرفت و توسعه‌ی آن کشور و یا حتی کشورهای دیگر تأثیر قابل توجهی بگذارند، همین تأثیری که مقالات علمی روی پیشرفت و توسعه‌ی کشورها می‌گذارند، دلیلی بر مهم بودن بررسی این مقالات و طبقه‌بندی آن‌هاست، به طوری که بتوان بین انواع مقالات (از لحاظ کیفیت محتوایی) تفاوت‌هایی قائل شد. مقالات علمی، دانش پیشرفته و مدرن اکتشافات علمی را ثبت می‌کنند، بنابراین، بررسی وضعیت آن‌ها می‌تواند شکاف‌های علمی را شناسایی کند (ونگ و همکاران^۱، ۲۰۱۵). همچنین با توجه به اینکه مقالات علمی شاخه‌ای از ادبیات تحقیق و شرح نتایج علمی در زمینه‌های مختلف علمی و دانشگاهی را نشان می‌دهد و همچنین ابزاری برای تبادلات علمی را نیز ارائه می‌دهد. از آنجا که آخرین اطلاعات فناوری‌ها اغلب در مقالات علمی مورد بحث و تبادل نظر قرار می‌گیرد، مقاله علمی یک حامل مهم اطلاعات در مورد فناوری و یک منبع مهم داده برای مطالعه توسعه و تغییر فناوری است (بهکامی و دیم^۲، ۲۰۱۲). با افزایش روزافزون مقالات علمی در هر یک از حوزه‌های علمی، شناسایی موضوعات و پژوهش‌های مطرح و مهم علمی یک حوزه خاص از ضرورت‌های علمی می‌باشد، همچنین با این حجم عظیم از مقالات منتشر شده در سراسر جهان، ارزیابی و بررسی تک‌تک مقالات تقریباً کاری غیرممکن یا بسیار سخت و زمان‌بر می‌باشد، لذا در اینجا استخراج خودکار دانش از این

امروزه در سرتاسر جهان مستندات علمی زیادی توسط پژوهشگران، اساتدان و دانشجویان دانشگاه‌های مختلف تولید می‌شوند که اغلب دربرگیرنده‌ی مطالب مهم و مفید بوده و در قالب‌های متنی منتشر می‌شوند. با توجه به افزایش روزافزون مقالات علمی و با توجه به حجم عظیم از مقالات منتشر شده، ارزیابی و بررسی تک‌تک مقالات و استخراج دستی اطلاعات و دانش از این حجم عظیم از متون، کاری طاقت‌فرسا یا غیرممکن است. با وجود این شناسایی الگوها و استخراج دانش بالقوه در حجم زیادی از داده‌های متنی، یک امر مهم در زمینه‌های علمی مختلف محسوب می‌شود (هاشمی و همکاران^۱، ۲۰۱۵). از این رو ارائه ابزارها و فنونی که با بررسی خودکار متون بتوانند تحلیلی روی آن‌ها انجام دهند منجر به شکل‌گیری حوزه‌ی متن‌کاوی شده است. متن‌کاوی به تحلیل هوشمند متن، داده‌کاوی متنی یا کشف دانش از متن نیز مشهور است و به‌طور کلی به فرایند استخراج دانش و اطلاعات مورد علاقه و مهم از مجموعه متنی غیر ساخت‌یافته اشاره دارد (رهبولز-شامن و همکاران^۲، ۲۰۱۲؛ جنسن و همکاران^۳، ۲۰۱۲؛ رودریگز-استبان و باندچاس^۴، ۲۰۱۶). همچنین شناسایی الگوها و استخراج دانش بالقوه در حجم زیادی از داده‌های متون علمی، خبری و پژوهش‌ها، یک امر مهم در زمینه‌های علمی مختلف محسوب می‌شود (کائو و پوتیت^۵، ۲۰۰۷؛ لی و همکاران^۶، ۲۰۰۸؛ هانگ و ژانگ^۷، ۲۰۱۲).

یکی از الگوریتم‌ها و فنون متن‌کاوی جهت کشف دانش از حجم زیادی از اسناد، استخراج کلمات کلیدی است (راجمان و بسانکن، ۱۹۹۸). استخراج کلیدواژه تکنیکی مهم برای ارزیابی اسناد، ارزیابی صفحات وب، دسته‌بندی اسناد، خلاصه‌سازی و جمع‌بندی، استخراج متن و غیره است. کلمات کلیدی، عناصر اصلی نمایش مفاهیم دانش در نظر گرفته می‌شوند و معمولاً برای نشان دادن ساختار دانش حوزه‌های تحقیقاتی مورد استفاده قرار می‌گیرند (سو و لی^۸، ۲۰۱۰).

کتابداری و اطلاع‌رسانی پزشکی یکی از تخصص‌های حوزه علوم کتابداری و اطلاع‌رسانی است که بیش از ۱۰۰ سال قدمت دارد. تأسیس انجمن کتابداری و اطلاع‌رسانی پزشکی در

1. Hashimi et al
2. Rebbholz-Schuhmann et al
3. Jensen et al
4. Rodriguez-Esteban & Bundschus
5. Kao & Poteet
6. Lee et al
7. Hung & Zhang
8. Su & lee

یک حوزه تحقیقاتی کمتر موفق هستند. به‌عنوان مثال، «کتابخانه»، «منابع اطلاعاتی» و «خدمات» کلمات کلیدی پرتکرار در حوزه کتابخانه دیجیتال هستند، اما در واقع این کلمات احتمالاً مفاهیم جهانی‌ای هستند که در بسیاری از حوزه‌های تحقیقاتی دیگر در کتابخانه و علوم اطلاعات نیز وجود دارند. بنابراین، آن‌ها نمایش خوبی از موضوعات تحقیقاتی کتابخانه دیجیتال نیستند (چن^۳، ۲۰۰۶). یکی از روش‌های بسیار مفیدی که برای شناسایی کلمات کلیدی مهم در متون مورد استفاده قرار می‌گیرد الگوریتم^۴ TF-IDF است. این الگوریتم روشی برای سنجش اهمیت یک اصطلاح در یک سند است. TF-IDF عدد آماری است که میزان اهمیت یک کلمه نسبت به یک سند در مجموعه‌ای از اسناد را نشان می‌دهد. در واقع هدف آن، نشان دادن اهمیت کلمه در متن است. مقدار TF-IDF به تناسب تعداد تکرار کلمه در سند افزایش می‌یابد و توسط تعداد اسنادی که در مجموعه هستند و شامل کلمه نیز می‌باشند متعادل می‌شود. به این معنی که اگر کلمه‌ای در بسیاری از متون ظاهر شود احتمالاً کلمه‌ای متداول است و ارزش چندانی در ارزیابی متن ندارد (کیم و دن، ۲۰۱۸؛ فریکس و بایز-ایتس^۵، ۱۹۹۲). الگوریتم TF-IDF دو جنبه از یک کلمه را ترکیب می‌کند: اهمیت یک کلمه برای یک سند و قدرت تمایز آن در کل مجموعه. این دو جنبه با نگرش عمومی مناسب بودن کلمات به‌عنوان کلمات کلیدی، به‌خوبی مطابقت دارند. با این حال، مقیاس TF-IDF به صورت خودکار فرض می‌کند که کلمات کاملاً مستقل هستند. استخراج کلمات کلیدی از یک متن ارتباط نزدیکی با رتبه‌بندی کلمات به وسیله میزان ارتباط آن‌ها با متن دارد (وارتنا و همکاران^۶، ۲۰۱۰).

با توجه به روند رو به رشد سریع علم و مقالات چاپ‌شده علمی، بررسی و تجزیه و تحلیل متون علمی منتشرشده برای سازمان‌ها، پژوهشگران و سیاست‌گذاران علمی از اهمیت ویژه‌ای برخوردار است. بنابراین استفاده از فنون متن کاوی و استخراج کلمات کلیدی می‌تواند به‌عنوان راه‌حلی برای استخراج دانش از حجم زیادی از متون علمی به کار رود.

پیشینه پژوهش

شناسایی روندهای نوظهور با استفاده از TF-IDF به محققان کمک می‌کند تا با پیشرفت‌های فعلی و تغییر در زمینه‌های

حجم عظیم مهم و ضروری می‌نماید. بنابراین تجزیه و تحلیل این منابع می‌تواند از اهمیت بسزایی برخوردار باشد. ارائه ابزارهایی که با بررسی متون بتواند تحلیلی روی این موارد انجام دهند منجر به شکل‌گیری حوزه‌ی متن کاوی شده است. این حوزه تمامی فعالیت‌هایی که به‌نوعی به دنبال استخراج دانش از متن هستند را شامل می‌گردد (سلیمانی نژاد و همکاران، ۱۳۹۷). همچنین تشخیص روندهای در حال ظهور یک مسئله و مبحث مهم تجزیه و تحلیل متن است. شناسایی روندهای نوظهور معمولاً به‌عنوان زمینه‌های موضوعی‌ای شناخته می‌شود که در طول زمان علاقه و کاربرد آن‌ها در حال افزایش است (کنتانتیس و همکاران^۱، ۲۰۰۴).

یکی از وظایف بسیار مهم برای شناسایی روندهای نوظهور، پیدا کردن روند در حال ظهور پژوهش در مجموعه‌ای از مقالات علمی است (لی^۲، ۲۰۰۵). بررسی و شناسایی مباحثی که اخیراً در یک حوزه علمی مورد توجه قرار گرفته است، برای پژوهشگران از اهمیت ویژه‌ای برخوردار است و سودمندی فراوانی برای آن‌ها دارد. بررسی دستی کلیه مقالات یک حوزه علمی جهت شناسایی موضوعات نوظهور، کاری وقت‌گیر و تقریباً غیرممکن می‌باشد. در این شرایط، تشخیص خودکار روندهای تحقیقاتی در حال ظهور می‌تواند محققان را در درک سریع وقایع و گرایش یک موضوع علمی کمک کند، از این‌رو استفاده از فنون متن کاوی می‌تواند در جهت شناسایی روندهای نوظهور مقالات استفاده شود (کنتانتیس و همکاران، ۲۰۰۴). با توجه به موارد مطرح شده بررسی متون علمی منتشرشده از اهمیت زیادی برای سازمان‌ها، پژوهشگران و سیاست‌گذاران علمی برخوردار است و با توجه به روند رشد سریع علم و مقالات چاپ‌شده علمی، شاهد حجم عظیمی از این داده‌ها هستیم بنابراین یکی از راه‌حلی‌هایی که می‌تواند برای دستیابی به این مشکل استفاده شود، تحلیل کلمات کلیدی مقالات است.

کلیدواژه‌های «محبوب» معمولاً مهم تلقی می‌شوند و بر اساس تکرار یا با استفاده از اقدامات شبکه مبتنی بر مرکزیت انتخاب می‌شوند، که اثبات شده هر دو کلیدواژه‌های بسیار مشابهی را انتخاب می‌کنند. با وجود این، در این روند انتخاب، مشکلی وجود دارد، کلمات کلیدی ممکن است مکرراً مورد استفاده قرار بگیرند، زیرا به موضوع کلیت می‌بخشند یا مضامین محبوب را نشان می‌دهند (سو و لی، ۲۰۱۰). کلمات کلیدی عمومی ممکن است در نشان دادن یک نمای کلی از یک رشته علمی مفید باشند، اما در نشان دادن موضوعات دقیق

3. Chen

4. Term Frequency–Inverse Document Frequency

5. Frakes & Baeza-Yates

6. Wartena et al

2. Kontostathis et al

2. Le

واژگان در دوره زمانی ۱۳۸۵ تا ۱۳۸۹ و واژگان دانشگاه، کتابخانه دانشگاهی و کتابخانه عمومی مهم‌ترین واژگان براساس بیشترین تکرار در مقالات کتابداری و اطلاع‌رسانی منتشر شده در ایران بوده است.

امامی و همکاران (۱۳۹۹)، نیز در بخشی از مطالعه خود مهم‌ترین واژگان حوزه تجهیزات پزشکی و آزمایشگاهی را براساس بیشترین تکرار نشان دادند، نتایج این پژوهش نشان داد واژگان وسایل پزشکی، پزشکی و آزمایشگاه پرتکرارترین واژگان بکار رفته در مقالات این حوزه موضوع است.

چانگ و همکاران^{۱۰} (۲۰۱۵)، به شناسایی تکامل پژوهش‌های کتابداری و اطلاع‌رسانی براساس تحلیل کلمات کلیدی، کتابشناختی و استنادات پرداختند که در این پژوهش کلمات مهم و پرکاربرد براساس بیشترین تکرار آن به دست آمده است. پژوهشگران مهم‌ترین واژگان براساس بیشترین تکرار را در چهار بازه زمانی پنج‌ساله بین سال‌های ۱۹۹۵ تا ۲۰۱۴ گزارش داده‌اند که مهم‌ترین واژگان بکار رفته در این بازه زمانی جستجو، شبکه جهان‌گستر وب، تحلیل استنادی و انتشارات پژوهشی بوده است.

بودری^{۱۱} (۲۰۱۵)، نیز تجزیه و تحلیل کلمات کلیدی براساس بیشترین تکرار در مقالات مرتبط با کاربرد وب ۲ در پزشکی بین سال‌های ۲۰۰۲ تا ۲۰۱۲ پرداخته است. تجزیه و تحلیل مهم‌ترین واژگان براساس بیشترین تکرار نشان داد که واژگان بلاک، ویکی، وب ۲ و رسانه‌های اجتماعی پرکاربردترین واژگان در ای مقالات بوده است.

ساهو و بوی^{۱۲} (۲۰۱۸)، در مطالعه خود با عنوان تحلیل روند پژوهش‌های کتابخانه عمومی هند، به تحلیل کتابشناسی پایان‌نامه‌های مرتبط با موضوع کتابخانه‌های عمومی در هند بین سال‌های ۱۹۷۷ تا ۲۰۱۶ پرداخت، در بخشی از این مطالعه به مهم‌ترین واژگان بکار رفته پرداخته است. که نتایج نشان داده است که واژگان کتابخانه، عمومی، مطالعه، توسعه و ناحیه مهم‌ترین واژگان بکار رفته براساس میزان تکرار در پایان‌نامه‌های مذکور بوده است. در این پژوهش تعداد و درجه هر واژگان توسط قانون زیف^{۱۳} به دست آمده بود. سالوم^{۱۴} و همکاران (۲۰۱۸)، نیز در قسمتی از مطالعه خود پرکاربردترین کلمات کلیدی در مقالات یادگیری موبایلی شناسایی کرده است.

مختلف علمی به‌روز شود (ابی و تسموتو^۱، ۲۰۱۲). این رویکرد می‌تواند برای انواع مختلف اسناد اعمال شود و توانایی ردیابی و پیش‌بینی مسیرهای تحقیق آینده را افزایش می‌دهد (ژو^۲ و همکاران، ۲۰۲۰).

در حالی که TF-IDF ابزاری قدرتمند برای تشخیص روندهای نوظهور است، اغلب با روش‌های دیگری مانند LDA و تجزیه و تحلیل همبازی تکمیل می‌شود تا درک جامع‌تری از تکامل موضوع ارائه دهد. این رویکردهای ترکیبی یک چارچوب قوی برای تجزیه و تحلیل چشم‌انداز پویا کتابداری و تحقیقات علوم اطلاعات ارائه می‌دهد.

در مطالعات گذشته، محققان عمدتاً بر شناسایی موضوعات تحقیقاتی (به‌عنوان مثال دسته‌بندی موضوعی پژوهش‌ها) و تفسیر نتایج آن متمرکز شده‌اند (سلیمانی نژاد، سلاجقه و طیبی نیا، ۲۰۱۹؛ لامبا و مدھوسدان^۳، ۲۰۱۲؛ فیگورلا، مارکو و پینتو^۴، ۲۰۱۷؛ کیم و دلن^۵، ۲۰۱۸؛ کیم و همکاران^۶، ۲۰۱۷). پژوهش‌هایی نیز از هم‌رخدادی واژگان و تجزیه و تحلیل شبکه‌ای برای شناسایی روندهای فناوری استفاده کرده‌اند (چوی^۷ و همکاران، ۲۰۱۱؛ لی^۸ و همکاران، ۲۰۱۶؛ وو^۹، ۲۰۱۶). همچنین به روند انتخاب کلمات کلیدی مناسب برای تجزیه و تحلیل‌ها کمتر توجه شده است. در ادامه برخی از پژوهش‌ها نیز در مطالعات خود به تجزیه و تحلیل واژگان مهم و پرکاربرد در پژوهش‌ها پرداخته‌اند ذکر می‌گردد.

مصطفوی و همکاران (۱۳۹۶)، در مطالعه خود مهم‌ترین واژگان به کار رفته در مقالات علم اطلاعات و دانش‌شناسی را در دوره زمانی ۱۹۹۳ تا ۱۹۹۷ و دوره زمانی ۲۰۰۹ تا ۲۰۱۳ نشان داده است، نتایج نشان داد که واژگان کتابخانه، اطلاعات و علم مهم‌ترین واژگان در دوره زمانی ۱۹۹۳ تا ۱۹۹۷ و واژگان اطلاعات وب و پژوهش مهم‌ترین واژگان در دوره زمانی ۲۰۰۹ تا ۲۰۱۳ بوده است.

سهیلی و همکاران (۱۳۹۸)، نیز در قسمتی از پژوهش خود پرتکرارترین کلیدواژه‌های مقالات علم اطلاعات و دانش‌شناسی منتشر شده در ایران در طی دو بازه زمانی پنج‌ساله از سال ۱۳۸۵ تا ۱۳۹۴ را نشان دادند. نتایج آن‌ها نشان داد واژگان دانشگاه، نشریات، پایگاه‌های اطلاعاتی مهم‌ترین

1. Abe & Tsumoto
2. Xu
3. Lamba & Madhusudhan
4. Figuerola et al
5. Kim & Delen
6. Kim et al
7. Choi
8. Li
7. Wu

10. Chang et al
11. Boudry
12. Sahoo, Bhui
13. zipf's
14. Salloum

محور است که جهت یافتن الگوها و روندها در مجموعه‌های داده‌های بزرگ استفاده می‌شود (هستی‌ه و همکاران، ۲۰۱۳). جامعه آماری، چکیده مقالات منتشر شده کتابداری و اطلاع‌رسانی پزشکی، مجلات تخصصی کتابداری و اطلاع‌رسانی پزشکی (مجلات) که مقالات حوزه کتابداری و اطلاع‌رسانی پزشکی را به‌عنوان یکی از محورهای اصلی انتشار می‌پذیرفتند) و همچنین کتابداری و اطلاع‌رسانی داخل کشور است. جهت جمع‌آوری داده‌های پژوهش، با مشاهده مستقیم وبسایت مجلات، مقالات مرتبط با موضوع کتابداری و اطلاع‌رسانی پزشکی انتخاب و عنوان و چکیده انگلیسی آن از سال ۱۳۷۶ دی‌ماه ۱۳۹۸ در قالب Bibtex استخراج گردیده و بعد در نرم‌افزار جاب رف (JabRef) به قالب CSV تبدیل شده است. در مجلاتی هم که قابلیت استخراج مقالات در قالب Bibtex را نداشتند، اطلاعات مقالات به‌طور مستقیم در فایل CSV ذخیره شده است. در پایان تمامی فایل‌های به دست آمده از مقالات مجلات جهت انجام عملیات متن‌کاوی بر روی آن، با هم در یک فایل نهایی در قالب CSV ادغام گردیدند. این فایل، داده‌های نهایی این پژوهش با تعداد ۹۸۲ رکورد عنوان و چکیده لاتین مقالات منتشرشده در مجلات فوق را شامل شده است.

بررسی پیشینه‌های پژوهش نشان می‌دهد که اکثر پژوهش‌ها در کنار سایر اهداف پژوهش خود پرتکرارترین واژگان مطالعات در حوزه‌های مختلف را نشان داده است. همچنین پژوهشی در زمینه شناسایی روندهای نوظهور و تغییرات واژگانی از فنون متن‌کاوی و الگوریتم TF-IDF مشاهده نشده است.

فرضیه‌های پژوهش

مهم‌ترین واژگان استفاده شده در مقالات کتابداری و اطلاع‌رسانی پزشکی منتشر شده در داخل ایران چه هستند؟ مهم‌ترین واژگان استفاده شده در مقالات کتابداری و اطلاع‌رسانی پزشکی منتشر شده در داخل ایران در دوره‌های زمانی مختلف چه هستند؟ واژگان نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی منتشر شده در داخل ایران چه هستند؟

روش انجام پژوهش

روش پژوهش حاضر متن‌کاوی است که با رویکرد اکتشافی انجام شده است. این مطالعه به دلیل استفاده از فنون متن‌کاوی یک مطالعه اکتشافی هست. متن‌کاوی یک روش اکتشافی داده

جدول ۱. لیست مجلات و تعداد مقالات انتخابی از هر مجله

تعداد مقاله	مجله
۷۲	اطلاع‌رسانی نوین پزشکی
۲۲	پژوهش‌نامه کتابداری و اطلاع‌رسانی
۳۲	پژوهش‌نامه پردازش و مدیریت اطلاعات
۷۹	پیاورد سلامت
۶	تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی
۳۸	تحقیقات کتابداری و اطلاع‌رسانی دانشگاهی
۳۳	تصویر سلامت
۳۰	دانش‌شناسی
۵	رهیافت
۶	سیاست علم و فناوری
۲۰	علم‌سنجی شاهد
۷۰	علم‌سنجی کاسپین
۱۶	علوم کتابداری و اطلاع‌رسانی
۲	گنجینه اسناد
۳۸۴	مدیریت اطلاعات سلامت
۱۱۶	مدیریت سلامت
۲۲	مطالعات کتابداری و علم اطلاعات
۲۹	مطالعات ملی کتابداری و سازمان‌دهی اطلاعات

متنوعی را برای کار با متون در اختیار کاربر قرار می‌دهد (رهورک و سچکا، ۲۰۱۰).

یافته‌ها

مهم‌ترین واژگان به کار رفته در مقالات کتابداری و اطلاع‌رسانی پزشکی

جدول شماره ۲، ده واژگان مهم مقالات منتشر شده در کتابداری اطلاع‌رسانی پزشکی منتشر شده در مجلات داخل کشور را نشان می‌دهد. شایان ذکر است که این واژگان قبل از استخراج با استفاده از الگوریتم ریشه‌یابی پورتر، ریشه‌یابی شده‌اند. همچنین در ادامه شکل شماره ۱ ابر واژگان، واژه‌های مهم به کار گرفته شده در مقالات منتشر شده در مجلات داخل کتابداری و اطلاع‌رسانی پزشکی منتشر شده در مجلات داخل کشور را نشان می‌دهد. در ابر واژگان، واژه‌های دارای اندازه بزرگ‌تر دارای اهمیت بیشتری نسبت با سایر موضوعات هستند، همچنین رنگ‌های جداگانه هر یک از واژه‌ها به جهت جداسازی و نمایش بهتر واژگان به کار رفته است. جدول شماره ۲ و شکل شماره ۱ نشان می‌دهد که واژگان کتابخانه با وزن ۱۲/۶۷، مجله با وزن ۱۲/۴۷ و اطلاعات با وزن ۱۲/۲۳ با اهمیت‌ترین واژگان در مقالات کتابداری و اطلاع‌رسانی پزشکی بوده است.

جدول ۲. مهم‌ترین واژگان در مقالات حوزه کتابداری و

اطلاع‌رسانی پزشکی با استفاده از الگوریتم TF-IDF

وزن	واژگان
۱۲/۶۷	کتابخانه
۱۲/۴۷	مجله
۱۲/۲۳	اطلاعات
۹/۹۰	بیمارستان
۹/۷۴	علمی
۹/۵۹	دانشگاه
۹/۴۹	استناد
۹/۳	سلامت
۹/۲۱	دانش
۶/۸۴	مدیریت

عملیات متن کاوی به کار رفته در این پژوهش مانند پژوهش ژانگ و همکاران^۱ ۲۰۱۵ شامل سه مرحله پیش پردازش متون، عملیات متن کاوی و پس‌پردازش بوده است. پیش‌پردازش متن شامل عملیات زیر است: انتخاب داده‌ها، دسته‌بندی، استخراج ویژگی، نرمال‌سازی و حذف کلمات زائد^۲. دومین مرحله متن کاوی در این پژوهش به‌کارگیری الگوریتم ریشه‌یابی، وزن‌دهی واژگان TF-IDF و بصری‌سازی داده است. در خلال مرحله سوم، تفسیر، تجزیه و تحلیل و بصری‌سازی دانش صورت گرفته است.

جهت ریشه‌یابی واژگان قبل از اجرای الگوریتم TF-IDF از الگوریتم ریشه‌یابی پورتر^۳ استفاده شده است. این الگوریتم پسوند کلمات هم‌ریشه را هرس کرده و در نتیجه اشکال مختلف کلمات و واژه‌هایی که از یک ریشه در متن وجود دارد را به یک صورت واحد تبدیل می‌نماید (ماینر^۴ و همکاران، ۲۰۱۲؛ ابوحی^۵ و همکاران، ۲۰۱۷؛ پورتر، ۱۹۸۰). به‌عنوان مثال برای واژه‌های Modeling، Models، Modeled را به واژه Model تغییر شکل می‌دهد.

همچنین جهت تصویرسازی داده‌ها از تصاویر ابر کلمه استفاده شده است، ابر کلمه یکی از رایج‌ترین روش‌های به کار رفته برای نمایش داده‌های متنی به صورت گرافیکی است که برای تجزیه و تحلیل اشکال مختلف داده‌های متنی از جمله مقالات و پاسخ‌های کوتاه یا نظرات داده شده به نظرسنجی و پرسشنامه مفید است (دی پائلو، ویلکینسون^۶، ۲۰۱۴). ابرهای کلمه یک راه منحصربه‌فرد برای خلاصه کردن محتوای اسناد متن ارائه می‌دهند (کیو و همکاران^۷، ۲۰۱۰). ابرهای کلمه توسط کلمات موجود در داده‌های متنی شکل گرفته‌اند. اندازه یک کلمه در ابر متناسب با تعداد دفعاتی یا اهمیتی (وزن کلمه) است، که کلمه در کل مجموعه داده استفاده می‌شود.

جهت اجرای الگوریتم‌های متن کاوی از زبان برنامه‌نویسی Python و کتابخانه‌های مرتبط استفاده شده است، زبان برنامه‌نویسی پایتون نحو^۸ ساده‌ای دارد، کم‌حجم و چندان‌منظوره است. به‌سادگی توسعه می‌یابد و کتابخانه‌های

1. Zhang et al
2. Stop Words
3. Porter
4. Miner
5. Abuhay
6. DePaolo & Wilkinson
7. Cui et al
8. Syntax



شکل ۱. ابر واژگان مهم‌ترین واژگان در مقالات حوزه کتابداری و اطلاع‌رسانی پزشکی با استفاده از الگوریتم TF-IDF

بزرگ‌تر، نشان از اهمیت آن واژه را می‌دهد. به‌عنوان مثال واژه رشته با وزن ۰/۲۶۹۰۴۷۹۰۷، داروساز با وزن ۰/۲۰۴۰۲۷۹۹۶، جهانی با وزن ۰/۱۹۶۷۶۶۳۸، مهم‌ترین در بازه زمانی ۱۳۷۹ تا ۱۳۸۴ بوده است، همچنین واژه مجله با وزن ۰/۱۱۰۸۵۳۵۴ و اطلاعات با وزن ۰/۳۵۵۷۸۶۲۹۳، مهم‌ترین واژگان در بازه زمانی ۱۳۹۴ تا ۱۳۹۹ بوده است.

مهم‌ترین واژگان به کار رفته در مقالات منتشر شده مجلات داخل کشور در طول زمان جدول شماره ۳، مهم‌ترین واژگان بکار رفته در مقالات کتابداری و اطلاع‌رسانی پزشکی را در بازه‌های زمانی به همراه وزن TF-IDF هرکدام از واژگان و شکل ابر واژگان آن‌ها را نشان می‌دهد. در شکل ابر واژگان هر بازه زمانی، پنجاه واژه مهم بکار رفته در مقالات منتشر شده کتابداری و اطلاع‌رسانی پزشکی نیز نشان داده شده است، که واژه‌های دارای اندازه

جدول ۳. مهم‌ترین واژگان به کار رفته در مقالات کتابداری و اطلاع‌رسانی پزشکی با استفاده از الگوریتم TF-IDF

ابَر واژگان	کلیدواژه‌ها	بازه زمانی
	رشته (۰/۲۶۹۰۴۷۹۰۷)	۱۳۷۹ تا ۱۳۸۴
	داروساز (۰/۲۰۴۰۲۷۹۹۶)	
	جهانی (۰/۱۹۶۷۶۶۳۸)	
	کتاب (۰/۱۹۶۵۱۵۱۸۴)	
	موضوع (۰/۱۸۱۵۶۴۵۵)	
	اطلاعات (۰/۱۶۳۸۹۷۰۹۷)	
	استناد (۰/۱۵۱۱۸۶۵۸)	
	پایان‌نامه (۰/۱۴۶۱۷۵۰۰۵)	
	دانشجو (۰/۱۳۹۶۶۴۳۳)	
	ملیت (۰/۱۳۹۰۶۴۸۴۷)	
	مجله (۱/۷۶۰۵۴۴۵۳۵)	۱۳۸۹ تا ۱۳۸۴
	کتابخانه (۱/۴۱۷۲۶۰۷۴۹)	
	استناد (۱/۲۳۰۴۰۸۹۲۳)	
	بیمارستان (۱/۱۵۷۱۰۴۱۵۱)	
	کتابداری (۱/۱۲۸۲۹۴۶۳۴)	
	سلامت (۱/۰۵۸۶۰۹۶۳۱)	
	جهانی (۱/۰۵۸۲۱۶۴۴)	
	مدیریت (۱/۰۰۱۶۶۲۷۱۶)	
	وب‌سایت (۰/۸۷۹۰۴۹۷۸)	
	سیستم (۰/۸۷۳۶۱۸۱۷۵)	

ادامه جدول ۳. مهم‌ترین واژگان به کار رفته در مقالات کتابداری و اطلاع‌رسانی پزشکی

با استفاده از الگوریتم TF-IDF

بازه زمانی	کلیدواژه‌ها	ابر واژگان
۱۳۸۹ تا ۱۳۹۴	کتابخانه (۵/۴۶۱۴۷۵۵۹۵)	
	دانش (۴/۵۲۶۴۸۱۸۶۴)	
	مجله (۴/۴۷۱۷۴۶۵۵۵۶)	
	بیمارستان (۴/۲۹۹۰۴۶۷۷۶)	
	علمی (۳/۷۰۸۱۸۶۱۴۳)	
	اطلاعات (۳/۶۷۷۰۸۳۶۵۶)	
	مدیریت (۳/۶۰۹۱۰۵۹۴۴)	
	استناد (۳/۵۳۹۳۲۹۷۱۶)	
	سیستم (۳/۳۱۳۷۸۷۳۰۷)	
جهانی (۳/۳۰۷۰۸۴۶۱۵)		
۱۳۹۴ تا ۱۳۹۹	مجله (۶/۱۱۰۸۵۳۵۴)	
	کتابخانه (۵/۴۲۷۷۵۳۱۵۶)	
	اطلاعات (۵/۳۵۵۷۸۶۲۹۳)	
	علمی (۵/۰۳۲۵۱۴۸۵۹)	
	جهانی (۵/۰۲۲۵۲۳۵۸)	
	سلامت (۴/۸۲۲۵۷۴۵۹۸)	
	مقاله (۴/۶۹۶۵۷۰۱۹۷)	
	استناد (۴/۵۹۴۴۹۹۳۳۵)	
	بیمارستان (۴/۲۸۸۳۰۲۱۹۱)	
وبسایت (۴/۱۵۹۴۱۲۰۴۵)		

وزن ۱/۱۰۰۵۶۵ و بابلیوترایی با وزن ۰/۹۴۶۴۵۶ مهم‌ترین واژگان نوظهور در مقالات منتشر شده کتابداری اطلاع‌رسانی پزشکی مجلات بین‌المللی در بازه زمانی ۲۰۱۰ تا ۲۰۱۵ بوده است.

تغییر واژگانی و واژگان نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی
جدول شماره ۴، ۱۵ واژه نوظهور در مقالات منتشر شده کتابداری اطلاع‌رسانی پزشکی مجلات بین‌المللی در بازه زمانی ۱۳۸۹ تا ۱۳۹۴ را نشان می‌دهد. داده‌های این جدول حاکی از آن است که واژگان لیب کوال با وزن ۱/۱۸۱۹۳۲، عملی با

جدول ۴. واژگان نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی

در بازه زمانی ۱۳۸۹ تا ۱۳۹۴

وزن TF-IDF		واژه
۱۳۹۴ تا ۱۳۹۹	۱۳۸۹ تا ۱۳۹۴	
۰/۲۸۱۳۶۲	۱/۱۸۱۹۳۲	لیب کوال
۱/۰۸۱۱۰۲	۱/۱۰۰۵۶۵	عملی
۰/۴۴۳۴۶۶	۰/۹۴۶۴۵۶	بابلیوترایی
۱/۰۳۰۰۵۷	۰/۸۹۸۹۳۷	مجازی
۰	۰/۷۸۹۰۳۶	آر. اف. ای. دی
۱/۰۶۴۸۴۱	۰/۷۳۳۳۲۵	تلفن هوشمند

ادامه جدول ۴. واژگان نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی

در بازه زمانی ۱۳۸۹ تا ۱۳۹۴

وزن TF-IDF		واژه
۱۳۹۹ تا ۱۳۹۴	۱۳۹۴ تا ۱۳۸۹	
۰/۲۴۳۹۳۹	۰/۶۹۵۶۸۴	اخلاق
۰/۶۰۳۰۴۳	۰/۵۵۶۲۷۱	بازاریابی
۰/۲۲۰۰۵۹	۰/۵۰۲۱۸۸	تلفن
۰/۰۳۸۸۳۷	۰/۴۸۰۲۹۱	بلاک
۰/۹۷۶۳۸۹	۰/۴۱۲۳۴۶	پزشکی از راه دور
۰/۰۳۰۰۳	۰/۴۰۰۹۳۷	هوشمند
۰/۱۵۸۵۰۲	۰/۳۸۱۴۶۹	سرقت ادبی
۰/۱۷۴۸۳۵	۰/۲۸۸۹۱۴	بازخورد
۰	۰/۲۷۸۵۴۶	سروکوال

مخزن با وزن ۱/۲۳۴۸۲۷ مهم‌ترین واژگان نوظهور در مقالات منتشر شده کتابداری اطلاع‌رسانی پزشکی در بازه زمانی ۲۰۱۵ تا ۲۰۲۰ بوده است.

جدول شماره ۵، ۱۵ واژه نوظهور در مقالات کتابداری اطلاع‌رسانی پزشکی در بازه زمانی ۱۳۹۴ تا ۱۳۹۹ را نشان می‌دهد. داده‌های این جدول حاکی از آن است که واژگان اختراع با وزن ۲/۱۱۹۷۰۷، آلت‌متریک با وزن ۱/۶۷۳۷۷۹ و واژه

جدول ۵. واژگان نوظهور در مقالات کتابداری و اطلاع‌رسانی پزشکی

در بازه زمانی ۱۳۹۴ تا ۱۳۹۹

وزن TF-IDF		واژه
۱۳۹۹ تا ۱۳۹۴	۱۳۹۴ تا ۱۳۸۹	
۲/۱۱۹۷۰۷	۰	اختراع
۱/۶۷۳۷۷۹	۰	آلت‌متریک
۱/۲۳۴۸۲۷	۰	مخزن
۰/۸۹۶۳۹۸	۰	ابر
۰/۷۰۷۳۰۹	۰	قابل خواندن
۰/۵۲۸۰۹۱	۰	یادگیرنده
۰/۴۷۹۵۵۱	۰	اینفودمیولوژی
۰/۳۳۶۴۳۸	۰	ریسرچ گیت
۰/۳۳۵۳۸۲	۰	توییت‌ر
۰/۳۳۴۲۲۲	۰	پریسما
۰/۳۳۳۱۰۴	۰	سلامت الکترونیک
۰/۳۲۳۱۶	۰	سویچ
۰/۳۰۶۲۴۸	۰	کلاس‌بندی
۰/۳۰۳۲۴۶	۰	کوهورت
۰/۲۸۹۲۱۳۰۷۷	۰	مندلی

استاد، سلامت، دانش و مدیریت، ده واژگان مهم و دارای بالاترین وزن TF-IDF هستند که بررسی این واژگان نمای کلی از مباحث اصلی مقالات منتشرشده کتابداری و

بحث و نتیجه‌گیری

نتایج حاصل از بررسی مهم‌ترین واژه‌ها نشان می‌دهد که واژگان کتابخانه، مجله، اطلاعات، بیمارستان، علمی، دانشگاه،

علم اطلاعات بوده است و بعد از ظهور وب نیز واژگان اطلاعات، وب و پژوهش مهم‌ترین واژگان به کار رفته در مطالعات علم اطلاعات بوده است. در پژوهش سهیلی، خاصه و کرانیان ۱۳۹۸ نیز مشخص شد که واژه نشریات از واژه‌های پرکاربرد در مقالات علم اطلاعات و دانش‌شناسی منتشر شده در ایران در بازه زمانی ۱۳۸۵ تا ۱۳۸۹ بوده است که این واژه در بازه زمانی ۱۳۹۰ تا ۱۳۹۴ بسیار کم‌رنگ شده است و در عوض واژگان وب‌سایت و فناوری اطاعات در دوره زمانی ۱۳۹۰ تا ۱۳۹۴ در مقالات علم اطلاعات و دانش‌شناسی منتشر شده در ایران حضور بیشتری دارد، که این نشان از توسعه فناوری اطلاعات و تحولات ایجاد شده فناوری در کتابداری و اطلاع‌رسانی بوده است.

نتایج تغییرات واژگان بکار رفته در طول زمان روند به‌کارگیری را نشان داده است. روند واژگان بکار رفته در مقالات کتابداری و اطلاع‌رسانی پزشکی نشان می‌دهد که مقالات کتابداری و اطلاع‌رسانی پزشکی با وجود اینکه قالب و موضوع اصلی این حوزه علمی را حفظ کرده است اما بررسی واژگان نوظهور این حوزه علمی نشان‌دهنده حرکت به سمت فناوری‌های اطلاعات است.

نتایج پژوهش حاضر در خصوص واژگان نوظهور در مقالات منتشر شده در مجلات کتابداری و اطلاع‌رسانی پزشکی نشان می‌دهد که واژگان لیب‌کوال، عملی، بایلیوتراپی، مجازی، آر.اف. ای. دی، تلفن هوشمند، اخلاق، بازاریابی، تلفن و بلاک در بازه زمانی ۱۳۸۹ تا ۱۳۹۴ برای اولین بار وارد مقالات و مطالعات حوزه کتابداری و اطلاع‌رسانی پزشکی در مجلات داخلی شده است. همچنین واژگان اختراع، آلت‌متریک، مخزن، ابر، قابل‌خواندن، یادگیرنده، اینفودمیولوژی، ریسرچ گیت، توئیتر، پریسما در بازه زمانی ۱۳۹۴ تا ۱۳۹۹ به‌تازگی وارد مقالات و مطالعات حوزه کتابداری و اطلاع‌رسانی پزشکی در مجلات داخلی شده است. که این نشان‌دهنده حرکت به سمت فناوری‌های اطلاعاتی و ارتباطی و مباحث جدید فناوری اطلاعات و دیجیتال است. در بین واژگان نوظهور در سال‌های اخیر واژگان مرتبط با فناوری وب ۲ و شبکه‌های اجتماعی نیز مشاهده می‌شود. در این خصوص واژگان بلاک از سال ۱۳۸۹ به بعد، و واژگان توئیتر و ریسرچ گیت از سال ۱۳۹۴ به بعد از در متون مقالات کتابداری و اطلاع‌رسانی پزشکی مشاهده شده است، بر همین اساس بوداری ۲۰۱۵ نیز نشان داده است که مقالات مرتبط با وب ۲ از سال ۲۰۰۲ تا ۲۰۱۲ در حوزه علوم پزشکی منتشر شده‌اند، واژه وبلاگ نیز پرتکرارترین واژه در بین مقالات منتشر شده وب ۲ در علوم پزشکی بوده است. طور

اطلاع‌رسانی پزشکی را نشان می‌دهد. فانک^۱ ۲۰۱۳ نیز نشان داده است که واژه‌های کتابخانه، اطلاعات، پزشکی، سلامت، کتابداری، کتابداران، خدمات، پژوهش، بیمارستان و زمان، پرتکرارترین واژگان بکار گرفته شده در مقالات منتشر شده مجلات بولتن انجمن کتابخانه‌های پزشکی ۲ و مجله انجمن کتابخانه‌های پزشکی ۳ بین سال‌های ۱۹۶۰ تا ۲۰۱۰ بوده‌اند. همچنین وی بیان می‌کند که واژه‌های این مقالات در دسته‌های، محیط، مدیریت، فناوری و پژوهش قرار می‌گیرند.

ساهو و بوی ۲۰۱۸ نیز نشان داده است که واژگان کتابخانه، عمومی، مطالعه، توسعه، مناطق و خدمات دارای بیشترین تکرار در پژوهش‌های کتابخانه‌های عمومی هند هستند. سهیلی، خاصه و کرانیان ۱۳۹۸ نیز نشان دادند واژه دانشگاه، ارزیابی، کتابداری و اطلاع‌رسانی و تحلیل اطلاعات پرتکرارترین واژه در مقالات علم اطلاعات و دانش‌شناسی منتشر شده در ایران در دوره زمانی ۱۳۸۴ تا ۱۳۹۴ هست، همچنین بررسی مهم‌ترین واژگان بکار رفته در مقالات منتشر شده کتابداری و اطلاع‌رسانی پزشکی مجلات داخلی در دوره‌های زمانی مختلف نشان از تغییراتی در به‌کارگیری واژه‌ها در دوره‌های زمانی مختلف می‌دهد و به‌کارگیری واژگان در مقالات در دوره‌های زمانی مختلف روند یکسانی را نداشته است. به‌عنوان نمونه در مقالات منتشر شده در بازه زمانی ۱۳۷۹ تا ۱۳۸۴ واژگان رشته، داروساز و جهانی مهم‌ترین واژگان بکار رفته در مقالات منتشر شده بوده‌اند. در بازه زمانی ۱۳۹۴ تا ۱۳۹۹ نیز واژه‌های مجله، کتابخانه و اطلاعات مهم‌ترین واژگان در مقالات منتشر شده بوده‌اند. این موضوع نشان می‌دهد، با گذشت زمان، واژگان بکار رفته در مقالات کتابداری و اطلاع‌رسانی تغییر کرده است. چانگ و همکاران ۲۰۱۵ نیز نشان داده‌اند که واژگان جستجو و بازاریابی اطلاعات آنلاین پرتکرارترین واژگان در بازه زمانی ۱۹۹۵ تا ۱۹۹۹، واژگان شبکه جهانی وب، بازاریابی اطلاعات آنلاین پرتکرارترین واژگان در بازه زمانی ۲۰۰۰ تا ۲۰۰۴، تحلیل استنادی، پرتکرارترین واژه در بازه زمانی ۲۰۰۵ تا ۲۰۰۹ و واژه انتشارات علمی نیز پرتکرارترین واژگان در بازه زمانی ۲۰۱۰ تا ۲۰۱۴ در پژوهش‌های کتابداری و اطلاع‌رسانی بوده است (چانگ همکاران، ۲۰۱۵). مصطفوی، عصاره و توکلی زاده ۱۳۹۶ نیز نشان دادند قبل از ظهور وب واژگان کتابخانه، اطلاعات و علم مهم‌ترین واژگان به کار رفته در مطالعات حوزه

1. Funk
2. Bulletin of the Medical Library Association
3. Journal of the Medical Library Association

واژگانی و موضوعی نشان داده است، تغییرات واژگانی و ظهور واژگان و موضوعات جدید در این حوزه علمی محسوس بوده است. روند تغییرات استفاده از واژگان مختلف در مقالات کتابداری و اطلاع‌رسانی پزشکی و همچنین شناسایی واژگان نوظهور این مقالات نشان داده است که همگام با ظهور و رشد فناوری‌های اطلاعاتی و ارتباطی، این رشته علمی نیز متناسب با نیازهای روز جامعه، دچار تغییر شده است.

پیشنهاد‌های پژوهش

- ✓ با استفاده از فنون متن کاوی و تجزیه و تحلیل واژگانی بکار رفته در این پژوهش، پژوهشگران می‌توانند روند رشد و تغییر حوزه‌های پژوهشی سایر حوزه‌ها و رشته‌های علمی را شناسایی نمایند، که این امر می‌تواند در سیاست‌های پژوهشی حوزه‌های مختلف علمی مفید واقع شود.
- ✓ فنون متن کاوی و تجزیه و تحلیل واژگانی به کار رفته در این پژوهش، می‌تواند در کنار سایر مطالعات علم‌سنجی جهت تقویت مطالعات و کسب نتایج بهتر استفاده شود.

کلی می‌توان گفت که ظهور فناوری‌های جدید اطلاعاتی، روند ذخیره، بازیابی، مدیریت و اشاعه اطلاعات سلامت را دگرگون ساخته و به دنبال آن، وظایف و خدمات سنتی کتابداران و اطلاع‌رسانان پزشکی را نیز متحول کرده است. به عبارت دیگر، امروزه حوزه فعالیت کتابداران و اطلاع‌رسانان پزشکی فراتر از کتابخانه‌ها می‌باشد و با توجه به توانایی‌ها و قابلیت‌هایی که دارند، قادر هستند تا در حیطه‌های مختلف سلامت مانند بیمارستان‌ها، کتابخانه‌ها، مراکز تحقیقاتی، مراکز انتشاراتی، مراکز علم‌سنجی، مجلات تخصصی، مراکز رشد و مراکز دانش‌بنیان به ارائه خدمت بپردازند (اشرفی‌ریزی و همکاران، ۲۰۱۷).

به‌طور کلی می‌توان گفت با گذر زمان موضوعات منتشر شده تمامی حوزه‌های علمی در حال تکامل است، موضوعات جدید پدیدار می‌شوند و موضوعات قدیمی نیز منسوخ می‌شوند. بر همین اساس نتایج این پژوهش نیز نشان داده است که حوزه علمی کتابداری و اطلاع‌رسانی پزشکی نیز از این امر مستثنی نبوده و همان‌طور که نتایج پژوهش حاضر در خصوص تغییرات

References

- Abe, H., & Tsumoto, S. (2012). Detection of research trends from bibliographical data. *International Journal of Data Mining, Modelling and Management*, 4(3), 255-266. Doi: [10.1504/IJDMMM.2012.048107](https://doi.org/10.1504/IJDMMM.2012.048107)
- Abuhay, T. M., Kovalchuk, S. V., Bochenina, K. O., Kampis, G., Krzhizhanovskaya, V. V., & Lees, M. H. (2017). Analysis of computational science papers from ICCS 2001-2016 using topic modeling and graph theory. *arXiv preprint arXiv:1705.02203*. Doi: [10.1016/j.procs.2017.05.183](https://doi.org/10.1016/j.procs.2017.05.183)
- Ashrafi-Rizi, H., Hodhodinezhad, N., Shahrzadi, L., & Soleymani M.R. (2017). A Study on the Novel Services of Medical Librarians in Health Information. *Journal of Health Information Management*, 13(3), 438-444. (In Persian)
- Behkami, N. A., & U. Daim, T. (2012). Research Forecasting for Health Information Technology (HIT), using technology intelligence. *Technological Forecasting and Social Change*, 79(3), 498-508. Doi: [10.1016/j.techfore.2011.08.015](https://doi.org/10.1016/j.techfore.2011.08.015)
- Boudry, C. (2015). Web 2.0 applications in medicine: trends and topics in the literature. *Medicine 2.0*, 4(1).
- Chang, Y.-W., Huang, M.-H., & Lin, C.-W. (2015). Evolution of research subjects in library and information science based on keyword, bibliographical coupling, and co-citation analyses. *Scientometrics*, 105(3), 2071-2087.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359-377. Doi: [10.1002/asi.20317](https://doi.org/10.1002/asi.20317)
- Choi, Sungchul, Janghyeok Yoon, Kwangsoo Kim, Jae Yeol Lee, and Cheol-Han Kim. "Sao Network Analysis of Patents for Technology Trends Identification: A Case Study of Polymer Electrolyte Membrane Technology in Proton Exchange Membrane Fuel Cells." *Scientometrics* 88, no. 3 (2011/06/12 2011): 863.
- Cui, W., Wu, Y., Liu, S., Wei, F., Zhou, M. X., & Qu, H. (2010). Context preserving dynamic word cloud visualization. *Paper presented at the 2010 IEEE Pacific Visualization Symposium (PacificVis)*.
- DePaolo, C. A., & Wilkinson, K. (2014). Get your head into the clouds: Using word clouds for analyzing qualitative assessment data. *TechTrends*, 58(3), 38-44.

- Emami, M., Riahinia, N., & Soheili, F. (2020). Mapping the Scientific Structure of Medical and Laboratory Equipment with Using the Co-occurrence Analysis, *Scientometric Research Journal*, 6(11), 41-56. (In Persian) Doi: [10.22070/rsci.2018.696](https://doi.org/10.22070/rsci.2018.696)
- Figuerola, C. G., García Marco, F. J., & Pinto, M. (2017). Mapping the evolution of library and information science (1978–2014) using topic modeling on LISA. *Scientometrics*, 112(3), 1507-1535. Doi: [10.1007/s11192-017-2432-9](https://doi.org/10.1007/s11192-017-2432-9)
- Frakes, W. B., & Baeza-Yates, R. (1992). Information retrieval: data structures and algorithms. *Prentice-Hall, Inc.*
- Funk, M. E. (2013). Our words, our story: a textual analysis of articles published in the Bulletin of the Medical Library Association/Journal of the Medical Library Association from 1961 to 2010. *Journal of the Medical Library Association: JMLA*, 101(1), 12. Doi: [10.3163/1536-5050.101.1.003](https://doi.org/10.3163/1536-5050.101.1.003).
- Gavgani, V. Z., & Mohan, V. V. (2008). Application of web 2.0 tools in medical librarianship to support medicine 2.0. *Webology*, 5(1).
- Hashimi, H., Hafez, A., & Mathkour, H. (2015). Selection criteria for text mining approaches. *Computers in Human Behavior*, 51, 729-733. <https://doi.org/10.1016/j.chb.2014.10.062>
- Hastie, T., Tibshirani, R., & Friedman, J. (2013). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media.
- Hung, J.-L., & Zhang, K. J. J. o. C. i. H. e. (2012). Examining mobile learning trends 2003–2008: A categorical meta-trend analysis using text mining techniques. 24(1), 1-17. Doi: [10.1007/s12528-011-9044-9](https://doi.org/10.1007/s12528-011-9044-9)
- Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics*, 13(6), 395.
- Kao, A., & Poteet, S. R. (2007). *Natural Language Processing and Text Mining*. Springer Science & Business Media.
- Kim, M. J., Ohk, K., & Moon, C. S. (2017). Trend Analysis by Using Text Mining of Journal Articles Regarding Consumer Policy. *New Physics: Sae Mulli*, 67(5), 555-561. Doi: [10.3938/NPSM.67.555](https://doi.org/10.3938/NPSM.67.555)
- Kim, Y.-M., & Delen, D. (2018). Medical informatics research trend analysis: A text mining approach. *Health informatics journal*, 24(4), 432-452.
- Lamba, M., & Madhusudhan, M. (2019). Mapping of topics in DESIDOC Journal of Library and Information Technology, India: a study. *Scientometrics*
- Lamba, M., & Madhusudhan, M. (2019). Mapping of topics in DESIDOC Journal of Library and Information Technology, India: a study. *Scientometrics*, 120(2), 477-505. Doi: [10.1007/s11192-019-03137-5](https://doi.org/10.1007/s11192-019-03137-5)
- Le, M. H., Ho, T. B., & Nakamori, Y. (2005). Detecting emerging trends from scientific corpora. *International Journal of Knowledge and Systems Sciences*, 2(2), 53-59.
- Lee, S., Lee, S., Seol, H., Park, Y. J. R., & Management, d. (2008). Using patent information for designing new product and technology: *keywordbased technology roadmapping*. 38(2), 169-188. <https://doi.org/10.1111/j.1467-9310.2008.00509>.
- Li, H., An, H., Wang, Y., Huang, J., & Gao, X. (2016). Evolutionary features of academic articles co-keyword network and keywords co-occurrence network: Based on two-mode affiliation network. *Physica A: Statistical Mechanics and its Applications*, 450, 657-669. <https://doi.org/10.1016/j.physa.2016.01.017>
- Miner, G. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press.
- Mostafavi, I., Osareh, F., & Tavakkolizadeh Ravari, M. (2017). Analysis on the Structure of KIS Scientific Words and Concepts Based on Social Network Analysis in the Web of Science Database (1993-1997 and 2009-2013), *Research on Information Scienc & Public Libraries*, 23(89), 237-264. (In Persian) Dor: [20.1001.1.26455730.1396.23.2.6.6](https://doi.org/20.1001.1.26455730.1396.23.2.6.6)
- Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130-137. <https://doi.org/10.1108/eb046814>
- Rajman, M., & Besançon, R. (1998). Text mining-knowledge extraction from unstructured textual data. In *Advances in data science and classification* (pp. 473-480): Springer.

- Rebholz-Schuhmann, D., Oellrich, A., & Hoehndorf, R. (2012). Text-mining solutions for biomedical research: enabling integrative biology. *Nature Reviews Genetics*, 13(12), 829.
- Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. Paper presented at the In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks.
- Rodriguez-Esteban, R., & Bundschuh, M. (2016). Text mining patents for biomedical knowledge. *Drug discovery today*, 21(6), 997-1002. <https://doi.org/10.1016/j.drudis.2016.05.002>
- Sahoo, S., & Bhui, T. (2018). Trend of Public library research in India: a bibliometric study. *Library Philosophy & Practice*.
- Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2018). Using text mining techniques for extracting information from research articles. In *Intelligent Natural Language Processing: Trends and Applications* (pp. 373-397): Springer.
- Soheili, F., Khasseh, A.A., & Koranian, P. (2019). Mapping Intellectual Structure of Knowledge and Information Science in Iran based on Co-word Analysis. *Journal of Information Processing and Management*, 34(4), 1905-1938. (In Persian) Doi: [10.35050/JIPM010.2019.026](https://doi.org/10.35050/JIPM010.2019.026)
- Soleimani Nezhad, A., Salajegheh, M., & Tayyebi Nia, E. (2019). Clustering scientific articles based on the k_means algorithm Case Study: Iranian Research Institute for information Science and Technology (IranDoc), *Journal of Information Processing and Management*, 34(2), 871-896. (IN PERSIAN) Doi: [10.35050/JIPM010.2019.060](https://doi.org/10.35050/JIPM010.2019.060)
- Su, H.-N., & Lee, P.-C. (2010). Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in Technology Foresight. *Scientometrics*, 85(1), 65-79. Doi: [10.1007/s11192-010-0259-8](https://doi.org/10.1007/s11192-010-0259-8)
- Wang, M. Y., Fang, S. C., & Chang, Y. H. (2015). Exploring technological opportunities by mining the gaps between science and technology: Microalgal biofuels. *Technological Forecasting and Social Change*, 92, 182-195. Doi: [10.1016/j.techfore.2014.07.008](https://doi.org/10.1016/j.techfore.2014.07.008)
- Wartena, C., Brussee, R., & Slakhorst, W. (2010, August). Keyword extraction using word co-occurrence. In *2010 workshops on database and expert systems applications* (pp. 54-58). IEEE.
- Wu, C.-C. (2016). Constructing a weighted keyword-based patent network approach to identify technological trends and evolution in a field of green energy: a case of biofuels. *Quality & Quantity*, 50(1), 213-235. Doi: [10.1007/s11135-014-0145-1](https://doi.org/10.1007/s11135-014-0145-1)
- Xu, M., Li, G., & Wang, X. (2020). Detecting emerging topics by exploiting probability burst and association rule mining: A case study of Library and Information Science. *Malaysian Journal of Library and Information Science*, 25(1), 47-66. Doi: [10.22452/mjlis.vol25no1.3](https://doi.org/10.22452/mjlis.vol25no1.3)
- Zhang, Y., Chen, M., & Liu, L. (2015, September). A review on text mining. In *2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS)* (pp. 681-685). IEEE.