

## تحلیل ترکیبی کتاب‌سنجی و متن‌کاوی تولیدات علمی حوزه پرونده الکترونیک سلامت در پایگاه PubMed

محبوبه شکوهیان<sup>۱</sup>، عاصفه عاصمی<sup>۱</sup>، احمد شعبانی<sup>۲</sup>، مظفر چشمه سهرابی<sup>۳</sup>

## مقاله پژوهشی

## چکیده

**مقدمه:** دسترسی به اطلاعات کامل بیمار، نقش مهمی در بهبود مراقبت‌های بالینی و کاهش اشتباهات پزشکی دارد. در این خصوص، پرونده الکترونیک سلامت، قسمت اصلی یک سیستم اطلاعات سلامت یکپارچه محسوب می‌شود. هدف از انجام پژوهش حاضر، تحلیل کتاب‌سنجی و متن‌کاوی تولیدات علمی منتشر شده در حوزه پرونده الکترونیک سلامت در پایگاه PubMed بود.

**روش بررسی:** این مطالعه به روش کتاب‌سنجی و متن‌کاوی در بازه زمانی سال‌های ۲۰۰۹ تا ۲۰۱۹ بر روی ۶۸۶۳ مقاله انجام شد. داده‌ها با استفاده از نرم‌افزارهای Excel و VOSviewer و ابزار Voyant مورد تجزیه و تحلیل قرار گرفت.

**یافته‌ها:** در حوزه مورد نظر، موضوعات پرونده الکترونیک سلامت، سلامت، مراقبت بهداشتی و سیستم‌های مراقبت بهداشتی اهمیت زیادی در پایگاه PubMed داشت. تولید مقالات در حوزه پرونده الکترونیک سلامت طی ده سال روندی صعودی را نشان داد و کشور آمریکا پر تولیدترین کشور در این حوزه بود. بیشترین مقالات به David Bates، Dean Sittig و Hardeep Singh اختصاص داشت.

**نتیجه‌گیری:** نقشه هم‌رخدادی واژگان برای هر کدام از واژه‌ها، نماینده یک مفهوم یا حوزه تحقیقاتی در سلامت می‌باشد. نتایج به دست آمده می‌تواند دید روشنی به منظور سیاست‌گذاری علمی این حوزه برای تأثیرگذاری بر تخصیص و توزیع منابع در فعالیت‌های علمی و فنی ارایه نماید. همچنین، می‌تواند به محققان در انتخاب موضوعات داغ و کسب بینش جامعی از چارچوب علمی حوزه مورد نظر کمک نماید.

**واژه‌های کلیدی:** PubMed؛ تولیدات علمی؛ پرونده‌های الکترونیک سلامت؛ کتاب‌سنجی؛ متن‌کاوی

**پیام کلیدی:** با توجه به اهمیت تولیدات علمی در حوزه سلامت به عنوان محور توسعه پایدار، توجه همه‌جانبه به روند رشد و موضوعات مطرح در آن ضروری به نظر می‌رسد. این امر موجب سیاست‌گذاری پژوهشی برای تولید علم به منظور ارتقا و حفظ سلامت افراد می‌شود.

تاریخ انتشار: ۱۳۹۸/۷/۱۵

پذیرش مقاله: ۱۳۹۸/۷/۲

دریافت مقاله: ۱۳۹۸/۶/۱

**ارجاع:** شکوهیان محبوبه، عاصمی عاصفه، شعبانی احمد، چشمه سهرابی مظفر. تحلیل ترکیبی کتاب‌سنجی و متن‌کاوی تولیدات علمی حوزه پرونده الکترونیک سلامت در پایگاه PubMed. مدیریت اطلاعات سلامت ۱۳۹۸؛ ۱۶ (۴): ۱۹۶-۱۹۰

علمی را با نشان دادن روابط بین موارد مهم در حوزه و چگونگی رشد و توسعه این حوزه‌ها در طی زمان نمایش می‌دهند. محققان می‌توانند از نقشه‌های در روش تحلیل هم‌واژگانی کلمات کلیدی، وجود واژه‌ها یا مفاهیم در کنار هم در

## مقدمه

فن‌آوری اطلاعات سلامت می‌تواند عامل بالقوه‌ای جهت افزایش کیفیت و اثربخشی مراقبت‌های بهداشتی باشد (۱). فن‌آوری‌های جدید مانند پرونده‌های الکترونیک سلامت، ابزارهایی را جهت سهولت به اشتراک‌گذاری اطلاعات ارزشمند سلامت و افزایش کیفیت مراقبت و بهبود سلامت جامعه در اختیار ارایه دهندگان مراقبت بهداشتی قرار می‌دهد. دسترسی به چنین اطلاعاتی، موجب تصمیم‌گیری بالینی بهتر بر اساس دیدگاه یکپارچه از داده‌های سلامت فرد، افزایش دسترسی به اطلاعات سلامت فرد و بهبود مراقبت بهداشتی یکپارچه برای بیماران می‌گردد (۲). از نظر Evans، پرونده الکترونیک سلامت، مجموعه اطلاعات مرتبط با سلامت افراد از پیش از تولد تا پس از مرگ است که به طور مداوم به شکل الکترونیکی توسط مراکز مراقبت‌های بهداشتی ذخیره و نگهداری می‌شود تا در صورت نیاز، تمام یا بخشی از آن با رعایت نکات امنیتی در دسترس افراد مجاز قرار گیرد (۳).

امروزه از نقشه‌های علمی به عنوان ابزار قدرتمندی به منظور مطالعه ساختار و پویایی زمینه‌های علمی استفاده می‌شود. نقشه‌ها ساختار یک رشته

مقاله حاصل پایان‌نامه دکتری تخصصی می‌باشد که با حمایت دانشگاه اصفهان انجام شده است.

- ۱- دانشجوی دکتری تخصصی، علم اطلاعات و دانش‌شناسی، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم تربیتی و روان‌شناسی، دانشگاه اصفهان، اصفهان، ایران
- ۲- دانشیار، علم اطلاعات و دانش‌شناسی، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم تربیتی و روان‌شناسی، دانشگاه اصفهان، اصفهان، ایران و دانشکده بی‌زیس انفورماتیک، دانشگاه کروویوس، بوداپست، مجارستان (نویسنده طرف مکاتبه)
- ۳- استاد، علم اطلاعات و دانش‌شناسی، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم تربیتی و روان‌شناسی، دانشگاه اصفهان، اصفهان، ایران
- ۴- دانشیار، علم اطلاعات و دانش‌شناسی، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم تربیتی و روان‌شناسی، دانشگاه اصفهان، اصفهان، ایران

Email: asefi.asefeh@uni-corvinus.hu

حوزه آموزش بهداشت و ارتقای سلامت طی سال‌های ۱۳۷۷ تا ۱۳۹۰ در پایگاه‌های SID Iran medex (Scientific Information Database)، MedLib و PubMed نشان دادند که رشد انتشارات در این حوزه در داخل و خارج روند مثبتی داشته است و موضوعات «سلامت روان، بهداشت باروری و تنظیم خانواده و بیماری‌های مزمن غیر واگیر» از جمله مهم‌ترین موضوعات تولیدات علمی ارتقای سلامت می‌باشند (۱۱).

با رشد گسترده داده‌های سلامت، محققان با انبوهی از اطلاعات مواجه هستند که برای استفاده از آن، به مدیریت اطلاعات سلامت و ابزارهای جدیدی برای کشف و مرور مجموعه‌های بزرگ ادبیات علمی نیاز دارند. مطالعه حاضر فرصت مناسبی را برای ارتقای کیفیت مراقبت از بیمار و بهبود سیستم‌های مراقبت بهداشتی فراهم می‌کند. پایگاه‌های استنادی، اطلاعاتی مانند افراد استنادکننده به تولیدات علمی، ردیابی فعالیت‌های پژوهشی افراد، ردیابی تاریخ و روند پیشرفت یک حوزه و... را به کاربران ارایه می‌دهند، اما پایگاه PubMed این امکان را برای کاربران فراهم نمی‌کند. هدف از انجام تحقیق حاضر، تحلیل کتاب‌سنجی و متن‌کاوی حوزه پرونده الکترونیک سلامت در پایگاه PubMed طی بازه زمانی ده سال اخیر بود تا ساختار مفهومی این حوزه و عناصر موضوعی آن مشخص گردد.

### روش بررسی

این پژوهش بر اساس مدل مفهومی از دو روش متن‌کاوی و کتاب‌سنجی استفاده نمود که موجب بهبود کارایی و غلبه بر برخی از محدودیت‌های این دو روش گردید. Janssens در پایان‌نامه خود بیان کرد که تلفیق الگوریتم و روش‌های متن‌کاوی و کتاب‌سنجی، موجب دسته‌بندی بهتر متون و حوزه‌بندی علمی آن می‌شود (۱۲). جامعه مورد مطالعه شامل مقالات علمی ده سال اخیر محققان حوزه پرونده الکترونیک سلامت نمایه شده در پایگاه PubMed بود. این پایگاه بزرگ‌ترین و معتبرترین پایگاه اطلاعات سلامت می‌باشد. در مرحله اول، برای تعیین موضوعات اصلی با استفاده از راهبرد جستجو پیشرفته زیر، نتایج به بازه زمانی ۲۰۱۹/۷/۲ تا ۲۰۱۹/۶/۲۹ و زبان انگلیسی محدود گردید.

"Electronic health record"[All Fields] OR "Health Care Information System"[All Fields] OR "Personal Health Record"[All Fields] AND ("2009/07/02"[PDat]:"2019/06/29"[PDat])

۶۸۶۳ رکورد از مقالات نمایه شده در پایگاه PubMed به عنوان جامعه تحقیق بازیابی گردید. سپس مقالات به دست آمده بررسی و موضوعات آن بر اساس عنوان، چکیده و کلیدواژه‌ها استخراج و در نرم‌افزار Excel وارد شد. برای به دست آوردن تعداد تکرار کلیدواژه‌ها و موضوعات، پیکره به دست آمده با استفاده از ابزار تحت وب Voyant به صورت فایل Word بارگذاری و تعداد تکرار واژگان کلیدی مشخص شد. سپس نقشه ابر کلمه موضوعات به دست آمد. پژوهشگران می‌توانند متن خود را در قالب‌های مختلف در ابزار Voyant برای تجزیه و تحلیل متن وارد کنند. سپس یک پنجره تجزیه و تحلیل و نمایش گرافیکی متن در ابزار Voyant، واژگان کلیدی و تعداد تکرار کلمات را مشخص می‌نماید (۱۳).

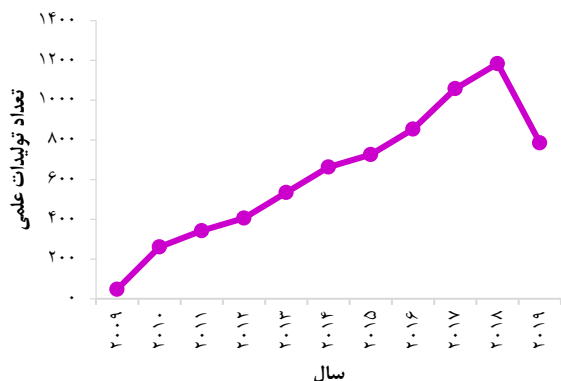
در مرحله بعد، ۶۸۶۳ رکورد بازیابی شده بر اساس معیارهای ورود و خروج و سؤالات تحقیق (تعیین روند رشد تولیدات، پرتولیدترین نویسندگان، پرتولیدترین

یک مدرک، نشان دهنده محتوای نزدیک آن‌ها به یکدیگر است که با کتاب‌شناختی برای به دست آوردن درک بهتر از حوزه‌ای که در آن کار می‌کنند، استفاده نمایند (۴). با اندازه‌گیری میزان این هم‌رخدادی، می‌توان ساختار مفهومی یک حوزه تحقیقاتی را ترسیم نمود. این شبکه مفهومی با شمارش تعداد دفعات وجود هر اصطلاح موضوعی و هم‌نشینی آن با موضوعات دیگر ترسیم می‌شود (۵). یک مقاله علمی می‌تواند شامل موضوعات مختلف و کلمات موجود در آن مقاله نیز می‌تواند منعکس‌کننده مجموعه خاصی از موضوعات مورد بحث باشد. با استفاده از روش متن‌کاوی و تحلیل محتوای اسناد، می‌توان به استخراج اطلاعات مفید برای کشف دانش جدید از متن دست یافت. در ادامه، به برخی فعالیت‌های پژوهشی مرتبط با پرونده الکترونیک سلامت اشاره شده است.

Chen و همکاران پژوهشی را با استفاده از هوش مصنوعی بر روی پرونده‌های الکترونیک سلامت در ایالات متحده آمریکا و چین طی سال‌های ۲۰۰۸ تا ۲۰۱۷ انجام دادند و به این نتیجه رسیدند که روند تولیدات برای هر دو کشور افزایشی می‌باشد. از طریق مدل‌سازی موضوعی و تجزیه و تحلیل خوشه‌بندی، ۳۵ موضوع شناسایی شده برای تحقیقات ایالات متحده آمریکا به ۹ موضوع اصلی شامل «لخته شدن خون، حریم خصوصی اطلاعات سلامتی و محرمانه بودن آن، عوارض جانبی دارو و واکسن، تصویربرداری، بیماری، عملکرد صوتی و تصویری، روش تجزیه و تحلیل داده‌های بالینی» و برای کشور چین نیز ۳۵ موضوع شناسایی شده به ۷ خوشه شامل «سرطان، تصویربرداری، حمایت از تصمیم‌گیری بالینی، دارو و رویداد خطر، طب چینی، دیابت بارداری و تکنیک‌های تجزیه و تحلیل داده‌های بالینی» طبقه‌بندی شدند (۶). Chen و همکاران به تحلیل کتاب‌شناختی پردازش زبان طبیعی در تحقیقات پزشکی در پایگاه PubMed طی سال‌های ۲۰۰۷ تا ۲۰۱۶ پرداختند. نتایج مطالعه آنان نشان داد که ایالات متحده آمریکا بیشترین تعداد انتشارات حوزه موضوعی را داشته است و موضوعات زیست‌شناسی محاسباتی، واژه‌کاوی، استخراج اطلاعات، طبقه‌بندی متن، رسانه اجتماعی و بازیابی اطلاعات، بیشترین تمرکز را در مقالات حوزه به خود اختصاص دادند (۷). Yao و همکاران به تحلیل علم‌سنجی سیستم‌های سلامت در پایگاه Web of Science از سال ۲۰۰۰ تا ۲۰۱۲ پرداختند و نتیجه‌گیری کردند که روند رشد تولیدات افزایشی بوده است و بیشتر محققان که سیستم‌های بهداشتی را مطالعه کرده‌اند، بر موضوعات سلامت، پاسخگویی به سیستم‌های بهداشتی، رهبری و حکمرانی، تأمین مالی و هزینه‌های بهداشتی، خدمات بهداشتی و کیفیت مراقبت، ارزیابی عملکرد سلامتی، سیستم‌های اطلاعات بهداشتی و نیروی کار درمانی تمرکز داشتند (۸).

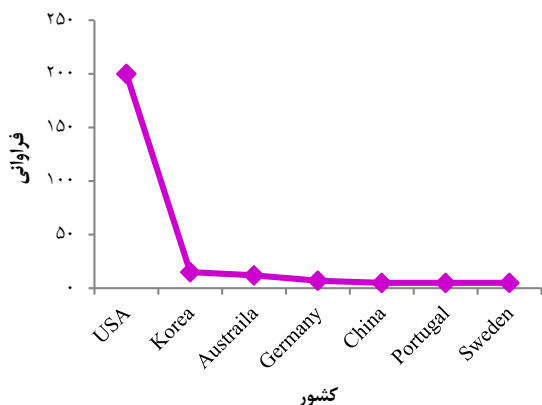
Han و همکاران در تحقیق خود به تجزیه و تحلیل موضوعی حوزه بیومدیکال مقالات پایگاه Web of Science طی سال‌های ۲۰۰۹ تا ۲۰۱۴ پرداختند و به این نتیجه رسیدند که موضوعات «سیستم‌های ضبط پزشکی، کامپیوتری، انفورماتیک پزشکی، پردازش زبان طبیعی، روش‌ها، سازمان و اداره، الگوریتم‌ها، ذخیره و بازیابی اطلاعات، تعامل کاربر و کامپیوتر، پیشگیری و کنترل» بیشترین هم‌رخدادی واژگان را داشتند (۹). Wen و همکاران با بررسی تولیدات علمی پرونده الکترونیک سلامت در پایگاه نمایه استنادی علوم طی سال‌های ۱۹۹۱ تا ۲۰۰۵، دریافته‌اند که تعداد زیاد مقالات و نشریات در حوزه پرونده الکترونیک سلامت بیانگر اهمیت روزافزون آن در جهان است. کشورهای ایالات متحده آمریکا و انگلستان بیشترین انتشارات حوزه را داشتند (۱۰). ایرانپور و همکاران در پژوهش کتاب‌سنجی و تحلیل محتوای برون‌داد پژوهشی ایران در

یافته‌ها نشان داد که روند رشد مقالات از سال ۲۰۰۹ تا ۲۰۱۹ صعودی بوده و هر سال به طور متوسط به صورت تصاعدی دو برابر شده است، اما به دلیل این که این گزارش در نیمه سال ۲۰۱۹ گرفته شد، تعداد تولیدات علمی حوزه سلامت در این سال کمتر بود (شکل ۲).



شکل ۲: روند تولیدات علمی حوزه پرونده الکترونیک سلامت در پایگاه PubMed طی سال‌های ۲۰۰۹ تا ۲۰۱۹

بررسی انجام شده در مورد کشورهای سرآمد در تولیدات علمی حوزه سلامت نشان داد که ایالات متحده آمریکا با اختلاف زیاد، پرتولیدترین کشور بود و پس از آن کشورهای کره جنوبی، استرالیا و آلمان جزء پرتولیدترین کشورها در امر تحقیقات حوزه سلامت بودند (شکل ۳).



شکل ۳: فراوانی پرتولیدترین کشورهای جهان در حوزه پرونده الکترونیک سلامت در پایگاه PubMed طی سال‌های ۲۰۰۹ تا ۲۰۱۹

برای تهیه شبکه همکاری نویسندگان در نرم‌افزار VOSviewer، حداقل تعداد مقالات منتشر شده توسط نویسندگان ۱۰ مورد در نظر گرفته شد و ۲۶۰۱۷ نویسنده مورد تجزیه و تحلیل قرار گرفتند که ۱۷۰ نفر معیار ورود را کسب کردند. نویسندگان David Bates، Dean Sittig و Hardeep Singh بیشترین مقالات را داشتند (جدول ۲).

کشورها، پربسامدترین واژه‌ها و تعیین هم‌رخدادی واژگان) به منظور تحلیل و مشاهده نقشه‌های کتاب‌شناختی (۱۴) به نرم‌افزار VOSviewer وارد گردید و برای ترسیم جداول و نمودار نیز از نرم‌افزار Excel استفاده شد.

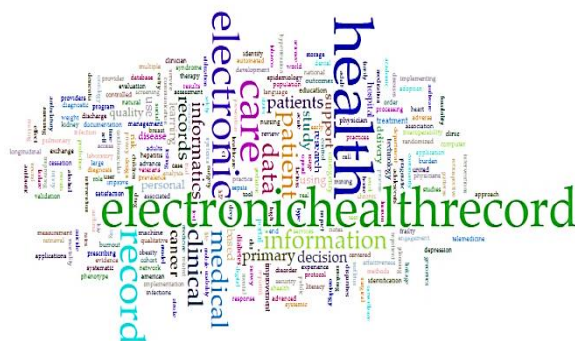
### یافته‌ها

یافته‌های مطالعه در دو قسمت متن‌کاوی و کتابسنجی ارایه گردید. موضوعات اصلی حوزه پرونده الکترونیک سلامت در پیکره مورد بررسی، با استفاده از روش متن‌کاوی در ابزار Voyant مشخص گردید. موضوعات اصلی فراوانی آن‌ها در جدول ۱ آمده است.

جدول ۱: موضوعات اصلی حوزه پرونده الکترونیک سلامت در مقالات پایگاه PubMed طی سال‌های ۲۰۰۹ تا ۲۰۱۹

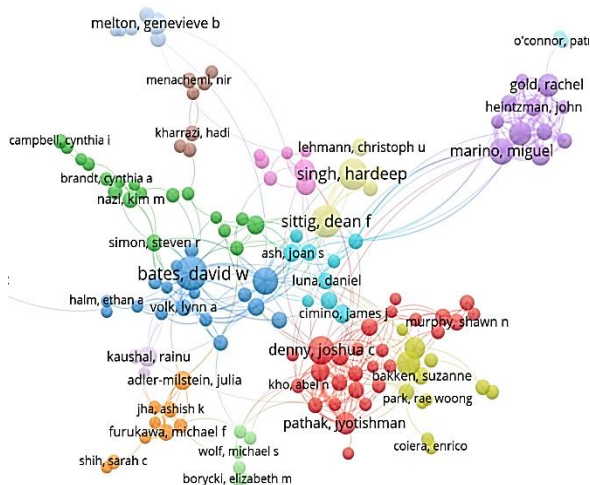
کلید واژه‌ها	فراوانی
Primary care	۶۶
Health record medical	۴۳
Patient portal	۴۱
Health information	۲۹
Electronic medical	۲۹
Clinical decision	۲۹
Machine learning	۲۸
Clinical decision support	۲۷
Electronic health record	۸۲۲
Health	۴۶۳
Health care	۱۴۶
Information	۱۳۰
Health care information system	۱۱۴
Medical	۹۸
Medical informatics	۹۴
Patient	۷۶

نقشه ابر کلمه موضوعات با ابزار Voyant در شکل ۱ نشان داده شده است.



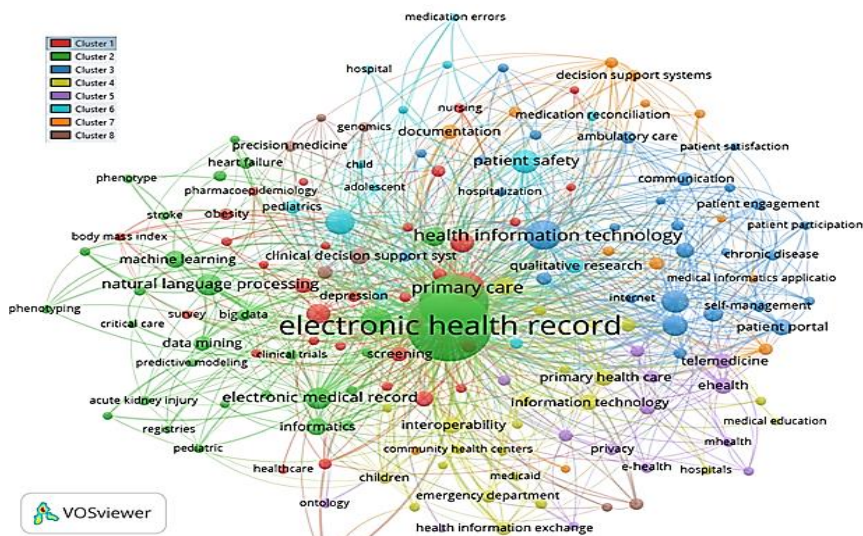
شکل ۱: نقشه ابر کلمه موضوعات حوزه پرونده الکترونیک سلامت در مقالات پایگاه PubMed طی سال‌های ۲۰۰۹ تا ۲۰۱۹

در نقشه بالا می‌رود و درک نقشه را مشکل می‌سازد. با در نظر گرفتن آستانه هم‌رخدادی مذکور برای واژه‌های مورد بررسی، نرم‌افزار ۱۶۱ واژه را در عناوین مدارک شناسایی نمود و نقشه هم‌واژگانی این تعداد واژه با استفاده از نرم‌افزار VOSviewer ترسیم گردید. خطوط بین مفاهیم در سطح نقشه نشان دهنده ارتباط بین مفاهیم می‌باشد و واژگان کلیدی که ارتباط بیشتری با هم دارند، در نقشه نزدیک به یکدیگر قرار می‌گیرند. همچنین، بزرگی و کوچکی دایره‌ها بیان‌کننده میزان دانش موجود در مورد هر مفهوم است. رنگ دایره‌ها نشان دهنده خوشه‌ای است که کلید واژه در آن واقع شده است. بیشترین تمرکز متون مورد بررسی به ترتیب بر روی موضوعات پرونده الکترونیک سلامت، مراقبت‌های اولیه، فن‌آوری اطلاعات سلامت و پرونده سلامت شخصی می‌باشد که دایره بزرگ‌تری نسبت به بقیه مفاهیم دارند (شکل ۵).



شکل ۴: شبکه همکاری نویسندگان در زمینه پرونده الکترونیک

سلامت طی سال‌های ۲۰۰۹ تا ۲۰۱۹



شکل ۵: نقشه هم‌رخدادی موضوعات حوزه پرونده الکترونیک سلامت بر اساس واژگان عنوان نویسنده

جدول ۲: فراوانی پرتولیدترین نویسندگان در حوزه پرونده الکترونیک

سلامت در پایگاه PubMed طی سال‌های ۲۰۰۹ تا ۲۰۱۹

نام نویسنده	تعداد مقاله	تعداد ارتباط با دیگر نویسندگان
David Bates	۷۱	۲۰
Dean Sittig	۶۴	۶۲
Hardeep Singh	۵۹	۳۹
Joshua Denny	۴۷	۲۶
Adam Wright	۴۷	۴۰
Miguel Marino	۴۲	۶۷
George Hripsal	۳۸	۱۵
Jennifer Devoe	۳۷	۶۹
Rachel Gold	۳۲	۵۵
Rovert Grundmeier	۳۱	۱۹
David Vavdrey	۳۱	۱۴
Jyotishman Pathak	۳۰	۲۵

در شکل ۴ شبکه همکاری نویسندگان پرتولید در زمینه سلامت با استفاده از نرم‌افزار VOSviewer ترسیم شده است. تراکم هر یک از نویسندگان بر اساس تعداد تولیدات علمی آن‌ها و موقعیت هر گره بر روی نقشه، نشان دهنده نزدیکی کار آن نویسندگان با یکدیگر است. تجزیه و تحلیل خوشه‌ای از شبکه همکاری نویسندگان نشان داد که این شبکه شامل ۱۴ خوشه در رنگ‌های مختلف می‌باشد. در تحلیل هم‌واژگانی، آستانه هم‌رخدادی برای کلید واژه‌ها تعیین می‌گردد. پژوهش حاضر حداقل هم‌رخدادی برای هر واژه در عنوان ۱۰ بار در نظر گرفته شد. موضوعاتی که واژه‌های آن‌ها آستانه هم‌رخدادی را کسب نکردند، از نتایج تحلیل حذف شدند. این مشکل در تحلیل هم‌واژگانی با پایین آوردن آستانه هم‌رخدادی کمتر می‌شود و کلید واژه‌های بیشتری با یابی می‌گردد، اما حجم واژه‌های کم‌ارزش



ایرانپور و همکاران (۱۱)، Yao و همکاران (۸)، Chen و همکاران (۷) و Wen و همکاران (۱۰) همسو بود. آن‌ها در نتایج خود به روند صعودی انتشار مقالات حوزه سلامت الکترونیک اشاره نمودند (۱۱، ۱۰، ۸، ۷).

کشور ایالات متحده آمریکا بیشترین انتشارات را به خود اختصاص داده بود. در نتایج تحقیقات Wen و همکاران (۱۰) و Yao و همکاران (۸) کشورهای ایالات متحده آمریکا و انگلستان و در مطالعه Chen و همکاران (۷) نیز کشور آمریکا به عنوان بیشترین کشور تولیدکننده انتشارات حوزه سلامت الکترونیک معرفی گردید که با نتایج بررسی حاضر همخوانی داشت. این یافته نشان می‌دهد که در کشورهای در حال توسعه به موضوع سلامت اهمیت زیادی داده می‌شود. آن‌ها بیشترین هزینه را صرف تحقیق و توسعه می‌کنند تا زمینه ارتقا و بهبود سلامت در جامعه را فراهم نمایند.

در پژوهش حاضر، بیشترین مقالات حوزه سلامت متعلق به David Bates، Dean Sittig و Hardeep Singh بود. David Bates استاد دانشگاه پزشکی هاروارد است که در زمینه‌های پشتیبانی از تصمیم‌گیری بالینی، انفورماتیک پزشکی و ایمنی بیمار فعالیت دارد (۱۵). Dean Sittig استاد انفورماتیک زیست پزشکی دانشگاه علوم پزشکی تگزاس می‌باشد که در زمینه‌های انفورماتیک پزشکی، تصمیم‌گیری بالینی، سلامت الکترونیک و پرونده الکترونیک سلامت فعالیت می‌کند (۱۶). Hardeep Singh رئیس برنامه سیاست‌گذاری سلامت، کیفیت و انفورماتیک مرکز نوآوری هوستون است که به موضوعات فن‌آوری اطلاعات سلامت و بهبود ایمنی بیمار از طریق پرونده الکترونیک سلامت علاقه دارد (۱۷). هر سه نویسنده پرتولید در حوزه سلامت از کشور آمریکا می‌باشند. در تحقیق Chen و همکاران، Denny از کشور آمریکا و Tang از کشور چین (۷) و در مطالعه Yao و همکاران نیز McKee از کشور انگلستان (۸) به عنوان پرتولیدترین نویسنده معرفی شدند. یافته‌ها نشان از عدم حضور نویسندگان ایرانی در جمع پرتولیدترین نویسندگان بود.

در پژوهش حاضر، مهم‌ترین موضوعات شامل واژه‌های «پرونده الکترونیک سلامت، مراقبت‌های اولیه، فن‌آوری اطلاعات سلامت، بهبود کیفیت، پشتیبانی تصمیم‌گیری بالینی، امنیت، پرونده سلامت شخصی، پرونده پزشکی الکترونیک، پردازش زبان طبیعی، اطلاعات پزشکی، اطلاع‌رسانی، یادگیری ماشین، سیستم‌های پشتیبانی تصمیم‌گیری بالینی، سلامت الکترونیک، پورتال بیمار، اپیدمیولوژی، فن‌آوری اطلاعات، تله‌مدیسن و مراقبت بیمار محور» بود. در تحقیق Chen و همکاران، موضوعات زیست‌شناسی محاسباتی، واژه‌کاوی، استخراج اطلاعات، طبقه‌بندی متن، رسانه اجتماعی و بازیابی اطلاعات بیشترین تمرکز در مقالات حوزه را داشتند (۷).

تحلیلی بر موضوعات کلیدی به دست آمده از دو روش متن‌کاوی و کتابسنجی نشان داد که مهم‌ترین موضوع در پرونده الکترونیک سلامت فرد، مراقبت خودمحوری با استفاده از امکاناتی که سیستم در اختیار او قرار داده است، می‌باشد. فرد برای انجام مراقبت‌های اولیه و حتی پیشگیری از بیماری، به صورت تله‌مدیسن نیاز به امکاناتی در سیستم دارد. مهم‌ترین این امکانات در درجه اول شامل تأمین امنیت و حفاظت از اطلاعات فرد و وجود خط‌مشی لازم، امکان برقراری ارتباطات اجتماعی با سایر بیماران و پزشکان در این محیط و امکان استفاده از سیستم‌های پیشرفته یادگیری ماشینی می‌باشد که کمک بسیار مؤثری در استفاده کارآمد از محیط شخصی‌سازی شده پرونده سلامت فرد می‌کند. همچنین، دسترسی به پایگاه‌های دانش و اطلاعات و فراهم آوردن امکان دسترسی آسان موضوعی به اطلاعات سلامت، از جمله مهم‌ترین

در تحلیل نقشه‌های هم‌رخدادی، محل قرار گرفتن هر مفهوم در سطح نقشه، فاصله نسبی هر مفهوم تا سایر مفاهیم می‌باشد که رابطه بین مفاهیم را مشخص می‌سازد. فاصله واژگان در نقشه‌ها مشخص می‌کند که متون موجود به چه میزان در مورد ارتباط و اثرات آن‌ها بر یکدیگر صحبت کرده‌اند. اگر مفهوم پرونده الکترونیک سلامت نزدیک به مفهوم پردازش زبان طبیعی قرار گیرد، به این معنی است که در متون موجود به اثرات آن‌ها بر یکدیگر توجه شده است، اما زمانی که فاصله زیاد پرونده الکترونیک سلامت از خطاهای پزشکی مشاهده می‌شود، بیانگر آن است که در ادبیات موجود به اثرات این واژگان بر یکدیگر کمتر توجه شده است. سطح نقشه از پراکندگی به نسبت خوبی برخوردار می‌باشد و این امر نشان می‌دهد که پژوهشگران حوزه سلامت به موضوعات متنوعی پرداخته‌اند.

پرکاربردترین واژه‌ها در حوزه پرونده الکترونیک سلامت بر اساس هم‌رخدادی واژگان عنوان در جدول ۳ آمده است.

جدول ۳: رخداد پرکاربردترین واژگان عنوان در حوزه پرونده

الکترونیک سلامت در پایگاه PubMed طی سال‌های ۲۰۰۹ تا ۲۰۱۹

کلید واژه‌ها	رخداد	تعداد ارتباط با دیگر مفاهیم
Electronic health records	۷۰۰	۳۹۹
Primary care	۱۳۴	۹۲
Health information technology	۱۲۹	۱۵۷
Quality improvement	۹۶	۸۳
Clinical decision support	۸۷	۸۵
Patient safety	۸۴	۹۶
Personal health records	۷۹	۶۲
Natural language processing	۶۸	۶۱
Medical informatics	۵۵	۷۳
Machine learning	۴۳	۳۸
Clinical decision support systems	۴۱	۴۷
Patient portal	۳۹	۵۴
Information technology	۳۷	۳۳
telemedicine	۳۷	۳۳

به طور کلی، ۱۶۱ آیتم در ۸ خوشه، نقشه هم‌رخدادی واژگان عنوان در حوزه پرونده الکترونیک سلامت را مشخص نمود (شکل ۵). بزرگ‌ترین خوشه شامل ۳۳ واژه و بزرگ‌ترین واژگان کلیدی آن پرونده الکترونیک سلامت بود که اهمیت آن در حوزه سلامت را نشان می‌دهد.

## بحث

نتایج مطالعه حاضر نشان داد که روند انتشار مقالات در حوزه سلامت الکترونیک و پرونده‌های سلامت طی چند سال اخیر رشد چشمگیری داشته است که نشان از توجه محققان و اهمیت آن دارد. نتایج به دست آمده با یافته‌های پژوهش‌های

### پیشنهادهای

پیشنهاد می‌شود محققان ایرانی نیز به انتشار مقاله در مجلات نمایه شده در پایگاه PubMed و در این حوزه اقدام نمایند تا با بهره‌گیری از پژوهش‌های هدفمند، به توسعه و ارتقای سلامت همگام با سایر کشورهای پیشرفته جهان دست یابند.

### تشکر و قدردانی

بدین وسیله از حمایت‌های مالی و معنوی دانشگاه اصفهان تقدیر و تشکر به عمل می‌آید.

### تضاد منافع

در انجام پژوهش حاضر، نویسندگان هیچ‌گونه تضاد منافی نداشته‌اند.

موضوعاتی است که باید در این محیط در نظر گرفته شود.

از جمله محدودیت‌های پژوهش حاضر این بود که همه پایگاه‌های اطلاعاتی مورد بررسی قرار نگرفت. بنابراین، احتمال دارد برخی از تولیدات علمی حوزه سلامت وارد پژوهش نشده باشد.

### نتیجه‌گیری

هر کدام از واژگان نقشه هم‌رخدادی واژگان می‌تواند نماینده یک مفهوم یا حوزه تحقیقاتی در سلامت باشد و واژه‌هایی که ارتباط نزدیکی با همدیگر دارند، در یک خوشه قرار می‌گیرند، اما گاهی اوقات ترکیب بعضی از این واژه‌ها حوزه جدیدی را ایجاد می‌کند. نتایج مطالعه حاضر ساختار موضوعی حوزه پرونده الکترونیک سلامت را نشان داد که می‌تواند دید روشنی را در سیاست‌گذاری علمی این حوزه به منظور تأثیرگذاری بر تخصیص و توزیع منابع برای فعالیت‌های علمی و فنی ارائه دهد. همچنین، به محققان در انتخاب موضوعات داغ و کسب بیش‌جامعی از چارچوب علمی حوزه مورد نظر کمک می‌نماید.

### References

- Fenton SH, Low S, Abrams KJ, Butler-Henderson K. Health information management: changing with time. Yearb Med Inform 2017; 26(1): 72-7.
- Ramzani Sharestani M. Verification of design parameters of health care information systems [MSc Thesis]. Rasht, Iran: University of Guilan; 2015. [In Persian].
- Evans RS. Electronic health records: Then, now, and in the future. Yearb Med Inform 2016; (Suppl 1): S48-S61.
- van Eck NJ, Waltman L, Noyons EC, Buter RK. Automatic term identification for bibliometric mapping. Scientometrics 2010; 82(3): 581-96.
- Callon M, Courtial JP, Turner WA, Bauin S. From translations to problematic networks: An introduction to co-word analysis. Social Science Information 1983; 22(2): 191-235.
- Chen X, Liu Z, Wei L, Yan J, Hao T, Ding R. A comparative quantitative study of utilizing artificial intelligence on electronic health records in the USA and China during 2008-2017. BMC Med Inform Decis Mak 2018; 18(Suppl 5): 117.
- Chen X, Xie H, Wang FL, Liu Z, Xu J, Hao T. A bibliometric analysis of natural language processing in medical research. BMC Med Inform Decis Mak 2018; 18(Suppl 1): 14.
- Yao Q, Chen K, Yao L, Lyu PH, Yang TA, Luo F, et al. Scientometric trends and knowledge maps of global health systems research. Health Res Policy Syst 2014; 12: 26.
- Han D, Wang S, Jiang C, Jiang X, Kim HE, Sun J, et al. Trends in biomedical informatics: Automated topic analysis of JAMIA articles. J Am Med Inform Assoc 2015; 22(6): 1153-63.
- Wen HC, Ho YS, Jian WS, Li HC, Hsu YH. Scientific production of electronic health record research, 1991-2005. Comput Methods Programs Biomed 2007; 86(2): 191-6.
- Iranpour A, Haghdoost A A, Bazrafshan A, Okhovati M, Sharifpoor E, Zare M, et al . Bibliometric and content analysis of scientific outputs relevant to health education and promotion in iran during 1998-2011. Health Develop J 2017; 6(2): 144-53. [In Persian].
- Janssens F. Clustering of scientific fields by integrating text mining and bibliometrics [PhD Thesis]. Leuven, Belgium: Catholic University of Leuven; 2007.
- Heimerl F, Lohmann S, Lange S, Ertl T. Word cloud explorer: Text analytics based on word clouds. Proceedings of 47<sup>th</sup> Hawaii International Conference on System Sciences; 2014 Jan 6-9; Waikoloa, HI, USA. p. 1833-42.
- van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010; 84(2): 523-38.
- Google Scholar. David Bates. Professor of Medicine, Harvard Medical School [online]. Available from: URL: <https://scholar.google.com/citations?user=fQFJcdQAAAAJ&hl=en&oi=ao>
- Google Scholar. Dean F. Sittig. Professor of Biomedical Informatics, University of Texas [Online]. Available from: URL: <https://scholar.google.com/citations?user=rrAE9l4AAAAJ&hl=en&oi=ao>
- Google Scholar. Hardeep Singh. Chief, Health Policy, Quality, and Informatics Program, Houston VA HSRD Center of Innovation [Online]. Available from: URL: <https://scholar.google.com/citations?user=quaszY0gAAAAJ&hl=en&oi=ao>

## Combined Bibliometric and Text-Mining Analysis of Scientific Productions in PubMed Database in the Field of Electronic Health Records

Mahboobeh Shokouhian<sup>1</sup>, [Asefeh Asemi](#)<sup>2</sup>, Ahmad Shabani<sup>3</sup>, Mozafar Cheshmeh-Sohrabi<sup>4</sup>

### Original Article

#### Abstract

**Introduction:** Access to patient's complete information is critical in improving clinical care and reducing medical errors. Electronic Health Record is a collection of individuals' health information, from prenatal to posthumous, which is stored electronically, is available at any center and at any time, and is an integral part of an integrated health information system. The purpose of the present study was bibliometric and text-mining analyze of scientific products in the field of Electronic Health Records in PubMed database.

**Methods:** This present study was carried out using bibliometric method and text mining. The study was conducted in the academic year of 2019 in PubMed database on the period of 2009-2019, and 6863 articles were selected for review. Excel, VOSviewer and Voyant were used for data analysis.

**Results:** In the studied field, issues of electronic health records, health, health care, information, health care systems were of great importance in PubMed. Developing articles in this field had been on the rise for ten years, and the United States was the most productive country in the field. David Bates, Dean Sittig, and Hardeep Singh had the most articles in the field of study.

**Conclusion:** Each item of co-occurring vocabulary map can represent a concept or research area in health. The findings can provide a clear insight to scientific policymaking of this field to influence the allocation and distribution of resources for scientific and technical activities. It can also help researchers in selecting the state-of-the-art topics and having a comprehensive insight into the academic context of the field.

**Keywords:** PubMed; Scientific Productions; Electronic Health Records; Bibliography; Text Mining

Received: 23 Aug., 2019

Accepted: 24 Sep., 2019

Published: 07 Oct., 2019

**Citation:** Shokouhian M, Asemi A, Shabani A, Cheshmeh-Sohrabi M. **Combined Bibliometric and Text-Mining Analysis of Scientific Productions in PubMed Database in the Field of Electronic Health Records.** Health Inf Manage 2019; 16(4):190-6

Article resulted from PhD thesis funded by University of Isfahan.

1- PhD Student, Knowledge and Information Science, Department of Knowledge and Information Science, School of Education and Psychology, University of Isfahan, Isfahan, Iran

2- Associate Professor, Knowledge and Information Science, Department of Knowledge and Information Science, School of Education and Psychology, University of Isfahan, Isfahan, Iran AND School of Business Informatics, Corvinus University, Budapest, Hungary (Corresponding Author) Email: asefeh.asefeh@uni-corvinus.hu

3- Professor, Knowledge and Information Science, Department of Knowledge and Information Science, School of Education and Psychology, University of Isfahan, Isfahan, Iran

4- Associate Professor, Knowledge and Information Science, Department of Knowledge and Information Science, School of Education and Psychology, University of Isfahan, Isfahan, Iran