

بسم الله الرحمن الرحيم



وزارت علوم، تحقیقات و فناوری
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

طرح پژوهشی

**مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی
پژوهشکده فناوری اطلاعات ایرانداک با موضوعات بین‌المللی**

مجری
آرمان ساجدی‌نژاد

همکار
محمد ربیعی

مرداد ۱۴۰۱

حقوق: پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی پژوهشگاه فناوری اطلاعات ایرانداک با موضوعات بین‌المللی پیرو قرارداد شماره ۳۳/۲۸۰۳ مورخ ۱۴۰۰/۴/۱۲ میان پژوهشگاه علوم و فناوری اطلاعات ایران (کارفرما) و آقای دکتر آرمان ساجدی نژاد اجرا شده است. گزارش حاضر یک جلد از مستندات این طرح است. این گزارش و تمامی حقوق مادی آن بر اساس قانون حمایت حقوق مولفان و مصنفان و هنرمندان، مصوب ۱۳۴۸ و اصلاحیه‌های بعدی آن و همچنین آیین‌نامه‌های اجرایی این قانون متعلق به پژوهشگاه علوم و فناوری اطلاعات ایران و هرگونه استفاده از تمامی یا پاره‌ای از آن، شامل: نقل قول، تکثیر، انتشار، کاربرد نتایج، تکمیل و مانند آنها به صورت چاپی، الکترونیکی یا وسایل دیگر فقط با اجازه کتبی پژوهشگاه امکان‌پذیر است. نقل قول در حد هزار واژه در انتشارات علمی مانند کتاب و مقاله با درج اطلاعات کامل کتابشناختی، نیازی به مجوز پژوهشگاه ندارد. صحت مندرجات گزارش بر عهده مجری طرح پژوهشی است.

شماره: ۳۳،۴۷۴۵
تاریخ: ۱۴۰۱/۵/۲۹
پیوست:

برنام خدا

وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران
ایرانداک



پژوهش و مدیریت دانش آموزش

گواهی

ایرانداک در نیم‌قرن دوم کار خود همچنان پویا و کوشا، همراه هر پژوهش، پژوهشگر، و سیاست‌گذار

۱۳۵۷-۱۴۰۱

انجام طرح پژوهشی:

◇ عنوان: مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی

پژوهشکده فناوری اطلاعات ایرانداک با موضوعات بین‌المللی

◇ پدیدآوران: دکتر آرمان ساجدی‌نژاد، دکتر محمد ربیعی

◇ شماره (و تاریخ) قرارداد: ۳۳/۲۸۰۳ (۱۴۰۰/۴/۱۲)

◇ تاریخ شروع: ۱۴۰۰/۴/۱۲

◇ تاریخ پایان (تصویب گزارش پایانی در شورای پژوهش): ۱۴۰۱/۵/۲۳ (نشست ۴۶۸)

◇ ناظر: دکتر آزاده محبی

در پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) گواهی می‌شود، از سوی

ایرانداک، از پدیدآوران و ناظر گرامی این طرح ارزشمند سپاسگزارم.

تندرست باشید
لیلا نامداریان
سرپرست معاونت پژوهش و آموزش
نامدار

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی پژوهشکده فناوری اطلاعات ایرانداک با موضوعات بین‌المللی

مدیر طرح پژوهشی: دکتر آرمان ساجدی نژاد، پژوهشگاه علوم و فناوری اطلاعات ایران.

نشانی: تهران، خیابان انقلاب، چهارراه فلسطین، شماره ۱۰۹۰، طبقه ۴، اتاق ۴۱۲

صندوق پستی: ۱۳۱۵۷۷۳۳۱۴ تلفن: ۰۲۱-۶۶۴۹۴۹۸۰ (داخلی ۴۳۲)

وبگاه: <http://irandoc.ac.ir/u/a-sajedinejad>

رایانامه: Sajedinejad@irandoc.ac.ir

همکار طرح پژوهشی: دکتر محمد ربیعی، پژوهشگاه علوم و فناوری اطلاعات ایران.

چکیده

در طرح پژوهشی «مقایسه روند پژوهش‌ها در حوزه فناوری اطلاعات در پایان‌نامه‌ها و رساله‌های کشور و روند جهانی آن با استفاده از روش تحلیل متون»، حوزه‌های فناوری اطلاعات در جهان به دقت مورد بررسی و پایش قرار گرفت و با تاکید بر روند کلیدواژه‌های مورد استفاده، محدوده و روند فناوری اطلاعات در جهان و ایران مقایسه شد (ساجدی‌نژاد، ۱۳۹۹). این پژوهش اما با رویکردی متفاوت سعی دارد از جنبه دیگر یعنی حوزه تمرکز موضوعات هر گروه پژوهشی ایرانداک، به بررسی شکاف پژوهشی (کمبودهای احتمالی) هر گروه پژوهشی به نسبت پژوهش‌های بین‌المللی (به تفکیک سه گروه پژوهشی) پردازد و تفکیکی میان همه حوزه‌های فناوری اطلاعات (طرح پژوهشی و مقالات) و سه گروه پژوهشی قایل گردد. با توجه به تغییرات سریع در حوزه فناوری و تغییرات در مطالعات پژوهشگران مطرح در این زمینه (بین‌المللی)، بازه ۱ ساله آخر مورد بررسی قرار خواهد گرفت.

با فرض درهم‌تنیدگی پژوهش‌های گروه‌های پژوهشی ایرانداک و ماموریت‌های آنها در حوزه‌های عملکردی‌شان، ایجاد راهنمای مناسب و نظام‌مند از فعالیت‌های پژوهشی مناسب و منطبق در هر یک از گروه‌های پژوهشی می‌تواند به عمق‌بخشی بیشتر ماموریت‌های ایرانداک منجر گردد.

هرچند ممکن است حوزه‌های موضوعی فناوری اطلاعات در گروه‌های پژوهشی پژوهشکده فناوری اطلاعات دارای کلیدواژه‌های مشترک (و یا موضوعات مشترک) زیادی باشند، اما اثر و اهمیت هر کلیدواژه (یا موضوع) ممکن است برای هر یک و نهایتاً هر پژوهشگر هر گروه متفاوت باشد. دلیل نگاشت پژوهش‌های بین‌المللی و گروه‌های پژوهشی در این طرح، استخراج موضوعات پژوهشی کمتر پرداخته شده در هر گروه پژوهشی و یا کشف خلا پژوهشی هر گروه پژوهشی پژوهشکده فناوری اطلاعات نسبت به وضعیت متغیر و با شتاب پژوهش‌ها در حوزه فناوری اطلاعات در سطح بین‌المللی است. از نتایج این طرح می‌توان به این نکته اشاره نمود که در بخش قابل توجهی از پژوهش‌ها و اولویت‌یابی پژوهش در پژوهش‌های پژوهشکده فناوری اطلاعات نسبتاً تحت تاثیر نیازهای بالادستی بوده و ساختارهای تصمیم‌گیر و سیاست‌گذار پژوهشی این تاثیر را بر روند پژوهش انجام داده‌اند. با این حال به نظر می‌رسد به ازای تقاضای بالای پژوهشی در برخی موضوعات، پژوهش کافی انجام نمی‌شود که شاید دلیل عدم تامین اعتبار کافی برای این پژوهش‌هاست. «سیستم‌های پیشنهاددهنده که در راستای نیاز سیاست‌گذاران باشد» یکی از مهم‌ترین موضوعاتی است که نیاز به تمرکز جدی در بعد پژوهش پژوهشکده و مورد نیاز در اسناد بالادستی است. در بررسی و مقایسه روند پژوهش‌های ایرانداک با روندهای بین‌المللی متاسفانه تطبیق زیادی مشاهده نشده است. در پژوهش‌های ایرانداک تمرکز ویژه بر پژوهش‌هایی شده که در روندهای بین‌المللی در حال افول است. از سوی دیگر حوزه‌های داده‌های بزرگ و اینترنت اشیا و محاسبات ابری که در تمرکز پژوهش‌های بین‌المللی هستند، در وضعیت نامطلوب در ایرانداک مشاهده شده‌اند.

کلیدواژه‌ها: مدل‌سازی موضوعی^۱، متن کاوی، ایرانداک، گروه پژوهشی، مدیریت فناوری اطلاعات، کسب‌وکار الکترونیکی، سیستم‌های اطلاعاتی

¹ Topic modeling

فهرست مطالب

عنوان	شماره صفحه
۱. مقدمه و اهداف	۱
۱-۱. مقدمه	۲
۲-۱. بیان مساله	۳
۳-۱. اهداف طرح	۶
۴-۱. تعریف واژه‌ها و مفاهیم	۷
۲. بررسی موضوعی سه حوزه فناوری اطلاعات	۸
۱-۲. مقدمه	۹
۲-۲. روش پژوهش	۱۱
۳-۲. انتخاب مجلات معتبر	۱۵
۱-۳-۲. جستجوی مجلات برتر در حوزه فناوری اطلاعات	۱۸
۲-۳-۲. انتخاب مجلات با استفاده از نظر تکمیلی خبرگان	۲۱
۴-۲. انتخاب کلیدواژگان منتخب	۲۴
۵-۲. تقسیم‌بندی کلیدواژه‌ها توسط خبرگان در گروه‌های پژوهشی	۲۹
۶-۲. تحلیل حوزه‌ها با بهره‌گیری از داده‌های جستجو	۳۹
۱-۶-۲. بررسی چگونگی کاربرد کلیدواژگان در مجلات بالای نیم درصد	۳۹
۲-۶-۲. بررسی ناشران فعال در سه حوزه	۴۵
۳-۶-۲. بررسی حوزه‌های مطرح در زمینه‌های سه‌گانه فناوری اطلاعات	۴۸
۴-۶-۲. بررسی محیط پژوهش مطرح در زمینه‌های سه‌گانه فناوری اطلاعات	۵۲
۵-۶-۲. بررسی کلیدواژه‌های مطرح در زمینه‌های سه‌گانه فناوری اطلاعات	۵۵
۳. مدل‌سازی موضوعی سه حوزه فناوری اطلاعات	۵۷
۱-۳. مقدمه و پیشینه	۵۸

۶۱.....	۲-۳. مدل سازی موضوعی
۶۶.....	۳-۳. خوشه بندی موضوعی
۶۹.....	۴-۳. دسته بندی حوزه ها با بهره گیری از فرکانس همزمانی استفاده از کلیدواژگان
۷۵.....	۵-۳. دسته بندی حوزه های علمی بر اساس مدل سازی موضوعی کلیدواژه ها
۸۳.....	۶-۳. جمع بندی و نتیجه گیری
۸۵.....	۴. تحلیل شکاف
۸۶.....	۴-۱. مقدمه
۸۸.....	۴-۲. استخراج عناوین فعالیت های پژوهشی
۱۱۶.....	۴-۳. تحلیل شکاف پژوهشی
۱۱۷.....	۴-۳-۱. بررسی میزان تطبیق
۱۲۱.....	۳-۲. مقایسه میزان تطبیق در هر گروه پژوهشی پژوهشکده فناوری اطلاعات
۱۲۶.....	۴-۳-۳. مقایسه میزان تطبیق در هر پژوهشگر
۱۲۷.....	۴-۳-۴. مقایسه حوزه های تناسب طرح ها
۱۳۲.....	۴-۴. تحلیل تطابق پژوهش ها با اسناد بالادستی
۱۴۰.....	۵. نتیجه گیری و تحقیقات آتی
۱۴۱.....	۵-۱. مقدمه
۱۴۱.....	۵-۲. تطبیق طرح های پژوهشی با نیاز بالادستی
۱۴۴.....	۵-۳. تطبیق طرح های پژوهشی با روندهای پژوهشی
۱۴۶.....	۵-۴. مقایسه پخش طرح ها و مقالات در موضوعات
۱۴۷.....	۵-۵. پیشنهاد برای مطالعات آتی
۱۴۹.....	۶. فهرست منابع
۱۵۰.....	۶-۱. فهرست منابع فارسی
۱۵۰.....	۶-۲. فهرست منابع انگلیسی
۱۵۳.....	۷. پیوست: دسته بندی پایگاه WOS از انتشارات علوم

فهرست شکل ها

عنوان	شماره صفحه
شکل ۱-۲- نمایش کلمات پرکاربرد در حوزه آمار (باتیستی و همکاران ۲۰۱۵).....	۱۰
شکل ۲-۲- میزان کاربرد یک کلیدواژه (باتیستی و همکاران ۲۰۱۵).....	۱۰
شکل ۳-۲- مراحل پژوهش جاری.....	۱۲
شکل ۴-۲- مدل پژوهش جاری.....	۱۴
شکل ۵-۲- نمای جستجو در پایگاه WOS.....	۳۲
شکل ۶-۲- جستجوی کلیدواژه گروه های پژوهشی در پایگاه.....	۳۳
شکل ۷-۲- کاربرد کلیدواژگان در مجلات بالای نیم درصد.....	۴۰
شکل ۸-۲- ناشرین فعال در زمینه کسب و کار الکترونیکی.....	۴۱
شکل ۹-۲- مجلات فعال در زمینه سیستم های اطلاعاتی.....	۴۲
شکل ۱۰-۲- انتشارات در حوزه گروه پژوهشی مدیریت فناوری اطلاعات.....	۴۵
شکل ۱۱-۲- انتشارات در حوزه گروه پژوهشی کسب و کار الکترونیک.....	۴۶
شکل ۱۲-۲- انتشارات در حوزه گروه پژوهشی سیستم های اطلاعاتی.....	۴۷
شکل ۱۳-۲- حوزه های ترکیبی کلیدی در زمینه مدیریت فناوری اطلاعات.....	۴۹
شکل ۱۴-۲- حوزه های ترکیبی کلیدی در زمینه کسب و کار الکترونیکی.....	۵۰
شکل ۱۵-۲- حوزه های ترکیبی کلیدی در زمینه سیستم های اطلاعاتی.....	۵۱
شکل ۱۶-۲- کلیدواژه پژوهش مطرح در زمینه های مدیریت فناوری اطلاعات.....	۵۳
شکل ۱۷-۲- کلیدواژه پژوهش مطرح در زمینه های کسب و کار الکترونیک.....	۵۴
شکل ۱۸-۲- کلیدواژه پژوهش مطرح در زمینه های کسب و کار الکترونیک.....	۵۵
شکل ۱۹-۲- کلیدواژه های پر کاربرد حوزه فناوری اطلاعات در هر گروه پژوهشی.....	۵۶
شکل ۱-۳- نمونه توسعه چارچوب فناوری اطلاعات (باتیستی و همکاران ۲۰۱۵).....	۶۵
شکل ۲-۳- مقایسه روش های مختلف خوشه بندی (یاو و همکاران ۲۰۱۵).....	۶۸
شکل ۳-۳- فلوچارت اجرای الگوریتم LDA.....	۷۱
شکل ۴-۳- نمای خروجی مدل سازی موضوعی.....	۷۳
شکل ۵-۳- میزان انحراف سرگشتگی با توجه به اندازه دسته بندی در گروه پژوهشی مدیریت فناوری اطلاعات.....	۷۴
شکل ۶-۳- میزان انحراف سرگشتگی با توجه به اندازه دسته بندی در گروه پژوهشی کسب و کار الکترونیک.....	۷۴
شکل ۷-۳- میزان انحراف سرگشتگی با توجه به اندازه دسته بندی در گروه پژوهشی سیستم های اطلاعاتی.....	۷۵
شکل ۱-۴- نحوه دسترسی به مشخصات طرح های پژوهشی در ایرانداک.....	۹۰
شکل ۲-۴- درصد انجام طرح های پژوهشی در سه گروه پژوهشی.....	۹۸

- شکل ۳-۴- تکرار کلیدواژگان در طرح‌های پژوهشی ایرانداک ۱۱۸
- شکل ۴-۴- میانگین ارتباط کلیدواژگان در طرح‌های پژوهشی با دسته‌بندی انجام شده ۱۲۱
- شکل ۵-۴- تعداد ارتباط (کم، متوسط و زیاد) کلیدواژگان در هر گروه با دسته‌بندی انجام شده ۱۲۲
- شکل ۶-۴- روند میانگین سالانه ارتباط کلیدواژگان هر گروه با دسته‌بندی انجام شده ۱۲۲
- شکل ۷-۴- درصد پخش مقالات پژوهشکده در هر گروه با دسته‌بندی انجام شده ۱۲۵
- شکل ۸-۴- میزان تطبیق کلیدواژگان طرح‌های پژوهشی هر پژوهشگر با دسته‌بندی انجام شده ۱۲۶
- شکل ۹-۴- ارتباط کلیدواژگان در طرح‌های پژوهشی گروه مدیریت فناوری اطلاعات با حوزه‌های دسته‌بندی ۱۲۸
- شکل ۱۰-۴- ارتباط کلیدواژگان در طرح‌های پژوهشی گروه کسب و کار الکترونیک با حوزه‌های دسته‌بندی ۱۲۹
- شکل ۱۱-۴- ارتباط کلیدواژگان در طرح‌های پژوهشی گروه سیستم‌های اطلاعاتی با حوزه‌های دسته‌بندی ۱۳۰
- شکل ۱۲-۴- ارتباط وزن‌دار کلیدواژگان در طرح‌های پژوهشی گروه مدیریت فناوری اطلاعات ۱۳۱
- شکل ۱۳-۴- ارتباط وزن‌دار کلیدواژگان در طرح‌های پژوهشی گروه کسب و کار الکترونیک با حوزه‌های دسته‌بندی ۱۳۱
- شکل ۱۴-۴- ارتباط وزن‌دار کلیدواژگان در طرح‌های پژوهشی گروه سیستم‌های اطلاعاتی با حوزه‌های دسته‌بندی ۱۳۲
- شکل ۱-۵- نواحی تطبیق کلیدواژگان طرح‌های پژوهشی با نیاز بالادستی ۱۴۳
- شکل ۲-۵- نواحی تطبیق کلیدواژگان طرح‌های پژوهشی با روندهای جهانی ۱۴۵
- شکل ۳-۵- درصد تطبیق کلیدواژگان طرح‌های پژوهشی با مقالات مستخرج ۱۴۶

فهرست جدول‌ها

عنوان	شماره صفحه
جدول ۱-۲- فهرست مجلات دارای ضریب تاثیر بالا در پایگاه WOS.....	۱۶
جدول ۲-۲- اطلاعات و حوزه پوشش مجلات منتخب.....	۱۸
جدول ۳-۲- کلمات پرکاربرد در مجلات دارای ضریب تاثیر بالا.....	۱۹
جدول ۴-۲- تکرار کلیدواژگان پرکاربرد در مجلات مورد تایید خبرگان.....	۲۱
جدول ۵-۲- انتخاب کلیدواژه از دیدگاه خبرگان برای تشکیل سبد واژگان.....	۲۴
جدول ۶-۲- فهرست واژگان پایگاه WOS منتخب در هر گروه پژوهش.....	۲۹
جدول ۷-۲- نمونه اطلاعات مقالات هر گروه پژوهشی.....	۳۳
جدول ۸-۲- نمونه کلیدواژگان مقالات یافت شده در جستجوی نهایی.....	۳۸
جدول ۹-۲- مجلات مطرح در هر سه زمینه فعالیت گروه‌های پژوهشی.....	۴۳
جدول ۱-۳- دسته‌بندی موضوعات با مدل‌سازی موضوعی گروه پژوهشی مدیریت فناوری اطلاعات.....	۷۵
جدول ۱-۴- رویکردهای مختلف نیازسنجی و اولویت‌یابی پژوهش (فتحی و اجارگاه ۱۳۸۹).....	۸۹
جدول ۲-۴- مشخصات طرح‌های پژوهشی ایرانداک.....	۹۰
جدول ۳-۴- مشخصات مقالات پژوهشی ایرانداک.....	۹۹
جدول ۴-۴- میزان ارتباط طرح‌های پژوهشی ایرانداک با موضوعات استخراج شده.....	۱۱۹
جدول ۵-۴- میزان ارتباط مقالات پژوهشی ایرانداک با موضوعات استخراج شده.....	۱۲۳
جدول ۶-۴- شمارش میزان ارتباط مقالات با هر موضوع.....	۱۲۵
جدول ۷-۴- تطبیق دسته‌بندی‌ها و مطالعات پژوهش‌کنده بر نیازهای بالادستی.....	۱۳۷
جدول ۱-۵- مقایسه کلیدواژه‌های پرکاربرد در دو بازه زمانی (ساجدی‌نژاد ۱۳۹۹).....	۱۴۴

قدردانی

در این قسمت جا دارد از زحمات همکاران خوبم بویژه جناب آقای دکتر علیدوستی رئیس محترم پژوهشگاه و دیگر همکاران گرامی که به عنوان خبرگان در این پژوهش سهمی بزرگی داشته‌اند و در این طرح کمک شایانی نموده‌اند و همچنین همکارانم در پژوهشکده مدیریت فناوری اطلاعات تشکر و قدردانی نمایم. همچنین از همکارم در این طرح پژوهشی جناب آقای دکتر محمد ربیعی تشکر دارم.

از خبرگان سهمیم در این پژوهش بویژه دکتر نعیمی صدیق و همچنین ناظر محترم سرکار خانم دکتر محبی کمال سپاس و قدردانی را دارم.

از سرکار خانم دین محمدی و اعضای محترم شورای پژوهش برای حمایت‌هایشان سپاسگزارم.

فصل اول

مقدمه و اهداف

۱-۱. مقدمه

پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) به‌عنوان تنها پژوهشگاه فعال در زمینه فناوری اطلاعات در وزارت عتف یکی از مراجعی است که با ساخت ده‌ها فرهنگ مرتبط جامع^۲، به عنوان یکی از مراجع مورد استناد پژوهش‌ها قرار گرفته است. با وجود انجام شدن پژوهش‌های فراوان در پژوهشگاه با هدف استخراج مدل‌های توسعه فناوری اطلاعات، به نظر می‌رسد کشف مرزهای حوزه فناوری اطلاعات و روندهای آن در سه گروه پژوهشی مدیریت فناوری اطلاعات، کسب و کار الکترونیک، سیستم‌های اطلاعاتی در پژوهشکده فناوری اطلاعات ایرانداک هنوز صورت پذیرفته است. از سوی دیگر ایرانداک با تجمع داده‌های فعالیت‌های علمی کشور ارزش افزوده فراوانی (مانند همانندجویی و رصد وضعیت پژوهش کشور و ...) ایجاد نموده است، اما روندهای علمی و پژوهشی به تفکیک پژوهشکده‌های در ایرانداک و همچنین به تفکیک گروه‌های پژوهشی آن تا کنون مورد توجه ویژه قرار نگرفته است. مدل‌سازی موضوعی یک رویکرد خوشه‌بندی اطلاعات متنی هستند که به‌منظور طبقه‌بندی موضوعات موجود در حوزه دانشی قابل استفاده است و با بکارگیری آن امکان ارائه اطلاعات علم‌سنجی مربوط به شاخه‌ها و زیرشاخه‌های یک حوزه دانشی فراهم می‌شود و ارائه یک ابزار عملی برای پژوهش در زمینه توسعه دانش است. به گفته وی و کرافت (۲۰۰۶) با استفاده از روش‌های مبتنی بر خوشه‌بندی و توانایی استفاده از مدل‌سازی موضوعی در تحلیل علم‌سنجی در آینده نزدیک روش‌های جدیدی برای تجزیه و تحلیل داده‌ها توسعه داده خواهد شد.

الگوریتم‌های مدل‌سازی موضوعی شامل تکنیک‌های آماری است که هدف آنها توصیف موضوعاتی است که در اسناد (منتشر شده) مانند مجلات منتشر شده‌اند. آنها بر اساس تجزیه و تحلیل آماری کلمات در آن اسناد، شناسایی خوشه‌هایی از کلمات مشترک، تشخیص موضوعات مورد بحث، روابط بین این موضوعات و تکامل آنها در طول زمان کشف می‌کنند. چندین مدل مانند تجزیه و

² Thesaurus

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۳

تحلیل معنایی نهفته^۳ (لاندرو و همکاران، ۱۹۹۸)، تجزیه و تحلیل معنایی احتمالی^۴ (هافمن ۱۹۹۹) و تخصیص نهفته دیریکله^۵ (بلی و همکاران ۲۰۰۳)؛ و همچنین برخی از مشتقات آنها، مانند پاچینو^۶ (لی و مک کالم ۲۰۰۶) و یا مدل‌سازی موضوعی رابطه‌ای (چانگ و بلی ۲۰۰۹) در این زمینه کاربرد دارند.

۲-۱. بیان مساله

در دهه‌های اخیر، شاهد افزایش چشم‌گیر تعداد پژوهش‌هایی هستیم که در زمینه‌های مختلف علوم صورت گرفته است، اما رشد کمی پژوهش‌ها، لزوماً با افزایش کیفیت آن‌ها و به‌روز بودنشان با سطح بین‌المللی همراه نبوده است و در بسیاری از موارد دستاوردهای این پژوهش‌ها منجر به رفع مشکلات جامعه نشده و حتی به‌ندرت مشاهده شده است که نتایج و یافته‌های پژوهشی برای بهبود اوضاع سیاسی، اجتماعی، اقتصادی و فرهنگی جامعه بکار گرفته شود. در این میان شاید رشته‌های مهندسی و پژوهش‌های کاربردی این گروه وضعیت بهتری را نسبت به دیگر گروه‌های آموزشی داشته باشند (قنادی‌نژاد و حیدری، ۱۳۹۶).

بر اساس تعریف راینسون^۷، کاستی پژوهشی، محدودیت اطلاعاتی در حوزه یا موضوعی است که مانع از نتیجه‌گیری در مورد یک سؤال پژوهشی می‌شود و نیاز پژوهشی محدودیت اطلاعاتی است که تصمیم‌گیری سیاست‌گذار را محدود می‌کند (Robinson, Saldanha and Mckoy 2011). به عبارتی بسیاری از پژوهش‌هایی که با عنوان تعیین یا شناسایی نیازهای پژوهشی مطرح شده‌اند و در فرایند شناسایی از تحلیل الگوهای جستجوی اطلاعات کاربران پایگاه‌های اطلاعاتی یا کاربران کتابخانه‌ها استفاده کرده‌اند ولی به نظر می‌رسد که نیازهای پژوهشی فراتر از خلأ پژوهشی است و حوزه‌های دانشی وجود دارد که ممکن است جزو نیازهای پژوهشی باشند اما بخشی از خلأ پژوهشی یا اطلاعاتی نباشند زیرا کمتر کسی به دنبال جستجوی این حوزه‌ها بوده است و در اصل تقاضایی برای این حوزه‌ها وجود نداشته و از این‌رو در خلأ پژوهشی جای نگرفته‌اند (Furner, 2015). همچنین بسیاری از نیازهای پژوهشی ممکن است به دلیل پیرو بودن پژوهشی رخ دهد و باعث ایجاد یک فاصله زمانی چند ساله میان پژوهش‌های داغ دنیا و در سطح ملی گردد.

³ LSA

⁴ PLSA

⁵ Dirichlet (LDA)

⁶ Pachinco

⁷ Robinson, Karen A

تشکیل و فعالیت گروه‌های پژوهشی دانشکده یا پژوهشکده، یک ابزار اصلی در اجرای اولویت‌های علمی دانشکده‌ها یا پژوهشکده‌ها است و پرداختن به موضوعات مورد پژوهش آنها یک سرمایه‌گذاری راهبردی برای پاسخگویی به انتظارات تعیین شده بلند مدت در برنامه استراتژیک هر دانشگاه یا پژوهشگاه خواهد بود. هدف گروه‌های پژوهشی دانشکده‌ها یا پژوهشکده‌ها، ارتقا کیفیت پژوهش است و این کار از طریق عمق بخشیدن به مطالعات تخصصی هر پژوهشگر در هر گروه پژوهشی، پوشش موضوعات آن گروه پژوهشی و نهایتاً از طریق تقویت همکاری پژوهشی در گروه‌های پژوهشی و پژوهشکده‌ها شکل می‌گیرد.

پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) دارای پژوهشکده‌ای با نام فناوری اطلاعات است که از سه گروه پژوهشی مدیریت فناوری اطلاعات، کسب و کار الکترونیک و سیستم‌های اطلاعاتی تشکیل شده است. با وجود پژوهش‌های انجام شده در پژوهشگاه مبتنی بر استخراج مدل‌های توسعه فناوری اطلاعات و همچنین کشف مرزهای فناوری اطلاعات و روندهای آن در زمینه فناوری اطلاعات، هنوز میزان تناسب موضوعات پژوهش‌های انجام شده در هر پژوهشکده با وضعیت بین‌المللی این پژوهش‌ها مورد بررسی قرار نگرفته است و درهم تنیدگی روندهای علمی و پژوهشی در حوزه‌های تخصصی فناوری اطلاعات در ایرانداک هنوز مورد تفکیک (در سه گروه پژوهشی) قرار نگرفته است.

در عموم طرح‌های بررسی شده (مانند طرح پژوهشی (ساجدی‌نژاد، ۱۳۹۹)) بازه بررسی روندهای فناوری اطلاعات در چندین سال مورد بررسی قرار گرفتند، زیرا به دست آوردن روند در هر حوزه نیازمند بررسی در بازه زمانی نسبت طولانی می‌بود. پژوهش جاری بیشتر به وضعیت کنونی فناوری اطلاعات در مطالعات بین‌المللی می‌پردازد. پس در این پژوهش، بازه کوتاه می‌تواند مورد استناد قرار گیرد. در این طرح بنا داریم در ابتدا مجموعه‌ای از کلیدواژه‌های مهم هر گروه پژوهشی را (با مصاحبه با نماینده هر گروه پژوهشی و در اختیار گذاردن کلیدواژه‌های فناوری اطلاعات مستخرج از طرح پژوهشی قبل (ساجدی‌نژاد، ۱۳۹۹)) استخراج کرده (سبد واژگان اولیه) و با استفاده از کلیدواژگان مقالات معتبر (سبد واژگان غنی شده)، مدل‌سازی موضوعی را در هر گروه پژوهشی پژوهشکده فناوری اطلاعات ایرانداک ارایه می‌دهیم. آنگاه با کشف تابع نزدیکی یک مجموعه کلیدواژه به دسته‌بندی‌های موضوعی یافت شده، سعی خواهیم نمود اثر و نزدیکی مقالات علمی پژوهشگران هر گروه پژوهشی (۵ سال اخیر) را بر موضوعات مدل‌سازی شده، کشف کنیم.

با توجه به توضیحات فوق، در این پژوهش، با رویکرد تشریح شده مدل‌سازی موضوعی بر روی کلیدواژه‌های پژوهش‌های مجموعه‌ای از افراد پیشرو در زمینه‌های جاری فناوری اطلاعات در مجلات

بین‌المللی، انجام پذیرفته و پس از پردازش نتایج، موضوعات در تقسیم‌بندی سه گروه پژوهشی پژوهشکده فناوری اطلاعات ایرانداک مورد بررسی قرار می‌گیرد. این پژوهش وضعیت موضوعات پژوهش‌های حوزه فناوری اطلاعات در سه پژوهشکده را با مدل موضوعی پژوهش‌های مربوط به فناوری اطلاعات در سطح جهان بررسی می‌کند. در تعریف یک مساله عملیاتی، یکی از مهمترین چالش‌ها، دسترسی مناسب به اطلاعات مورد نیاز است. منابع اصلی این پژوهش مجلات پایگاه WOS و همچنین سایت ایرانداک (برای بررسی پژوهش‌های سه گروه پژوهشی پژوهشکده) هستند.

بنابراین در این پژوهش اقدامات اساسی زیر صورت خواهد پذیرفت:

- استخراج موضوعات پژوهشی هر گروه پژوهشی پژوهشکده فناوری اطلاعات با دانش به‌روز در جهان با (مطالعه بین‌المللی و خبرگی)
 - استخراج الگوی نزدیکی هر مجموعه از کلیدواژگان هر مقاله به هر موضوع (مدل‌سازی شده موضوعی)
 - مقایسه دستاوردهای فعلی هر پژوهشکده با حوزه‌های بین‌المللی (تحلیل آماری)
- در این پژوهش از زبان برنامه‌نویسی R استفاده خواهد شد. مهم‌ترین ویژگی‌هایی که موجب انتخاب این زبان برنامه‌نویسی در این پژوهش عبارت‌اند از:
- یک زبان علمی است و قابلیت‌ها و ابزار فراوانی در مدیریت داده‌ها، توابع آماری و ریاضیاتی، توابع داده‌کاوی، متن‌کاوی و پردازش زبان طبیعی را داراست و سطح بالایی از همخوانی بین مفاهیم علمی و آکادمیک در حوزه‌های پردازش و تحلیل داده و داده‌کاوی با ابزارها و توابع توسعه داده‌شده برای این محیط وجود دارد، به‌نحوی که بسته‌های نرم‌افزاری، توابع، ابزارها و نحوه کار آن‌ها عمده‌تاً در مقالات علمی و به زبان آکادمیک تشریح شده و در پایگاه این زبان در دسترس است (Gru & Hornik, 2011).
 - به دلیل متن‌باز بودن این امکان را فراهم می‌کند تا اجزای داخلی الگوریتم‌ها، مدل‌ها و توابع را مشاهده کرده و امکان تغییر آن‌ها را فراهم می‌کند.
 - امکان بسط، توسعه و بومی‌سازی بسته‌ها و توابع در این زبان فراهم است.
 - به دلیل ساختار ماتریسی که برای متغیرها در نظر می‌گیرد تطابق زیادی با نگاشت‌های مفهومی در فرایندهای متن‌کاوی و داده‌کاوی دارد و انجام عملیات کشف دانش در آن ملموس است.

- به دلیل متن‌باز بودن گستره زیادی از توابع، ابزارها و راه‌حل‌های موردنیاز در حوزه‌های آماری، پردازش متن و مدل‌سازی را می‌توان در فضای وب جستجو کرده و از آن‌ها استفاده نمود.
- امکان انجام پردازش‌های موازی و توزیع‌شده برای کار با حجم انبوه داده‌ها و پردازش‌های سنگین را دارد که این قابلیت برای داده‌ها و پردازش‌های مربوط به این پژوهش بسیار حیاتی است.
- دارای راهنمای داخلی بسیار خوبی است و در فضای وب گروه‌های علمی زیادی پاسخگوی پرسش‌ها و نیازهای احتمالی خواهند بود.
- رایگان بوده و قابلیت نصب روی سیستم‌عامل‌های گوناگون را دارد و امکان تبادل داده و کد را روی نسخه‌های مختلف داراست.

۳-۱. اهداف طرح

از اهداف نهایی طرح جاری می‌توان به موارد زیر اشاره نمود:

- استخراج موضوعات پژوهشی در هر گروه پژوهشی پژوهشکده فناوری اطلاعات
 - استخراج الگوی نزدیکی هر کلیدواژه به هر موضوع
 - مقایسه شکاف بین پژوهش‌های فناوری اطلاعات در ایرانداک و موضوعات جهانی و تحلیل و مقایسه مطالعات و پژوهش‌های هر گروه پژوهشی به موضوعات آن گروه
- پرسش‌های پژوهش جاری در سه دسته طبقه‌بندی می‌شوند:
- موضوعات پژوهشی و سبد واژگان حوزه‌های سه گانه (گروه‌های پژوهشی) فناوری اطلاعات شامل چه واژگانی است؟
 - چگونه می‌توان میزان ارتباط یک کلیدواژه را به هر یک از موضوعات هر گروه‌های پژوهشی یافت؟
 - پژوهش‌های هر گروه پژوهشی در حوزه فناوری اطلاعات با طبقه‌بندی بین‌المللی چقدر تطابق داشته است و کمبودهای پژوهشی مطالعات پژوهش‌های فناوری اطلاعات در ایرانداک نسبت به پژوهش‌های بین‌المللی در کدام موضوعات است؟
- لازم به ذکر است این طرح پژوهشی با طرح‌های پژوهشی پیشین پژوهشگر
- مقایسه روند پژوهش‌ها در حوزه فناوری اطلاعات در پایان‌نامه‌ها و رساله‌های کشور و روند جهانی آن با استفاده از روش تحلیل متون، آرمان ساجدی نژاد، ۱۳۹۹
 - طرح پژوهشی بررسی کارکردهای مدیریت اطلاعات علمی و فناوریانه در بافت دولت الکترونیکی و دولت همراه، آرمان پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)،

آرمان ساجدی نژاد، ۱۳۹۶

- طراحی چارچوب مفهومی بومی تعالی فناوری اطلاعات سازمانی، آرمان ساجدی نژاد، ۱۳۹۷
- بازطراحی کارکرد مدیریت اطلاعات علمی و فناوری پژوهشگاه علوم و فناوری اطلاعات ایران در بافت دولت الکترونیکی، آرمان ساجدی نژاد، ۱۳۹۵

و همچنین با تعدادی از پژوهش‌های پیشین مانند

- معجبی، آ. (۱۳۹۶). طرح پژوهشی تدوین نقشه راه فناوری اطلاعات پژوهشگاه با تمرکز بر پیشرفتهای نوین در حوزه تجزیه و تحلیل اطلاعات. پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
 - نامداریان، ل. (۱۳۹۵). طرح پژوهشی بررسی و تبیین جایگاه پژوهشگاه علوم و فناوری اطلاعات در نظام سیاستگذاری علم و فناوری کشور. پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
 - طراحی و تدوین منشور پژوهش، پژوهشکده‌ها، و گروه‌های پژوهشی پژوهشگاه علوم و فناوری اطلاعات ایران، عمار جلالی منش، ۱۳۹۹
- ارتباط و پیوند خواهد داشت.

۴-۱. تعریف واژه‌ها و مفاهیم

- وابستگی پیشرو: افرادی که از لحاظ ضریب تاثیر در حوزه یاد شده دارای بالاترین رتبه‌ها هستند
- کلیدواژه: واژه‌های کلیدی مقالات منتخب در مجلات پیشرو یا پژوهش‌های داخلی و یا انتخاب نمایندگان هر گروه پژوهشی
- سبد واژگان (اولیه و غنی شده): مجموعه واژگان استخراج شده از خبرگان و یا عناوین و کلیدواژه مقالات مورد مطالعه در حوزه فناوری اطلاعات
- موضوع: مجموعه‌ای از کلیدواژه‌ها که می‌توانند نماینده یک حوزه علمی در زمینه فناوری اطلاعات باشند
- مدل‌سازی موضوعی: مدل‌سازی کلمات در موضوعات مختلف متون علمی با کشف درجه احتمال پراکندگی
- پارسا: پایان نامه و رساله‌ی ثبت شده در زمینه فناوری اطلاعات در سامانه گنج و دارای کلیدواژه‌های انگلیسی
- مدل‌سازی موضوعی: مدل‌سازی کلمات در متون علمی با کشف درجه احتمال پراکندگی

فصل دوم

بررسی موضوعی سه حوزه فناوری اطلاعات

۱-۲. مقدمه

همانطور که در فصل قبل ذکر شد، در مدل‌سازی موضوعی، ارتباط میان کلمات و زمینه‌های معنایی نهفته در اسناد استخراج می‌شود (بلی و لفرتی ۲۰۰۷). در ادبیات (هال و همکاران ۲۰۰۸؛ بلی ۲۰۱۲) همواره بین نیاز به تعداد زیادی از واژگان برای پوشش همه زمینه‌های یک حوزه علمی در مجموعه اسناد موضوعی چالش وجود دارد. (گرین و هورنیک ۲۰۱۱). همانطور که در همه رشته‌ها رایج است، مجلات زیادی فناوری اطلاعات را از دیدگاه‌های مختلف بررسی می‌کنند. رایج‌ترین معیارهای بخش‌بندی ادبیات و مجلات عبارتند از: تعریفی یا نظری بودن مطالعات آنها، ارتباط نام مجله با موضوعات، مرزها و زیر شاخه‌های زمینه مورد نظر (بروک ۲۰۰۷)، بکلند ۲۰۱۲ و فرنر ۲۰۱۵). بسیاری از مطالعات منتشر شده در مجلات دامنه محدودی دارند، که بر منابع خاص، موضوع یا تخصصی ویژه متمرکز شده‌اند. بنابراین، بسیاری از پژوهش‌ها در زمینه کشف حوزه‌های فعالیت بر روی یک مجله متمرکز شده است (به‌عنوان نمونه مطالعه هارت و هوتن (۱۹۹۲) و تأکید وابستگی‌های نهادی نویسندگان یا (سوین و همکاران ۲۰۱۲)). در عین حال، برخی دیگر از مطالعات شناخت حوزه‌های مفهومی (مانند فناوری اطلاعات) بر مجموعه‌ای از کنفرانس‌ها تمرکز می‌کنند (مانند مقاله اسمتون و همکاران (۲۰۰۲) در مورد بازیابی اطلاعات، و یا مدل‌سازی موضوعی با استفاده از تکنیک‌های تجزیه و تحلیل شبکه اجتماعی (SNA) توسط ساگیموتو و مک‌کین (۲۰۱۰)). گاهی اوقات، نویسندگان مجبورند در مدل‌سازی موضوعی فقط به نوعی از مجلات یا یک مجله خاص متمرکز شوند، بدون اینکه تلاش کنند تا کل زمینه علمی را به طور جامع پوشش دهند. چنین موردی در مورد مطالعه موخر جی (۲۰۰۹) با تجزیه و تحلیل ۱۷ مجله دسترسی آزاد در حوزه مربوطه صورت پذیرفت. در موارد دیگر تمرکز بر روی نوع پژوهش است؛ بطور مثال پژوهش کیم و جونگ (۲۰۰۶) در مورد پژوهش‌های نظری است. یکی دیگر از دیدگاه‌های کمی متمایز ارائه شده در مطالعات، استفاده از روش‌های تجزیه و تحلیل محتوا است. این روشها بر تجزیه و تحلیل کلمات کلیدی متمرکز هستند و کلمات در عنوان یا چکیده و یا حتی متن کامل را مورد تحلیل قرار می‌دهند (مانند فعالیت

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۱۰

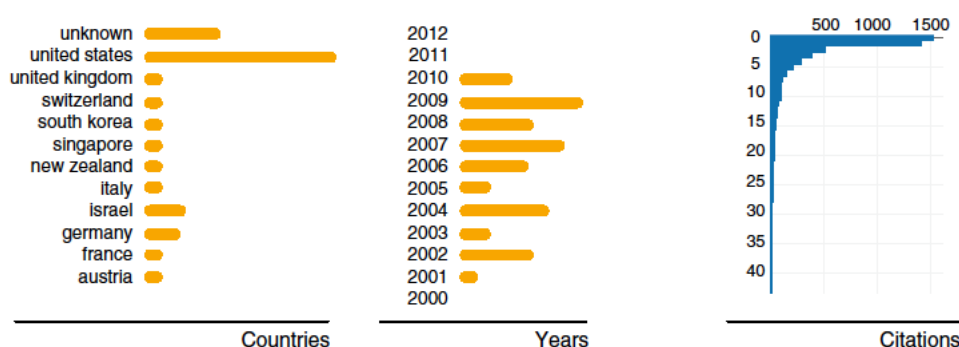
علمی آستروم (2002). اما همانطور که ذکر گردید، فعالیت مشابه در زمینه شناسایی حوزه فناوری اطلاعات در ادبیات انجام در این طرح پژوهشی یافت نشده است.

به منظور ارائه نمونه‌ها از توصیف موضوع، به ارایه نمونه‌ای از مدل‌سازی موضوعی در حوزه آمار اشاره می‌نمایم. با توجه به رویکرد توصیف شده در بخش‌های قبلی این طرح، در این پژوهش توصیفگرهای موضوع با تعیین کلیدی‌ترین کلمات کلیدی و منبع انتشار، سال و توزیع در کشورها، محاسبه شده‌اند. دو مجله مورد بررسی این پژوهش در بازه ۱۰ ساله به طور خاص، سالنامه آمار و JASA بوده‌اند. توصیفگرها همچنین می‌توانند برای ارائه یک نمایش گرافیکی از یک موضوع، با توجه به رویکرد بیان شده استفاده شوند. یک موضوع به عنوان یک متن نمایش داده می‌شود که بزرگی و قطر آن متناسب با تعدادی از مقالات شامل آن موضوع است (شکل زیر).



شکل ۲-۱- نمایش کلمات پر کاربرد در حوزه آمار (باتیستی و همکاران ۲۰۱۵)

کلمه گسترش و کشف (Discoveries) در طی سال‌ها نوسان داشته و در طول زمان با افزایش اوج خود در سال ۲۰۰۹ افزایش یافته است. بیشترین فعالیت در این زمینه کشور ایالات متحده است (شکل ۲-۲).



شکل ۲-۲- میزان کاربرد یک کلیدواژه (باتیستی و همکاران ۲۰۱۵)

با استفاده روزافزون از پایگاه داده‌های متنی ساختار یافته، رویکردهای تازه در زمینه اندازه‌گیری علم بیشتر بر روش‌های کمی تکیه می‌کنند که اغلب به عنوان کتابشناختی نامیده می‌شوند (برگمن و فرنر

(2002). در این پژوهش‌ها و مطالعات از ابزارهای متدولوژیکی مختلف برای استخراج اطلاعات از پایگاه‌های داده استفاده شده و پژوهشگران تلاش دارند تا ساختارهای زیرین را درون مجموعه‌ای از داده‌ها کشف کنند (دایم و همکاران 2006). مطالعات با استفاده از روش‌های مختلف ابتدایی (سومینن 2013) تا با بهره‌گیری از روش‌های مختلف برای تعیین تعداد کشورها، نویسندگان یا فناوری‌های ذکر شده در یک سال توسعه یافته و این متغیرها با روش‌های متن‌کاوری تحلیل می‌شوند (گلیسون و همکاران 2005). در این طرح پژوهشی، با بهره‌گیری از روش مدل‌سازی موضوعی که یکی از روش‌های آماری تحلیل متون^۸ است، به مدل‌سازی موضوعی کلیدواژه‌های ثبت شده در بروندهای پژوهشی بین‌المللی از نوع مقالات با ابزارهای متن‌کاوی پرداخته و با پردازش نتایج مدل‌سازی موضوعی در زمینه‌های سه‌گانه تخصصی فناوری اطلاعات (سه عنوان گروه‌های پژوهشی)، میدان تحقیقات منابع معتبر بین‌المللی را ارائه می‌دهیم.

۲-۲. روش پژوهش

فاز اول این طرح پژوهشی استخراج محدوده و مرزهای حوزه فناوری اطلاعات در قالب سه گروه پژوهشی است که با مصاحبه با خبرگان و نمایندگان سه گروه پژوهشی شکل می‌گیرد. پس از تکمیل سبد اولیه و با روش مطالعه میدانی از طریق مطالعات بین‌المللی، مهمترین وابستگی‌های مورد استناد سبد اولیه واژگان هر گروه پژوهشی مورد جستجو قرار گرفته و با استفاده از کلیدواژه‌های پژوهش‌های ایشان، کشف موضوعات هر گروه پژوهشی انجام شده و استخراج مدل نزدیکی به هر موضوع با مدل‌سازی موضوعی انجام می‌شود. مشکل اصلی بهره از مدل‌سازی موضوعی، محدود بودن استفاده از اسناد مشاهده شده (مثلاً ۵ سال مقالات در این پژوهش) برای به دست آوردن متغیرهای نهفته است. خروجی مدل‌سازی موضوعی مدل‌های احتمالی هستند که در آن داده‌ها یک نتیجه از یک فرایند آماری هستند که شامل متغیرهایی نهفته است. این فرایند توزیع احتمالی مشترک را در هر دو متغیر تصادفی مشاهده شده و نهفته تعریف می‌کند.

شمارنده توزیع شرطی توزیع مشترک تمام متغیرهای تصادفی است که می‌تواند به راحتی محاسبه شود و از لحاظ تئوری، می‌توان آن را با تجمیع توزیع مشترک تمامی ساختارهای نهفته، محاسبه کرد. عملاً، به دلیل اینکه تعداد ساختارهای موضوع احتمالاً زیاد است، این محاسبه دشوار و زمان‌بر است. الگوریتم‌های مدل‌سازی موضوعی به دو دسته تقسیم می‌شوند که توزیع‌های جایگزین دیگری را پیشنهاد می‌دهند تا تقریباً الگوریتم‌های مبتنی بر نمونه‌برداری را بهبود بخشند:

- گروه اول ارائه یک زنجیره مارکوف (یک توالی از متغیرهای تصادفی)،

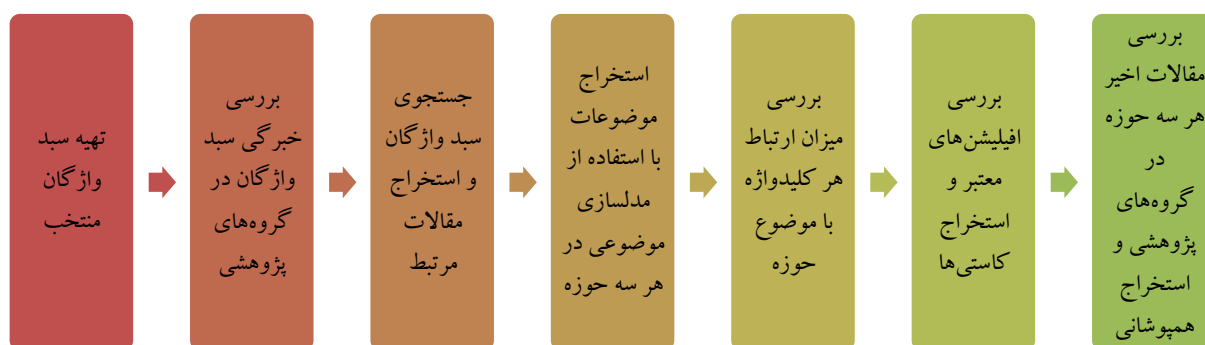
⁸ Text Mining

- گروه دوم، الگوریتم‌های مبتنی بر نمونه‌گیری (Steyvers, et al., 2004) (VEM).

با توجه به نوع داده‌های انتخاب شده (عنوان مقالات و/یا کلیدواژگان) در این طرح، الگوریتم VEM انتخاب خواهد شد، زیرا الگوریتم VEM برای هر مقاله، همپوشانی بین موضوعات مهم را ارایه می‌دهد و کشف همبستگی‌های موضوعی برای کلیدواژه‌های برگزیده هر گروه پژوهشی یکی از اهداف اصلی این طرح (کشف مرزهای حوزه فناوری اطلاعات) است.

در فاز نهایی تحلیل و مقایسه مطالعات و پژوهش‌های هر گروه به موضوعات آن گروه از روش‌های تحلیل آماری صورت گرفته و نتایج حاصل از فاز قبل، برای تمامی گروه‌های پژوهشی پژوهشکده فناوری اطلاعات بررسی و تحلیل می‌شود. این بررسی با دو روش صورت می‌گیرد. ابتدا براساس فاز قبل، دامنه اهمیت کلیدواژه‌های هر موضوع برای هر گروه پژوهشی پژوهشکده پدیدار می‌شود. اگر موضوعی نو در این میان مورد کشف قرار گرفته باشد، احتمالاً به آینده پژوهشی آن گروه پژوهشی کمک خواهد نمود. در فاز بعد سعی می‌شود با اندازه‌گیری میزان نزدیکی هر فعالیت پژوهشی گذشته در هر گروه پژوهشی، میزان انطباق هر پژوهش را با دسته‌بندی موضوعی کشف شده استخراج نمود. شاید این مساله راهنمایی برای پژوهشگران در تمرکز بر فعالیت‌های اصیل هر گروه پژوهشی باشد.

روش پژوهش جاری در فاز اول به مراحل زیر انجام خواهد پذیرفت:



شکل ۲-۳- مراحل پژوهش جاری

به طور خاص، چند مرحله متفاوت در این پژوهش پیاده سازی می‌شود:

- (۱) کلیدواژه‌های تخصصی هر گروه پژوهشی از خبرگان این حوزه استخراج می‌شود. این مطالعه میدانی از طریق یا بوسیله خبرگی یک نماینده از هر گروه پژوهشی انجام می‌پذیرد. این انتخاب تأثیری در نتیجه پژوهش نخواهد گذاشت زیرا سبد واژگان اولیه تشکیل شده توسط کلیدواژه‌های پژوهش‌های مطرح بین‌المللی تکمیل خواهند شد.
- (۲) به منظور مطالعه موضوعات فناوری اطلاعات در سطح بین‌المللی، موضوعات با تحلیل کلیدواژه‌های تمامی مقالات مهمترین وابستگی‌های پیشرو در سال اخیر شناسایی می‌شوند.

(۳) سبد واژگان این پژوهش از کلیدواژه‌های مقالات منتخب پایگاه‌های مورد پژوهش (مجلات بین‌المللی WOS) بند قبل استخراج و غنی می‌شود. این قسمت، جبرانی برای تعداد احتمالی اندک کلیدواژه‌های کشف شده از مطالعه میدانی خبرگان انجام خواهد بود. بدین معنی که کلیدواژه‌های اشاره شده توسط ایشان در مطالعات جمع‌آوری شده و کلیدواژه‌های دیگر کشف شده از آن مطالعات به سبد افزون می‌شود. در این فاز تمامی کلیدواژه‌های مطرح شده فارغ از سقف یا کف تعداد آنها در فاز بعد مدل‌سازی می‌شوند (در نظر گرفته می‌شوند). لازم به ذکر است سبد واژگان اولیه در مرحله ۱ بصورت جداگانه برای هر گروه پژوهشی استخراج شده و سبد واژگان غنی شده در این قسمت توسط مطالعات بین‌المللی کامل می‌شود. این سبد واژگان، مستقل از کلیدواژه‌های مقاله‌های همکاران در پژوهشکده فناوری اطلاعات است و در مراحل بعدی مقایسه بین این دو صورت خواهد پذیرفت.

(۴) به منظور مطالعه تکامل و روند پژوهش‌ها در حوزه‌های گروه‌های پژوهشی فناوری اطلاعات، مقالات استخراج شده با توجه به کلیدواژه‌ها، تحلیل موضوعی و محتوایی (مدل‌سازی موضوعی) می‌شوند.

(۵) پژوهش‌های هر گروه پژوهشی (کلیدواژه‌های انگلیسی ثبت شده) مجدداً مورد بررسی قرار می‌گیرند و میزان تطبیق و نزدیکی آنها با موضوعات استخراج شده بین‌المللی (کشف شده از طریق مدل‌سازی موضوعی) استخراج می‌شود.

(۶) مقایسه و تحلیل موضوعات و نقاط کمتر پرداخته شده پژوهشی برای هر گروه تحلیل می‌شود.

مدل پژوهش جاری در شکل زیر نمایش داده شده است:

داده

با بهره‌گیری از دایره‌ی کلیدواژگان
مستخرج از مجلات علمی

انتخاب داده

با استفاده از نظرات خبرگان و
انتخاب کلیدواژه‌های مختص هر
گروه پژوهشی

پالایش داده

جستجو در پایگاه، پاکسازی داده،
حذف تکرار و ...

آماده‌سازی داده

تبدیل داده به فرمت قابل تبدیل برای پردازش در نرم‌افزار
و تهیه رکوردهای مستقل
ایجاد ماتریس کلیدواژگان (عنوان سطر و ستون شامل
کلیدواژه‌های یافت شده و احتمال وجود هر کلیدواژه با

داده‌کاوی (مدل‌سازی موضوعی)

- تعیین تعداد تکرار واژه
- دسته‌بندی به ازای واژه‌های پر تکرار و خویشاوندی
- ارزیابی کیفیت دسته‌بندی
- تهیه دسته‌های مناسب و احتمال حضور هر کلیدواژه در دسته

تحلیل

- بررسی تاثیر کلیدواژه‌ها در موضوع
- بررسی مقالات جاری و تاثیر هر یک در هر گروه پژوهشی

شکل ۲-۴- مدل پژوهش جاری

به طور خاص، چند رویکرد متفاوت دیگر نیز در این پژوهش پیاده سازی می‌شود:

- (۱) به منظور مطالعه موضوعات فناوری اطلاعات در سطح بین‌المللی در حال حاضر، موضوعات با تحلیل کلیدواژه‌های تمامی مقالات چند مجله پیشرو در چند سال اخیر شناسایی می‌شوند.
- (۲) در این فاز سبد واژگان شکل می‌گیرد. این سبد شامل کلمات مطرح شده در عنوان مقالات و کلیدواژه‌های مقالات منتخب پایگاه‌های مورد پژوهش در زمینه فناوری اطلاعات است. در مرحله اولیه سبد واژگان شامل کلمات عمومی (و غیر مرتبط با فناوری اطلاعات مانند a , the , an , ...) نیز می‌باشد که باید پاکسازی گردند. تمام کلمات این سبد اولیه واژگان مجدداً باید مورد جستجو در پایگاه‌ها قرار گیرد تا سبد کاملتر و غنی‌تر و میزان کاربرد و روند استفاده هر واژه در سبد یافت شود.

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۱۵

همانگونه که ذکر شد، سبد واژگان اولیه (که نشان‌دهنده بخش بزرگی از حوزه مربوطه هستند و می‌توانند پژوهش‌های حوزه فناوری اطلاعات را بازتابی نمایند) در اختیار خبرگان قرار گرفته و واژگان اصیل منتخب می‌شوند.

(۳) به‌منظور مطالعه تکامل و روند پژوهش‌ها در حوزه فناوری اطلاعات، کلیدواژه‌ها با خبرگی به گروه‌های پژوهشی تخصیص داده می‌شوند.

(۴) به‌منظور مطالعه تکامل و روند پژوهش‌ها در حوزه فناوری اطلاعات، مقالات با کلیدواژه‌های فوق استخراج شده و با توجه به کلیدواژه‌ها تحلیل موضوعی و محتوایی (مدل‌سازی موضوعی) می‌شوند.

(۵) مقایسه و تحلیل حوزه و مطالعات گروه‌های پژوهشی صورت می‌پذیرد.

۳-۲. انتخاب مجلات معتبر

در فاز اول از برنامه پژوهش تشریح شده در بخش قبل، انتخاب مجلات معتبر برای استخراج اولیه سبد واژگان انجام می‌پذیرد. در این پژوهش استفاده از کلیدواژگان مورد استفاده در مجلات برتر در حوزه فناوری اطلاعات به‌عنوان روشی برای تشکیل سبد اولیه در نظر گرفته شد. شاخص‌های متعددی برای انتخاب این مجلات پیش روی محققین قرار دارد. یکی از انواع شاخص‌ها، استفاده از شاخص‌های فعالیت مجلات مانند اچ‌ایندکس، سابقه، میزان ضریب تاثیر و ... است. نوع دیگر انتخاب این مجلات، استفاده از دانش خبرگی متخصصین حوزه فناوری اطلاعات است. ما در این پژوهش از هر دو روش بهره‌گیری نموده‌ایم.

با توجه به میزان استناد به مجلات در حوزه‌های دسته‌بندی شده فناوری اطلاعات پایگاه www.scimagojr.com، مجلات دارای ضریب اچ بالاتر، استخراج و در اختیار خبرگان قرار گرفتند. میزان ضریب اچ این مجلات در این پایگاه مجموع ضریب اچ برای مقالات هر مجله در بازه ۲ ساله است.

مجلات معتبر در حوزه‌های فناوری اطلاعات در جدول زیر آمده است. لازم به ذکر است ردیف‌های دارای رنگ سبز مجلاتی است که در هر حوزه فناوری اطلاعات برای جستجوی پژوهش‌ها در گام بعد، توسط خبرگان انتخاب شده‌اند.

جدول ۱-۲- فهرست مجلات دارای ضریب تاثیر بالا در پایگاه WOS

تکرار	Total Cites (3years)	Total Docs. (3years)	H index	عنوان	رتبه در H	زمینه
4	11286	577	326	IEEE Transactions on Pattern Analysis and Machine Intelligence	1	AI
4	7777	1061	180	Pattern Recognition	2	
2	3440	726	173	Journal of Machine Learning Research	3	
2	4345	278	172	International Journal of Computer Vision	4	
2	4245	457	170	IEEE Transactions on Fuzzy Systems	5	
2	14186	2405	335	Bioinformatics	1	Computational theory and mathematics
4	11286	577	326	IEEE Transactions on Pattern Analysis and Machine Intelligence	2	
2	4245	457	170	IEEE Transactions on Fuzzy Systems	3	
1	2247	183	154	IEEE Transactions on Evolutionary Computation	4	
2	4124	691	148	IEEE Transactions on Knowledge and Data Engineering	5	
2	13081	1371	242	IEEE Transactions on Image Processing	1	Computer graphics and computer aided design
1	5251	734	173	ACM Transactions on Graphics	2	
1	2874	623	118	IEEE Transactions on Visualization and Computer Graphics	3	
1	3208	368	113	Medical Image Analysis	4	
1	1495	433	110	Journal of Mechanical Design - Transactions of the ASME	5	
2	12402	1106	213	IEEE Communications Magazine	1	Computer network and communication
1	8083	741	211	IEEE Journal on Selected Areas in Communications	2	
1	3446	707	157	IEEE/ACM Transactions on Networking	3	
1	905	192	152	Computer Communication Review	4	
1	14929	2388	146	IEEE Transactions on Vehicular Technology	5	
2	14186	2405	335	Bioinformatics	1	Computer science application
1	9981	1495	260	IEEE Transactions on Automatic Control	2	
2	5123	1391	249	IEEE Transactions on Information Theory	3	
1	25178	2627	236	IEEE Transactions on Industrial Electronics	4	
2	12402	1106	213	IEEE Communications Magazine	5	
1	10379	2088	226	European Journal of Operational Research	1	Computer science (miscellaneous)
1	2212	946	189	Communications of the ACM	2	
1	4208	548	149	Computers and Education	3	
1	3119	685	133	Computers and Operations Research	4	
1	3048	275	132	ACM Computing Surveys	5	
4	11286	577	326	IEEE Transactions on Pattern Analysis and Machine Intelligence	1	Computer vision and pattern
4	7777	1061	180	Pattern Recognition	2	

تکرار	Total Cites (3 years)	Total Docs. (3 years)	H index	عنوان	رتبه در H	زمینه
2	4345	278	172	International Journal of Computer Vision	3	recognition
1	1767	892	144	Journal of the Optical Society of America A: Optics and Image Science, and Vision	4	
2	3110	836	139	Pattern Recognition Letters	5	
1	3522	874	154	Computer Physics Communications	1	Hardware and architecture
1	4268	824	120	IEEE Transactions on Parallel and Distributed Systems	2	
1	561	155	117	Journal of the ACM	3	
1	2342	262	111	IEEE Network	4	
1	3496	772	110	IEEE Transactions on Computers	5	
1	13748	2250	137	Computers in Human Behavior	1	Human computer interaction
1	1245	401	119	Cyberpsychology, Behavior, and Social Networking	2	
1	4643	735	111	IEEE Transactions on Systems, Man, and Cybernetics: Systems	3	
1	1058	249	109	International Journal of Human Computer Studies	4	
1	1206	267	104	IEEE Transactions on Human-Machine Systems	5	
2	5123	1391	249	IEEE Transactions on Information Theory	1	Information systems
1	1750	148	195	MIS Quarterly: Management Information Systems	2	
2	4124	691	148	IEEE Transactions on Knowledge and Data Engineering	3	
1	955	151	145	Information Systems Research	4	
1	1650	239	142	Information and Management	5	
1	9099	1459	233	IEEE Transactions on Signal Processing	1	Signal processing
4	7777	1061	180	Pattern Recognition	2	
1	1969	282	155	IEEE Signal Processing Magazine	3	
2	3110	836	139	Pattern Recognition Letters	4	
1	7738	1271	134	Mechanical Systems and Signal Processing	5	
4	11286	577	326	IEEE Transactions on Pattern Analysis and Machine Intelligence	1	Software
2	13081	1371	242	IEEE Transactions on Image Processing	2	
1	5845	686	195	IEEE Transactions on Medical Imaging	3	
4	7777	1061	180	Pattern Recognition	4	
2	3440	726	173	Journal of Machine Learning Research	5	

ردیف تکرار نشان‌دهنده تعداد حضور هر یک از مجلات در بین ۵ مجله برتر (همین جدول) در هر حوزه است.

۲-۳-۱. جستجوی مجلات برتر در حوزه فناوری اطلاعات

بر اساس جدول ۲-۲، مجلات پر کاربرد در هر دسته انتخاب بر حسب شاخص اچ مرتب‌سازی شده و توسط محققین انتخاب شدند. این انتخاب به نحوی صورت گرفت که مجلاتی که در چند حوزه در فهرست ۵ مجله برتر قرار گرفته‌اند انتخاب گردند. همچنین همه حوزه‌ها حداقل یک مجله به‌عنوان نماینده برای جستجو داشته باشند. در نتیجه در مرحله اول یافتن مجلاتی که در ۵ مجله تاثیرگذار هر حوزه گنجانده و همچنین در زمینه‌های مختلف فعالیت دارند در اولویت انتخاب قرار گرفت. در جدول زیر فهرست مجلات ۹ گانه انتخاب شده و حوزه فعالیت آنها مشخص شد. لازم به ذکر است دیگر مشخصات این مجلات در جدول بخش قبل آمده است.

جدول ۲-۲- اطلاعات و حوزه پوشش مجلات منتخب

نام مجله	حوزه‌ی پوشش دهنده	ناشر	کشور	اولین انتشار
Bioinformatics	Computational theory and Mathematics	Elsevier Ltd.	United Kingdom	1985
	Computer Science Application			
Computer Physics Communications	Hardware and Architecture	Elsevier BV	Netherlands	1969
Computers in Human Behavior	Human Computer Interaction	Elsevier Ltd.	United Kingdom	1985
European Journal of Operational Research	Computer Science (miscellaneous)	Elsevier BV	Netherlands	1977
IEEE Communications Magazine	Computer Network and Communication	Institute of Electrical and Electronics Engineers	United States	1979
	Computer Science Application			
IEEE Transactions on Image Processing	Computer Graphics and Computer Aided Design	Institute of Electrical and Electronics Engineers	United States	1992
	Software			
IEEE Transactions on Information Theory	Computer Science Application	Institute of Electrical and Electronics Engineers	United States	1969
	Information Systems			
IEEE Transactions on Pattern Analysis and Machine Intelligence	AI	Institute of Electrical and Electronics Engineers	United States	1978
	Computational Theory and Mathematics			

نام مجله	حوزه‌ی پوشش دهنده	ناشر	کشور	اولین انتشار
	Computer Vision and Pattern Recognition	Engineers		
	Software			
Pattern Recognition	AI	Elsevier Ltd.	United Kingdom	1968
	Signal Processing			

با توجه به انتخاب مجموعه فوق و با کد زیر، تعداد ۸۴۵۱ رکورد شامل مشخصات، کلیدواژه، عنوان و چکیده در ۳ سال اخیر دریافت شدند.

OR SO=(BIOINFORMATICS) OR SO=(Computer Physics Communications)
 SO=(Computers in Human Behavior) OR SO=(European Journal of Operational Research)
 OR SO=(IEEE Communications Magazine) OR SO=(IEEE Transactions on Image Processing)
 OR SO=(IEEE Transactions on Information Theory) OR SO=(IEEE Transactions on Pattern Analysis and Machine Intelligence) OR
 SO=(Pattern Recognition)

بر این اساس ۱۷۴۰۳ کلیدواژه منحصر به فرد از مجموع ۲۶۲۶۸ کلیدواژه مجموعه مقالات استخراج شده شناسایی شده و واژگان بالای ۱۰ تکرار در جدول زیر آمده است:

جدول ۲-۳- کلمات پر کاربرد در مجلات دارای ضریب تاثیر بالا

ردیف	واژه	تکرار
۱	Deep Learning	171
۲	Convolutional Neural Network (CNN)	145
۳	Supply Chain Management	70
۴	Social Media	66
۵	Data Envelopment Analysis	62
۶	Game Theory	62
۷	Combinatorial Optimization	42
۸	Integer Programming	42
۹	Optimization	41
۱۰	Feature Extraction	40
۱۱	Health Services	39
۱۲	Machine Learning	38
۱۳	Scheduling	38
۱۴	Image Segmentation	37

ردیف	واژه	تکرار
۱۵	Routing	35
۱۶	Classification	34
۱۷	Sparse Representation	34
۱۸	Transportation	34
۱۹	Task Analysis	33
۲۰	Inventory	31
۲۱	Logistics	31
۲۲	Person Re-Identification	31
۲۳	Energy	29
۲۴	Visual Tracking	29
۲۵	Finance	28
۲۶	Clustering	26
۲۷	Heuristics	26
۲۸	Social Media	26
۲۹	Dynamic Programming	25
۳۰	Object Detection	25
۳۱	Facebook	24
۳۲	Dictionary Learning	22
۳۳	Feature Selection	22
۳۴	Semi-Supervised Learning	22
۳۵	Simulation	22
۳۶	Stochastic Programming	22
۳۷	Training	22
۳۸	Decision Analysis	21
۳۹	Metric Learning	21
۴۰	Neural Networks	21
۴۱	Visualization	21
۴۲	Image Classification	20
۴۳	Image Retrieval	20
۴۴	Robust Optimization	20
۴۵	Super-Resolution	20

۲-۳-۲. انتخاب مجلات با استفاده از نظر تکمیلی خبرگان

با توجه به روش پژوهش مشروح در بخش سوم این فصل، از ۳ خبره در زمینه فناوری اطلاعات برای معرفی مجلات معتبر و مورد نظر ایشان نیز بهره گرفته شد. براین اساس، علاوه بر مجلات انتخاب شده در بخش قبل، مجلات زیر نیز توسط خبرگان معرفی شدند. خبرگان همکار در این طرح پژوهشی شامل اعضای هیات علمی در پژوهشکده فناوری اطلاعات ایرانداک و از هر سه گروه پژوهشی بصورت منتخب انتخاب شدند. همچنین برای اطمینان از کفایت تعداد خبرگان، مقالات اخیر ایشان و همراهی آن با کلیدواژه‌های استخراج شده مورد بررسی قرار گرفت.

- IEEE Intelligent Systems
- IEEE Transactions on Information Theory
- IEEE Transactions on Information Forensics And Security
- ACM Transactions on Computer Systems
- IBM Journal of Research And Development
- Journal of Visual Languages And Computing
- Computational Linguistics
- Information Systems
- IEEE Transactions on Neural Networks And Learning Systems
- International Journal of Human-Computer Studies
- IEEE Transactions on Knowledge And Data Engineering
- Applied Intelligence
- AI Magazine

جستجوی مجلات فوق در بازه ۵ ساله با کد زیر در پایگاه WOS انجام پذیرفت:

so=(ieee intelligent systems) or so=(ieee transactions on information theory) or so=(ieee transactions on information forensics and security) or so=(acm transactions on computer systems) or so=(ibm journal of research and development) or so=(journal of visual languages and computing) or so=(computational linguistics) or so=(information systems) or so=(ieee transactions on neural networks and learning systems) or so=(international journal of human-computer studies) or so=(ieee transactions on knowledge and data engineering) or so=(applied intelligence) or so=(ai magazine)

با استفاده از مجلات معرفی شده توسط خبرگان در طول ۵ سال اخیر ۷۲۴۱ رکورد شامل مشخصات، کلیدواژه، عنوان و چکیده دریافت شدند. فهرست واژگان بالای ۲۰ تکرار جستجوی فوق در جدول زیر آمده است:

جدول ۲-۴- تکرار کلیدواژگان پر کاربرد در مجلات مورد تایید خبرگان

ردیف	واژه	تکرار
۱	Neural Network(s) (NN(s))	۱۶۵
۲	Machine Learning	۹۸
۳	Classification	۸۴

ردیف	واژه	تکرار
۴	Feature Selection	۸۰
۵	Deep Learning	۷۹
۶	Clustering	۷۲
۷	Data Mining	۶۸
۸	Compressed Sensing	۶۲
۹	Security	۶۰
۱۰	Optimization	۵۹
۱۱	Social Network (s)	۵۹
۱۲	Capacity	۵۸
۱۳	Physical Layer Security	۵۴
۱۴	Privacy	۵۲
۱۵	Network Coding	۴۷
۱۶	Reinforcement Learning (RL)	۴۷
۱۷	Cloud Computing	۴۵
۱۸	Channel Capacity	۴۲
۱۹	Feature Extraction	۳۶
۲۰	Online Learning	۳۶
۲۱	Synchronization	۳۴
۲۲	Adaptive Control	۳۲
۲۳	Game Theory	۳۲
۲۴	Mutual Information	۳۲
۲۵	Supervised Learning	۳۲
۲۶	Robustness	۳۱
۲۷	Authentication	۲۹
۲۸	Optimal Control	۲۹
۲۹	Stability	۲۹
۳۰	Big Data	۲۸
۳۱	Information Theory	۲۷
۳۲	Unsupervised Learning	۲۶
۳۳	Quantization	۲۴
۳۴	Recommender Systems	۲۴
۳۵	Source Coding	۲۴
۳۶	Crowdsourcing	۲۳
۳۷	Dynamic Programming	۲۳

ردیف	واژه	تکرار
۳۸	Error Exponent	۲۳
۳۹	Face Recognition	۲۳
۴۰	Side Information	۲۳
۴۱	Feedback	۲۲
۴۲	Hypothesis Testing	۲۲
۴۳	Interference Channel	۲۲
۴۴	Query Processing	۲۲
۴۵	Wiretap Channel	۲۲
۴۶	Interference Alignment	۲۱
۴۷	Mimo	۲۱
۴۸	Nonlinear Systems	۲۱
۴۹	Particle Swarm Optimization	۲۱
۵۰	Regularization	۲۱
۵۱	Biometrics	۲۰
۵۲	Convolutional Neural Network	۲۰
۵۳	Entropy	۲۰
۵۴	Image Classification	۲۰
۵۵	Polar Codes	۲۰

کشف این تکرار با توجه به کد نرم‌افزاری توسعه یافته در این طرح پژوهشی در نرم‌افزار R به دست آمده است و گام‌های زیر را انجام می‌دهد:

- مرتب‌سازی کلیدواژگان از نظر " " اضافی که ممکن است ایجاد شده باشد
- کوچک‌سازی همه حروف به منظور کشف یکسانی واژگان
- جداسازی کلیدواژه‌های یک مقاله که توسط "؛" در فایل خروجی پایگاه از هم تفکیک شده‌اند
- ایجاد سطر متفاوت برای هر کلیدواژه (ساخت ماتریس ۱ در n)
- شمارش تکرار هر کلیدواژه

۴-۲. انتخاب کلیدواژگان منتخب

با توجه به فهرست بالای واژگان پرتکرار در دو جستجوی مجلات پیشرو و مجلات منتخب خبرگان، یافتن سبد واژگان منتخب فناوری اطلاعات با قوانینی (به‌عنوان توصیه به خبرگان) در این گام مورد توجه قرار گرفت. قوانین زیر برای اجماع خبرگان در نظر گرفته شد:

- واژه مختص به فناوری اطلاعات باشد
- واژه شامل ابزار حل مساله نباشد
- واژه‌های علمی عمومی مانند مدل و ... در سبد واژگان قرار نگیرد
- تعداد واژه پرتکرار باشد (در این پژوهش ۱۰ تکرار)
- واژه شامل کاربرد فناوری اطلاعات در زمینه دیگر علمی نباشد

به منظور در اختیار گذاشتن تمامی واژگان یافت شده در دو جستجوی اخیر، پاکسازی واژگان نیز در دستور کار قرار گرفت. با توجه به نوع نوشتار برخی از واژه‌ها در حالت اختصار و ... در ابتدا واژگان بر اساس تعداد تکرار مرتب شده و سپس ۳۰۰ واژه اول (دارای تکرار بالا) در میان باقی کلیدواژگان جستجو و در صورت نوشتار با حالت اختصار و ... به تعداد تکرار آن واژه افزوده شد. بطور مثال استفاده از کلیدواژگان در مجلات مختلف ممکن است یکسان نباشد. بطور مثال واژه IOT و یا Internet of Things ممکن است در کلیدواژگان مقالات استفاده شوند در صورتی که به یک مفهوم اشاره دارند. در جدول زیر کلیدواژگان پرتکرار در اختیار خبرگان قرار گرفته و در لیست زیر کلیدواژگان منتخب آنها انتخاب شدند.

جدول ۲-۵- انتخاب کلیدواژه از دیدگاه خبرگان برای تشکیل سبد واژگان

ردیف	واژه	تایید خبره	ردیف	واژه	تایید خبره
1	Deep Learning	*	13	Clustering	
2	Machine Learning	*	14	Channel Capacity	
3	Supply Chain Management		15	Combinatorial Optimization	
4	Convolutional Neural Network		16	Integer Programming	
5	Social Media	*	17	Compressed Sensing	
6	Data Envelopment Analysis	*	18	Feature Extraction	*
7	Security	*	19	Health Services	
8	Optimization		20	Data Mining	*
9	Feature Selection	*	21	Neural Network	
10	Privacy	*	22	Scheduling	
11	Classification		23	Image Segmentation	*
12	Network Coding	*	24	Online Learning	*

ردیف	واژه	تایید خبره
61	Dictionary Learning	
62	Feedback	
63	Hypothesis Testing	
64	Interference Channel	*
65	Query Processing	*
66	Reinforcement Learning	
67	Simulation	
68	Stochastic Programming	
69	Training	
70	Wiretap Channel	
71	Cnn	
72	Decision Analysis	*
73	Interference Alignment	
74	Metric Learning	
75	Mimo	
76	Nonlinear Systems	
77	Particle Swarm Optimization	
78	Regularization	
79	Visualization	*
80	Biometrics	*
81	Entropy	
82	Image Classification	*
83	Image Retrieval	*
84	Polar Codes	
85	Robust Optimization	
86	Super-Resolution	
87	Broadcast Channel	*
88	Channel Coding	*
89	Dimensionality Reduction	
90	Distributed Storage	
91	Kernel Methods	*
92	Matrix Factorization	
93	Pattern Recognition	*
94	Regenerating Codes	
95	Segmentation	
96	Anomaly Detection	

ردیف	واژه	تایید خبره
25	Capacity	
26	Game Theory	
27	Routing	
28	Sparse Representation	
29	Synchronization	
30	Transportation	
31	Task Analysis	
32	Adaptive Control	*
33	Mutual Information	
34	Physical Layer Security	*
35	Supervised Learning	*
36	Inventory	
37	Logistics	
38	Person Re-Identification	*
39	Robustness	
40	Authentication	*
41	Energy	
42	Optimal Control	
43	Stability	
44	Visual Tracking	
45	Big Data	*
46	Finance	
47	Information Theory	*
48	Heuristics	
49	Dynamic Programming	
50	Object Detection	*
51	Cloud Computing	*
52	Facebook	
53	Quantization	
54	Recommender Systems	*
55	Social Network	*
56	Source Coding	*
57	Crowdsourcing	*
58	Error Exponent	*
59	Face Recognition	*
60	Side Information	*

ردیف	واژه	تایید خبره	ردیف	واژه	تایید خبره
132	State Estimation		97	Graph Theory	
133	Transfer Learning		98	Image Reconstruction	*
134	Analytics		99	Multiple Criteria Analysis	
135	Efficiency		100	Secrecy Capacity	
136	Global Optimization		101	Sparsity	
137	Joint Source-Channel Coding		102	Support Vector Machine	
138	Kernel		103	User Experience	*
139	Linear Codes		104	Artificial Intelligence	*
140	Marketing		105	Collaborative Filtering	
141	Reliability		106	Computational Modeling	*
142	Sampling		107	Convex Optimization	
143	Semantics	*	108	Degrees Of Freedom	
144	Semisupervised Learning		109	Genetic Algorithm	
145	Uncertainty		110	Gpu	
146	Column Generation		111	Image Restoration	*
147	Community Detection	*	112	Learning Systems	*
148	Cryptography	*	113	Multiple Objective Programming	
149	Cyberbullying	*	114	Multi-Task Learning	*
150	Differential Privacy		115	Pricing	
151	Image Denoising		116	Quantum Entanglement	
152	Information Theoretic Security	*	117	Recurrent Neural Network	
153	Linear Programming		118	Regression	
154	Mapreduce		119	Saliency Detection	*
155	Multiple Access Channel		120	Time Delay	
156	Network Security	*	121	Action Recognition	*
157	Physical-Layer Security		122	Adolescents	
158	Private Information Retrieval	*	123	Community Operational Research	
159	Representation Learning		124	Computer Vision	*
160	Semidefinite Programming		125	Convergence	
161	Trust		126	Ensemble Learning	*
162	Usability		127	Internet Of Things	*
163	Wireless Networks	*	128	List Decoding	*
164	5g Mobile Communication	*	129	Location	
165	Adaptive Dynamic Programming (Adp)		130	Renyi Divergence	
166	Capacity Region		131	Scalability	

ردیف	واژه	تایید خبره
203	Nonconvex Optimization	
204	Org	
205	Process Mining	*
206	Project Scheduling	
207	Reed-Solomon Codes	
208	Resource Allocation	
209	Second-Order Asymptotics	
210	Secrecy	
211	Sentiment Analysis	*
212	Spectral Clustering	*
213	Superposition Coding	*
214	Swarm Intelligence	
215	System Identification	*
216	Visual Analytics	
217	Wireless Sensor Networks	*
218	Zero-Error Capacity	
219	Artificial Neural Networks	
220	Boolean Functions	
221	Change Detection	
222	Computer Security	
223	Concept Drift	
224	Cooperation	
225	Correlation	
226	Cuda	
227	Data Models	*
228	Energy Harvesting	
229	Environment And Climate Change	
230	Error Correction Codes	*
231	Group Testing	
232	Homomorphic Encryption	*
233	Human-Computer Interaction	*
234	Iris Recognition	
235	Large Deviations	
236	Location Privacy	
237	Molecular Dynamics	

ردیف	واژه	تایید خبره
167	Channel State Information	
168	Computational Complexity	
169	Discrete Optimization	
170	Diversity	
171	Domain Adaptation	
172	Hashing	*
173	Image Forensics	
174	Index Coding	*
175	Information Entropy	*
176	Maintenance	
177	Manifold Learning	
178	Maritime Industry	
179	Monte Carlo	
180	Parallel Computing	*
181	Random Forest	
182	Self-Esteem	
183	Sparse Coding	*
184	Www	*
185	Attention	
186	Benders Decomposition	
187	Branch-And-Price	
188	Cloud Storage	*
189	Data Privacy	*
190	Deep Neural Network	
191	Differential Evolution	
192	Dimension Reduction	
193	Estimation	
194	Evolutionary Algorithm	
195	Forecasting	
196	Gaussian Mixture Model	
197	Generative Adversarial Networks	
198	Information Security	*
199	Iterative Decoding	*
200	Learning Analytics	
201	Mds Codes	*
202	Motivation	

ردیف	واژه	تایید خبره
270	Intrusion Detection	
271	Link Prediction	*
272	Locally Repairable Codes	
273	Loneliness	
274	Low-Density Parity-Check (Ldpc) Codes	
275	Metaheuristics	
276	Mixed Integer Programming	
277	Multi-View Learning	*
278	Natural Language Processing	
279	Natural Resources	
280	Networks	
281	Parameter Estimation	
282	Personalization	
283	Quantum Mechanics	
284	Random Coding	*
285	Redundancy	
286	Secrecy Outage Probability	
287	Social Support	
288	Spiking Neural Network (Snn)	
289	Stability Analysis	
290	Stochastic Geometry	
291	Strong Converse	
292	Subspace Clustering	
293	Subspace Codes	
294	Sustainability	
295	Text Mining	*
296	Time-Varying Delay	
297	Topic Modeling	*
298	Twitter	
299	Upper Bounds	
300	Verification	

ردیف	واژه	تایید خبره
238	Multi-Objective Optimization	
239	Ontology	
240	Pose Estimation	
241	Production	
242	Rate-Distortion Theory	
243	Remote Sensing	
244	Renyi Entropy	
245	Revenue Management	
246	Risk Management	
247	Salient Object Detection	
248	Shannon Theory	
249	Smart Grid	
250	Subspace Learning	
251	Video Games	
252	Virtual Reality	
253	Wireless Communication	
254	Admm	
255	Children	
256	Competition	
257	Compression	
258	Convolutional Neural Network (Cnn)	
259	Data Streams	
260	Decision Processes	
261	Decision Support Systems	
262	Dispersion	
263	Extreme Learning Machine (Elm)	
264	Gaussian Processes	
265	High-Dimensional Data	
266	Image Coding	
267	Image Fusion	
268	Image Recognition	*
269	Internet Addiction	

از مجموع ۳۰۰ واژه با تکرار مناسب، ۸۸ واژه توسط خبرگان برای قرارگیری در سبد انتخاب شدند.

۵-۲. تقسیم‌بندی کلیدواژه‌ها توسط خبرگان در گروه‌های پژوهشی

فهرست شامل ۸۸ کلیدواژه بالا به خبرگان در ۳ گروه پژوهشی ارائه شده و از ایشان درخواست گردید که کلیدواژه‌های مختص به گروه پژوهشی خود را انتخاب کنند. در هر سه گروه پژوهشی از سه خبره شامل یک عضو هیات علمی آن گروه پژوهشی، محققین این پژوهش و یک عضو هیات علمی خارج از موسسه استفاده شده و کلیدواژه‌های مورد اجماع هر کدام به عنوان منبع اصلی جستجوی کلیدواژگان انتخاب شدند. در جدول زیر می‌توان نتیجه انتخاب اجماع همه کلیدواژه‌ها را در هر سه گروه پژوهشی مشاهده کرد. با توجه به اینکه اجماع خبرگان (رای دادن همه ایشان به هر کلیدواژه منتخب) مورد توجه قرار گرفته است، ممکن است خطایی شامل از دست رفتن کلیدواژه‌ی خاص در جستجو ایجاد گردد که در بخش بعدی پژوهش و بررسی وابستگی‌های مرتبط بازمینی خواهد شد.

جدول ۲-۶- فهرست واژگان پایگاه WOS منتخب در هر گروه پژوهش

ردیف	واژه	گروه کسب و کار اینترنتی	گروه سیستم‌های اطلاعاتی	گروه مدیریت فناوری اطلاعات
۱	Deep Learning			
۲	Machine Learning			
۳	Social Media	*		
۴	Data Envelopment Analysis			
۵	Security			*
۶	Feature Selection			
۷	Privacy			*
۸	Network Coding			
۹	Feature Extraction			
۱۰	Data Mining	*		
۱۱	Image Segmentation			
۱۲	Online Learning	*		
۱۳	Adaptive Control			
۱۴	Physical Layer Security			*
۱۵	Supervised Learning		*	
۱۶	Person Re-Identification			
۱۷	Authentication			
۱۸	Big Data			
۱۹	Information Theory			*

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۳۰

ردیف	واژه	گروه کسب و کار اینترنتی	گروه سیستم‌های اطلاعاتی	گروه مدیریت فناوری اطلاعات
۲۰	Object Detection			
۲۱	Cloud Computing			*
۲۲	Recommender Systems		*	
۲۳	Social Network			
۲۴	Source Coding			
۲۵	Crowdsourcing		*	
۲۶	Error Exponent			
۲۷	Face Recognition			
۲۸	Side Information			
۲۹	Interference Channel			
۳۰	Query Processing		*	
۳۱	Decision Analysis	*		
۳۲	Visualization			
۳۳	Biometrics			
۳۴	Image Classification			
۳۵	Image Retrieval			
۳۶	Broadcast Channel			
۳۷	Channel Coding			
۳۸	Kernel Methods		*	
۳۹	Pattern Recognition			
۴۰	Image Reconstruction			
۴۱	User Experience	*		
۴۲	Artificial Intelligence		*	
۴۳	Computational Modeling			
۴۴	Image Restoration			
۴۵	Learning Systems			
۴۶	Multi-Task Learning			
۴۷	Saliency Detection			
۴۸	Action Recognition			
۴۹	Computer Vision			
۵۰	Ensemble Learning			
۵۱	Internet Of Things		*	
۵۲	List Decoding			

ردیف	واژه	گروه کسب و کار اینترنتی	گروه سیستم‌های اطلاعاتی	گروه مدیریت فناوری اطلاعات
۵۳	Semantics		*	
۵۴	Community Detection		*	
۵۵	Cryptography			
۵۶	Cyberbullying			
۵۷	Information Theoretic Security			*
۵۸	Network Security			
۵۹	Private Information Retrieval		*	
۶۰	Wireless Networks			*
۶۱	5g Mobile Communication			
۶۲	Hashing			
۶۳	Index Coding			
۶۴	Information Entropy		*	
۶۵	Parallel Computing			
۶۶	Sparse Coding			
۶۷	Www	*		
۶۸	Cloud Storage			*
۶۹	Data Privacy			*
۷۰	Information Security			*
۷۱	Iterative Decoding			
۷۲	Mds Codes			
۷۳	Process Mining		*	
۷۴	Sentiment Analysis			
۷۵	Spectral Clustering			
۷۶	Superposition Coding			
۷۷	System Identification			
۷۸	Wireless Sensor Networks			
۷۹	Data Models			
۸۰	Error Correction Codes			
۸۱	Homomorphic Encryption			
۸۲	Human-Computer Interaction			
۸۳	Image Recognition		*	

ردیف	واژه	گروه کسب و کار اینترنتی	گروه سیستم‌های اطلاعاتی	گروه مدیریت فناوری اطلاعات
۸۴	Link Prediction		*	
۸۵	Multi-View Learning		*	
۸۶	Random Coding			
۸۷	Text Mining		*	
۸۸	Topic Modeling			

کد دستوری زیر (مطابق شکل زیر) برای جستجوی هر کلیدواژه برای استخراج مفاهیم مرتبط با گروه پژوهشی کسب و کار الکترونیکی در پایگاه مورد استفاده قرار گرفت:

TS=(Www) OR TS=(User Experience) OR TS=(Social Media) OR TS=(Online Learning) OR TS=(Decision Analysis) OR TS=(Data Mining)

شکل ۲-۵- نمای جستجو در پایگاه WOS

همچنین سال انتشار مقالات برای مقایسه پذیر بودن ۲۰۲۱ انتخاب شدند. همچنین کدهای زیر برای استخراج مطالعات مرتبط با گروه پژوهشی سیستم‌های اطلاعاتی مورد استفاده قرار گرفت:

TS=(Text Mining) OR TS=(Supervised Learning) OR TS=(Semantics) OR TS=(Recommender Systems) OR TS=(Query Processing) OR TS=(Process Mining) OR TS=(Private Information Retrieval) OR TS=(Multi-View Learning) OR TS=(Link Prediction) OR TS=(Kernel Methods) OR TS=(Internet Of Things) OR TS=(Information Entropy) OR TS=(Image Recognition) OR TS=(Crowdsourcing) OR TS=(Community Detection) OR TS=(Artificial Intelligence)

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۳۳

و در نهایت کدهای زیر برای استخراج مطالعات مرتبط با گروه پژوهشی مدیریت فناوری اطلاعات مورد استفاده قرار گرفت:

TS=(Wireless Networks) OR TS=(Security) OR TS=(Privacy) OR TS=(Physical Layer Security) OR TS=(Information Theory) OR TS=(Information Theoretic Security) OR TS=(Information Security) OR TS=(Data Privacy) OR TS=(Cloud Storage) OR TS=(Cloud Computing)

از بین مقالات مورد جستجو در سال ۲۰۲۱ بیش از ۲۰۰ هزار مقاله بازیابی شدند (شکل زیر). از بین این مقالات و به ترتیب میزان نزدیکی به جستجو در هر زمینه (۳ گانه) ۱۰ هزار مقاله مورد استخراج قرار گرفت.

Search History:

Set	Results	Save History / Create Alert	Open Saved History	Edit Sets	Combine Sets AND OR Combine	Delete Sets Select All Delete
# 19	80,202	(TS=(Www) OR TS=(User Experience) OR TS=(Social Media) OR TS=(Online Learning) OR TS=(Decision Analysis) OR TS=(Data Mining)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI Timespan=Year to date		Edit	☐	☐
# 18	61,948	(TS=(Wireless Networks) OR TS=(Security) OR TS=(Privacy) OR TS=(Physical Layer Security) OR TS=(Information Theory) OR TS=(Information Theoretic Security) OR TS=(Information Security) OR TS=(Data Privacy) OR TS=(Cloud Storage) OR TS=(Cloud Computing)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI Timespan=Year to date		Edit	☐	☐
# 17	66,181	(TS=(Text Mining) OR TS=(Supervised Learning) OR TS=(Semantics) OR TS=(Recommender Systems) OR TS=(Query Processing) OR TS=(Process Mining) OR TS=(Private Information Retrieval) OR TS=(Multi-View Learning) OR TS=(Link Prediction) OR TS=(Kernel Methods) OR TS=(Internet Of Things) OR TS=(Information Entropy) OR TS=(Image Recognition) OR TS=(Crowdsourcing) OR TS=(Community Detection) OR TS=(Artificial Intelligence)) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI Timespan=Year to date		Edit	☐	☐

شکل ۲-۶- جستجوی کلیدواژه گروه‌های پژوهشی در پایگاه

داده‌های مسخرج از هر مقاله بسیار متنوع و وسیع و شامل تمام مشخصات مقاله است. در جدول زیر یک نمونه از اطلاعات مقالات مستخرج برای هر گروه پژوهشی آمده است:

جدول ۲-۷- نمونه اطلاعات مقالات هر گروه پژوهشی

گروه پژوهشی	مدیریت فناوری اطلاعات	کسب و کار الکترونیک	سیستم‌های اطلاعاتی
Publication Type	J	J	J
Authors	Nguyen, VL; Lin, PC; Cheng, BC; Hwang, RH; Lin, YD	Zhao, L; Zhang, J; Wang, QH; Chen, ZK	Zhao, L; Zhang, J; Wang, QH; Chen, ZK
Author Full Names	Van-Linh Nguyen; Lin, Po-Ching; Cheng, Bo-Chao; Hwang, Ren-Hung; Lin, Ying-Dar	Zhao, Liang; Zhang, Jie; Wang, Qiuhao; Chen, Zhikui	Zhao, Liang; Zhang, Jie; Wang, Qiuhao; Chen, Zhikui
Article Title	Security and Privacy for 6G: A Survey on Prospective Technologies and Challenges	Dual Alignment Self-Supervised Incomplete Multi-View Subspace Clustering Network	Dual Alignment Self-Supervised Incomplete Multi-View Subspace Clustering Network
Source Title	IEEE COMMUNICATIONS SURVEYS AND TUTORIALS	IEEE SIGNAL PROCESSING LETTERS	IEEE SIGNAL PROCESSING LETTERS
Language	English	English	English
Document Type	Article	Article	Article
Author	Security; 6G mobile	Data models; Manifolds;	Data models; Manifolds;

گروه پژوهشی	مدیریت فناوری اطلاعات	کسب و کار الکترونیک	سیستم‌های اطلاعاتی
Keywords	communication; Privacy; 5G mobile communication; Computer architecture; Systematics; Standards; 6G; security and privacy; AI security; physical layer security; connection security; service security	Kernel; Decoding; Clustering algorithms; Software; Semantics; Incomplete multi-view clustering; dual alignment constraints; self-expression; auto-encoders	Kernel; Decoding; Clustering algorithms; Software; Semantics; Incomplete multi-view clustering; dual alignment constraints; self-expression; auto-encoders
Keywords Plus	Physical-Layer Security; Nonorthogonal Multiple-Access; Visible-Light Communication; Pilot Contamination Attack; Secret Key Generation; Wireless Communications; Comprehensive Survey; Intrusion Detection; Tactile Internet; Spoofing Attack		
Abstract	Sixth-generation (6G) mobile networks will have to cope with diverse threats on a space-air-ground integrated network environment, novel technologies, and an accessible user information explosion. However, for now, security and privacy issues for 6G remain largely in concept. This survey provides a systematic overview of security and privacy issues based on prospective technologies for 6G in the physical, connection, and service layers, as well as through lessons learned from the failures of existing security architectures and state-of-the-art defenses. Two key lessons learned are as follows. First, other than inheriting vulnerabilities from the previous generations, 6G has new threat vectors from new radio technologies, such as the exposed location of radio stripes in ultra-massive MIMO systems at Terahertz bands and attacks against pervasive intelligence. Second, physical layer protection,	Incomplete multi-view clustering has attracted much attention in decade years. To date, most of the remarkable achievements, however, exploit shallow models to learn shared feature representations based on incomplete views. Although some deep learning methods have been proposed to solve this issue, the existing ones still have the following problems: 1) The consistency between views is ignored, which will have serious negative impacts on incomplete multi-view learning. 2) The learned features do not have sufficient cluster-friendliness, that is, the tightness within clusters and the repulsiveness between clusters are not fully considered. To tackle the above shortcomings, we propose a Dual Alignment Self-supervised Incomplete	Incomplete multi-view clustering has attracted much attention in decade years. To date, most of the remarkable achievements, however, exploit shallow models to learn shared feature representations based on incomplete views. Although some deep learning methods have been proposed to solve this issue, the existing ones still have the following problems: 1) The consistency between views is ignored, which will have serious negative impacts on incomplete multi-view learning. 2) The learned features do not have sufficient cluster-friendliness, that is, the tightness within clusters and the repulsiveness between clusters are not fully considered. To tackle the above shortcomings, we propose a Dual Alignment Self-supervised Incomplete

گروه پژوهشی	مدیریت فناوری اطلاعات	کسب و کار الکترونیک	سیستم‌های اطلاعاتی
	deep network slicing, quantum-safe communications, artificial intelligence (AI) security, platform-agnostic security, real-time adaptive security, and novel data protection mechanisms such as distributed ledgers and differential privacy are the top promising techniques to mitigate the attack magnitude and personal data breaches substantially.	paper. Specifically, the manifold alignment constraint and consistency alignment constraint are integrated with the autoencoder to preserve the compact inherent local structure within the view and the consistency semantics between incomplete views, respectively. Moreover, a self-expression layer coupled with a spectral clustering module is designed to naturally separate different types of data, leveraging the current clustering results to supervise subspace learning, which excludes inter-cluster. Experimental results on several datasets show that our algorithm outperforms all compared state-of-the-arts.	Multi-view Subspace Clustering network (DASIMSC) in this paper. Specifically, the manifold alignment constraint and consistency alignment constraint are integrated with the autoencoder to preserve the compact inherent local structure within the view and the consistency semantics between incomplete views, respectively. Moreover, a self-expression layer coupled with a spectral clustering module is designed to naturally separate different types of data, leveraging the current clustering results to supervise subspace learning, which excludes inter-cluster. Experimental results on several datasets show that our algorithm outperforms all compared state-of-the-arts.
Addresses	[Van-Linh Nguyen; Lin, Po-Ching; Hwang, Ren-Hung] Natl Chung Cheng Univ, Dept Comp Sci & Informat Engn, Minhsung 621301, Chiayi, Taiwan; [Van-Linh Nguyen] Thai Nguyen Univ Informat & Commun Technol, Dept Informat Technol, Thai Nguyen 25000, Vietnam; [Cheng, Bo-Chao] Natl Chung Cheng Univ, Dept Commun Engn, Minhsung 621301, Chiayi, Taiwan; [Hwang, Ren-Hung] Natl Chung Cheng Univ, Adv Inst Mfg High Tech Innovat, Minhsung 621301, Chiayi, Taiwan;	[Zhao, Liang; Zhang, Jie; Wang, Qiuhaio; Chen, Zhikui] Dalian Univ Technol, Sch Software Technol, Dalian 116620, Peoples R China; [Zhao, Liang; Zhang, Jie; Wang, Qiuhaio; Chen, Zhikui] Dalian Univ Technol, Key Lab Ubiquitous Network & Serv Software Liaoni, Dalian 116620, Peoples R China	[Zhao, Liang; Zhang, Jie; Wang, Qiuhaio; Chen, Zhikui] Dalian Univ Technol, Sch Software Technol, Dalian 116620, Peoples R China; [Zhao, Liang; Zhang, Jie; Wang, Qiuhaio; Chen, Zhikui] Dalian Univ Technol, Key Lab Ubiquitous Network & Serv Software Liaoni, Dalian 116620, Peoples R China
Reprint Addresses	Hwang, RH (corresponding author), Natl Chung Cheng	Zhao, L (corresponding author), Dalian Univ	Zhao, L (corresponding author), Dalian Univ

گروه پژوهشی	مدیریت فناوری اطلاعات	کسب و کار الکترونیک	سیستم‌های اطلاعاتی
	Univ, Dept Comp Sci & Informat Engn, Minhsung 621301, Chiayi, Taiwan.; Hwang, RH (corresponding author), Natl Chung Cheng Univ, Adv Inst Mfg High Tech Innovat, Minhsung 621301, Chiayi, Taiwan.	Technol, Sch Software Technol, Dalian 116620, Peoples R China.	Technol, Sch Software Technol, Dalian 116620, Peoples R China.
<i>Email Addresses</i>	nvlinh@cs.ccu.edu.tw; pclin@cs.ccu.edu.tw; bcheng@ccu.edu.tw; rhhwang@cs.ccu.edu.tw; ydlin@cs.nctu.edu.tw	liangzhao@dlut.edu.cn; zhangj812@mail.dlut.edu.cn; MeanderW@163.com; zkchen@dlut.edu.cn	liangzhao@dlut.edu.cn; zhangj812@mail.dlut.edu.cn; MeanderW@163.com; zkchen@dlut.edu.cn
<i>ORCIDs</i>	Lin, Ying-Dar/0000-0002-5226-4396		
<i>Funding Orgs</i>	Ministry of Science and Technology (MOST) of Taiwan; Ministry of Science and Technology, Taiwan [110-2811-E-194-501-MY2, 108-2221-E-194-022-MY3, 108-2221-E-194-019-MY3]; Advanced Institute of Manufacturing with High-Tech Innovations (AIM-HI) through the Featured Areas Research Center Program by the Ministry of Education (MOE) in Taiwan; Ministry of Education and Training (MOET) of Vietnam [B2021-TNA-02]; Thai Nguyen University [B2021-TNA-02]	National Natural Science Foundation of China; National Natural Science Foundation of China (NSFC) [61906030]; Fundamental Research Funds for the Central Universities; Fundamental Research Funds for the Central Universities [DUT20RC(4)009]; Natural Science Foundation of Liaoning Province; Natural Science Foundation of Liaoning Province [2020-BS-063]	National Natural Science Foundation of China; National Natural Science Foundation of China (NSFC) [61906030]; Fundamental Research Funds for the Central Universities; Fundamental Research Funds for the Central Universities [DUT20RC(4)009]; Natural Science Foundation of Liaoning Province; Natural Science Foundation of Liaoning Province [2020-BS-063]
<i>Funding Text</i>	This work was supported in part by the Ministry of Science and Technology (MOST) of Taiwan under Grant 110-2811-E-194-501-MY2, Grant 108-2221-E-194-022-MY3, and Grant 108-2221-E-194-019-MY3; in part by the Advanced Institute of Manufacturing with High-Tech Innovations (AIM-HI) through the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan; and in part by the Ministry of Education and Training (MOET) of Vietnam and	This work was supported in part by the National Natural Science Foundation of China under Grant 61906030, in part by the Fundamental Research Funds for the Central Universities under Grant DUT20RC(4)009, and in part by the Natural Science Foundation of Liaoning Province under Grant 2020-BS-063.	This work was supported in part by the National Natural Science Foundation of China under Grant 61906030, in part by the Fundamental Research Funds for the Central Universities under Grant DUT20RC(4)009, and in part by the Natural Science Foundation of Liaoning Province under Grant 2020-BS-063.

گروه پژوهشی	مدیریت فناوری اطلاعات	کسب و کار الکترونیک	سیستم‌های اطلاعاتی
	Thai Nguyen University under Grant B2021-TNA-02.		
Cited Reference Count	251	30	30
Times Cited, WoS	1	0	0
Core Times Cited, All Databases	1	0	0
180 Day Usage Count	4	2	2
Since 2013 Usage Count	4	2	2
Publisher	IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC	IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC	IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC
Publisher City	PISCATAWAY	PISCATAWAY	PISCATAWAY
Publisher Address	445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA	445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA	445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA
ISSN	1553-877X	1558-2361	1558-2361
eISSN	1553-877X	1558-2361	1558-2361
Journal Abbreviati on	IEEE COMMUN SURV TUT	IEEE SIGNAL PROC LET	IEEE SIGNAL PROC LET
Journal ISO Abbreviati on	IEEE Commun. Surv. Tutor.	IEEE Signal Process. Lett.	IEEE Signal Process. Lett.
Publicatio n Year	2021	2021	2021
Volume	23	28	28
Issue	4		
Start Page	2384	2122	2122
End Page	2428	2126	2126
DOI	10.1109/COMST.2021.310 8618	10.1109/LSP.2021.31203 11	10.1109/LSP.2021.31203 11
Number of Pages	45	5	5
WoS Categories	Computer Science, Information Systems; Telecommunications	Engineering, Electrical & Electronic	Engineering, Electrical & Electronic
Research Areas	Computer Science; Telecommunications	Engineering	Engineering
IDS Number	XE7SH	WR0LI	WR0LI

گروه پژوهشی	مدیریت فناوری اطلاعات	کسب و کار الکترونیک	سیستم‌های اطلاعاتی
UT (Unique WOS ID) Open Access Designations Date of Export	WOS:000723582700013	WOS:000714200500011	WOS:000714200500011
	Green Submitted		
	2021-12-25	2021-12-25	2021-12-25

داده‌های مستخرج شده از هر مقاله در هر سه گروه پژوهشی (نمونه کلیدواژگان یافت شده در مقالات یافت شده) در جدول زیر آمده است:

جدول ۲-۸- نمونه کلیدواژگان مقالات یافت شده در جستجوی نهایی

ردیف	گروه پژوهشی	کلیدواژه مقالات
۱	مدیریت فناوری اطلاعات	Security; 6G Mobile Communication; Privacy; 5G Mobile Communication; Computer Architecture; Systematics; Standards; 6G; Security And Privacy; AI Security; Physical Layer Security; Connection Security; Service Security
۲		Cloud Computing; Authorization; Data Privacy; Indexes; Secure Storage; Servers; Authorized Access; Cloud Storage; Integrity Auditing; Sensitive Information Hiding
۳		Emperor-Imperial Technique; HS-DRT (High-Security Distribution And Rake Technology); Mobile Cloud Computing (MCC); Quantum Key Fibo Privacy Approach (QKFPA); Trefoil Congruity Framework
۴		Access Control; Cloud Computing; Data Privacy; Encryption; Medical Services; Cryptography-Based Access Control; Data Security And Privacy; Attribute-Based Encryption; Secure Deduplication
۵		Security; Cloud Computing; Games; Servers; Forgery; Computer Science; Data Privacy; Cloud Computing; Provable Data Possession; Storage-As-A-Service; Privacy Preserving; Security
۶		Bloc Kchain&Nbsp; Privacy&Nbsp; Protection; Smart Contract; Public Encryption&Nbsp; Keyword Search
۷		Pk-Anonymization; Cloud Computing; Data Security; Key Generation; Map Reduce; Secure Data Storage
۸		Privacy Preservation; Message Encoding; Encryption; Decryption; Homomorphic; Message Decoding
۹	کسب و کار الکترونیک	Data models; Manifolds; Kernel; Decoding; Clustering algorithms; Software; Semantics; Incomplete multi-view clustering; dual alignment constraints; self-expression; auto-encoders
۱۰		Data classification; dynamic decay adjustment; machine learning; multilayer perceptron; radial basis function; UCI machine learning repository
۱۱		Multi-view representation learning; Weighted deep metric learning; Complementarity; Label propagation
۱۲		Redundancy; Feature extraction; Measurement; Machine learning; Taxonomy; Semisupervised learning; Semi-supervised multi-view deep representation learning; consensus and complementarity; redundancy; adversarial training; Siamese network; density clustering

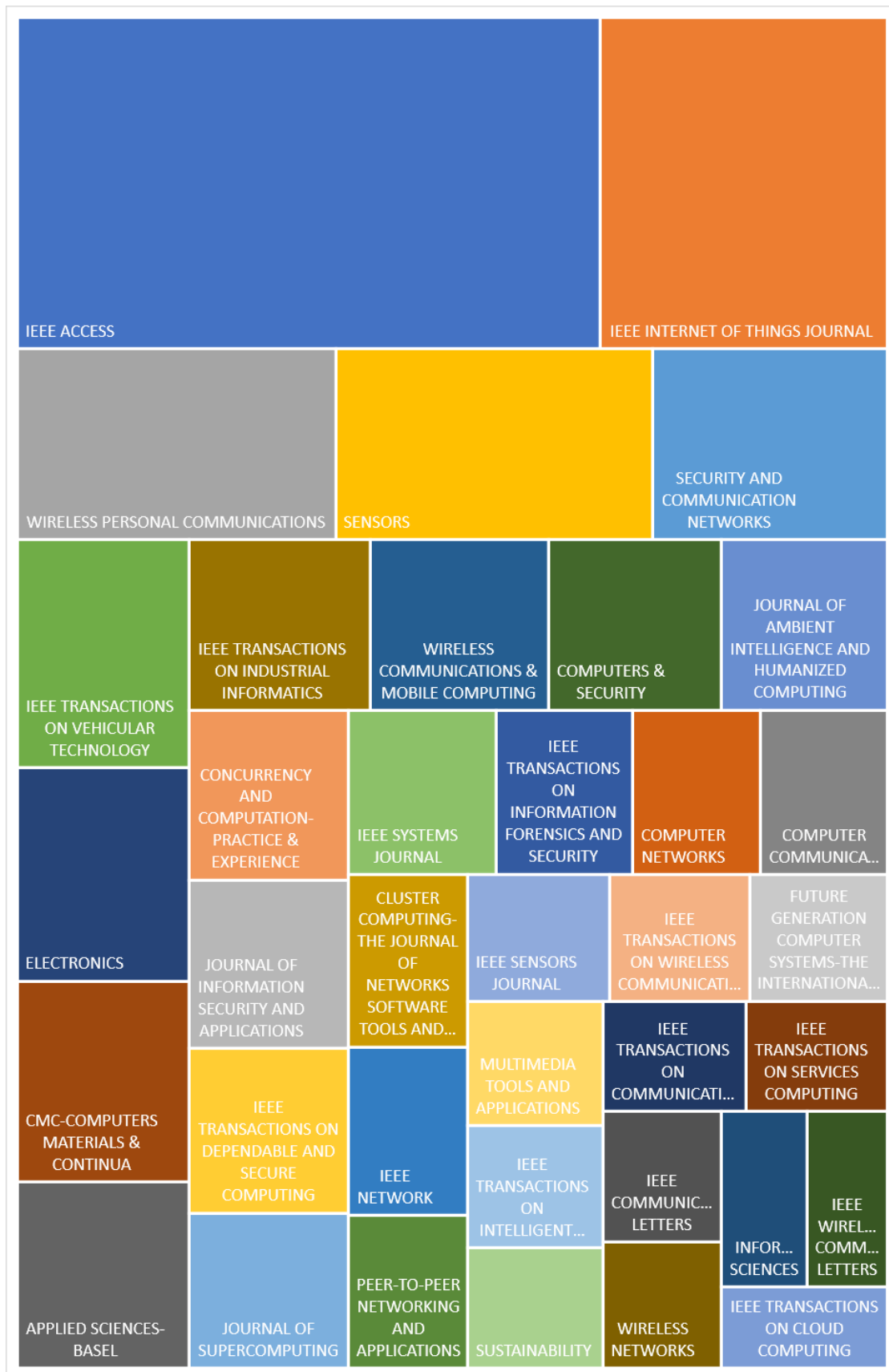
ردیف	گروه پژوهشی	کلیدواژه مقالات
۱۳		Weighted parameter; Multi-view learning; Common subspace learning; Supervised learning
۱۴		Multi-view graphs; Label propagation; Semi-supervised classification; Optimization; Redundancy
۱۵		Image retrieval; Training; Semantics; Feature extraction; Kernel; Adaptation models; Binary codes; Social image retrieval; partial multi-view hashing; discrete optimization; adaptive weights
۱۶		Drug-target interactions; Bipartite network; Multi-view learning; Link prediction; Graph regularized model
۱۷	سیستم‌های اطلاعاتی	Data models; Manifolds; Kernel; Decoding; Clustering algorithms; Software; Semantics; Incomplete multi-view clustering; dual alignment constraints; self-expression; auto-encoders
۱۸		Data classification; dynamic decay adjustment; machine learning; multilayer perceptron; radial basis function; UCI machine learning repository
۱۹		Multi-view representation learning; Weighted deep metric learning; Complementarity; Label propagation
۲۰		Redundancy; Feature extraction; Measurement; Machine learning; Taxonomy; Semisupervised learning; Semi-supervised multi-view deep representation learning; consensus and complementarity; redundancy; adversarial training; Siamese network; density clustering
۲۱		Weighted parameter; Multi-view learning; Common subspace learning; Supervised learning
۲۲		Multi-view graphs; Label propagation; Semi-supervised classification; Optimization; Redundancy
۲۳		Image retrieval; Training; Semantics; Feature extraction; Kernel; Adaptation models; Binary codes; Social image retrieval; partial multi-view hashing; discrete optimization; adaptive weights
۲۴		Drug-target interactions; Bipartite network; Multi-view learning; Link prediction; Graph regularized model

۶-۲. تحلیل حوزه‌ها با بهره‌گیری از داده‌های جستجو

با توجه به جستجوی بیش از ۲۰۰ هزار مدرک علمی در این پژوهش، تصمیم بر آن شد که بر اساس ارتباط پژوهش‌ها با هر حوزه جستجو ۱۰ هزار سند مرتبط انتخاب شوند. این انتخاب می‌تواند بر اساس سال انتشار، تعداد کلمات موجود در عنوان در جستجو، میزان ارجاع و شباهت باشد. معیار شباهت بر اساس تعداد کلمات مورد جستجو در عنوان و کلیدواژه با وزن بیشتر و در چکیده و کلیدواژه‌های تکمیلی با وزن کمتر است (Web of Science 2020).

۶-۲-۱. بررسی چگونگی کاربرد کلیدواژگان در مجلات بالای نیم درصد

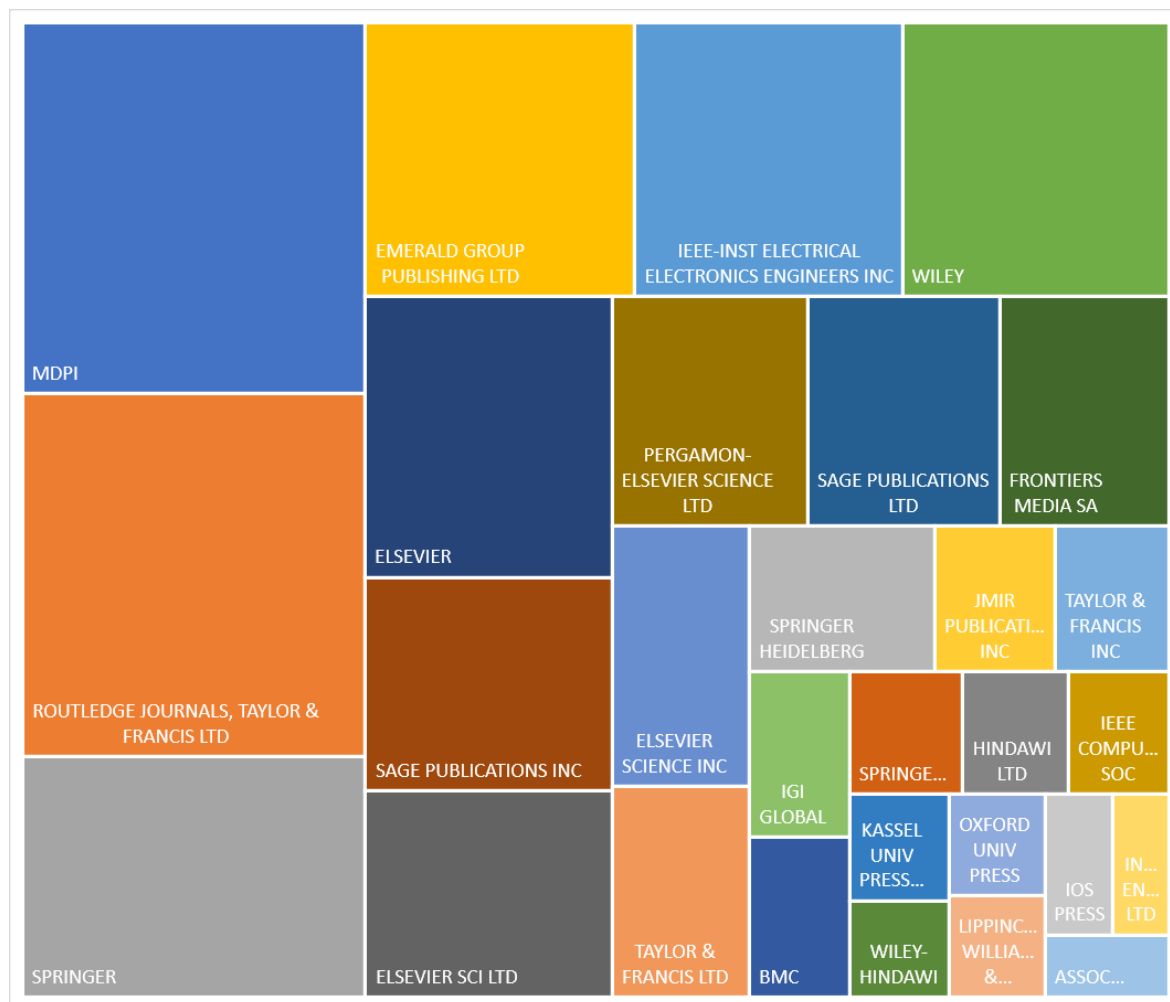
در جستجوهای انجام پذیرفته، در زمینه اطلاعات مرتبط با گروه مدیریت فناوری اطلاعات مجلات زیر بیش از نیم درصد نمونه مورد بازبینی را به خود اختصاص دادند:



شکل ۲-۷- کاربرد کلیدواژگان در مجلات بالای نیم درصد

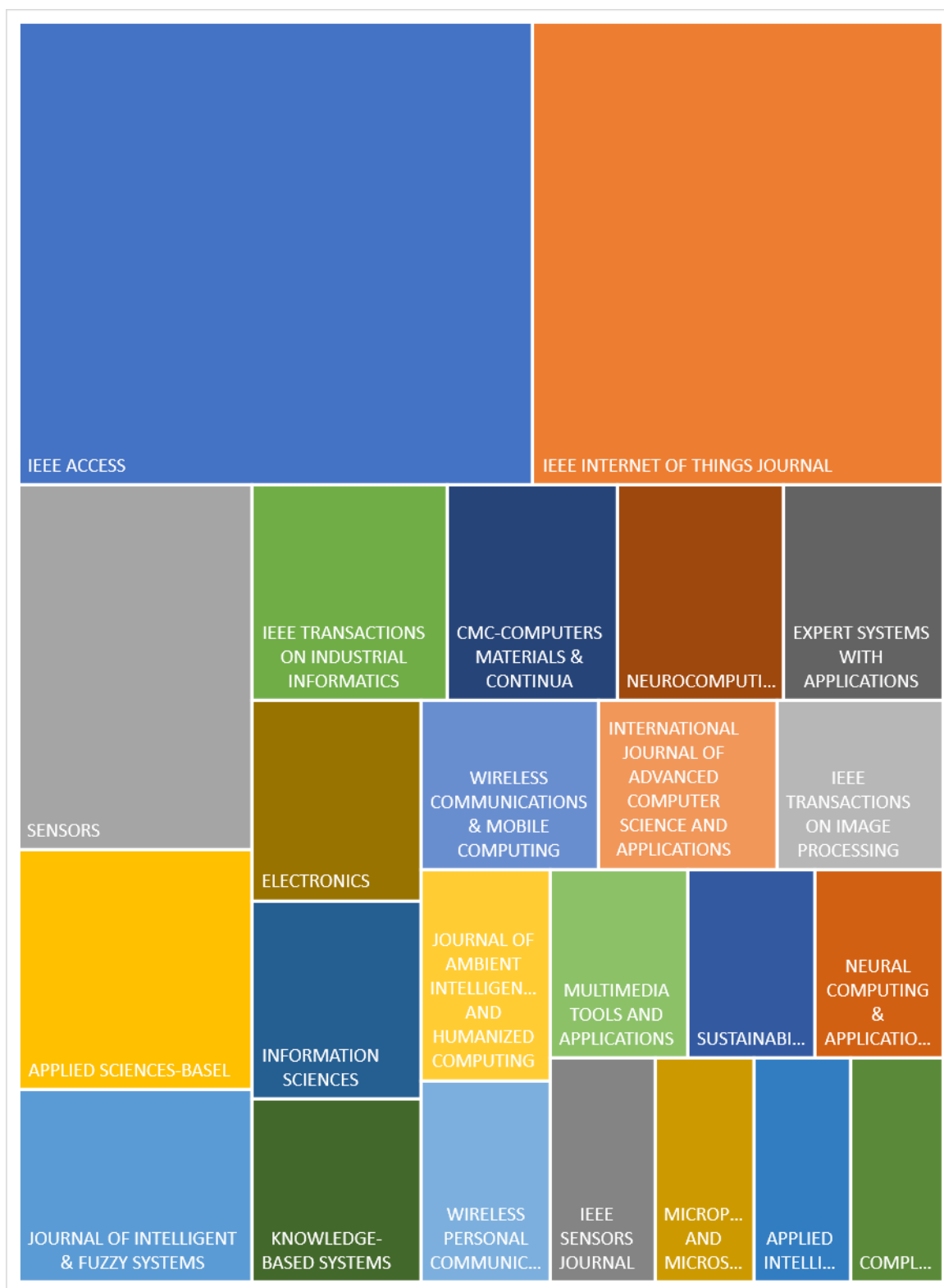
مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۴۱

همچنین مجلات فعال در زمینه کسب‌وکار الکترونیکی به شرح زیر (بالای نیم درصد نتایج نمونه) بودند:



شکل ۲-۸- ناشرین فعال در زمینه کسب‌وکار الکترونیکی

و نمودار مشابه برای مجلات مرتبط با سیستم‌های اطلاعاتی در زیر آمده‌اند:



شکل ۲-۹- مجلات فعال در زمینه سیستم‌های اطلاعاتی

با توجه به سه جدول فوق می‌توان مجلات مطرح در هر سه زمینه گروه‌های پژوهشی را مشخص نمود:

جدول ۲-۹- مجلات مطرح در هر سه زمینه فعالیت گروه‌های پژوهشی

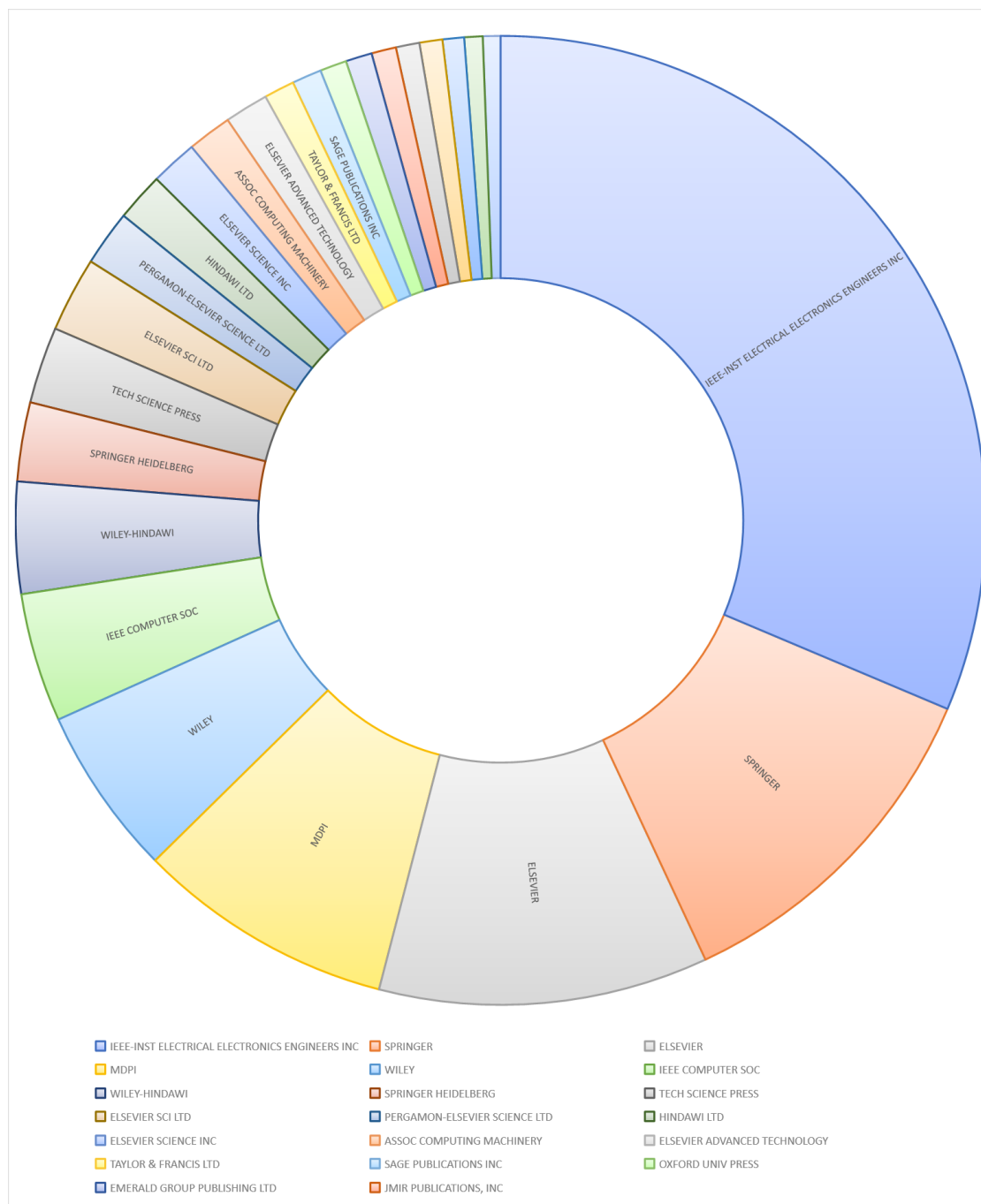
مجموع	سیستم‌های اطلاعاتی	کسب و کار الکترونیک	مدیریت فناوری اطلاعات	مجله / تعداد در نمونه گروه پژوهشی
1519	195	738	586	IEEE ACCESS
831		363	468	IEEE INTERNET OF THINGS JOURNAL
440		230	210	SENSORS
295		231	64	WIRELESS PERSONAL COMMUNICATIONS
260		122	138	APPLIED SCIENCES-BASEL
223		140	83	ELECTRONICS
222		132	90	CMC-COMPUTERS MATERIALS & CONTINUA
221		118	103	IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS
200	200			SUSTAINABILITY-BASEL
190		116	74	WIRELESS COMMUNICATIONS & MOBILE COMPUTING
175		108	67	JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING
170		170		SECURITY AND COMMUNICATION NETWORKS
150		150		IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY
139		57	82	INFORMATION SCIENCES
136	136			FRONT PSYCHOL
128		64	64	MULTIMEDIA TOOLS AND APPLICATIONS
127		69	58	IEEE SENSORS JOURNAL
121		62	59	SUSTAINABILITY
113		113		COMPUTERS & SECURITY
111			111	JOURNAL OF INTELLIGENT & FUZZY SYSTEMS
105	105			INT J ENV RES PUB HE
103		103		CONCURRENCY AND COMPUTATION- PRACTICE & EXPERIENCE
102		102		JOURNAL OF INFORMATION SECURITY AND APPLICATIONS
100		100		IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING
94		94		JOURNAL OF SUPERCOMPUTING
92		92		IEEE SYSTEMS JOURNAL
88			88	NEUROCOMPUTING
85		85		IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY
85	85			APPL SCI-BASEL
84			84	EXPERT SYSTEMS WITH APPLICATIONS

مجموع	سیستم‌های اطلاعاتی	کسب و کار الکترونیک	مدیریت فناوری اطلاعات	مجله / تعداد در نمونه گروه پژوهشی
80		80		COMPUTER NETWORKS
79		79		COMPUTER COMMUNICATIONS
78		78		CLUSTER COMPUTING-THE JOURNAL OF NETWORKS SOFTWARE TOOLS AND APPLICATIONS
77	77			EDUC INF TECHNOL
76			76	KNOWLEDGE-BASED SYSTEMS
76		76		IEEE NETWORK
75	75			J MED INTERNET RES
74			74	INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS
69			69	IEEE TRANSACTIONS ON IMAGE PROCESSING
69		69		PEER-TO-PEER NETWORKING AND APPLICATIONS
68		68		IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS
68	68			INT J EMERG TECHNOL
67		67		FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE
66	66			COMPUT HUM BEHAV
63		63		IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS
60		60		IEEE TRANSACTIONS ON COMMUNICATIONS
59			59	NEURAL COMPUTING & APPLICATIONS
59		59		IEEE TRANSACTIONS ON SERVICES COMPUTING
59		59		IEEE COMMUNICATIONS LETTERS
57		57		WIRELESS NETWORKS
57	57			EDUC SCI
55			55	MICROPROCESSORS AND MICROSYSTEMS
54		54		IEEE WIRELESS COMMUNICATIONS LETTERS
54	54			NEW MEDIA SOC
53			53	APPLIED INTELLIGENCE
52	52			INT J ADV COMPUT SC
51			51	COMPLEXITY
51		51		IEEE TRANSACTIONS ON CLOUD COMPUTING

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۴۵

۲-۶-۲. بررسی ناشران فعال در سه حوزه

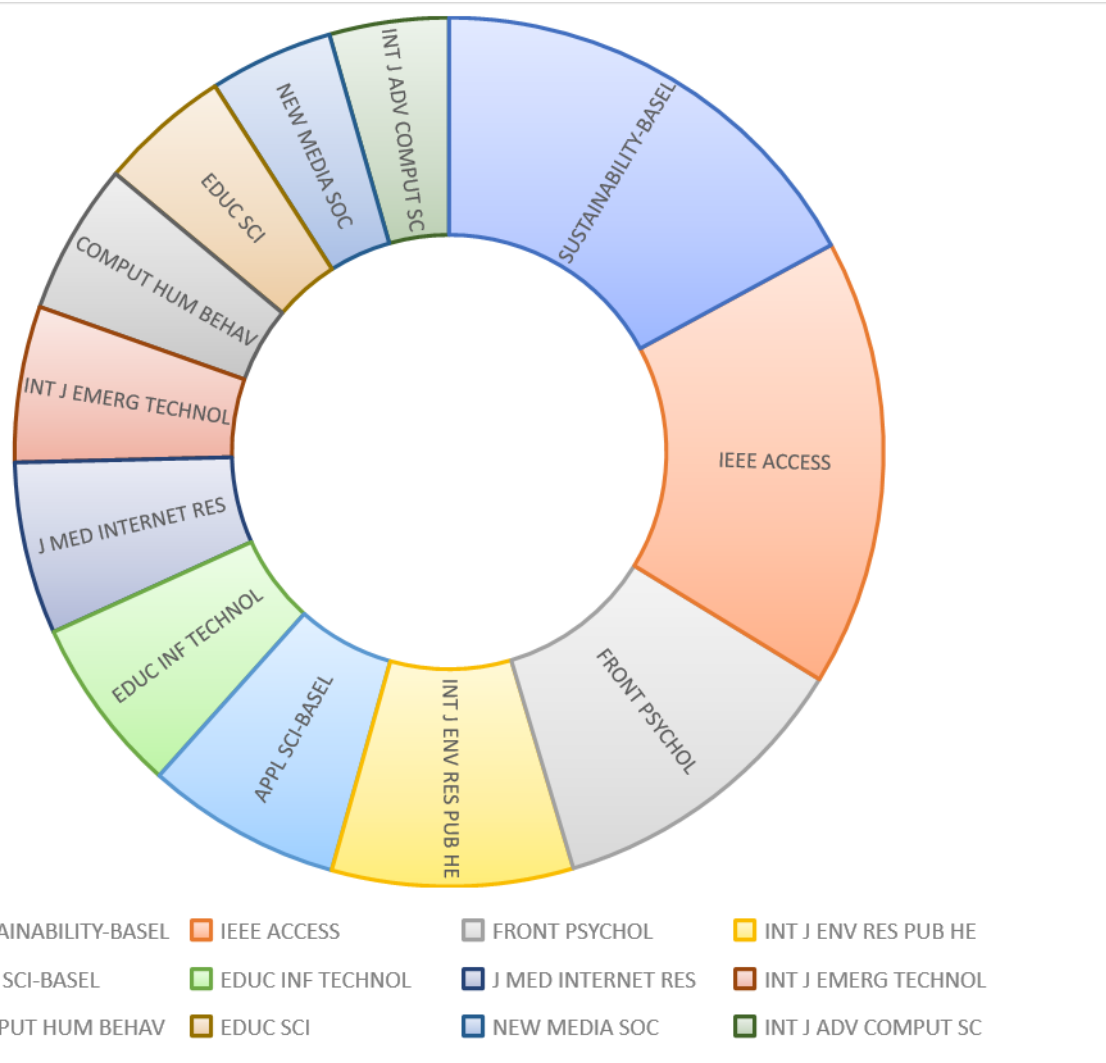
بر اساس نمونه ۱۰ هزار تایی از انتشارات در حوزه هر سه گروه پژوهشی فناوری اطلاعات، ناشران در هر گروه برای نشر بیش از نیم درصد نمونه مشخص شدند. بنابر نتایج بدست آمده ناشران برتر در حوزه فعالیت گروه مدیریت فناوری اطلاعات به شرح زیر استخراج شدند:



شکل ۲-۱۰- انتشارات در حوزه گروه پژوهشی مدیریت فناوری اطلاعات

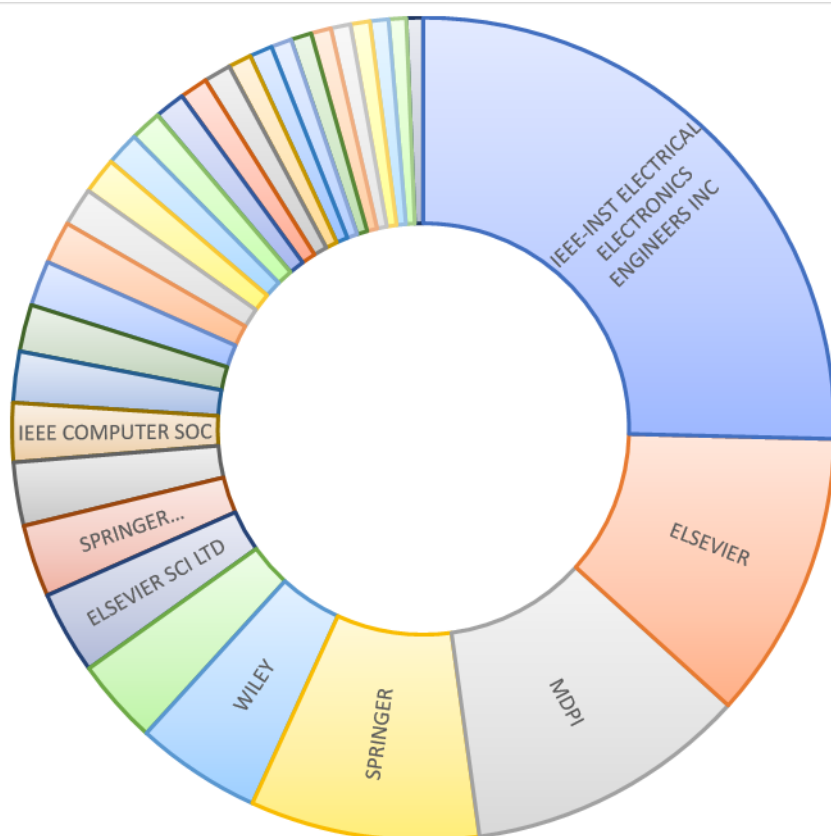
مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۴۶

همچنین بنا بر نتایج بدست آمده ناشران برتر در حوزه فعالیت گروه کسب‌وکار الکترونیکی به شرح زیر استخراج شدند:



شکل ۲-۱۱- انتشارات در حوزه پژوهشی کسب‌وکار الکترونیک

و بنا بر نتایج بدست آمده ناشران برتر در حوزه فعالیت گروه سیستم‌های اطلاعاتی به شرح زیر استخراج شدند:



- IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC
- ELSEVIER
- MDPI
- SPRINGER
- WILEY
- PERGAMON-ELSEVIER SCIENCE LTD
- ELSEVIER SCI LTD
- SPRINGER HEIDELBERG
- ELSEVIER SCIENCE INC
- IEEE COMPUTER SOC
- HINDAWI LTD
- WILEY-HINDAWI
- IOS PRESS
- TECH SCIENCE PRESS
- TAYLOR & FRANCIS LTD
- FRONTIERS MEDIA SA
- ASSOC COMPUTING MACHINERY
- EMERALD GROUP PUBLISHING LTD

شکل ۲-۱۲- انتشارات در حوزه گروه پژوهشی سیستم‌های اطلاعاتی

۲-۶-۳. بررسی حوزه‌های مطرح در زمینه‌های سه‌گانه فناوری اطلاعات

بر اساس یافته‌های انجام شده حوزه‌های ترکیبی پرکاربرد در زمینه فناوری اطلاعات به شرح زیر قرار گرفته است. این حوزه‌ها توسط پایگاه به هر یک از مقالات اختصاص یافته و به منظور استانداردسازی جستجو و بازیابی اطلاعات در پایگاه WoS استفاده می‌شود. ذکر این تحلیل در اینجا بدین معناست که بتوان حوزه‌های پرکاربردتر از نظر پایگاه و ترکیب آنها را مشاهده نمود. کلیه حوزه‌های پایگاه (WoS Categories) در پیوست نمایش داده شده است.

حوزه‌های ترکیبی کلیدی در زمینه مدیریت فناوری اطلاعات به شکل زیر هستند:



شکل ۲-۱۳- حوزه‌های ترکیبی کلیدی در زمینه مدیریت فناوری اطلاعات

حوزه‌های ترکیبی زیر در زمینه کسب و کار الکترونیکی بیشترین فعالیت را رصد نمودند:

- Education & Educational Research
- Communication
- Business

حوزه‌های ترکیبی کلیدی در زمینه سیستم‌های اطلاعاتی به شکل زیر هستند:



شکل ۲-۱۵- حوزه‌های ترکیبی کلیدی در زمینه سیستم‌های اطلاعاتی

حوزه‌های ترکیبی زیر در زمینه سیستم‌های اطلاعاتی بیشترین فعالیت را رصد نمودند:

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۵۲

- Computer Science, Information Systems; Engineering, Electrical & Electronic; Telecommunications
- Computer Science, Artificial Intelligence
- Computer Science, Information Systems

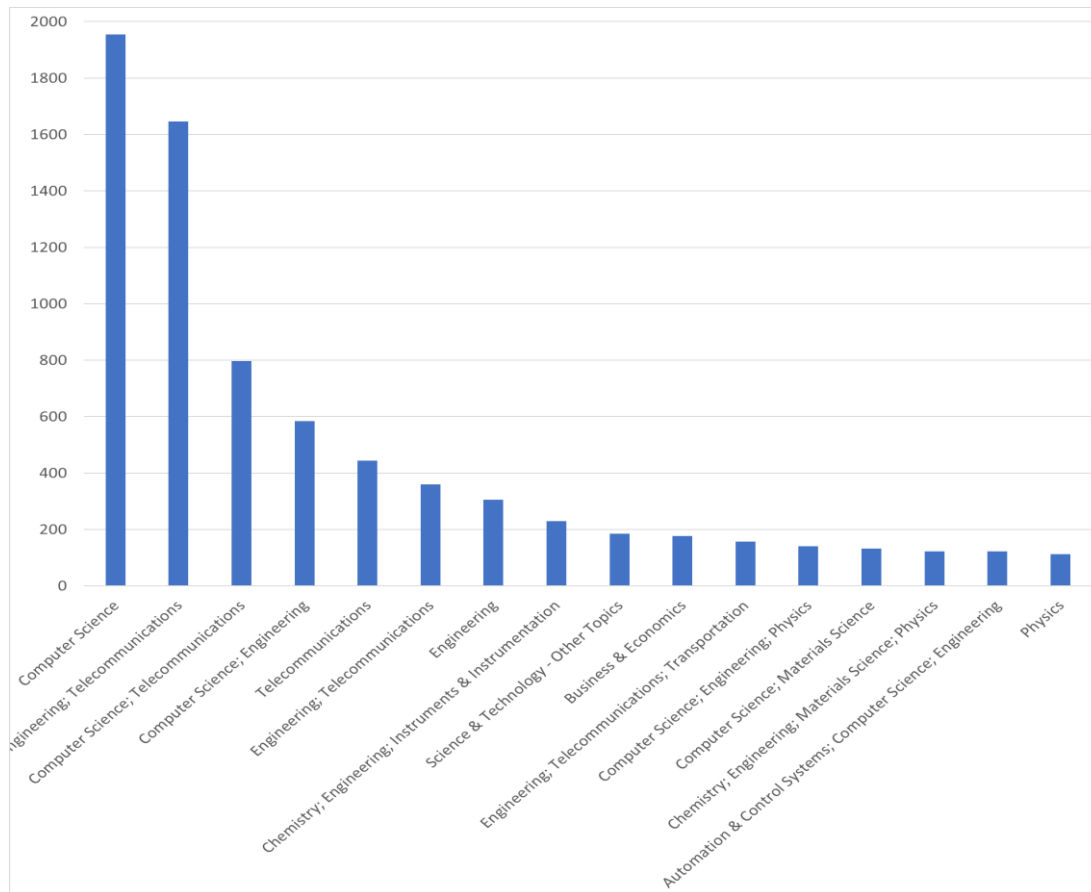
۴-۶-۲. بررسی محیط پژوهش مطرح در زمینه‌های سه‌گانه فناوری اطلاعات

بر اساس یافته‌های استخراج مطالعات بیش از ۱ درصد مطالعات در محیط پژوهش مدیریت فناوری اطلاعات به شرح زیر بودند:

- Computer Science
- Computer Science; Engineering; Telecommunications
- Computer Science; Telecommunications
- Computer Science; Engineering
- Telecommunications
- Engineering; Telecommunications
- Engineering
- Chemistry; Engineering; Instruments & Instrumentation
- Science & Technology - Other Topics
- Business & Economics
- Engineering; Telecommunications; Transportation
- Computer Science; Engineering; Physics
- Computer Science; Materials Science
- Chemistry; Engineering; Materials Science; Physics
- Automation & Control Systems; Computer Science; Engineering
- Physics

که برای مقایسه به نمودار زیر می‌رسیم:

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۵۳

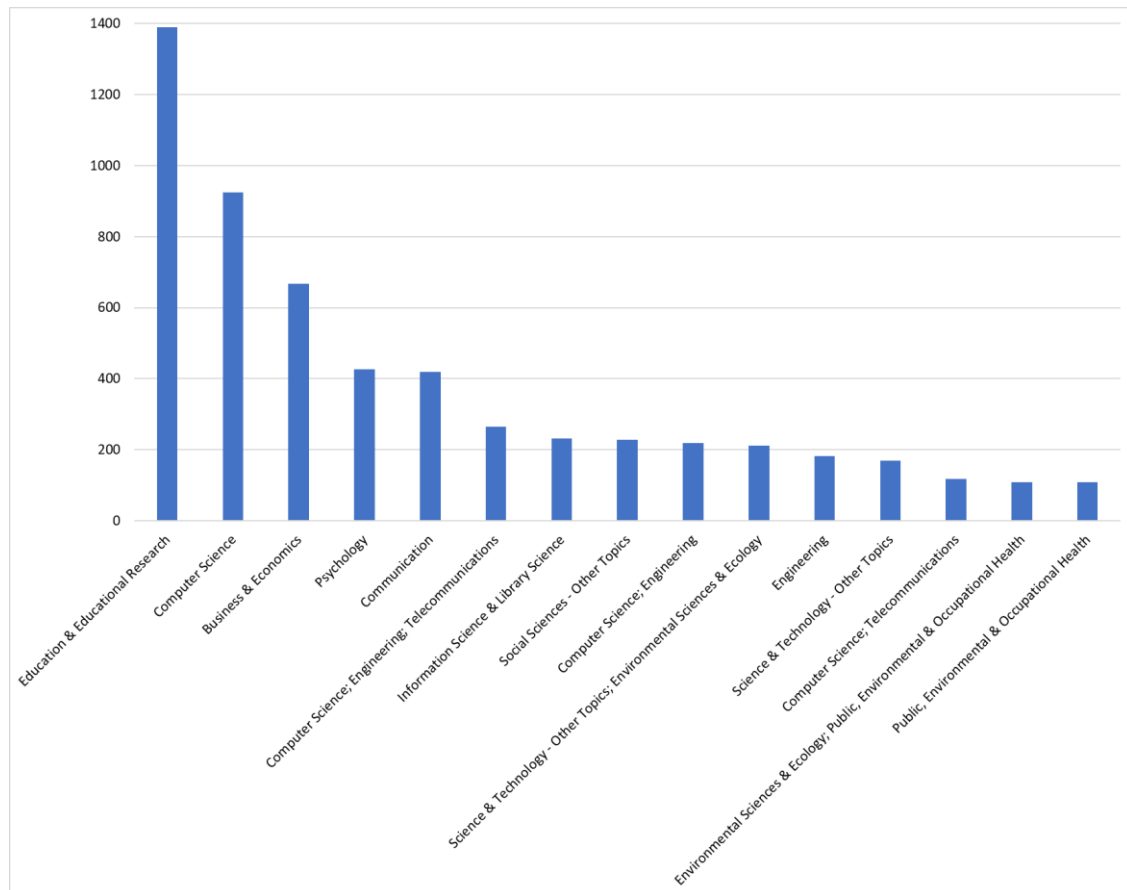


شکل ۲-۱۶- کلیدواژه پژوهش مطرح در زمینه‌های مدیریت فناوری اطلاعات

همین مقایسه برای گروه پژوهشی کسب و کار الکترونیک به فهرست:

- Education & Educational Research
- Communication
- Business
- Computer Science, Information Systems; Engineering, Electrical & Electronic; Telecommunications
- Computer Science, Information Systems
- Information Science & Library Science
- Psychology, Multidisciplinary
- Green & Sustainable Science & Technology; Environmental Sciences; Environmental Studies
- Computer Science, Artificial Intelligence
- Multidisciplinary Sciences
- Social Sciences, Interdisciplinary
- Management
- Public, Environmental & Occupational Health
- Environmental Sciences; Public, Environmental & Occupational Health

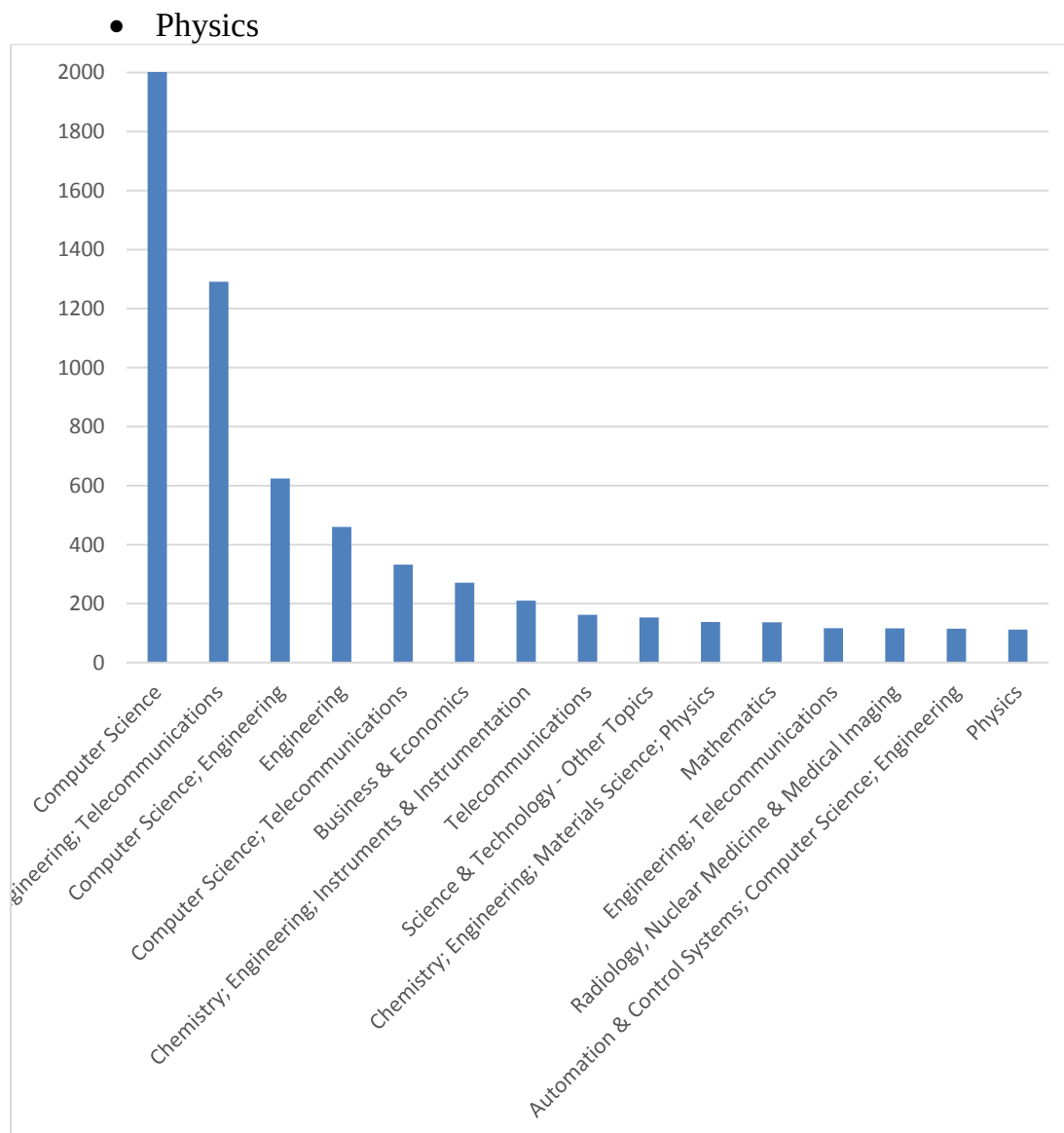
و نمودار زیر خواهد رسید:



شکل ۲-۱۷- کلیدواژه پژوهش مطرح در زمینه‌های کسب و کار الکترونیک

و در زمینه سیستم‌های اطلاعاتی داریم: (لازم به ذکر است تکرار واژگانی مانند Engineering و یا Computer Science نشان‌دهنده ترکیب کاربرد این حوزه‌ها با دیگر حوزه‌هاست و از نظر پایگاه در آن مقالات دو یا بیشتر برچسب مورد نیاز بوده است)

- Computer Science
- Computer Science; Engineering; Telecommunications
- Computer Science; Engineering
- Engineering
- Computer Science; Telecommunications
- Business & Economics
- Chemistry; Engineering; Instruments & Instrumentation
- Telecommunications
- Science & Technology - Other Topics
- Chemistry; Engineering; Materials Science; Physics
- Mathematics
- Engineering; Telecommunications
- Radiology, Nuclear Medicine & Medical Imaging
- Automation & Control Systems; Computer Science; Engineering

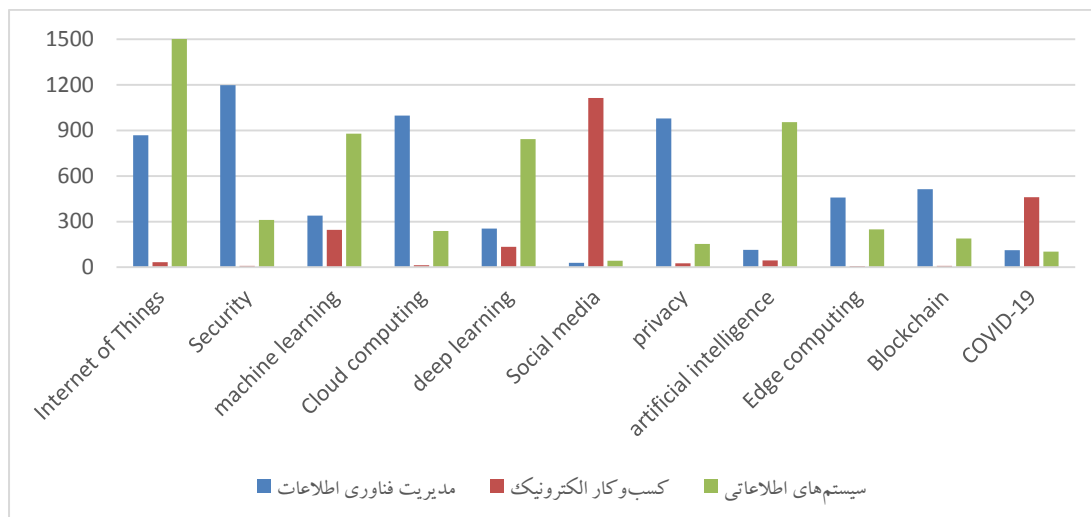


شکل ۲-۱۸- کلیدواژه پژوهش مطرح در زمینه‌های کسب‌وکار الکترونیک

۲-۶-۵. بررسی کلیدواژه‌های مطرح در زمینه‌های سه‌گانه فناوری اطلاعات

در گام بعدی ازین تحلیل، به نظر رسید دسته‌بندی مطالعات فناوری اطلاعات در دسته فوق می‌تواند به این سوال منجر شود که هر دسته فوق در نمونه مورد بررسی بیشتر از چه کلیدواژگانی بهره می‌جوید. در این جستجو ۲۳۳ هزار کلیدواژه در زمینه مدیریت فناوری اطلاعات، ۱۱۷ هزار کلیدواژه در زمینه کسب‌وکار الکترونیک و ۲۱۰ هزار کلیدواژه در زمینه سیستم‌های اطلاعاتی استخراج شدند. با توجه به فرکانس کلیدواژه‌های مقالات در ۱۰ هزار مجله مرتبط با جستجو، برای ترسیم ۱۵ کلیدواژه برتر به نمودار زیر می‌رسیم:

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۵۶



شکل ۱۹-۲- کلیدواژه‌های پرکاربرد حوزه فناوری اطلاعات در هر گروه پژوهشی

فصل سوم

مدل‌سازی موضوعی سه حوزه فناوری اطلاعات

۳-۱. مقدمه و پیشینه

فاصله یا شکاف میان پژوهش‌های دانشگاهی به‌ویژه پایان‌نامه‌ها، با نیازهای جامعه و در سطح جهانی یکی از چالش‌های اساسی نظام آموزش عالی در ایران است. گسترش کمی دانشگاه‌ها و افزایش پذیرش دانشجویان تحصیلات تکمیلی باعث شده است تا از یک طرف شاهد انتقادهای و واکنش اندیشمندان و خبرگان آموزش و پژوهش باشیم و از طرف دیگر کارفرمایان صنایع و اقتصاد از توانمندی اندک دانش‌آموختگان در حل مسائل واقعی کشور گله‌مند باشند (فتاحی، ۱۳۹۰). همچنین عرصه پژوهش شاهد دوباره‌کاری‌های بسیاری است که دلیل عمده آن، عدم توانایی استادان و دانشجویان در یافتن مسئله پژوهش در حوزه‌های مختلف موضوعی است و در بسیاری از موارد جامعه دانشگاهی بجای «مسئله‌یابی» به دنبال «مسئله‌سازی» شده است.

سیاست‌گذاران پژوهش پس از اطلاع از خلأ پژوهشی و نیز پژوهش‌هایی که تقاضای قابل توجهی در جامعه پژوهشی برای آن‌ها وجود ندارد و با بهره‌گیری از دانش و اطلاعات دیگری که در اختیار دارند و نیز در نظر گرفتن سیاست‌های کلان پژوهشی، نیازهای پژوهشی را شناسایی و تعیین می‌کنند. در این مرحله سیاست‌گذار ممکن است سیاست‌ها و برنامه‌هایی را اعمال کند که تقاضای پژوهشی خاصی در جامعه پژوهشی ایجاد شود و این نیاز به پژوهشگران منتقل گردد. در نهایت سیاست‌گذاران و مدیران بر اساس معیارهای مختلف و با توجه به منابع موجود اولویت‌های پژوهشی را در حوزه‌های مختلف تعیین می‌کنند.

در فعالیتی که توسط دکتر جلالی‌منش و همایون‌والا در ایرانداک (۲۰۱۲) صورت پذیرفته بود، ۶۵۰ عنوان پایان‌نامه در زمینه مهندسی صنایع با الگوریتم‌های متن‌کاوی و تحلیل شبکه اجتماعی با هدف استخراج حوزه مهندسی صنایع، مورد پایش قرار گرفت. در این مطالعه ۱۵۰۰ واژه به‌عنوان واژه‌های مرتبط با مهندسی صنایع انتخاب شدند. در این مطالعه از روشی متفاوت با فعالیت این طرح بهره گرفته شده و در یک مرحله واژگان استخراج شده‌اند (فرایند طراحی شده در این پژوهش ۲ مرحله‌ای است)، ولی می‌توان از تجربیات مرتبط با پاکسازی فهرست لغات استخراج شده و تجربیات مرتبط با

مصورسازی نتایج بهره جست (Jalalimanesh & Homayoon Vala, 2012). در همین سال صدیقی و جلالی‌منش در پژوهشی با استفاده از روش‌های علم‌سنجی و کتاب‌سنجی، در صدد پرداختن به این مساله بودند که مجموعه تولیدات علمی در حوزه مدیریت دانش در سطح بین‌المللی در بازه زمانی ۱۰ ساله (۲۰۰۱-۲۰۱۰)، چه رویکردی داشته و به چه حوزه‌هایی پرداخته است. از نظر ایشان به لحاظ موضوعی، بیشترین گرایش به برخی فناوری‌های اطلاعاتی پشتیبان، به‌ویژه «داده‌کاوی» و نیز مفاهیم و جنبه‌های سازمانی و اجتماعی حوزه مدیریت دانش بوده است (صدیقی و جلالی‌منش، ۱۳۹۱). در ادامه این پژوهش، پژوهشی با عنوان «طراحی و تدوین منشور پژوهش، پژوهشکده‌ها، و گروه‌های پژوهشی پژوهشگاه علوم و فناوری اطلاعات ایران» توسط ایشان انجام پذیرفت که در آن زمینه‌های پژوهشی توسط تحلیل محتوای اسناد بالادست، تحلیل محتوای برنامه‌ریزی راهبردی و اساسنامه پژوهشگاه، نمونه‌کاوی سازمان‌ها و مراکز پژوهشی مشابه و دانش اساتید انجام شد (جلالی‌منش، ۱۳۹۹). همچنین در طرح پژوهشی «تدوین نقشه راه فناوری اطلاعات پژوهشگاه با تمرکز بر پیشرفت‌های نوین در حوزه تجزیه و تحلیل اطلاعات» معماری سه لایه متشکل از لایه پیش‌ران‌ها (اهداف و محرک‌ها)، پیش‌برنده‌ها (فناوری) و خدمات اطلاعاتی برای تدوین نقشه راه استفاده شد. پیش‌ران‌ها و پیش‌برنده‌ها عواملی هستند که به ترتیب چرایی و چگونگی یک خدمت اطلاعاتی را توصیف می‌کنند. براساس شناسایی اجزای لایه پیش‌ران‌ها، از پنل‌های تخصصی، نظرات مدیران ارشد سازمان، و اسناد برنامه‌ریزی راهبردی استفاده شد. برای شناسایی اجزای لایه پیش‌برنده‌ها، یا همان روندهای فناوری اطلاعات، نمودارهای هایپ سایکل گارتر همراه با گزاره‌های پیش‌بینی موسسات تحقیقاتی برتر در این حوزه در دنیا تحلیل و فناوری‌های اصلی در حوزه تحلیل اطلاعات شناسایی شدند. برای این منظور ابتدا حوزه‌های اصلی فناوری شناسایی شده، سپس جایگاه هر یک از فناوری‌ها در هر حوزه مشخص گشت. برای تعیین اجزای لایه خدمات اطلاعاتی، پیشنهاد شد که مخاطبین و گروه‌های آنها، به همراه نیازهایشان شناسایی شوند. سپس، یک روش جدید بر مبنای رویکرد چابک برای توسعه سیستم، برای تعیین خدمات اطلاعاتی براساس پیشنهاد گردید. در نهایت با توجه به روش‌های پیشنهادی نقشه راه فناوری اطلاعات پژوهشگاه برای خدمات اطلاعاتی تا افق ۱۴۰۴ مشخص شدند (مجبی، ۱۳۹۵).

همانگونه که پیشتر اشاره شد یکی از شاخه‌های مهم داده‌کاوی، متن‌کاوی است که هدف آن استخراج دانش از داده‌های متنی غیر ساختاریافته یا نیمه ساختاریافته است. از متن‌کاوی با عناوین دیگری چون داده‌کاوی متنی و کشف دانش از پایگاه‌های متنی نیز نام برده شده است. کاربردهای عمده متن-کاوی در شناخت متون مشابه، خلاصه‌سازی متون، ترجمه ماشینی متون، انتخاب کلیدواژه

یا برچسب‌زنی متون، رده‌بندی اطلاعات متنی و ... است که در فرایند آن از علوم مختلف، مخصوصاً پردازش زبان طبیعی استفاده می‌شود (Choudhary, et al., 2009).

هدف از متن‌کاوی استنتاج دانش ضمنی^۹ نهان در متن و ارائه آن به شکل دانش صریح^{۱۰} است. برای دستیابی به این هدف معمولاً چهار فعالیت مهم در فرایند متن‌کاوی انجام می‌شود که عبارت‌اند از: بازیابی اطلاعات^{۱۱}، استخراج اطلاعات^{۱۲}، کشف دانش^{۱۳} و فرضیه‌سازی^{۱۴}. سیستم بازیابی اطلاعات درصدد است تا متن‌های متناسب با حوزه موردنظر را به دست آورد درحالی‌که استخراج اطلاعات به دنبال یافتن اطلاعاتی با ساختار از پیش تعیین‌شده (مانند استخراج روابط) از این متون است. سیستم کشف دانش موجب دستیابی به دانش جدید از اطلاعات استخراج‌شده می‌گردد و درنهایت فرضیه‌سازی به تفسیر دانش کشف‌شده در حوزه مربوطه می‌پردازد. در مرحله بازیابی اطلاعات با استفاده از تکنیک‌های مختلف ازجمله به کارگیری اصطلاح‌نامه^{۱۵}، بسط جستجو^{۱۶}، به کارگیری شبکه واژگان^{۱۷} تخصصی حوزه مدنظر و ... تلاش می‌شود تا پایگاهی از متون مرتبط ایجاد شود و در مراحل بعدی از این پایگاه به‌عنوان منبع اطلاعاتی متن‌کاوی استفاده شود. رویکرد موضوعی و محتوایی نیز بر ماهیت و ساختار دانش موجود در منابع موردبررسی دلالت دارد. به عبارتی این رویکرد بر مبنای شاخه‌ها و زیرشاخه‌های علمی شناخته‌شده در حوزه موردنظر بنا شده است. این شاخه‌ها می‌توانند از منابع موردبررسی استخراج شده باشند و یا از شاخه‌های موضوعی رشته موردنظر در دیگر منابع (مانند دانشگاه‌ها یا پایگاه‌های علمی) به دست آمده باشند.

هرچند که استفاده از روش‌های ماشینی و تکیه بر تحلیل حجم انبوه اطلاعات، اخیراً در فرایندهای تصمیم‌گیری و برنامه‌ریزی موردنظر قرار گرفته‌اند اما کاربست این روش‌ها در منابع زبان فارسی بسیار اندک بوده و الگوریتم‌ها و ابزارهای توسعه داده‌شده در این روش‌ها نیز بیشتر مبتنی بر متون عمومی مانند روزنامه و وب‌سایت‌های خبری و ... است و کمتر به متون علمی و پژوهشی پرداخته شده است. پیشتر اشاره شد که مدل‌سازی موضوعی از روش‌های کارای تحلیل متن است که هدف آن گروه‌بندی مجموعه‌ای از متون مختلف به‌صورت زیرمجموعه‌های معنادار که موضوع^{۱۸} نامیده می‌شوند است.

^۹ Implicit Knowledge

^{۱۰} Explicit Knowledge

^{۱۱} Information Retrieval

^{۱۲} Information Extraction

^{۱۳} Knowledge Discovery

^{۱۴} Hypothesis Generation

^{۱۵} Thesaurus

^{۱۶} Query Expansion

^{۱۷} WordNet

^{۱۸} Topic

الگوریتم‌های روش‌های مدل‌سازی موضوعی با حداقل دخالت انسان این فرایند را انجام می‌دهند و همین نکته باعث جذابیت و استفاده روزافزون آن در تحلیل حجم عظیمی از متون شده است. در این روش بجای استفاده از عناوین از قبل تعیین شده، تنها با تعیین تعداد موضوعات، زیرمجموعه‌های موضوعی استخراج شده و میزان ارتباط هر سند با هر زیرمجموعه نیز تعیین می‌شود. در این فصل به مدل‌سازی موضوعی و نقش آن در فرایند پژوهش در این طرح پژوهشی پرداخته می‌شود.

۲-۳. مدل‌سازی موضوعی

مدل‌سازی موضوعی بر اساس این ایده تشکیل شده است که اسناد، ترکیبی از مباحث هستند و موضوع (عنوان و کلیدواژه یا فقط عنوان) آنها با یک توزیع احتمالی بیش از دیگر کلمات مورد استفاده قرار داده می‌شود. در این مدل‌ها موضوعات (و توزیع آنها) به عنوان ساختار نهفته یا متغیرهای نهفته در نظر گرفته می‌شوند. الگوریتم‌های مدل‌سازی موضوعی روش‌های آماری هستند که کلمات اصلی اسناد را تجزیه و تحلیل می‌کنند تا مباحثی را که از طریق آنها رایج می‌شود تحلیل نموده و چگونه اتصال این موضوعات را به یکدیگر و چگونگی تغییر آنها را با گذر زمان کشف نمایند. یک موضوع به عنوان یک توزیع در یک واژگان ثابت تعریف شده است. به عنوان مثال موضوع اینترنت اشیا دارای کلمات در حوزه فناوری اطلاعات با احتمال بالا است. این مدل فرض می‌کند که موضوعات قبل از اسناد (مقالات) تولید می‌شوند (Yau, et al., 2014).

الگوریتم‌های مدل‌سازی موضوعی شامل تکنیک‌های آماری است که هدف آنها توصیف موضوعاتی است که در اسناد منتشر شده مانند مجلات منتشر شده‌اند. آنها بر اساس تجزیه و تحلیل آماری کلمات در آن اسناد، شناسایی خوشه‌هایی از کلمات مشترک، تشخیص موضوعات مورد بحث، روابط بین این موضوعات و تکامل آنها در طول زمان کشف می‌کنند. چندین مدل مانند تجزیه و تحلیل معنایی نهفته، تجزیه و تحلیل معنایی بالقوه احتمالی^{۱۹} و تخصیص نهفته دیریکله^{۲۰} (Blei, Andrew, & Jordan, 2003)؛ و همچنین برخی از مشتقات آنها، مانند پاچینو^{۲۱} و یا مدل‌سازی موضوعی رابطه‌ای (Blei, 2009 & Chang) در این زمینه کاربرد دارند.

¹⁹ PLSA

²⁰ Dirichlet (LDA)

²¹ Pachinco

مدل‌سازی موضوعی نوعی از مدل آماری برای کشف موضوعات نهفته^{۲۲} است که در مجموعه‌ای از اسناد از طریق یادگیری ماشین رخ می‌دهد. در حال حاضر، تخصیص نهفته دیریکله یک رویکرد مدل‌سازی کارا و رایج در این زمینه است.

از کاربردهای استفاده از مدل‌سازی ذکر شده می‌توان به ارائه سطح بالا از مفاهیم، کلمات کلیدی آن حوزه و زمینه‌هایی که مربوط به موضوع مورد نظر در یک دوره زمانی برای یک زمینه خاص پژوهش مانند فناوری اطلاعات است و ارائه یک نقشه از مجموعه واژگان کلیدی آن حوزه اشاره کرد. به‌طور مثال می‌توان زمینه‌های علمی مربوط به ۱۰ سال گذشته محدود به چند مجله پیشرو در حوزه فناوری اطلاعات شناسایی و تحلیل نمود. برای به دست آوردن این نتیجه، ابتدا باید هر موضوع با یک توصیفگر مرتبط شود که این توصیف می‌تواند شامل اطلاعات مربوط به کلمات کلیدی مرتبط، برخی اطلاعات آماری مفید درباره انتشار، توزیع مجلات مقالات در هر موضوع، توزیع مقالات موضوعی در هر سال، کشور و تعداد استنادات باشد.

روش‌های استخراج موضوع بر مبنای این فرض عمل می‌کنند که کلمات موجود در یک متن از نظر معنایی به یکدیگر وابسته هستند و معنایی که در مستندات یک موضوع مسنجم وجود دارد از مجموعه کلمات این مستندات به دست آمده است. به عبارتی صرف‌نظر از معانی یا مکان قرارگیری متن، هم‌رخدادی کلمات مورد تحلیل قرار گرفته و اسناد به‌صورت مجموعه‌ای از کلمات در نظر گرفته می‌شود. در مدل‌سازی موضوعی مفروضات زیر در نظر گرفته می‌شود (H. Wallach 2006):

۱. ترتیب حضور مستندات در پیکره اهمیتی ندارد؛
 ۲. ترتیب قرار گرفتن کلمات در هر مستند اهمیتی ندارد؛
 ۳. موضوعات توزیع چندجمله‌ای از کلمات موجود در یک واژه‌نامه هستند؛
 ۴. کلمات یک مستند از تعدادی موضوع پنهان نشأت گرفته‌اند.
- دو فرض اول که ترتیب مستندات در یک پیکره و ترتیب کلمات در یک مستند را نادیده می‌گیرند منجر به خلق مفهومی به نام سبد واژگان^{۲۳} می‌شود که نماینده‌ای از یک مستند خواهد بود و اصلی‌ترین ورودی الگوریتم‌های مدل‌سازی موضوعی است. فرض سوم و چهارم تبیین‌کننده مفهوم موضوع است که توزیعی از کلمات موجود در یک واژه‌نامه است. در صورتی که نیاز به تعیین نام یا برچسب برای یک موضوع وجود داشته باشد، معمولاً از چند کلمه اولی که بیشترین ارتباط را به آن موضوع دارند استفاده می‌شود (H. Wallach 2006).

²² Latent

²³ Bag of Words

همانگونه که در ابتدا نیز اشاره شد، در سال ۱۳۹۹ و در طرح پژوهشی «مقایسه روند پژوهش‌ها در حوزه فناوری اطلاعات در پایان‌نامه‌ها و رساله‌های کشور و روند جهانی آن با استفاده از روش تحلیل متون»، حوزه‌های فناوری اطلاعات در جهان به دقت مورد بررسی و پایش قرار گرفت و روند فناوری اطلاعات در جهان و ایران مقایسه شد. در این طرح نیز همانند طرح پیشین از روش مدل‌سازی موضوعی بهره گرفته می‌شود ولی دسته‌بندی کلیدواژه‌ها به تفکیک گروه‌های پژوهشی ایرانداک، مورد توجه قرار می‌گیرد.

پژوهش‌های قبلی مرتبط با شناسایی کاستی‌های پژوهشی، عمدتاً از روش کیفی و استفاده از نظر خبرگان استفاده کرده‌اند و محدودیت‌هایی چون تعداد اندک افراد خبره‌ای که در پژوهش مشارکت می‌کنند، سوگیری و تأثیر بالای نظر پژوهشگر در مصاحبه و جمع‌بندی نتایج و هزینه و زمان بالای انجام این پژوهش‌ها را پذیرفته‌اند. هرچند که استفاده از نظر خبرگان در تعیین اولویت‌ها و شناسایی نیازهای پژوهشی این پژوهش نیز لازم به نظر می‌رسد اما خبرگان در شناسایی دقیق کاستی‌های پژوهشی، هر یک لزوماً نظر شخصی و محدود به دانش خود را از محیط بیان خواهند کرد.

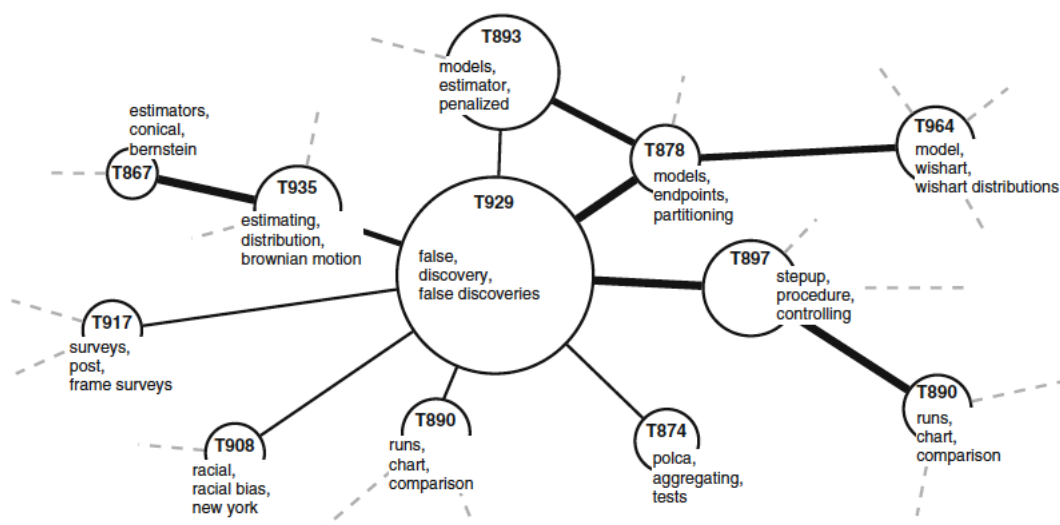
از طرف دیگر مطالعات گذشته عمدتاً صرف‌نظر از تکنیک و روش انجام کار، نشان می‌دهد که در چه حوزه‌هایی و به چه میزان کار پژوهشی انجام شده است، اما مشخص نمی‌کند که چه حوزه‌هایی مغفول مانده‌اند و در آن‌ها پژوهش قابل ملاحظه‌ای صورت پذیرفته است.

همچنین در بسیاری از پژوهش‌های پیشین، تعداد محدودی از منابع اطلاعاتی یا کاربران و یا خبرگان را به عنوان جامعه پژوهشی در بازه زمانی محدود در نظر گرفته‌اند و با مصاحبه و یا پرسش از آن‌ها به تعیین کاستی‌های پژوهشی پرداخته‌اند و روش‌های مورد استفاده نیز تنها قابلیت استفاده در فرایندهای گردآوری و تحلیل نظرات گروه‌های محدود و یا تعداد افراد محدود کاربرد دارد. حال آنکه نیاز به استفاده از روش‌های ماشینی و مستقل از خبرگان وجود دارد تا بتوان حجم انبوه اطلاعات را مورد تحلیل قرار داد و اطلاعات مختلف مرتبط با یک محقق پیشرو را در همه مطالعاتش و در بازه زمانی بلندتر مورد بررسی قرار داد که در این طرح سعی می‌شود به این مهم پرداخته شود.

یکی از کاربردهای استفاده از مدل‌سازی ذکر شده به منظور ارائه سطح بالا از مفاهیم، کلمات کلیدی آن حوزه و زمینه‌هایی که مربوط به موضوع مورد نظر در یک دوره زمانی برای یک زمینه خاص پژوهش مانند فناوری اطلاعات است، ارائه یک نقشه از مجموعه واژگان کلیدی آن حوزه است. بطور مثال می‌توان زمینه‌هایی مربوطه را در ۱۰ سال گذشته محدود به چند مجله پیشرو در حوزه فناوری اطلاعات شناسایی تحلیل نمود. برای به دست آوردن این نتیجه، ابتدا باید هر موضوع با یک توصیفگر مرتبط شده که این توصیف می‌تواند شامل اطلاعات مربوط به کلمات کلیدی مرتبط، برخی اطلاعات

آماري مفید درباره انتشار، توزیع مجلات مقالات در هر موضوع، توزیع مقالات موضوعی در هر سال، کشور و تعداد استنادات باشد. در فعالیتی که توسط دکتر جلالی‌منش در ایرانداک (۲۰۱۲) صورت پذیرفته بود، ۶۵۰ عنوان پایان‌نامه در زمینه مهندسی صنایع با الگوریتم‌های متن‌کاوی و تحلیل شبکه اجتماعی با هدف استخراج حوزه مهندسی صنایع، مورد پایش قرار گرفت. در این مطالعه ۱۵۰۰ واژه به‌عنوان واژه‌های مرتبط با مهندسی صنایع انتخاب شدند. در این مطالعه از روشی متفاوت با فعالیت این طرح بهره گرفته شده و در یک مرحله واژگان استخراج شده‌اند (فرایند طراحی شده در این پژوهش ۲ مرحله‌ای است)، ولی می‌توان از تجربیات مرتبط با پاکسازی فهرست لغات استخراج شده و تجربیات مرتبط با مصورسازی نتایج بهره جست.

به منظور ارائه دیدگاه جامع از مباحث علمی مطرح شده در حوزه فناوری اطلاعات توسط مجموعه‌ها از مقالات، مفهوم نقشه موضوعی مطرح می‌شود. نمونه این ارائه در ادبیات مرتبط در حوزه فناوری توسط نویسندگان این طرح یافت نشده است. در صورت مشخص بودن حوزه و روند بین‌المللی در ادبیات، این طرح می‌توانست از نتایج آن استفاده و فقط مقایسه را در راستای پژوهش‌های ملی انجام دهد. اما به دلیل فقدان موضوع و همچنین سرعت رشد بالای موضوع فناوری اطلاعات (امکان ایجاد تغییرات در زمان کوتاه)، در این طرح ملزم به بررسی و استخراج در هر دو حوزه داخلی و بین‌المللی هستیم. یک نقشه موضوعی یک گراف است که گره‌ها، موضوع‌ها را نشان می‌دهند و ارتباطات نمایانگر روابط بین موضوعات می‌باشند. علاوه بر این، هر گره از نقشه موضوعی می‌تواند به صورت گرافیکی توسط یک دایره که حاوی مقادیر مناسب از کلیدواژگان موضوع مربوطه است، نمایان شود. قطر دایره، متناسب با تعداد رخدادهای آن واژگان در موضوع و حوزه و متناسب با تعداد مقالاتی است که در این موضوع ارائه داده شده است. همچنین موقعیت گرافیکی گره‌های مربوطه، طوری مشخص می‌شود تا موضوعات مشابه نزدیک به یکدیگر نشان داده شوند. عرض خط ارتباطی متناسب با میزان شباهت دو موضوع است. به عنوان نمونه می‌توان به شکل زیر اشاره نمود.



شکل ۲۰- نمونه توسعه چارچوب فناوری اطلاعات (باتیستی و همکاران ۲۰۱۵)

با توجه به توضیحات فوق، در این پژوهش، مدل‌سازی موضوعی بر روی مجموعه‌ای از مجلات بین‌المللی، با رویکرد تشریح شده انجام پذیرفته و پس از پردازش نتایج توصیفات، عمق فعالیت‌های پژوهش‌های مربوطه از این زمینه در طول سال‌های مختلف مورد بررسی قرار می‌گیرد. این پژوهش روند کلی پژوهش‌های حوزه فناوری اطلاعات در پایان‌نامه‌های و رساله‌های تحصیلات تکمیلی کشور (برگرفته از سامانه گنج) را با روند کلی پژوهش‌های مربوط به فناوری اطلاعات در سطح جهان بررسی نموده و به مقایسه تفکیکی دانشگاه‌های کشور با دیگر تفکیک‌ها پرداخته نمی‌شود. یکی از دلایل انتخاب این مساله دسترسی به این منابع در بعد داخلی و بین‌المللی است. در تعریف یک مساله عملیاتی، یکی از مهمترین چالش‌ها، دسترسی مناسب به اطلاعات مورد نیاز است. پژوهشگران این طرح به پایان‌نامه‌های بین‌المللی مناسب در زمینه فناوری اطلاعات دسترسی ندارند حال اینکه به مجلات پایگاه ذکر شده دسترسی دارند. این مساله خللی در اجرای این پژوهش ایجاد نمی‌نماید زیرا یکی از بزرگترین دستاوردهای این پژوهش انجام روشمند این مدل در پژوهشگاه می‌باشد. در صورتی که پژوهشگاه به بانک اطلاعات غنی‌تری در آینده دسترسی داشته باشد می‌تواند این پژوهش را در آن پایگاه تکرار نماید. در بعد خروجی نیز مقایسه بین تحقیقات ملی و مقالات بین‌المللی بسیار ارزشمند است. بطور مثال می‌توان فاصله زمانی (شکاف) میان مطالعات ملی را با پژوهش‌های مجلات پیشرو کشف کنیم.

۳-۳. خوشه‌بندی موضوعی

رویکرد استخراج متن به دنبال یافتن کلمات یا عباراتی است که می‌تواند محتوا و ساختار (روابط) موجود در داده‌ها را توضیح دهد. شناسایی مشترکات در متن، متمرکز بر تجزیه و تحلیل داده‌ها و توزیع آنها در متن و ارتباطات با استفاده از رویکردهای مختلف خوشه‌بندی است (فلدمن و سانجر ۲۰۰۶). در فعالیت هافمن (۱۹۹۹) ادعا شده است که مدل‌سازی موضوعی ابزار مفید برای استخراج اطلاعات از داده‌های متنی است. بعدها، وی و کرافت (۲۰۰۶) نشان دادند که در بازیابی اطلاعات، مدل‌سازی‌های موضوعی بیشتر از روش‌های سنتی، مبتنی بر خوشه‌ها عمل می‌کنند.

اگر چه بخشی از ادبیات در حال حاضر در زمینه مدل‌سازی موضوع توسعه یافته است، اما این ادبیات عمدتاً بر متدولوژی و توسعه آن متمرکز است. پژوهشگران بازیابی اطلاعات در اواخر دهه ۱۹۹۰، اولین مدل احتمالی با نام احتمال نامگذاری مستقل معکوس (pLSI) را پیشنهاد دادند. مدل‌های شاخص نامتقارن احتمالی، هر کلمه در یک سند را به عنوان نمونه‌ای از یک مدل بزرگ مطرح می‌کند که اجزای آن متغیرهای تصادفی چندجمله‌ای هستند که می‌توانند به عنوان نمایه‌ای از موضوعات در نظر گرفته شوند. مشکل در الگوریتم pLSI این است که تعداد پارامترها به صورت خطی با تعداد موضوع رشد می‌کنند.

برای حل این مشکل، یک مدل بهبود یافته، تخصیص نهایی دیریکله^{۲۴} پیشنهاد شده است (بلی و همکاران، ۲۰۰۳). این رویکرد یک مدل سه بعدی است که نقاط ضعف pLSI را برطرف می‌کند و در حال حاضر به طور کاربردی گسترده‌ای در مسایل مختلف مانند استخراج متن، بیوانفورماتیک و پردازش تصویر مورد استفاده قرار می‌گیرد. علاوه بر این، مدل‌سازی‌های موضوعی ارتباطی (CTM)، تخصیص سلسله مراتبی نهفته دیریکله (LDA سلسله مراتبی) و فرایند سلسله مراتبی دیریکله (HDP) در سال‌های اخیر مطرح و توسعه داده شده است.

مدل‌سازی موضوعی نیز برای رسیدن به سری زمانی در متن (یان و همکاران ۲۰۱۲؛ استیورز و همکاران ۲۰۰۴) گسترش یافته است و روش‌های عملی برای تحلیل ساختارهای متون علمی ارائه داده‌اند. نی و همکاران (۲۰۱۲)، مدعی شده‌اند که کارایی این روش‌ها به شدت به توانایی الگوریتم مدل‌سازی موضوعی برای تمایز بین موضوعات مختلف بستگی دارد. اخیراً شاهد آن هستیم که مدل‌سازی موضوعی در حال تبدیل شدن به روشی متداول است و پژوهشگران بسیاری از این الگوریتم‌ها استفاده می‌کنند.

²⁴ - Dirichlet (LDA)

در مقاله‌ای دیگر از یاو و همکاران (۲۰۱۴)، روش‌هایی از جمله LDA برای جداسازی مجموعه‌ای از انتشارات علمی به چندین خوشه بررسی می‌شود. برای ارزیابی نتایج، مجموعه‌ای از اسناد که حاوی مقالات علمی از چندین حوزه مختلف است، ایجاد شده و بررسی می‌شود که آیا مقالات در همان حوزه با یکدیگر همخوانی خواهند داشت. برنامه‌های کاربردی علم‌سنجی از جمله توانایی تجزیه و تحلیل متن در این پژوهش مورد بررسی قرار می‌گیرد. در این مقاله، مسئله جداسازی یک مجموعه سند علمی به چند دسته بر اساس محتویات آنها بر اساس مدل‌سازی موضوع در نظر گرفته شده است. برای این منظور، دو نمونه از داده‌ها توسط پژوهشگر ساخته شده است. هدف این است که از مدل‌سازی موضوعی به عنوان یک روش نسبتاً خودکار استفاده شود که خوشه‌های مستندات حاوی اسناد مشابه ایجاد کند و سپس نتایج با ساختار واقعی داده‌ها مقایسه شود. با انجام این کار، توانایی روش مدل‌سازی موضوعی برای تشخیص ساختارها از متن و سنجش موفقیت آن با دو عامل دقت و فراخوانی ارزیابی می‌شود. در زمینه یادگیری ماشین، روش ساده‌ای برای جمع‌آوری داده‌ها، استفاده از الگوریتم‌های طبقه‌بندی شده است. با این حال، از آنجا که نشانه‌گذاری مجموعه‌ای از اسناد می‌تواند زمان‌بر و ناکارآمد باشد، در این مقاله ترجیح داده شده است تا مدل‌سازی موضوعی برای حل این مشکل استفاده شود. داده‌ها قبل از مدل‌سازی موضوعی پردازش می‌شوند. در ابتدا، داده‌ها تمیز می‌شود تا فقط از عنوان، کلیدواژه و چکیده‌ی هر سند استفاده شود. از دو فرایند مختلف برای تمیز کردن داده‌ها استفاده می‌شود:

(۱) داده‌ها به صورت نشانه‌های متشکل از کلمات استفاده شده، یا

(۲) استخراج نشانه با استفاده از پردازش زبان طبیعی (NLP).

بر این اساس ابتدا هر کلمه را به عنوان یک نشانه در نظر گرفته و سپس کلمات در لیست توقف نوشتاری ارائه شده در نرم افزار MALLET (مک کالم ۲۰۰۲) پالایش می‌شوند (نیومن و همکاران ۲۰۱۲). مزیت این روش پیش پردازش این است که می‌تواند شرایط معنی دارتری را تولید کند، به عنوان مثال، سلول خورشیدی نباید به یک واژه مجزای خورشیدی و سلول تقسیم شود و باید به عنوان یک واژه‌ی دو کلمه‌ای مورد استفاده قرار گیرد. نتایج هر دو فرایند به طور جداگانه با الگوریتم‌های مدل‌سازی موضوعی مورد تجزیه و تحلیل قرار می‌گیرد. همانطور که نتایج پژوهش‌ها نشان می‌دهد، خوشه‌بندی K-means نیز مانند الگوریتم مدل‌سازی موضوع عمل می‌نماید. با این حال، مدل‌سازی موضوعی نتایج با ثبات‌تری تولید می‌کند و در نتیجه بهتر عمل می‌کند. این نتایج در شکل زیر خلاصه شده است (یاو و همکاران ۲۰۱۵).

	<i>F</i> score average	<i>F</i> score st. dev.
K-means	0.66	0.30
LDA	0.71	0.04
CTM	0.72	0.07
hLDA	Not computed	Not computed
HDP	0.90	0.04

شکل ۲۱- مقایسه روش‌های مختلف خوشه‌بندی (یاو و همکاران ۲۰۱۵)

نتایج نشان می‌دهد که مدل‌سازی موضوعی در مرحله فعلی بلوغ، نیاز به توسعه دارد تا دقت را بهبود بخشد (یاو و همکاران ۲۰۱۵).

مشکل اصلی بهره از مدل‌سازی موضوعی، نیاز به بالا بودن تعداد رکوردها و کلیدواژه‌ها برای بالاتر بودن صحت مدل است. بدین‌منظور در پژوهش جاری به استفاده از اسناد ۱۰ سال مقالات ۵ مجله پیشرو (ذکر شده در پیشنهاد) برای به دست آوردن متغیرهای نهفته محدود نشده و مجلات بیشتری با توجه به نظر خبرگان انتخاب شده‌اند. خروجی مدل‌سازی موضوعی مدل‌های احتمالی هستند که در آن داده‌ها یک نتیجه از یک فرایند آماری هستند که شامل متغیرهایی نهفته است. این فرایند توزیع احتمالی مشترک را در هر دو متغیر تصادفی مشاهده شده و نهفته تعریف می‌کند.

شمارنده توزیع شرطی توزیع مشترک تمام متغیرهای تصادفی است که می‌تواند به راحتی محاسبه شود و از لحاظ تئوری، می‌توان آن را با تجمیع توزیع مشترک تمامی ساختارهای نهفته، محاسبه کرد. عملاً، به دلیل اینکه تعداد ساختارهای موضوع احتمالاً زیاد است، این محاسبه دشوار و زمان‌بر است. الگوریتم‌های مدل‌سازی موضوعی به دو دسته تقسیم می‌شوند که توزیع‌های جایگزین دیگری را پیشنهاد می‌دهند تا تقریباً الگوریتم‌های مبتنی بر نمونه‌برداری را بهبود بخشند:

- گروه اول ارایه یک زنجیره مارکوف (یک توالی از متغیرهای تصادفی)،
- گروه دوم، الگوریتم‌های مبتنی بر نمونه‌گیری (VEM)^{۲۵}.

با توجه به نوع داده‌های انتخاب شده (عنوان مقالات و/و یا کلیدواژگان) در این طرح، الگوریتم VEM انتخاب خواهد شد، زیرا الگوریتم VEM برای هر مقاله، همپوشانی بین موضوعات مهم را ارایه می‌دهد و کشف همبستگی‌های موضوعی یکی از اهداف اصلی این طرح (کشف مرزهای حوزه فناوری اطلاعات) است.

²⁵ - Variational expectation-maximization (VEM) algorithm

۴-۳. دسته‌بندی حوزه‌ها با بهره‌گیری از فرکانس همزمانی استفاده از کلیدواژگان

در بخش «بررسی چگونگی کاربرد کلیدواژگان»، بر اساس دسته‌بندی‌های انجام شده پایگاه، کلیدواژه‌های مرتبط با هر زمینه فناوری اطلاعات مشخص شد. با توجه به اینکه این حوزه‌های مفهومی توسط خود پایگاه اختصاص می‌یابد، میزان دقت و یا درستی این دسته‌بندی مشخص نیست. در ادامه این طرح پژوهشی، با در نظر نگرفتن تمامی دسته‌بندی‌های موضوعی پایگاه و فقط در نظر گرفتن روابط بین کلیدواژه‌ها (رابطه وجود در یک مقاله)، به بررسی دسته‌بندی بر اساس کمترین احتمال پراکندگی با روش مدل‌سازی موضوعی می‌پردازیم. مدل‌سازی موضوعی از روش‌های کارای تحلیل متن است که هدف آن گروه‌بندی مجموعه‌ای از متون مختلف به صورت زیرمجموعه‌های معنادار که موضوع نامیده می‌شوند است. الگوریتم‌های روش‌های مدل‌سازی موضوعی با حداقل دخالت انسان این فرایند را انجام می‌دهند و همین نکته باعث جذابیت و استفاده روزافزون آن در تحلیل حجم عظیمی از متون شده است. در این روش بجای استفاده از عناوین از قبل تعیین شده، تنها با تعیین تعداد موضوعات، زیرمجموعه‌های موضوعی استخراج شده و میزان ارتباط هر سند (شامل مقالات منتشر شده در مجلات پایگاه) با هر زیرمجموعه نیز تعیین می‌شود (John W and Petko 2013).

روش‌های استخراج موضوع بر مبنای این فرض عمل می‌کنند که کلمات موجود در یک متن از نظر معنایی به یکدیگر وابسته هستند و معنایی که در مستندات یک موضوع مسنجم وجود دارد از مجموعه کلمات این مستندات به دست آمده است. به عبارتی صرف نظر از معانی یا مکان قرارگیری متن، هم رخدادی کلمات مورد تحلیل قرار گرفته و اسناد به صورت مجموعه‌ای از کلمات در نظر گرفته می‌شود. بنابر این و مطابق با ادبیات در مدل‌سازی موضوعی مفروضات زیر در نظر گرفته می‌شود (Blei, Andrew and Jordan 2003):

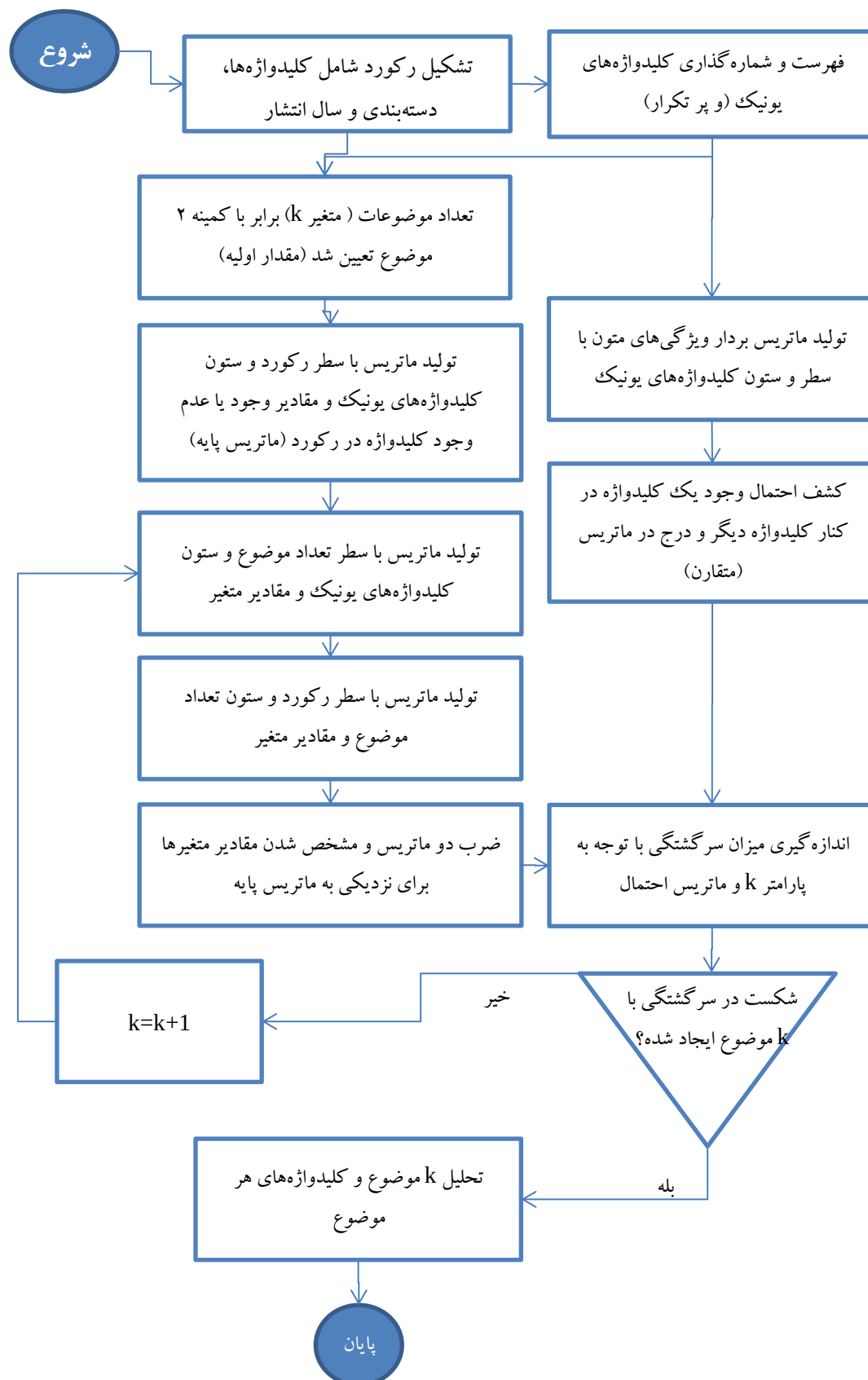
۱. ترتیب حضور مستندات در پیکره اهمیتی ندارد؛
 ۲. ترتیب قرار گرفتن کلمات در هر مستند اهمیتی ندارد؛
 ۳. موضوعات توزیع چندجمله‌ای از کلمات موجود در یک واژه‌نامه هستند؛
 ۴. کلمات یک مستند از تعدادی موضوع پنهان نشأت گرفته‌اند.
- پیدا کردن موضوعات یک سند حتی برای اسنادی که توسط خبرگان نمایه‌سازی شده و موضوعی به آن‌ها اختصاص یافته است، مورد توجه است. هر سند از تعدادی واژه تشکیل شده است و میزان ارتباط یک سند با یک موضوع به میزان تکرار کلمات مرتبط با آن موضوع در سند وابسته است. با عبارتی اگر میزان اهمیت یا وزن یک کلمه در یک سند را w نشان دهیم و یک سند شامل m واژه باشد که

هر یک وزن و اهمیت خاص خود را برای آن سند دارند و با فرض اینکه پیکره متنی موردنظر شامل n مستند است. آنگاه LDA می‌تواند تعداد k موضوع را شناسایی کند به نحوی که (D. M. Blei 2012):

$$\begin{matrix} & \text{Doc}_1 & \text{Doc}_2 & \dots & \text{Doc}_n \\ \text{Term}_1 & w_{1,1} & w_{1,2} & \dots & w_{1,n} \\ \text{Term}_2 & w_{2,1} & w_{2,2} & \dots & w_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Term}_m & w_{m,1} & w_{m,2} & \dots & w_{m,n} \end{matrix} \xrightarrow{\text{LDA(BoW(Term,Doc), k)}}$$

$$\begin{matrix} & \text{Topic}_1 & \text{Topic}_2 & \dots & \text{Topic}_k \\ \text{Term}_1 & \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,k} \\ \text{Term}_2 & \beta_{2,1} & \beta_{2,2} & \dots & \beta_{2,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Term}_m & \beta_{m,1} & \beta_{m,2} & \dots & \beta_{m,k} \end{matrix} * \begin{matrix} & \text{Doc}_1 & \text{Doc}_2 & \dots & \text{Doc}_n \\ \text{Topic}_1 & \theta_{1,1} & \theta_{1,2} & \dots & \theta_{1,n} \\ \text{Topic}_2 & \theta_{2,1} & \theta_{2,2} & \dots & \theta_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Topic}_k & \theta_{k,1} & \theta_{k,2} & \dots & \theta_{k,n} \end{matrix}$$

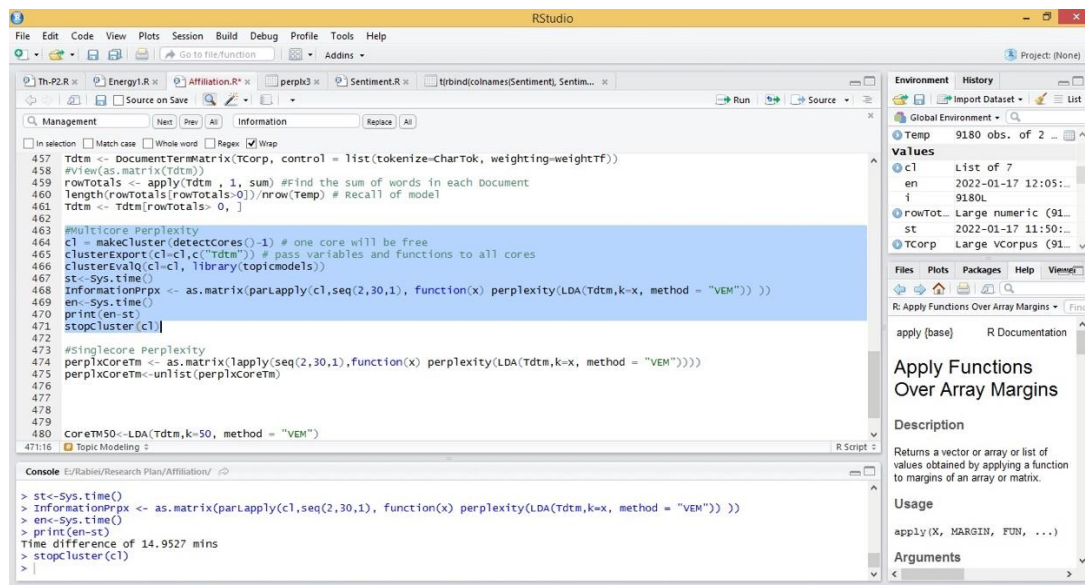
در زیر الگوریتم اجرای روش به تصویر کشیده شده است:



شکل ۲۲- فلوجارت اجرای الگوریتم LDA

کد این روش نیز در نرم‌افزار R توسعه داده شده است. مهم‌ترین ویژگی‌هایی که موجب انتخاب این زبان برنامه‌نویسی در این پژوهش شد، عبارت‌اند از:

- یک زبان علمی است و قابلیت‌ها و ابزار فراوانی در مدیریت داده‌ها، توابع آماری و ریاضیاتی، توابع داده‌کاوی، متن‌کاوی و پردازش زبان طبیعی را داراست و سطح بالایی از همخوانی بین مفاهیم علمی و آکادمیک در حوزه‌های پردازش و تحلیل داده و داده‌کاوی با ابزارها و توابع توسعه داده‌شده برای این محیط وجود دارد، به‌نحوی که بسته‌های نرم‌افزاری، توابع، ابزارها و نحوه کار آن‌ها عمدتاً در مقالات علمی و به زبان آکادمیک تشریح شده و در پایگاه این زبان در دسترس است.
 - به دلیل متن‌باز بودن این امکان را فراهم می‌کند تا اجزای داخلی الگوریتم‌ها، مدل‌ها و توابع را مشاهده کرده و امکان تغییر آن‌ها را فراهم می‌کند.
 - امکان بسط، توسعه و بومی‌سازی بسته‌ها و توابع در این زبان فراهم است.
 - به دلیل ساختار ماتریسی که برای متغیرها در نظر می‌گیرد تطابق زیادی با نگاشت‌های مفهومی در فرایندهای متن‌کاوی و داده‌کاوی دارد و انجام عملیات کشف دانش در آن ملموس است.
 - به دلیل متن‌باز بودن گستره زیادی از توابع، ابزارها و راه‌حل‌های موردنیاز در حوزه‌های آماری، پردازش متن و مدل‌سازی را می‌توان در فضای وب جستجو کرده و از آن‌ها استفاده نمود.
 - امکان انجام پردازش‌های موازی و توزیع‌شده برای کار با حجم انبوه داده‌ها و پردازش‌های سنگین را دارد که این قابلیت برای داده‌ها و پردازش‌های مربوط به این پژوهش بسیار حیاتی است.
 - دارای راهنمای داخلی بسیار خوبی است و در فضای وب گروه‌های علمی زیادی پاسخگوی پرسش‌ها و نیازهای احتمالی خواهند بود.
- رایگان بوده و قابلیت نصب روی سیستم‌عامل‌های گوناگون را دارد و امکان تبادل داده و کد را روی نسخه‌های مختلف داراست. با توجه به میزان نسبتاً وسیع تصویر زیر حاوی بخشی از کد است که بصورت موازی از تمام کره‌های سی‌پی‌یوی کامپیوتر برای انجام موازی فرایند بهره می‌گیرد:



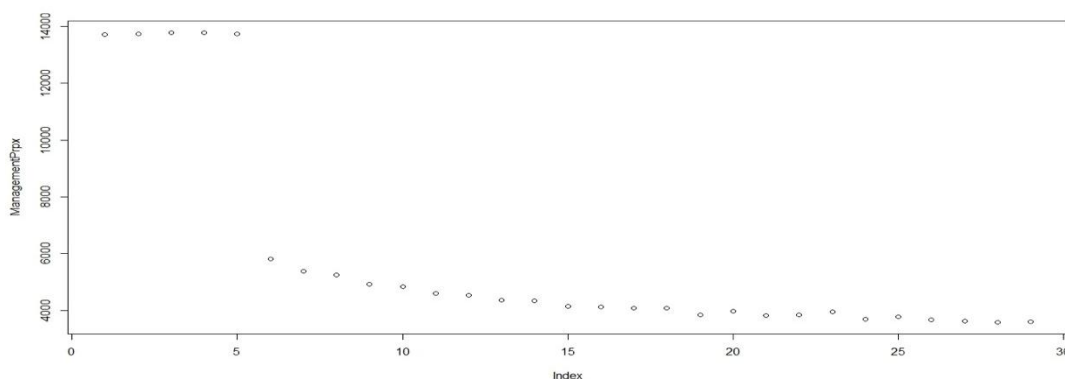
شکل ۲۳- نمای خروجی مدل سازی موضوعی

همانگونه که در شکل قابل نمایش است، مدل سازی با تمامی مقادیر نمونه مورد بررسی در این پژوهش انجام شده است. بر اساس پژوهشهای پیشین، میزان درست یا مناسب دسته ها می تواند از ملاک های زیادی پیروی کند. یکی از این معیارها میزان سرگشتگی (Perplexity) جواب در تعداد دسته های انتخابی است. پارامتر ورودی مهم دیگر LDA، مقدار k است که تعداد موضوعات مدنظر را نشان می دهد. روش های مختلفی برای تعیین تعداد موضوعات وجود دارد (Hall, Jurafsky and Manning 2008). مزیت انتخاب تعداد موضوعات زیاد، این است که تمامی حوزه های موضوعی پژوهش پوشش داده خواهد شد، از طرفی مزیت تعداد حوزه های محدود این است که خبرگان و سیاست گذاران درک بهتری از تعداد موضوعات محدود خواهند داشت و تفسیر و تحلیل موضوعات راحت تر خواهد بود (De Battisti, Ferrara and Salini 2015). از این رو انتخاب معیاری برای تعیین تعداد موضوعات مناسب بسیار ضروری است. برای این منظور از معیار سرگشتگی استفاده می شود. سرگشتگی یک معیار اندازه گیری در مدل های آماری است که برای بررسی میزان مناسب بودن یک توزیع احتمالی یا پیش بینی مدل احتمالی یک نمونه مورد استفاده قرار می گیرد و می توان از آن برای مقایسه مدل های احتمالی استفاده کرد. برای ارزیابی مدل سازی موضوعی از سرگشتگی اسناد مشاهده نشده استفاده می شود که تابعی نزولی از لگاریتم احتمال وقوع اسناد مشاهده نشده است و هر چه مقدار آن کمتر باشد نشان می دهد که مدل بهتر است و قابلیت تعمیم پذیری بالاتری دارد (D. M. Blei 2012) و مقدار بالاتر سرگشتگی نشان دهنده این است که عبارت ها در زمان یادگیری مدل به موضوعات مناسبی تخصیص نیافته اند (De Battisti, Ferrara and Salini 2015). بدین منظور

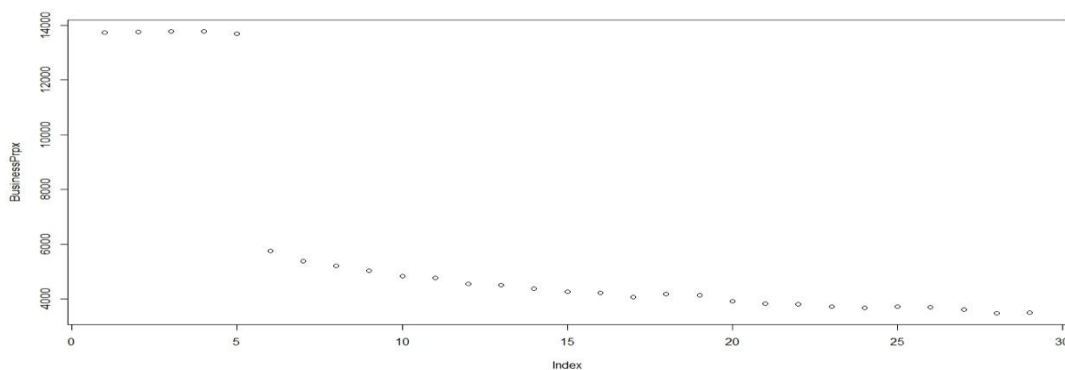
مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۷۴

کمترین میزان به صورت معمول برای انتخاب اندازه دسته مناسب برگزیده می‌شود. در شکل‌های زیر میزان شاخص برای دسته‌های ۲ تا ۳۰ برای داده‌های مختص به هر گروه پژوهشی نشان می‌دهد. پارامتر ورودی مهم LDA، مقدار k است که تعداد موضوعات مدنظر را نشان می‌دهد و از این رو انتخاب معیاری برای تعیین تعداد موضوعات مناسب بسیار ضروری است. برای این منظور از معیار سرگشتگی (Perplexity) استفاده می‌شود.

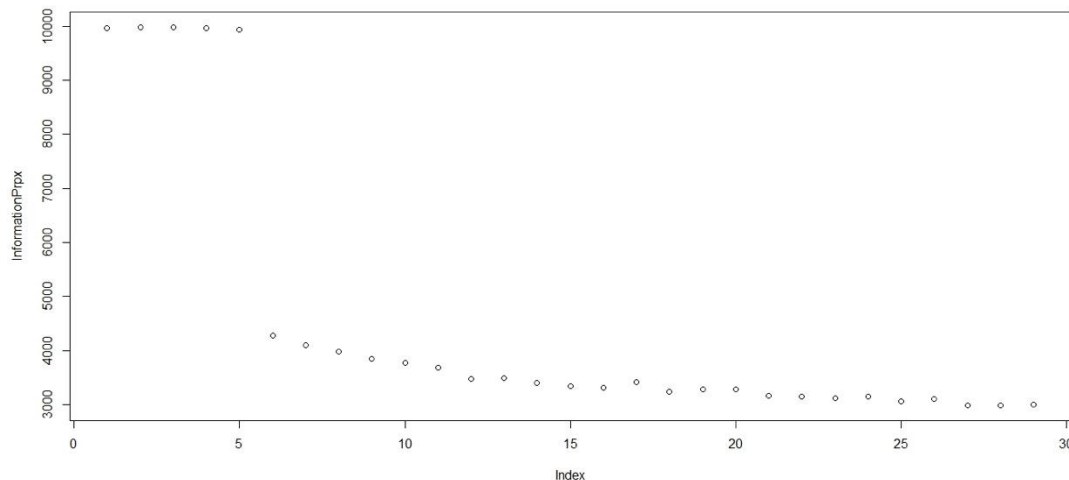
در صورتی که برای این آستانه، مقدار پایینی انتخاب شود، هر مستند توسط موضوعات زیادی پوشش داده خواهد شد و نتیجه این خواهد شد که اکثر موضوعات توسط اکثر مستندات تشریح خواهند شد و اگرچه عمده مستندات موجود در پیکره دست کم در یک موضوع قرار خواهند گرفت اما موضوع و مستند با ارتباطی بسیار ضعیف به یکدیگر متصل خواهند شد. در عوض در صورتی که این آستانه خیلی سخت گیرانه انتخاب شود، برخی از موضوعات با تعداد بسیار کمی از مستندات تشریح خواهند شد و بسیاری از مستندات نیز هیچ موضوعی را تشریح نخواهند کرد. برای انتخاب مقدار مناسب k ، مدل با استفاده از مقادیر مختلف k اجرا شد و مقدار سرگشتگی برای هر اجرا محاسبه شد. نتایج این انتخاب در زیر آمده است.



شکل ۲۴- میزان انحراف سرگشتگی با توجه به اندازه دسته‌بندی در گروه پژوهشی مدیریت فناوری اطلاعات



شکل ۲۵- میزان انحراف سرگشتگی با توجه به اندازه دسته‌بندی در گروه پژوهشی کسب و کار الکترونیک



شکل ۲۶- میزان انحراف سرگشتگی با توجه به اندازه دسته‌بندی در گروه پژوهشی سیستم‌های اطلاعاتی به صورت اتفاقی هر سه گروه پژوهشی دارای ۶ حوزه تفکیک شده خواهند بود که در بخش بعد توصیف می‌شوند.

۳-۵. دسته‌بندی حوزه‌های علمی بر اساس مدل‌سازی موضوعی کلیدواژه‌ها

همانگونه که از شکل‌های بخش فوق بر می‌آید، تعداد دسته‌های مورد انتظار از ۶ به بعد با افزایش تعداد به ندرت باعث کاهش در شاخص می‌شود. این کاهش تدریجی به علت تنوع بالا و حجم زیاد کلیدواژه‌های مورد استفاده است. زیرا با تعداد دسته‌بندی بالا می‌توان حجم زیادی از آنها را در یک محیط پراکنده در یک گروه قرار داد. بر این اساس، تعداد گروه ۶ که نقطه شکست نمودار است، به عنوان تعداد دسته‌های مدل‌سازی موضوعی انتخاب می‌شود. بر این اساس ۶ دسته ارتباطات کلیدواژه‌های فناوری اطلاعات برای تشکیل ۶ موضوع در جدول زیر (نمایش ده کلیدواژه پر احتمال برای هر موضوع) ترسیم می‌شوند. نامگذاری دسته‌ها با توجه به کلیدواژه‌های استخراج شده و با نظر خبرگان انتخاب شده است. لازم به ذکر است در فصل قبل کلیات دسته‌ها از پایگاه استخراج ولی افزاز آنها به گروه‌های پژوهشی انجام پذیرفت.

جدول ۳-۱۰- دسته‌بندی موضوعات با روش مدل‌سازی موضوعی گروه پژوهشی مدیریت فناوری اطلاعات

گروه پژوهشی	دسته	نام دسته	کلیدواژه	وزن نرمال شده
مدیریت فناوری اطلاعات	۱	IOT	internet of things	6.76%
			privacy	6.73%
			blockchain	6.71%
			physical layer security	6.69%
			cryptography	6.68%
			quality of service	6.68%
			authentication	6.68%

گروه پژوهشی	دسته	نام دسته	کلیدواژه	وزن نرمال شده
			resource management	6.66%
			cloud computing	6.65%
			wireless communication	6.65%
			wireless sensor networks	6.64%
			wireless networks	6.63%
			servers	6.62%
			information security	6.61%
			sensors	6.61%
	۲	cloud computing	cloud computing	6.79%
			machine learning	6.73%
			edge computing	6.72%
			data privacy	6.69%
			data models	6.67%
			computational modeling	6.67%
			resource management	6.66%
			task analysis	6.66%
			heuristic algorithms	6.66%
			fog computing	6.65%
			differential privacy	6.63%
			privacy preservation	6.63%
			intrusion detection	6.62%
			computer architecture	6.61%
			monitoring	6.61%
	۳	wireless communication	privacy	6.80%
			wireless communication	6.72%
			data privacy	6.70%
			optimization	6.69%
			authentication	6.68%
			security	6.67%
			task analysis	6.66%
			wireless networks	6.66%
			physical layer security	6.66%
			tools	6.64%
			wireless sensor networks	6.63%
			encryption	6.62%
			information security	6.62%
			computer architecture	6.62%
			delays	6.62%

گروه پژوهشی	دسته	نام دسته	کلیدواژه	وزن نرمال شده
	۴	modelling – (service quality – industrial optimization)	deep learning	6.73%
			protocols	6.73%
			machine learning	6.70%
			communication system security	6.69%
			computational modeling	6.69%
			data models	6.67%
			routing	6.67%
			task analysis	6.66%
			differential privacy	6.65%
			resource allocation	6.64%
			optimization	6.64%
			training	6.64%
			quality of service	6.63%
			neural networks	6.63%
			load balancing	6.63%
	۵	security	privacy	6.82%
			encryption	6.74%
			authentication	6.73%
			security	6.73%
			data privacy	6.68%
			5g mobile communication	6.66%
			authentication	6.65%
			network security	6.65%
			privacy preserving	6.63%
			data security	6.63%
			cryptography	6.63%
			image encryption	6.62%
			surveillance	6.61%
			security and privacy	6.61%
			privacy protection	6.60%
	۶	machine learning (application – standardization)	machine learning	6.73%
			wireless sensor networks	6.72%
			protocols	6.70%
			fog computing	6.69%
			data models	6.69%
			optimization	6.68%
			cloud computing	6.68%
			computer architecture	6.67%
			training	6.66%

گروه پژوهشی	دسته	نام دسته	کلیدواژه	وزن نرمال شده
کسب و کار الکترونیک			energy harvesting	6.65%
			access control	6.64%
			reliability	6.63%
			task analysis	6.63%
			artificial intelligence	6.62%
			smart cities	6.61%
	۱	online learning	social media	6.86%
			online learning	6.79%
			sentiment analysis	6.69%
			online teaching	6.69%
			distance learning	6.69%
			data mining	6.67%
			big data	6.66%
			twitter	6.65%
			online education	6.63%
			online	6.62%
			facebook	6.62%
			blended learning	6.61%
			media	6.61%
			text mining	6.61%
			blogs	6.61%
	۲	Health	mental health	6.72%
			survey	6.69%
			prediction algorithms	6.68%
			artificial intelligence	6.67%
			medical education	6.67%
			students	6.66%
			technology	6.66%
			anxiety	6.66%
			gender	6.66%
			public health	6.66%
			sentiment analysis	6.66%
			natural language processing	6.66%
			university students	6.65%
			optimization	6.65%
			content analysis	6.65%
	۳	user experience (media – NLP)	data mining	6.80%
			user experience	6.73%
			sentiment analysis	6.72%
			text mining	6.67%
			classification	6.66%

گروه پژوهشی	دسته	نام دسته	کلیدواژه	وزن نرمال شده
			fake news	6.66%
			content analysis	6.66%
			task analysis	6.65%
			decision tree	6.65%
			data mining	6.64%
			media	6.64%
			social networks	6.63%
			natural language processing	6.63%
			feature extraction	6.63%
			artificial intelligence	6.62%
	۴	augmented reality	internet	6.75%
			data mining	6.74%
			text mining	6.67%
			quality of experience	6.67%
			heuristic algorithms	6.67%
			augmented reality	6.66%
			feature extraction	6.66%
			reinforcement learning	6.66%
			sentiment analysis	6.66%
			satisfaction	6.66%
			decision tree	6.66%
			emergency remote teaching	6.64%
			user experience	6.64%
			knowledge	6.64%
			virtual reality	6.63%
	۵	social media	social media	6.84%
			data mining	6.74%
			engagement	6.67%
			facebook	6.67%
			sentiment analysis	6.67%
			instagram	6.66%
			twitter	6.66%
			motivation	6.65%
			media	6.65%
			big data	6.65%
			data mining	6.63%
			adolescents	6.63%
			self-regulated learning	6.63%
			twitter	6.63%

گروه پژوهشی	دسته	نام دسته	کلیدواژه	وزن نرمال شده
	۶	deep learning	china	6.63%
			covid-19	6.79%
			deep learning	6.73%
			data mining	6.72%
			sentiment analysis	6.70%
			user experience	6.69%
			natural language processing	6.68%
			task analysis	6.65%
			data models	6.65%
			training	6.64%
			feature extraction	6.64%
			reinforcement learning	6.63%
			optimization	6.63%
			machine learning	6.63%
			online learning	6.61%
			pandemic	6.61%
سیستم‌های اطلاعاتی	۱	IOT	internet of things	6.89%
			edge computing	6.75%
			machine learning	6.69%
			cloud computing	6.69%
			optimization	6.67%
			deep learning	6.67%
			wireless sensor networks	6.65%
			resource management	6.65%
			sensors	6.65%
			wireless communication	6.64%
			servers	6.63%
			delays	6.61%
			security	6.61%
			computer architecture	6.61%
			artificial intelligence (ai)	6.60%
	۲	AI (big data – deep learning)	artificial intelligence	6.91%
			deep learning	6.77%
			machine learning	6.73%
			semi-supervised learning	6.71%
			supervised learning	6.70%
			machine learning	6.70%

گروه پژوهشی	دسته	نام دسته	کلیدواژه	وزن نرمال شده
			convolutional neural networks	6.65%
			internet of things	6.64%
			sentiment analysis	6.62%
			deep learning	6.62%
			big data	6.61%
			image processing	6.60%
			classification	6.59%
			transfer learning	6.59%
			fuzzy logic	6.58%
	۳	machine learning	text mining	6.79%
			machine learning	6.77%
			deep learning	6.71%
			recommender systems	6.71%
			natural language processing	6.67%
			crowdsourcing	6.65%
			ontology	6.65%
			social media	6.65%
			collaborative filtering	6.64%
			supervised machine learning	6.64%
			random forest	6.63%
			data mining	6.63%
			semi-supervised learning	6.63%
			clustering	6.62%
			trust	6.62%
	۴	security	security	6.80%
			blockchain	6.77%
			privacy	6.76%
			authentication	6.72%
			servers	6.69%
			encryption	6.67%
			protocols	6.64%
			security	6.64%
			access control	6.64%
			cryptography	6.63%
			data privacy	6.63%
			federated learning	6.61%
			computer architecture	6.61%
			authentication	6.60%
			peer-to-peer	6.59%

وزن نرمال شده	کلیدواژه	نام دسته	دسته	گروه پژوهشی		
	computing					
6.78%	feature extraction	semantics	۵			
6.75%	task analysis					
6.71%	semantics					
6.68%	data models					
6.67%	image recognition					
6.67%	convolutional neural network					
6.66%	neural networks					
6.66%	semi-supervised learning					
6.65%	feature extraction					
6.65%	image segmentation					
6.64%	kernel					
6.63%	visualization					
6.62%	computational modeling					
6.62%	predictive models					
6.61%	correlation					
6.76%	link prediction				recommender system	۶
6.72%	recommender system					
6.70%	topic modeling					
6.69%	feature selection					
6.69%	attention mechanism					
6.67%	multi-view learning					
6.67%	natural language processing					
6.66%	genetic algorithm					
6.65%	multi-view clustering					
6.64%	complex networks					
6.64%	transfer learning					
6.64%	community detection					
6.63%	algorithms					
6.63%	heuristic algorithms					
6.61%	medical services					

۶-۳. جمع‌بندی و نتیجه‌گیری

در این فصل به بررسی و مطالعه حوزه‌های (موضوعات) فناوری اطلاعات در گروه‌های پژوهشی پرداخته شد. در این بررسی از روش‌های مختلفی برای دستیابی به نتایج استفاده شد. در بخش اول و در انتخاب مجلات پیشرو در زمینه فناوری اطلاعات، با وجود ذکر در پیشنهاد بر استفاده از مجلات پر استناد و پیشرو (اچ ایندکس بالا) برای تشکیل سبد واژگان کلیدی اولیه، از نظرات خبرگان نیز استفاده شده و ۱۱ مورد مجله دیگر به فهرست جستجوی اولیه افزون شدند. این گام باعث شد سبد واژگان اولیه معتبر گردد. همچنین در فاز دوم علی‌رغم وجود دسته‌بندی موضوعی در پایگاه مورد مطالعه (WOS) و دسته‌بندی کلیدواژه‌های پرکاربرد در این دسته‌ها در این پژوهش، روش مدل‌سازی موضوعی نیز مورد توجه قرار گرفت و محققین فارغ از نتایج بدست آمده در توزیع کلیدواژه‌ها در دسته‌بندی‌های مذکور، مجدد دسته‌بندی را بر اساس ادبیات ذکر شده در بخش قبل به انجام رساندند. با توجه به افت ضریب سرگشتگی در عدد ۶ در هر گروه پژوهشی، ۶ دسته‌بندی برای موضوعات ترکیبی کلیدواژه‌های مرسوم انتخاب شده و این دسته‌بندی در بخش قبل به نمایش درآمد.

بدین ترتیب در هر گروه پژوهشی دسته‌های زیر بعنوان حوزه‌های پژوهش در سال اخیر منتخب شدند:

• گروه پژوهشی مدیریت فناوری اطلاعات

- اینترنت اشیا
- رایانش ابری
- ارتباطات بی‌سیم
- مدل‌سازی (کیفیت خدمات و بهینه‌یابی صنعتی)
- امنیت
- یادگیری ماشین (استانداردسازی و پیاده‌سازی)

• گروه پژوهشی کسب و کار الکترونیک

- یادگیری برخط
- سلامت
- تجربه‌ی کاربری (رسانه و پردازش زبان طبیعی)
- واقعیت مجازی
- شبکه‌های اجتماعی
- یادگیری عمیق

• گروه پژوهشی سیستم‌های اطلاعاتی

- اینترنت اشیا

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۸۴

○ هوش مصنوعی

○ یادگیری ماشین

○ امنیت

○ سمانتک

○ سیستم‌های پیشنهاددهنده

در فصل بعد تعداد مدارک بازیابی شده در هر دسته در هر سه گروه پژوهشی مورد مطالعه قرار خواهد گرفت.

فصل چہارم

تحلیل شکاف

۴-۱. مقدمه

در دهه‌های اخیر، شاهد افزایش چشم‌گیر تعداد پژوهش‌هایی هستیم که در زمینه‌های مختلف علوم صورت گرفته است، اما رشد کمی پژوهش‌ها، لزوماً متناسب نبوده است و در بسیاری از موارد دستاوردهای این پژوهش‌ها منجر به رفع مشکلات جامعه نشده و به‌ندرت مشاهده شده است که نتایج و یافته‌های پژوهشی برای بهبود اوضاع سیاسی، اجتماعی، اقتصادی و فرهنگی جامعه به کار گرفته شود. در این میان شاید رشته‌های مهندسی و پژوهش‌های کاربردی این گروه وضعیت بهتری را نسبت به دیگر گروه‌های آموزشی داشته باشند (قنادی نژاد، حیدری و چینی پرداز ۱۳۹۷).

پیش از این پژوهشگران برای بررسی و تحلیل آثار پژوهشی یک دوره زمانی، ملزم به استفاده از روش‌های دشوار کیفی بودند ولی با ظهور فناوری‌های نوین تحلیل اطلاعات و نیز توسعه پایگاه‌های اطلاعات علمی این امکان برای پژوهشگران پدید آمد تا با پژوهش‌های کارآمد و متنوعی روی داده‌های وسیع انجام دهند و بتوانند بینش مناسبی از تعاریف، مرزبندی‌ها، روندها، کاستی‌ها و نیازهای پژوهشی یک حوزه علمی ارائه نمایند (خاصه و سهیلی، ترسیم چشم‌انداز پژوهش در علم‌سنجی و حوزه‌های سنجشی وابسته ۱۳۹۷). استفاده پژوهشگران از این پایگاه‌ها به‌عنوان مرجعی برای شناسایی آخرین تحقیقات علمی انجام شده در حوزه‌های مختلف علوم گسترش یافته است. از این‌رو کاربرد داده‌کاوی و متن‌کاوی به‌عنوان زیرشاخه‌ی مهمی از آن به‌منظور کشف دانش از منابع موجود و ابزاری برای به دست آوردن الگوهای جدیدی از اطلاعات روزبه‌روز گسترده‌تر شده است. از آنجاکه کاربران این پایگاه‌ها عمدتاً پژوهشگران و افراد متخصص هستند، در کنار محتوای اطلاعاتی پایگاه‌های علمی، رفتار کاربران در تعامل با این پایگاه‌ها نیز از اهمیت ویژه‌ای برخوردار است و تحلیل رفتار جستجوگران اطلاعات می‌تواند اطلاعات ارزشمندی از نیازها و تقاضاهای واقعی آن‌ها

را در اختیار قرار دهد. با شناخت دقیق این تقاضاها و بررسی و تحلیل آن و نیز مقایسه آن با اطلاعات و دانش موجود می‌توان کمبودهای منابع اطلاعاتی را شناسایی نموده و در جهت تأمین آن اقدام نمود. مطالعات گسترده‌ای در حوزه مدیریت پژوهش با به‌کارگیری ابزارهای متن‌کاوی یا داده‌کاوی صورت پذیرفته است که هر یک اهداف خاص خود را دنبال کرده‌اند، این پژوهش‌ها با استفاده از تحلیل داده‌های پایگاه اطلاعاتی در تلاش هستند تا وضعیت فعلی پژوهش‌ها را رصد و تحلیل نموده و روندهای پژوهش در آینده را مشخص کنند یا با رصد حوزه‌های مختلف پژوهشی ارتباطات ایجاد شده بین حوزه‌ها، رشد یا زوال حوزه‌های مختلف، تأثیر انتشارات علمی در یک حوزه روی شکل‌های دیگر انتشارات همان حوزه، بخش‌بندی و طبقه‌بندی حوزه‌های بزرگ‌تر به شاخه‌های فرعی و ... را شناسایی کنند. سؤالات و مفاهیم مهمی که در حوزه مدیریت پژوهش وجود دارد و داده‌کاوی و متن‌کاوی امکان پاسخگویی بالقوه به این سؤالات را دارند در زیر ادامه آمده است:

- شناسایی علوم، فناوری و پژوهش‌هایی که در حال انجام هستند،
- شناسایی افرادی که این پژوهش‌ها را انجام می‌دهند،
- شناسایی زیردامنه‌ها و موضوعات که در یک حوزه پژوهشی مطرح هستند،
- شناسایی مکان‌هایی که این پژوهش‌ها در آنجا انجام می‌شود،
- شناسایی علوم، فناوری و پژوهش‌هایی که مغفول مانده است،
- شناسایی حوزه‌هایی که تمرکز اصلی روی آن‌هاست و تعیین اینکه چه ارتباطی بین این حوزه‌ها وجود دارد،
- شناسایی سمت و سوی پژوهش‌هایی که در آینده انجام خواهد شد،
- و ...

تحلیل و بررسی پژوهش با ابزارها و روش‌های گوناگون اهداف مختلفی را مدنظر دارد که از مهم‌ترین آن‌ها می‌توان به تحلیل روند پژوهش‌ها، شناسایی کاستی‌های پژوهش، نیازسنجی و اولویت‌بندی پژوهش‌ها اشاره کرد. سیاست‌گذاران پژوهشی و صاحبان صنعت با شناخت کامل از کاستی پژوهشی این امکان را خواهند داشت تا در برنامه‌ریزی و تصمیم‌گیری‌های کلان خود حوزه‌هایی که دچار کمبود منابع علمی و پژوهشی هستند را مدنظر قرار داده و برای پر کردن این کاستی چاره‌اندیشی کنند.

در این فصل با استفاده از دسته‌بندی‌های انجام پذیرفته در فصل قبل و همچنین بررسی فعالیت‌های پژوهشی انجام پذیرفته در هر گروه پژوهشی، سعی می‌کنیم تطبیق محتوایی بین مطالعات انجام شده و دسته‌بندی‌های استخراج شده (در فصل قبل) داشته باشیم.

در نهایت تلاش داریم نقاط کمتر پرداخته شده در هر گروه پژوهشی را مشخص نموده و گذری بر دلایل این اختلاف داشته باشیم.

۴-۲. استخراج عناوین فعالیت‌های پژوهشی

همان‌طور که پیش‌تر نیز بیان شد، شناسایی کاستی‌های پژوهش اولین قدم در نیازسنجی پژوهش است و در بسیاری از پژوهش‌ها منظور از این نیازسنجی، شناسایی کاستی‌های حوزه موردنظر است. به‌طور کلی رویکردهای مطرح در زمینه شناسایی نیازهای پژوهشی و اولویت‌بندی آن‌ها متعدد است که از مهم‌ترین این رویکردها می‌توان به رویکرد آسیب‌شناختی، رویکرد موضوعی/ محتوایی، رویکرد تقاضامحور، رویکرد عرضه‌محور، رویکرد نظام‌مند و رویکرد اداره‌گری/مدیریتی اشاره کرد. این رویکردها لزوماً در تضاد با یکدیگر نیستند و هر یک از آن‌ها از یک بعد به مسئله نیازسنجی و اولویت‌یابی پژوهش نگریسته‌اند. در این میان رویکرد موضوعی/محتوایی بر ماهیت و ساختار دانش موجود در منابع موردبررسی دلالت دارد. به عبارتی این رویکرد بر مبنای شاخه‌ها و زیرشاخه‌های علمی شناخته‌شده در حوزه موردنظر بنا شده است. این شاخه‌ها می‌توانند از منابع مورد بررسی استخراج شده باشند و یا از شاخه‌های موضوعی رشته موردنظر در دیگر منابع (مانند دانشگاه‌ها یا پایگاه‌های علمی) به دست آمده باشند. این رویکرد غنای هر یک از موضوعات و شاخه‌ها را بررسی و تحلیل می‌کند (فتحی‌واجارگاه ۱۳۸۹).

جدول ۴-۱۱- رویکردهای مختلف نیازسنجی و اولویت‌یابی پژوهش (فتحی و اجارگاه ۱۳۸۹)

رویکرد نیازسنجی	مراحل کلی	نقاط قوت	نقاط ضعف
۱- سبب-نتیجه-رابطه	شناسایی مشکلات و آسیب‌های جامعه تعیین اولویت‌های پژوهشی برای حل مسئله	بهره‌گیری از روش حل مسئله کمک به رفع مشکلات و نارسایی‌ها	نیازمند نظر خبرگان برای شناسایی مشکلات و آسیب‌ها غیرقابل ارزیابی در کوتاه‌مدت
۲- موضوعی/محتوایی	شناسایی و تحلیل شاخه‌ها و زیرشاخه‌ها تعیین زیرشاخه‌های غنی یا فقیر پژوهشی	مبتنی بر محتوای پژوهشی و واقعیت‌ها امکان تحلیل کلی و جزئی زیرشاخه‌ها	مبتنی بر پژوهش‌های گذشته نیازمند پایگاه اطلاعاتی جامع و بروز پیچیدگی و مشکلات پردازش زبان طبیعی
۳- فناصم‌محور	انتخاب گروهی از متقاضیان یا کاربران پژوهش نظرسنجی متقاضیان و کاربران تعیین نیازها و اولویت‌ها بر مبنای نتایج نظرسنجی	مقبولیت بالا از نظر متقاضیان پژوهش شناسایی نیازهای متقاضیان جلوگیری از انجام پژوهش‌های کم‌کاربرد	نیازمند سطح بالایی از دانش در کاربران نامناسب برای اولویت‌یابی بلندمدت پژوهش عدم توجه کافی به پژوهش‌های بنیادی
۴- ریشه‌محور	محاسبه میزان پژوهش‌های انجام شده در شاخه‌ها و زیرشاخه‌های مختلف تعیین نیازها و اولویت‌ها بر مبنای تحلیل پژوهش‌های موجود و نظر خبرگان	مبتنی بر محتوای پژوهشی و واقعیت‌های موجود مناسب برای برنامه‌ریزی‌های کلان	نگاه کاملاً کمی و عدم توجه به نیاز کاربران مبتنی بر پژوهش‌های گذشته نیازمند پایگاه اطلاعاتی جامع و بروز وابستگی به نظر خبرگان
۵- فناصم‌مند	تعیین معیارها و شاخص‌های اثربخشی پژوهش‌ها ارزیابی پژوهش‌ها بر اساس این معیارها تعیین اولویت‌های پژوهشی بر مبنای بازده پژوهش‌ها	توجه به کارایی و اثربخشی پژوهش سنجش با شاخص‌های دقیق کارایی پژوهش	عدم امکان کمی کردن برخی از برون‌دادهای پژوهشی پیچیدگی وزن‌دهی به معیارها عدم توجه کافی به پژوهش‌های بنیادی
۶- اداره‌گری/مدیریتی	تعیین ساختار و سلسله‌مراتب مدیریتی تعیین اولویت‌های کلان توسط مدیران رده‌بالا تعیین اقدامات و طرح‌های پژوهشی بر مبنای اولویت‌ها	ساختارمند و توجه به برنامه‌های کلان امکان کنترل و نظارت بالا	عدم توجه به نیازهای پژوهشگران عدم توجه به کیفیت و کارایی پژوهش انعطاف‌پذیری پایین

مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۹۰

طرح‌های پژوهشی همکاران ایرانداک به تفکیک پژوهشکده‌های مختلف پس از تصویب پیشنهاد در سایت و در اختیار همگان قرار می‌گیرد. در شکل زیر نحوه دسترسی به این اطلاعات آمده است. با کلیک کردن بر روی تصویر هر طرح پژوهشی مشخصات چکیده و کلیدواژه‌های آن در اختیار محقق قرار می‌گیرد:

شکل ۲۷- نحوه دسترسی به مشخصات طرح‌های پژوهشی در ایرانداک

با توجه به تصویر فوق، طرح‌های پژوهشی زیر و با توجه به پیشنهاد این طرح، از سال ۱۳۹۵ تا ۱۴۰۱ مورد استخراج قرار گرفته است. در جداول زیر می‌توان مشخصات هر طرح پژوهشی را مشاهده نمود:

جدول ۱۲- مشخصات طرح‌های پژوهشی ایرانداک

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
مدیریت فناوری اطلاعات	A01	طراحی مدل کسب و کار پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)	حمیدرضا خدمتگزار	۱۳۹۸	مدل کسب و کار، ارزش ادراک شده، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، تحلیل ذینفعان

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	A02	ارائه برنامه عملیاتی ارتقاء و توسعه بهره‌وری در پژوهشگاه علوم و فناوری اطلاعات ایران (ایرنداک)	محمدجواد ارشادی	۱۴۰۰	بهره‌وری، برنامه عملیاتی، عارضه‌یابی، توسعه بهره‌وری، بهبود مستمر، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرنداک)
	A03	اندازه‌گیری و تحلیل شاخص‌های کیفیت داده در فرایند ثبت پایان‌نامه / رساله‌های دانش‌آموختگان داخل کشور	محمدجواد ارشادی	۱۳۹۸	کیفیت داده، نمودار کنترل، سامانه ثبت پایان‌نامه/رساله، قابلیت فرایند
	A04	طراحی و مستندسازی فرایندهای پژوهشگاه علوم و فناوری اطلاعات ایران در حوزه‌های آموزش، پژوهش، جذب، ترفیع، و ارتقا	محمدجواد ارشادی	۱۳۹۷	رویکرد فرایندی، مدیریت فرایندها، مستندسازی، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرنداک)
	A05	طراحی چارچوب مفهومی بومی تعالی فناوری اطلاعات سازمانی	آرمان ساجدی‌نژاد	۱۳۹۷	مدیریت فناوری اطلاعات، چارچوب بهبود فناوری اطلاعات، چارچوب مفهومی، تعالی فناوری اطلاعات سازمانی
	A06	برنامه‌ریزی منابع انسانی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرنداک) ۱۳۹۷-۱۴۰۰	آرمان ساجدی‌نژاد	۱۳۹۷	برنامه‌ریزی نیروی انسانی، مدیریت استراتژیک، مدیریت توسعه منابع انسانی، طرح‌های منابع انسانی
	A07	طراحی مدل کیفیت فراداده در سامانه ثبت پایان‌نامه/رساله‌های دانش‌آموختگان داخل و خارج کشور	محمدجواد ارشادی	۱۳۹۶	فراداده، کنترل کیفیت، طرح کیفیت، سامانه‌های اطلاعاتی تحقیقاتی، سامانه ثبت پایان‌نامه/رساله
	A08	بازطراحی کارکرد مدیریت اطلاعات علمی و فناورانه پژوهشگاه علوم و فناوری اطلاعات ایران در بافت دولت الکترونیکی	آرمان ساجدی‌نژاد	۱۳۹۶	منابع اطلاعاتی، دولت الکترونیک، خدمات الکترونیکی، مدیریت اطلاعات، زنجیره ارزش پژوهش

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	A09	بررسی کارکردهای الکترونیکی مدیریت اطلاعات علمی و فناورانه	آرمان ساجدی‌نژاد	۱۳۹۵	فناوری اطلاعات و ارتباطات، مدیریت اطلاعات، مدیریت منابع
	A10	طرح‌ریزی و استقرار نظام تضمین کیفیت در سامانه‌های گردآوری و ثبت، سازمان‌دهی، و اشاعه اطلاعات پایان‌نامه‌ها/ رساله‌های دانش‌آموختگان داخل کشور	محمدجواد ارشادی	۱۳۹۵	تضمین کیفیت، سامانه‌های اطلاعاتی تحقیقاتی، ممیزی کیفیت
	A11	بررسی روند پژوهش در زمینه عوامل موثر بر پذیرش و کاربرد فناوری اطلاعات در ایران	سیروس علیدوستی	۱۳۹۷	چارچوب تحلیل و بازنمایی اطلاعات پایان‌نامه‌ها، رساله‌ها و مقاله‌ها، عوامل موفقیت پروژه‌های فناوری اطلاعات، عوامل شکست پروژه‌های فناوری اطلاعات، تحلیل روند
	A12	ارائه یک چارچوب استاندارد برای ثبت و تحلیل آماری مدارک نا هم‌خوان در سامانه ثبت پایان‌نامه‌ها و رساله‌ها	محمدجواد ارشادی	۱۳۹۹	کیفیت داده، کنترل کیفیت، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، سامانه ثبت پایان‌نامه/رساله، تحلیل نا هم‌خوانی
	A13	مقایسه روند پژوهش‌ها در حوزه فناوری اطلاعات در پایان‌نامه‌ها و رساله‌های کشور و روند جهانی آن با استفاده از روش تحلیل متون	آرمان ساجدی‌نژاد	۱۳۹۹	مدل‌سازی موضوعی، متن کاوی، پایگاه اطلاعات علمی ایران (گنج)، فناوری اطلاعات، ایرانداک، روند پژوهش
	A14	ارائه طرح نمونه‌برداری پذیرش در فرایند ویراستاری و نمایه‌سازی پایان‌نامه‌ها و رساله‌ها	محمدجواد ارشادی	۱۳۹۹	نمایه‌سازی، کنترل کیفیت، کیفیت داده، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، نمونه‌برداری، سازماندهی و تحلیل اطلاعات

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	A15	طراحی مدل کسب و کار سامانه‌ها و خدمات پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) در سطوح پنج‌گانه مأموریتی	حمیدرضا خدمتگزار	۱۴۰۰	مدل کسب و کار، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، سامانه‌ها و خدمات، سطوح مأموریتی
	A16	تدوین سند برنامه راهبردی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) ۱۴۰۴-۱۴۰۰	آرمان ساجدی‌نژاد	۱۳۹۹	برنامه‌ریزی راهبردی، ایرانداک، استراتژی پدیداری، استراتژی آگاهانه، برنامه‌ریزی استراتژیک، استراتژی نوآورانه
	A17	مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی پژوهشکده فناوری اطلاعات ایرانداک با موضوعات بین‌المللی	آرمان ساجدی‌نژاد	۱۴۰۱	مدل‌سازی موضوعی، ایرانداک، مدیریت فناوری اطلاعات، کسب و کار الکترونیک، سیستم‌های اطلاعاتی
	A18	طراحی چارچوب تامین سامانه‌ها، خدمات و تجهیزات فناوری اطلاعات پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)	حمیدرضا خدمتگزار	۱۴۰۱	ایرانداک، سامانه‌ها، خدمات، تجهیزات فناوری اطلاعات، چارچوب تامین
	A19	ارائه راهکارهای اجرایی برای بهبود کیفیت اطلاعات پایان‌نامه/رساله‌ها در سامانه ثبت بر پایه تکنیک‌های داده‌کاوی	محمدجواد ارشادی	۱۴۰۱	کیفیت داده، کیفیت اطلاعات، کنترل کیفیت، بهبود مستمر، داده‌کاوی، خوشه‌بندی
کسب و کار الکترونیک	B01	طراحی چارچوب ارزیابی جایزه فناوری اطلاعات برتر در همایش ملی مدیران فناوری اطلاعات	علی نعیمی صدیق	۱۳۹۸	جایزه فاب، همایش مدیران فناوری اطلاعات، چارچوب ارزیابی، روش گروه کانونی، روش‌های تصمیم‌گیری چند معیاره
	B02	تحلیل عرضه و تقاضای پایان‌نامه‌ها و رساله‌های حوزه انرژی در پایگاه گنج	محمد ربیعی	۱۴۰۰	انرژی، عرضه و تقاضای پژوهش، مدل‌سازی موضوعی، متن‌کاوی، پایگاه گنج، ایرانداک

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	B03	ارائه چارچوب ارزیابی شرکت‌های تحول‌آفرین در حوزه فناوری اطلاعات	علی نعیمی صدیق	۱۴۰۰	همایش ملی مدیران فناوری اطلاعات، شرکت‌های تحول‌آفرین، روش تصمیم‌گیری چند معیاره، دسته‌بندی شرکت‌ها
	B04	ارائه راهکارهایی برای بهبود پشتیبانی فنی از مشتریان سامانه‌های پرکاربرد پژوهشگاه علوم و فناوری اطلاعات ایران	فرهاد شیرانی	۱۴۰۰	رضایت کاربران، ایرانداک، پشتیبانی فنی از مشتریان، سامانه گنج، سامانه همانندجو، سامانه ثبت، سامانه پیشینه پژوهش، ارتباط با مشتری، پشتیبانی کاربران
	B05	پیشنهاد پرتفوی دوره‌های آموزشی حضوری و مجازی پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)	محمد ربیعی	۱۴۰۱	آموزش مجازی، آموزش حضوری، پرتفوی دوره‌های آموزشی، رضایت دانش‌پذیران
	B06	طراحی مدل کسب و کار و قیمت‌گذاری خدمات همانندجو با استفاده از نظریه بازی	علی نعیمی صدیق	۱۳۹۷	مدل کسب و کار، قیمت‌گذاری، سامانه همانندجو، تقاضای وابسته به قیمت، نظریه بازی، بازار انحصاری، بازار چندقطبی
	B07	مطالعه تطبیقی قیمت‌گذاری سامانه خدمات همانندجو	علی نعیمی صدیق	۱۳۹۶	قیمت‌گذاری، سامانه همانندجو، خدمات نرم‌افزاری
	B08	طراحی، ساخت و پیاده‌سازی سامانه عرضه و تقاضای (ساعت) پژوهش	محمد ربیعی	۱۳۹۵	عرضه و تقاضای پژوهش، پایان‌نامه‌ها و رساله‌ها، نیاز پژوهشی، ارتباط صنعت و دانشگاه
	C01	بررسی موتور جست‌وجوی گنج به منظور بهبود بازیابی اطلاعات در زبان فارسی	مرضیه زرین‌بال ماسوله	۱۳۹۸	بازیابی اطلاعات، سامانه گنج، موتور جست‌وجو، آپاچی سولر، آپاچی لوسن
	C02	طراحی و پیاده‌سازی سامانه دیداری‌سازی تعاملی داده‌های تاریخی علوم، مورد مطالعه: تاریخچه حوزه هوش مصنوعی	مرضیه زرین‌بال ماسوله	۱۳۹۹	دیداری کردن نتایج جست‌وجو، دیداری‌سازی تعاملی، هوش مصنوعی، داده‌های تاریخی، سامانه مبتنی بر وب
سیستم‌های اطلاعاتی					

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	C03	ارائه چارچوبی برای ارزیابی پایگاه وب پژوهشگاه‌ها از منظر معماری اطلاعات: مطالعه موردی وبگاه پژوهشگاه علوم و فناوری اطلاعات ایران	امیرحسین صدیقی	۱۳۹۸	معماری اطلاعات، ارزیابی پایگاه وب، فناوری اطلاعات، مطالعه موردی
	C04	تهیه مجموعه داده استاندارد فارسی برای مساله استخراج خودکار کلیدواژه از اسناد علمی	آزاده محبی	۱۴۰۰	استخراج خودکار کلیدواژه، یادگیری ماشین، مجموعه داده، بازیابی اطلاعات، روش‌های رتبه‌بندی کلیدواژه‌ها، روش RAKE، روش TEXTRANK، روش TFIDF
	C05	طراحی و پیاده‌سازی افزونه پیش‌بینی جایگاه دانشگاه‌ها، پژوهشگاه‌ها، و مؤسسه‌های آموزش عالی، پژوهشی، و فناوری در نظام‌های فراگیر و موضوعی رتبه‌بندی جهانی و پیشنهاد راهکار برای سامانه جایگاه علم، فناوری، و نوآوری ایران در جهان (نما)	سمیه فتاحی	۱۴۰۰	نظام رتبه‌بندی دانشگاهی، داده‌های سری زمانی، پیش‌بینی، سامانه نما
	C06	طراحی یک روش برای برچسب‌زدن تصاویر استخراج‌شده از پارساهای موجود در پایگاه اطلاعات علمی ایران (گنج)	آزاده فخرزاده	۱۴۰۰	برچسب‌زدن تصویر، متن‌کاوی، استخراج کلیدواژه، پردازش زبان طبیعی
	C07	توسعه رویکرد خوشه‌بندی برای پیشنهاد اسناد علمی فارسی بر مبنای داده‌های فارسی پایگاه اطلاعات علمی ایران (گنج)	مرضیه زرین‌بال ماسوله	۱۴۰۰	خوشه‌بندی، یادگیری ماشین، اسناد علمی فارسی، آنتروپی نسبی، خوشه‌بندی عمیق

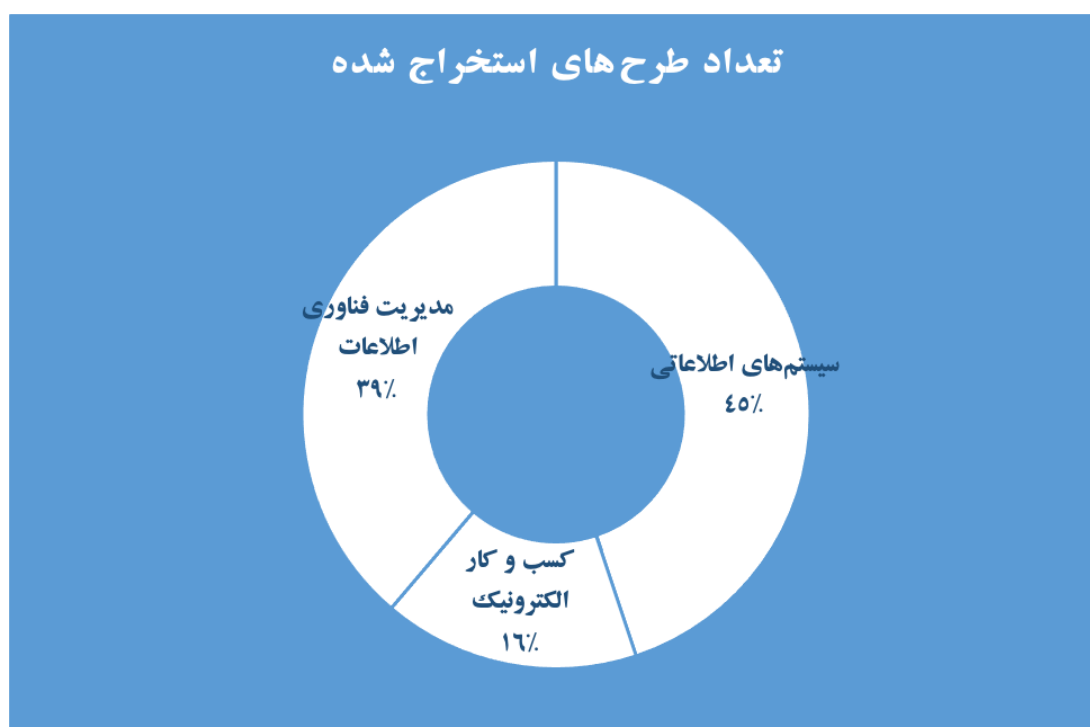
پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	C08	ارزیابی پیکربندی امنیتی در زیرساخت شبکه پژوهشگاه علوم و فناوری اطلاعات ایران	امیرحسین صدیقی	۱۴۰۰	امنیت سایبری، ارزیابی آسیب‌پذیری، ارزیابی پیکربندی امنیتی، زیرساخت شبکه، فناوری اطلاعات، مطالعه موردی
	C09	توسعه سیستم پیشنهاددهنده کلیدواژه برای مستندات علمی فارسی بر پایه اطلاعات اصطلاح‌نامه‌های تخصصی	آزاده محبی	۱۴۰۱	استخراج خودکار کلیدواژه، اصطلاح‌نامه‌های تخصصی، یادگیری ماشین، نمایه‌سازی متن
	C10	استخراج و تحلیل الگوهای رفتاری کاربران با استفاده از لاگ‌کاوی بر روی موتور جستجوی پایگاه اطلاعات علمی ایران (گنج)	سمیه فتاحی	۱۴۰۱	پایگاه اطلاعات علمی ایران (گنج)، ایرانداک، سامانه اطلاعاتی، موتور جستجو، لاگ‌کاوی، دنباله‌های رفتاری
	C11	دسته‌بندی تصویرها در مدارک پایگاه اطلاعات علمی ایران (گنج) بر اساس یادگیری عمیق	آزاده فخرزاده	۱۴۰۱	بازیابی تصویرهای علمی، یادگیری عمیق، برچسب زدن تصویرها، دسته‌بندی تصویرهای علمی
	C12	پیشنهاد همکاری پژوهشی با استفاده از روش‌های تحلیل شبکه و الگوریتم‌های پیش‌بینی یال و تکمیل گراف: موردکاوی شبکه همکاری‌های پژوهشی پژوهشگاه علوم و فناوری اطلاعات ایران	عمار جلالی‌منش	۱۴۰۰	همکاری علمی، پیش‌بینی یال، تکمیل گراف، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، تحلیل شبکه
	C13	تهیه، تدوین و استقرار کامل نظام رصد اجتماعی	عمار جلالی‌منش	۱۳۹۹	نظام رصد اجتماعی، شاخص‌های اجتماعی، آسیب‌های اجتماعی، داشبورد مدیریتی، نرم‌افزار تحلیل داده، رصد مدل مبنا
	C14	طراحی و پیاده‌سازی سامانه برخط ارزیابی رضایت کاربران پایگاه اطلاعات علمی ایران "سامانه گنج"	سمیه فتاحی	۱۳۹۸	پایگاه اطلاعات علمی ایران (گنج)، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، سامانه اطلاعاتی، موتور جست‌وجو، رضایت کاربران، کیفیت خدمات سامانه‌های تحت وب

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	C15	طراحی و تدوین منشور پژوهش، پژوهشکده‌ها، و گروه‌های پژوهشی پژوهشگاه علوم و فناوری اطلاعات ایران	عمار جلالی‌منش	۱۳۹۹	ایرانداک، منشور پژوهشکده‌ها، گروه کانونی، الگو برداری، اولویت‌های پژوهش، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
	C16	ارائه روشی هوشمند برای استخراج کلیدواژه از مستندات علمی زبان فارسی بر اساس سیستم‌های پیشنهاددهنده	آزاده محبی	۱۳۹۷	سیستم‌های پیشنهاددهنده، استخراج کلیدواژه، استدلال نمونه محور، بازیابی اطلاعات، یادگیری ماشین، روش بازنمایی کلمه-به-بردار
	C17	طراحی مدل مفهومی برای معنابخشی به خدمات ارائه شده در سامانه گنج	مرضیه زرین‌بال ماسوله	۱۳۹۷	معنابخشی، تعامل انسان و رایانه، تعامل انسان و اطلاعات، الگوی مفهومی، سامانه گنج، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
	C18	مصورسازی جغرافیایی نتایج جست‌وجو در پایان‌نامه‌ها و رساله‌های موجود در پایگاه داده گنج پژوهشگاه علوم و فناوری اطلاعات ایران	امیرحسین صدیقی	۱۳۹۷	سیستم اطلاعات جغرافیایی (جی‌آی‌اس)، مصورسازی داده، اطلاعات مکانی، فراداده
	C19	تحلیل رفتار جست‌وجوی اطلاعات پژوهشگران در موتور جست‌وجوی سامانه گنج	سمیه فتاحی	۱۳۹۶	رفتار جست‌وجوی اطلاعات، لاگ کاوی، موتور جست‌وجو، سامانه اطلاعاتی، تحلیل رفتار، سامانه گنج
	C20	طراحی و توسعه الگوی مفهومی روش بازی‌وارسازی به منظور بازطراحی سامانه ثبت پایان‌نامه‌ها و رساله‌ها	مرضیه زرین‌بال ماسوله	۱۳۹۵	بازی‌وارسازی، الگوی مفهومی، مشارکت کاربران، سامانه ثبت، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
	C21	مرحله آغازین طرح «ایجاد و توسعه روبات‌ک: ربات پژوهشیار»	آزاده محبی	۱۳۹۵	ربات اجتماعی، تشخیص چهره، پردازش و تشخیص گفتار، جست‌وجو و بازیابی اطلاعات، مصورسازی اطلاعات

پژوهشکده	کد	طرح	مجری	سال	کلیدواژگان
	C22	تدوین نقشه راه فناوری اطلاعات پژوهشگاه با تمرکز بر پیشرفت‌های نوین در حوزه تجزیه و تحلیل اطلاعات	آزاده محبی	۱۳۹۵	نقشه راه، فناوری اطلاعات، تجزیه و تحلیل اطلاعات، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

لازم به ذکر است کد A برای طرح‌های پژوهشی گروه مدیریت فناوری اطلاعات، کد B برای طرح‌های پژوهشی گروه کسب‌وکار الکترونیک و کد C برای طرح‌های پژوهشی گروه سیستم‌های اطلاعاتی در نظر گرفته شده است.

در این میان تعداد طرح‌های پژوهشی (که مجموعاً ۴۹ طرح در این بازه زمانیست) در گروه‌های پژوهشی متفاوت پژوهشکده‌ی فناوری اطلاعات استخراج شده است. درصد انجام (یا در دست انجام بودن) طرح‌های پژوهشی در هر گروه پژوهشی به شرح شکل زیر است:



شکل ۲۸- درصد انجام طرح‌های پژوهشی در سه گروه پژوهشی

همچنین در جدول زیر می‌توان مشخصات هر مقاله علمی پژوهشگران پژوهشکده را نیز در ۵ سال اخیر مشاهده نمود:

جدول ۱۳- مشخصات مقالات پژوهشی ایرانداک

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
مدیریت فناوری اطلاعات	D01	Multi-objective economic-statistical design of simple linear profiles using a combination of NSGA-II, RSM, and TOPSIS	نگار روشن‌بین، محمدجواد ارشادی، سید تقی اخوان نیاکی	۱۴۰۱	multi-objective, optimization, NSGA-II algorithm, TOPSIS, RSM
	D02	شناسایی و رتبه‌بندی عوامل مؤثر بر قیمت‌گذاری کالاهای اطلاعاتی در مراکز اطلاع‌رسانی ایران	لیلا نامداریان، حمیدرضا خدمتگزار	۱۴۰۱	قیمت‌گذاری، کالاهای اطلاعاتی، عوامل مؤثر، مراکز اطلاع‌رسانی ایران
	D03	ارائه یک روش ساختارمند در اندازه‌گیری و تحلیل کیفیت اطلاعات پایان‌نامه‌ها و رساله‌های درون کشور	محمدجواد ارشادی، معصومه نبی‌زاده	۱۴۰۱	کیفیت داده، کیفیت اطلاعات، کنترل کیفیت، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، سامانه ثبت پایان‌نامه/رساله تحلیل، ناهمخوانی
	D04	مفهوم‌سازی برند کتابخانه‌ها بر پایه مرور نظام‌مند پژوهش‌ها	فرزانه سهلی، سیروس علیدوستی	۱۴۰۰	برند، برندسازی، کتابخانه، مرور نظام‌مند، مفهوم برند
	D05	Selection and Performance Estimation of Green Lean Six Sigma Projects: A Hybrid Approach of Technology Readiness Level, Data Envelopment Analysis, and ANFIS	محمدجواد ارشادی، امید قنادی تقی‌زاده، محمد حاجی مولانا	۱۴۰۰	Green Lean Six Sigma (GLS), project Data envelopment analysis (DEA), Technology readiness level (TRL), Adaptive neuro-fuzzy inference system (ANFIS)
	D06	Investigating the Role of Social Networking Technology on the Organizational Agility: A Structural Equation Modeling Approach	شادی احمدی، محمدجواد ارشادی	۱۴۰۰	Social Networks, organizational agility, social networking technology, Structural Equation modeling, organizational performance

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D07	LVTIA: A new method for keyphrase extraction from scientific video lectures	حمید حسنی، محمدجواد ارشادی، آزاده محبی	۱۴۰۰	Multimedia indexing, Video lecture indexing, Text mining, Keyword extraction, Keyphrase extraction
	D08	توسعه راهکارهای خطاناپذیرسازی سیستم‌های اطلاعاتی به کمک روش بهترین-بدترین	آنیثا حاجی‌زاده، محمدجواد ارشادی، محمدرضا نبات‌چیان	۱۴۰۰	پوکایوکه، خطاناپذیرسازی، سیستم‌های اطلاعاتی، روش بهترین-بدترین، تصمیم‌گیری با معیارهای چندگانه
	D09	Development and validation of a measuring instrument for digital library maturity	فاطمه شیخ‌شعاعی، نادر نقشینه، سیروس علیدوستی، مریم ناخدا، حسین دهداری‌راد	۱۴۰۰	Digital library, Instrument development, Maturity model, Validation Management
	D10	Multidisciplinary analysis of international environments based on impacts of Covid-19: State of art	زینب رحیمی‌ریسه، محمد مهدی ارشادی، محمدجواد ارشادی	۱۴۰۰	Covid-19 pandemic, public-health systems, Coronavirus lockdowns, GHG, emissions fall E, conomic crisis
	D11	بررسی روند پژوهش در زمینه عوامل موثر بر پذیرش و کاربرد فناوری اطلاعات در ایران	سیروس علیدوستی	۱۴۰۰	چارچوب تحلیل و بازنمایی اطلاعات پایان‌نامه‌ها، رساله‌ها و مقاله‌ها، عوامل موفقیت پروژه‌های فناوری اطلاعات، عوامل شکست پروژه‌های فناوری اطلاعات، تحلیل روند

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D12	شناسایی و اولویت‌بندی فناوری‌های نوین در مدیریت کتابخانه‌های عمومی ایران	حمیدرضا خدمتگزار، سارا منصفی، مهدی علیپور حافظی، محمدجواد ارشادی، عصمت مومنی	۱۴۰۰	فناوری‌های نوین مدیریت، کتابخانه عمومی، اولویت‌بندی، تصمیم‌گیری با معیارهای چندگانه
	D13	شناخت و تحلیل سیستمی متدولوژی‌های کیفیت داده و ارائه یک چارچوب جامع (با استفاده از روش فراترکیب)	بابک سهرابی، حمیدرضا یزدانی، محمدجواد ارشادی، سوده دوروش	۱۴۰۰	کیفیت داده، متدولوژی کیفیت داده، رویکرد سیستمی فراترکیب
	D14	Developing a Fuzzy Goal Programming Model for Health, Safety, and Environment Risks Based on Hybrid Fuzzy FMEA-VIKOR Method	معصومه نبی‌زاده، محمد خلیل‌زاده، سعدالله ابراهیم‌نژاد، محمدجواد ارشادی	۱۴۰۰	Health Safety and Environment (HSE) risk, FMEA method, Fuzzy VIKOR, Fuzzy Delphi, Fuzzy Goal Programming
	D15	Data Quality Strategy Selection in CRIS: Using a Hybrid Method of SWOT and BWM	اتمان آزرول، محمدجواد ارشادی، امیر عزیزی، ملیکاسادات بنی‌هاشمی، رضا ادريس آبادی	۱۴۰۰	current research information systems (CRIS), data quality (DQ), knowledge management (KM), SWOT, multi-criteria decision-making (MCDM), best-worst method (BWM)
	D16	The impact of perceived risks on internet banking adoption in Iran: a longitudinal survey	حمیدرضا خدمتگزار	۱۴۰۰	Perceived risk theory , Internet banking (PRT) Adoption, Longitudinal survey

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D17	ارائه مدل ترکیبی ارزیابی تأثیر نوآوری کارکنان و کیفیت خدمات فناوری اطلاعات بر رضایت کاربران دانشگاه جامع علمی کاربردی	محمدجواد ارشادی، محمد مهدی ارشادی، آرمین آهنربای	۱۳۹۹	کیفیت خدمات فناوری اطلاعات، نوآوری، کارکنان سازمانی، رضایت کاربران، فرایند تحلیل شبکه، تصمیم‌گیری و اولویت‌بندی
	D18	Simulation based optimization of multi-product supply chain under a JIT system	آرمان ساجدی‌نژاد، عرفان حسن‌نایی، محمدساویر اسدی‌لاری	۱۳۹۹	Supply chain management, genetic algorithm, Optimization, just in time, discrete event simulation
	D19	ارزیابی تأثیر عوامل فرهنگی و اجتماعی بر کارآفرینی زنان در خانه‌های کارآفرینی شهرداری تهران	محمدجواد ارشادی، نسا رضایی، رضا باقری حسین آبادی	۱۳۹۹	کارآفرینی، کارآفرینی زنان، خانه‌های کارآفرینی، خانه‌های کارآفرینی شهرداری
	D20	ارائه راهکارهای اجرایی برای بهبود کیفیت اطلاعات پایان‌نامه/رساله‌ها در سامانه ثبت بر پایه تکنیک‌های داده‌کاوی	محمدجواد ارشادی	۱۳۹۹	کیفیت داده، کیفیت اطلاعات، کنترل کیفیت، بهبود مستمر، داده‌کاوی، خوشه‌بندی
	D21	Digital identifiers in scientific publishing and e-health	حمیدرضا خدمتگزار	۱۳۹۹	Digital identifier, scientific publishing, e-health

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D22	مفهوم‌سازی سایه‌نویسی: نگاهی به بدرفتاری‌های علمی در بافت آموزش عالی ایران	بهروز رسولی، مرتضی نبی‌میدی، حمیدرضا مختاری اسکی، مجید نبوی، مریم نظری، سیروس علیدوستی	۱۳۹۹	بدرفتاری علمی، کسب‌وکار پایان‌نامه و مقاله‌نویسی، مطالعه بافتی، سایه نویسی مشتری، سایه‌نویسی
	D23	پایش شاخص‌های کیفیت داده‌های پژوهشی به کمک نمودار کنترل مشاهده‌های انفرادی با دامنه متحرک	آیناز اشتریان اصفهانی، محمدجواد ارشادی، امیر عزیزی	۱۳۹۹	کیفیت داده، نمودار کنترل، شاخص‌های کلیدی عملکرد، نمودار I-MR، سامانه گنج
	D24	A Framework to Enhance the Procurement Processes Based on Hybrid Approach of FMEA-DEA	محمدجواد ارشادی، الهه آتشک ممان، الناز آتشک ممان	۱۳۹۹	Business Process Reengineering, (BPR), Failure Mode and Effect Analysis (FMEA), Data envelopment analysis (DEA), Procurement Process, Process Management
	D26	اولویت‌بندی شاخص‌های کیفیت داده‌های پژوهشی بر پایه تکنیک‌های تصمیم‌گیری با معیارهای چندگانه	هدیه کرم‌نای فر، محمدجواد ارشادی، فرزاد موحدی سبحانی	۱۳۹۹	سامانه‌های اطلاعاتی، کیفیت داده‌ها، دیمتل فازی، ANP فازی، FDANP
	D27	A Structural Equation Method for Improving Customer Satisfaction of E-Commerce Websites Using the Hybrid Approach of QFD and WebQual	محمدجواد ارشادی، مجتبی صادقی، امیر عزیزی	۱۳۹۹	Structural Equation Method (SEM), E- Commerce, QFD, WebQual, Case study

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D28	مدل‌سازی مساله انتخاب و تخصیص سفارش به تامین‌کنندگان بر پایه برنامه‌ریزی آرمانی و رویکرد ترکیبی QFD و ANP	میلاد توکلیان، محمدجواد ارشادی، امیر عزیزی	۱۳۹۹	گسترش عملکرد کیفیت (QFD)، فرآیند تحلیل شبکه (ANP)، برنامه‌ریزی آرمانی، انتخاب تأمین‌کننده
	D29	The Impact of Knowledge Management on Organizational Performance: A Structural Equation Modeling Study	لیلا نامداریان، آرمان ساجدی‌نژاد، سمیه بهانسته	۱۳۹۹	Knowledge management, Organizational performance, Structural equation modeling, Commercial companies
	D30	منطق درآمدزایی شبکه‌های اجتماعی: ارزیابی تطبیقی	مرتضی نبی‌میلدی، حمیدرضا خدمتگزار، سیروس علیدوستی	۱۳۹۹	شبکه‌های اجتماعی، دسته‌بندی، شبکه‌های اجتماعی ارزش آفرین، درآمدزایی، مدل‌های کسب‌وکار
	D31	An integrated HFMEA-DES model for performance improvement of general hospitals: a case study	محمد مهدی ارشادی، محمدجواد ارشادی، سید تقی اخوان نیاکی	۱۳۹۹	Healthcare, Failure Modes and Effects, Analysis, discrete event simulation, Case study, General Hospital, Healthcare in Iran
	D32	مدل تلفیقی کارت امتیازی متوازن بهبودیافته و روش دیماتل برای ارزیابی عملکرد پژوهشگاه‌ها و مراکز پژوهشی	آرمان ساجدی‌نژاد، عرفان حسن‌نایی، احمد گنجی	۱۳۹۹	مدیریت راهبردی، عملکرد پژوهشگاه‌ها، مدل معادلات ساختاری، ارزیابی عملکرد متوازن، چهارچوب ارزیابی عملکرد

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D33	Simulation-optimization framework for train rescheduling in rapid rail transit	عرفان حسن نایی، آرمان ساجدی نژاد، علی کاردانیا، مسعود شکیبایی فر، حسین جعفری، احسان منصوری	۱۳۹۸	Disturbance management, event-driven simulation, control strategy, uncertain recovery time, waiting times
	D34	Strategic Alignment of Project Management with Health, Safety and Environmental Management	محمدجواد ارشادی، رضا ادریس آباد، عقیله شکوری	۱۳۹۸	Strategic convergence, Project management, HSE management, Structural Equation modeling
	D35	توسعه شاخص‌های کیفیت داده به منظور ارزیابی سامانه‌های اطلاعاتی تحقیقاتی: یک مطالعه موردی	آیناز اشتریان اصفهانی، محمدجواد ارشادی، امیر عزیزی	۱۳۹۸	کیفیت داده، شاخص‌های کلیدی عملکرد، سامانه‌های اطلاعاتی تحقیقاتی، مطالعه موردی
	D36	بررسی نقش ابعاد مدیریت دانش در اثربخشی و بهبود عملکرد معلمان شرکت‌کننده در دوره‌های آموزشی: یک مطالعه موردی	محمدجواد ارشادی، رضا رجبعلی بگلو، احسان نجفی عرب	۱۳۹۸	مدیریت دانش، بازدهی معلمان، توسعه و نگهداری دانش، توزیع و ترکیب دانش
	D37	چارچوب مفهومی موضوع‌های راهبردی فناوری اطلاعات (تحلیل اسناد راهبردی ملی و جهانی)	منصور شیدائی، سیروس علیدوستی، مرتضی نبی‌میلدی	۱۳۹۸	فناوری اطلاعات، اسناد راهبردی، راهبردهای فناوری اطلاعات، موضوع‌های راهبردی، تحلیل محتوای کیفی

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D38	ارائه مدلی جهت ارزیابی ریسک پروژه‌های EPC در صنعت پتروشیمی مبتنی بر تکنیک‌های ANP و FMEA	فاطمه مهدیان موغاری، ناصر صفایی، محمدجواد ارشادی	۱۳۹۸	ارزیابی ریسک پروژه‌های EPC، تکنیک FMEA و ANP
	D39	طراحی مدل کیفیت فراداده: موردکاوی سامانه ثبت پایان‌نامه/رساله در پژوهشگاه علوم و فناوری اطلاعات ایران	محمدجواد ارشادی، عمار جلالی‌منش، جلال‌الدین نصیری	۱۳۹۸	فراداده، کنترل کیفیت، طرح کیفیت، سامانه‌های اطلاعاتی تحقیقاتی، سامانه ثبت، پایان‌نامه/رساله
	D40	Evaluation and Improvement of Service Quality in Information Technology Department of a Detergent Production Company Using the SERVQUAL Approach	محمدجواد ارشادی، سید تقی اخوان نیاکی، راضیه صادقی	۱۳۹۸	Service quality, Information technology, SERVQUAL, Gap analysis
	D41	Analysing the Factors Affecting Environmental Risks of Projects Using Hybrid approach of DEMATEL-ANP, Artificial Neural Network: A Case Study	محمدجواد ارشادی، فاطمه کریمی آشتیانی	۱۳۹۸	Environmental Risk Management, DEMATEL, Fuzzy Hierarchical Analysis, Neural Network, Algorithm Optimization
	D42	Measuring the Impact of Soft and Hard Total Quality Management Factors on Customer Behavior Based on the Role of Innovation and Continuous Improvement	محمدجواد ارشادی، نفیسه نجفی، پریا سلیمانی	۱۳۹۸	soft factors, hard factors, total quality management, customer behavior, Innovation, continuous improvement, construction industry

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D43	توسعه مدیریت اطلاعات علمی در بستر دولت الکترونیکی: مطالعه موردی سامانه‌های پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)	آرمان ساجدی‌نژاد، امیرحسین صدیقی	۱۳۹۸	دولت الکترونیکی، مدیریت اطلاعات علمی، اسناد بالادستی، خدمات علم و فناوری پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
	D44	ارزیابی تطبیقی کارایی ساختار نحو نظام‌های شناسگر دیجیتالی	حمیدرضا خدمتگزار، مهدی علیپور حافظی	۱۳۹۸	شناسگر دیجیتال، ساختار نحو، رشته شناسگر، ارزیابی تطبیقی، کارایی
	D45	ارائه مدل ارزیابی میزان رضایت کارکنان از سیستم مدیریت آموزش الکترونیکی (LMS)	محمدجواد ارشادی، سیما صوفی نیارکی، علی نعیمی صدیق	۱۳۹۸	آموزش الکترونیکی، ارزیابی رضایت کارکنان، ای‌کوال، سیستم آموزش الکترونیکی (LMS)، وب‌کوال
	D46	An Efficient Economic-Statistical Design of Simple Linear Profiles Using a Hybrid Approach of Data Envelopment Analysis, Taguchi Loss Function, and MOPSO	مریم فاضلی مقدم، محمدجواد ارشادی، سید تقی اخوان‌نیاکی	۱۳۹۸	MOPSO economic-statistical design, Linear profiles, Quadratic loss function, Data envelopment analysis (DEA), Response Surface Methodology (RSM)
	D47	مدلی برای بهینه‌سازی زمان‌بندی نگهداری و تعمیرات پیشگیرانه برای سیستم‌های چند جزئی با استفاده از الگوریتم ژنتیک	آرمان ساجدی‌نژاد، میثم لطفی	۱۳۹۸	نت پیشگیرانه، قابلیت دسترسی، قابلیت اطمینان، هزینه، الگوریتم ژنتیک
	D48	تعیین اهمیت معیارهای کیفیت وب‌سایت از دیدگاه کاربران بر پایه روش‌های ای‌کوال و فرایند تحلیل شبکه فازی	محمدجواد ارشادی، زهرا مقدم	۱۳۹۸	کیفیت خدمات، وب‌کوال، ای‌کوال، پایگاه گنج، مطالعه موردی، ANP

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D49	ارزیابی ریسک‌های نرم‌افزاری امنیت اطلاعات سامانه اطلاعاتی تحقیقاتی با استفاده از روش ترکیبی تجزیه تحلیل حالات بالقوه خرابی فازی و تصمیم‌گیری با معیارهای چندگانه فازی	مهرداد فروزنده، محمدجواد ارشادی، مهدی کرباسیان	۱۳۹۷	سامانه اطلاعاتی تحقیقاتی، مدیریت ریسک، امنیت اطلاعات، تجزیه و تحلیل، حالات خرابی و آثار آن
	D50	Design of a digital library (DLMM)maturity model	فاطمه شیخ‌شعاعی، نادر نقشینه، سیروس علیدوستی، مریم ناخدا	۱۳۹۷	Iran, Digital libraries, Delphi technique, Meta-synthesis, Maturity models, Digital library maturity model
	D51	کاربرد رویکرد ترکیبی DEA-FMEA در باز مهندسی فرآیندهای حوزه تأمین در صنعت پیمانکاری نفتی، مطالعه موردی در یک شرکت پیمانکاری حوزه نفت (OEID)	محمدجواد ارشادی، امین کمالی، الهه آتشک ممان	۱۳۹۷	باز مهندسی فرآیندها BPR، تجزیه تحلیل حالات شکست و اثرات آن FMEA، تحلیل پوششی داده‌ها DEA
	D52	عوامل موثر بر موفقیت شبکه اجتماعی حرفه‌ای- تخصصی: یافته‌های مطالعه کیفی در ایران	مرتضی نبی‌میدی، سیروس علیدوستی، مریم نظری	۱۳۹۷	سیاست‌گذار، شبکه اجتماعی حرفه‌ای- تخصصی، عوامل موفقیت، کاربر، مطالعه کیفی
	D53	Root Cause Analysis in Quality Problem Solving of Research Information Systems: A Case Study	محمدجواد ارشادی، روزبه ایاسی، شیرین کاظمی	۱۳۹۷	Root Cause Analysis (RCA), Research Information Systems (RIS), Quality Problem, Case study
	D54	Implementation of failure modes and effects analysis in detergent production companies: A case study	محمدجواد ارشادی، محمد مهدی ارشادی	۱۳۹۷	Failure modes effects analysis, Case study, Risk management, Detergent production

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	D55	The role of dimensions of perceived risk in adoption of corporate internet banking by customers in Iran	حمیدرضا خدمتگزار، آرزو شهنازی	۱۳۹۷	Corporate internet banking, Adoption of technology, Corporate clients, Perceived risk theory
	D56	طراحی و ساخت پایگاه وب منابع اطلاعات شاخص‌های پایش و ارزیابی علم، فناوری، و نوآوری	لیلا نامداریان، سیروس علیدوستی	۱۳۹۷	شاخص علم، فناوری، و نوآوری پایش و ارزیابی منابع اطلاعات شاخص‌ها نهادهای پایش و ارزیابی پایگاه وب پایش و ارزیابی
	D57	Multi-Criteria Supplier Selection Decisions in Supply Chain Networks: A Multi-Objective Optimization Approach	آرمان ساجدی‌نژاد، سید کمال چهارسوقی	۱۳۹۷	Supply Chain, Supplier Selection, Linguistic Variables, Fuzzy AHP, genetic algorithm
	D58	دروازه اطلاعات علمی، پژوهشی، و فناوریانه ایران: خدمتی نوین برای پژوهشگران ایرانی	رویا پورنقی، بهروز رسولی، سیروس علیدوستی	۱۳۹۷	خدمات اطلاعاتی، دروازه اطلاعات، دروازه موضوعی اطلاعات علمی، اطلاعات پژوهشی اطلاعات فناوریانه
	D59	طراحی مدل ارزیابی کیفیت خدمات در ادارات خدمات مشترکین شرکت توزیع برق تهران بزرگ بر پایه مدل سروکوال و ANP	محمدجواد ارشادی، محمدمهدی ارشادی، علیرضا چرسنج	۱۳۹۷	ارزیابی کیفیت خدمات، مدل سروکوال، روش تحلیل شبکه‌ای، تصمیم‌گیری و اولویت‌بندی
کسب و کار الکترونیک	E01	Digital Transformation in the Value Chain Disruption of Banking Services	علی نعیمی صدیق، طیه عسگری، محمد ربیعی	۱۴۰۱	Digital transformation, Digital Disruption, Value Chain, Banking Services, Innovation
	E02	Possibilistic cooperative advertising and pricing games for a two-echelon supply chain	مرضیه مظفری، علی نعیمی صدیق، امیرحسین صدیقی	۱۴۰۰	Cooperative, Advertising, Pricing possibilistic, Game Theory, Nash Game, Stackelberg game

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	E03	Bi-level pricing and inventory strategies for perishable products in a competitive supply chain	ناصر رجبی، مرضیه مظفری، علی نعیمی صدیق	۱۴۰۰	supply chain coordination perishable products inventory decisions Pricing bi- level programming Stackelberg-Nash- Cournot Equilibrium
	E04	Optimal Pricing And Advertising Decisions With Suppliers' Oligopoly Competition: Stakelberg-Nash Game Structures	علی نعیمی صدیق، سید کمال چهارسوقی، مرضیه مظفری	۱۴۰۰	Pricing, Advertising, Production-Inventory, Stackelberg-Nash, Equilibrium, Oligopoly Competition
	E05	هماهنگی تصمیم‌های قیمت‌گذاری و بازاریابی محصول جدید با استفاده از نظریه بازی	علی نعیمی صدیق، سمیه فتاحی	۱۳۹۹	رقابت، توسعه محصول جدید، بازی استکلبرگ، بازی نش بازی با همکاری
	E06	مسئله پوشش کامل در حالت حداکثرسازی پراکندگی مکانی با در نظر گرفتن تسهیلات موجود، محدودیت ظرفیت و هزینه متغیر انتقال	علی نعیمی صدیق، امیر امامی، مرضیه مظفری	۱۳۹۹	مسئله پوشش کامل، پراکندگی تسهیلات، الگوریتم ژنتیک، روش تاگوچی
	E07	توسعه راهکارهایی برای دسترسی آزاد به تمام‌متن پایان‌نامه‌ها و رساله‌ها در بافت موسسات آموزشی و پژوهشی ایران	فرهاد شیرانی، هانیه هدائی، لیلا نامداریان	۱۳۹۹	اشتراک دانش، دسترسی آزاد پایان‌نامه رساله، روش دلفی، مؤسسات آموزشی و پژوهشی، قوانین و مقررات تأثیر اجتماعی فناوری اطلاعات
	E08	Strategic Planning for Commercialization of Research Findings in Iran: The Case Study of IranDoc	لیلا نامداریان، علی نعیمی صدیق	۱۳۹۹	Commercialization, Strategic planning, SWOT, Research institutes, IranDoc

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	E09	استفاده از چارچوب کارت امتیازی متوازن برای رتبه‌بندی واحدهای استراتژیک کسب‌وکارهای الکترونیکی	علی نعیمی صدیق، زهرا شفیعی نیک‌آبادی، مهدیه اخباری	۱۳۹۹	کسب‌وکار الکترونیک، عملکرد نوآورانه، کارت امتیازی متوازن، روش دیمتل فازی، روش ANP فازی
	E10	تعادل استکلبرگ-نش در بازی مکان‌یابی رقابتی تسهیلات میان یک امتیازدهنده و دو سرمایه‌گذار	عباس کاری مجیدآباد، مرضیه مظفری، علی نعیمی صدیق	۱۳۹۸	مکان‌یابی رقابتی، مسئله امتیازدهنده سرمایه‌گذار، نظریه بازی، تعادل استکلبرگ، تعادل نش
	E11	A game theoretical for coordination of pricing, recycling, and green product decisions in the supply chain	سحر پارسایی‌فر، علی بزرگی امیری، علی نعیمی صدیق، محمدصادق سنگری	۱۳۹۸	Three-echelon supply chain, Competition, Pricing, Recycling, Green products, Game theory
	E12	تدوین راهبردهای نوین در صنعت بانکداری با تکیه بر فناوری‌های دیجیتال	طیبه عسگری، علی نعیمی صدیق، محمد عبدالشاه	۱۳۹۸	راهبرد دیجیتال فناوری‌های نوین صنعت بانکداری خدمات نوآوری ارزش
	E13	رتبه‌بندی تأمین‌کنندگان در زنجیره تأمین پایدار با استفاده از تکنیک تحلیل پوششی داده‌ها (مطالعه موردی: شرکت آب و فاضلاب منطقه شش شهر تهران)	محمد کاظم عقا، علی نعیمی صدیق	۱۳۹۷	مدیریت زنجیره تأمین، پایداری، تحلیل پوششی داده‌ها، شرکت آب و فاضلاب
	E14	Towards an understanding of the commercialization drivers of research findings in Iran	لیلا نامداریان، علی نعیمی صدیق	۱۳۹۷	Commercialization, drivers, grounded theory, Iran research

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	E15	بررسی عوامل موثر بر درآمد شهرداری‌ها و ارائه راهبردهای مناسب به منظور ایجاد درآمد پایدار	سمیه مسگری، علی نعیمی صدیق، محمد عبدالشاه	۱۳۹۷	درآمد پایدار، منطق فازی، منابع متفرقه، نظام شهرداری، شهرداری کرج
	E16	Barriers to Commercialization of Research Findings in Humanities in Iran	لیلا نامداریان، علی نعیمی صدیق	۱۳۹۷	commercialization barriers, humanities, Delphi method, grounded theory
	E17	تأثیر محدودیت‌های اطلاعاتی بر کاربری اطلاعاتی اینترنت در بین جوانان شهر تهران	احمد گنجی، محمد ربیعی، آذین رحمتی	۱۳۹۷	کاربران اینترنت، کاربری اطلاعاتی، محدودیت‌های اطلاعاتی، سابقه کاربری، نوجوانان و جوانان تهران
سیستم‌های اطلاعاتی	F01	ارائه روشی برای برچسب زدن تصاویر موجود در متون علمی فارسی با استفاده از روش‌های پردازش متن	آزاده فخرزاده، محدثه رهنما، جلال‌الدین نصیری	۱۴۰۱	برچسب زدن تصویر، نشانه‌گذاری تصویر، بازیابی تصویر، پردازش متن، استخراج فراداده، فناوری اطلاعات
	F02	بهبود الگوریتم RAKE برای استخراج کلیدواژه از متون علمی فارسی؛ مطالعه موردی: پایان‌نامه‌ها و رساله‌های فارسی	الهه محرابی، آزاده محبی، عباس احمدی	۱۴۰۰	استخراج کلیدواژه، الگوریتم RAKE، برچسب‌گذاری دستوری، پردازش زبان طبیعی، مستندات علمی فارسی
	F03	ارائه چارچوبی برای ارزیابی وب‌گاه‌ها از منظر معماری اطلاعات	امیرحسین صدیقی	۱۴۰۰	معماری اطلاعات، محیط اطلاعاتی، ارزیابی وب‌گاه، سیاهه واریسی، فناوری اطلاعات
	F04	ارزیابی معماری اطلاعات وب‌گاه دانشگاه تهران	سمیرا دانیالی، امیرحسین صدیقی	۱۴۰۰	معماری اطلاعات، وب‌گاه دانشگاه تهران، ارزیابی وب‌گاه، مطالعه موردی
	F05	ارزیابی معماری وب‌گاه کتابک از منظر معماری اطلاعات	پریناز بابایی، امیرحسین صدیقی	۱۳۹۹	معماری اطلاعات، ارزیابی وب‌گاه، سازماندهی، برچسب‌زنی، پیمایش

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	F06	پردازش نظام‌مند قانون: مطالعه موردی قانون پیشگیری و مقابله با تقلب در تهیه آثار علمی	رضا حسان، امیرحسین صدیقی، رحمان شریف‌زاده	۱۳۹۹	قانون پردازش قواعد، تحلیل نهادی، ابزار گرامر نهادی، فناوری اطلاعات
	F07	روش‌شناسی پژوهشی علم طراحی به‌مثابه یک روش‌شناسی راه‌حل محور	رضا حسان، رحمان شریف‌زاده، امیرحسین صدیقی	۱۳۹۹	علم طراحی، روش‌شناسی پژوهش، روش‌شناسی کیفی، روش‌شناسی کمی، فناوری اطلاعات
	F08	فراروش سازمان‌نگر معماری اطلاعات در وبگاه سازمانی مطالعه موردی: وبگاه سازمان بیمه مرکزی جمهوری اسلامی ایران	سمیه قویدل، امیرحسین صدیقی	۱۳۹۹	معماری اطلاعات، سازماندهی، برچسب‌زنی، پیمایش، جست‌وجو، سازمان بیمه مرکزی جمهوری اسلامی ایران
	F09	ارائه روشی ساختارمحور برای ایجاد پایگاه داده از تصاویر مستخرج از اسناد علمی: مورد مطالعه پایگاه اطلاعات علمی ایران (گنج)	آزاده فخرزاده، امیرحسین صدیقی	۱۳۹۹	پردازش تصویر، استخراج تصویر، استخراج فراداده، فناوری اطلاعات
	F10	A new subject-based retrieval and search result visualization approach for scientific digital libraries	سید محمود بخشایش، عباس احمدی، آزاده محبی	۱۳۹۹	Digital libraries searching, Document retrieval, Scientific Documents, Search Result, Visualization, Subject-based Retrieval
	F11	استفاده از مدل LRFM برای خوشه‌بندی کاربران براساس تحلیل رفتار جستجو (مورد مطالعه: پایگاه اطلاعات علمی ایران «گنج»)	سمیه فتاحی، محمد ربیعی	۱۳۹۹	خوشه‌بندی، مدل LRFM، ماتریس ارزش مشتری مارکوس، تحلیل رفتار کاربران سامانه گنج
	F12	ارائه رویکردی مبتنی بر یادگیری عمیق برای کشف کلاهبرداری در سرویس‌های پرداخت مالی	امیرحسین صدیقی، آرمان ساجدی‌نژاد	۱۳۹۸	یادگیری عمیق، کشف کلاهبرداری، یادگیری ماشینی، هوش مصنوعی، فناوری اطلاعات

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	F13	ارزیابی وبگاه دانشگاه علوم پزشکی شیراز از منظر معماری اطلاعات در سال ۱۳۹۸	مرضیه مروتی، امیرحسین صدیقی	۱۳۹۸	معماری اطلاعات، طراحی وبگاه، مدیریت اطلاعات ایران
	F14	ارائه یک معماری پیشنهادی برای مدیریت دانش در بنگاه ۰/۲ با تمرکز بر رویکرد نوآوری باز	فرشته احمدی، آزاده محبی، لیلا نامداریان	۱۳۹۸	معماری مدیریت دانش جریان دانش بنگاه ۲,۰ نوآوری باز نظریه داده بنیاد
	F15	ارزیابی رضایت کاربران موتور کاوش پایان‌نامه‌ها و رساله‌ها در پایگاه اطلاعات علمی ایران (گنج): بر پایه مدل «ایکوال»	سمیه فتاحی، محمدجواد ارشادی	۱۳۹۸	پایگاه اطلاعات علمی ایران گنج، پژوهشگاه علوم و فناوری اطلاعات ایران ایرانداک، سامانه اطلاعاتی، موتور کاوش، رضایت کاربر ارزیابی، کیفیت ایکوال
	F16	دسته‌بندی سبک‌های یادگیری با استفاده از ویژگی‌های رفتاری و ماشین‌بردار پشتیبان دوقلو	جلال‌الدین نصیری، امیرمحمود میر، سمیه فتاحی	۱۳۹۸	آموزش الکترونیکی، سبک یادگیری، ماشین‌بردار پشتیبان
	F17	An experimental study on an adaptive e-learning environment based on learner's personality and emotion	سمیه فتاحی	۱۳۹۸	Adaptive e-learning, Personality, Emotion, MBTI, OCC
	F18	هماندجویی در متون فارسی بازنویسی شده با استفاده از روش‌های معنایی و احتمالاتی	نصرااله پاک‌نیت، آزاده محبی	۱۳۹۸	سرقت علمی، هماندجویی، هماندجویی معنایی، هماندجویی احتمالاتی، متون بازنویسی شده
	F19	A new reliable performance evaluation model: IFB-IER-DEA	شیوا مسلمی، حمیدرضا ایزدبخش، مرضیه زرین‌بال ماسوله	۱۳۹۸	Performance evaluation, DEA, Evidential, Reasoning approach, IFBIER approach, Anderson Peterson

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	F20	Integrating Fuzzy Inference System, Image Processing and Quality Control to Detect Defects and Classify Quality Level of Copper Rods	محمد مهدی دهدار، مصطفی جهانگشای رضائی، مرضیه زرین‌بال ماسوله، حمیدرضا ایزدبخش	۱۳۹۸	Image processing, Quality control, Fuzzy inference system, Features from accelerated segment test, Copper rods
	F21	طراحی مدل مفهومی پیاده‌سازی معنابخشی به خدمات ارائه شده در سامانه گنج پژوهشگاه علوم و فناوری اطلاعات ایران	مرضیه زرین‌بال ماسوله	۱۳۹۸	معنابخشی، تعامل انسان و اطلاعات، سامانه گنج، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
	F22	ارزیابی سامانه اطلاعاتی هوشمند بر اساس ویژگی‌های شناختی کاربر (مطالعه موردی: محیط یادگیری الکترونیکی)	سمیه فتاحی، منوچهر مرادی سبزواری	۱۳۹۷	شخصیت، احساس خوشایندی، سامانه اطلاعاتی
	F23	An empirical study of using sequential behavior pattern mining approach to predict learning styles	سمیه فتاحی، فایزه شعبانعلی فمی، هادی مرادی	۱۳۹۷	Keywords Sequential pattern mining, MBTI, Learning styles, Sequential behavior patterns
	F24	Low-dose exposure to Bisphenol A during development has limited effects on male reproduction in midpubertal and aging Fischer 344 rats	الینور اسپورندلی نیز، جولی بوبرگ، الیزابت اکستد، لئا هلم، آزاده فخرزاده	۱۳۹۷	Bisphenol A, Developmental exposure, Low dose, Male reproduction, Puberty, Aging, Histopathology, Endocrine disrupting chemical
	F25	ارائه مدلی ریاضی برای قیمت‌گذاری سوخت با در نظر گرفتن تصمیمات طراحی شبکه زنجیره تامین	امیرحسین صدیقی، علی نعیمی صدیق	۱۳۹۷	زنجیره تامین سوخت قیمت‌گذاری سوخت بازی استکلب‌رگ بازی نش طراحی زنجیره تامین

گروه پژوهشی	کد	مقاله	نویسندگان	سال	کلیدواژگان
	F26	A hierarchical stochastic modelling approach for reconstructing lung tumour geometry from 2D CT images	پرنیان افشار، عباس احمدی، آزاده محبی، محمدحسین فاضل زرنندی	۱۳۹۷	Lung cancer tumour recognition, volume reconstruction, computed tomography, stochastic modelling
	F27	ارائه راهکاری خودکار بر اساس متن کاوی برای شناخت و تحلیل روند تحقیقات حوزه‌های علمی	اشکان خطیر، آزاده محبی، سهیل گنج‌فر	۱۳۹۷	تحلیل روند، بلوغ حوزه علمی، ماتریس هم‌رخدادی کلمات، مرکزیت حوزه علمی

۴-۳. تحلیل شکاف پژوهشی

خلاً یا کاستی دانش فاصله موجود میان دانش موردنیاز یک سازمان برای انجام وظایف و رسیدن به اهداف تعیین شده با دانشی که سازمان از طریق منابع مختلف دانشی خود از آن بهره‌مند است را بیان می‌کند (Jens-Erik, Donald and Given 2016). کاستی دانش هم می‌تواند در مورد دانش موجود و آرشیو شده در یک سازمان، نهاد یا جامعه صدق کند و هم می‌تواند دانش قابل استقرار و کارآمد را مدنظر داشته باشد پس کاستی دانش درکی از میزان اختلاف بین دانش موجود و گذشته با دانش موردنیاز یک سیستم یا فرد برای رسیدن به اهداف خود است (Weidong and Dai 2015). همین تعریف برای بیان کاستی پژوهشی نیز کاربرد دارد و به عبارتی جای خالی پژوهش‌های موردنیاز یک سازمان یا جامعه را در بین منابع پژوهشی در دسترس سازمان یا جامعه را شامل می‌شود. این در حالی است که خلاً یا شکاف پژوهش در مطالعه یک موضوع مبین کاستی‌ها و نواقص پژوهش‌های پیشین در آن موضوع خاص است و پژوهشگر سعی دارد تا آن کاستی‌ها را برطرف نماید.

فاصله یا شکاف میان پژوهش‌های دانشگاهی، به‌ویژه پایان‌نامه‌ها یا طرح‌های پژوهشی با نیازهای جامعه یکی از چالش‌های اساسی نظام آموزش عالی در ایران است. گسترش کمی دانشگاه‌ها و افزایش پذیرش دانشجویان تحصیلات تکمیلی باعث شده است تا از یک طرف شاهد انتقادات و واکنش اندیشمندان و خبرگان آموزش و پژوهش باشیم و از طرف دیگر کارفرمایان صنایع و اقتصاد از توانمندی اندک دانش‌آموختگان در حل مسائل واقعی کشور گله‌مند باشند (فتاحی، حکایت استفاده کاربردی از پژوهش‌ها در ایران ۱۳۹۰). همچنین عرصه پژوهش شاهد دوباره کاری‌های بسیاری است که دلیل عمده آن، عدم توانایی استادان و دانشجویان در یافتن مسئله پژوهش در حوزه‌های مختلف

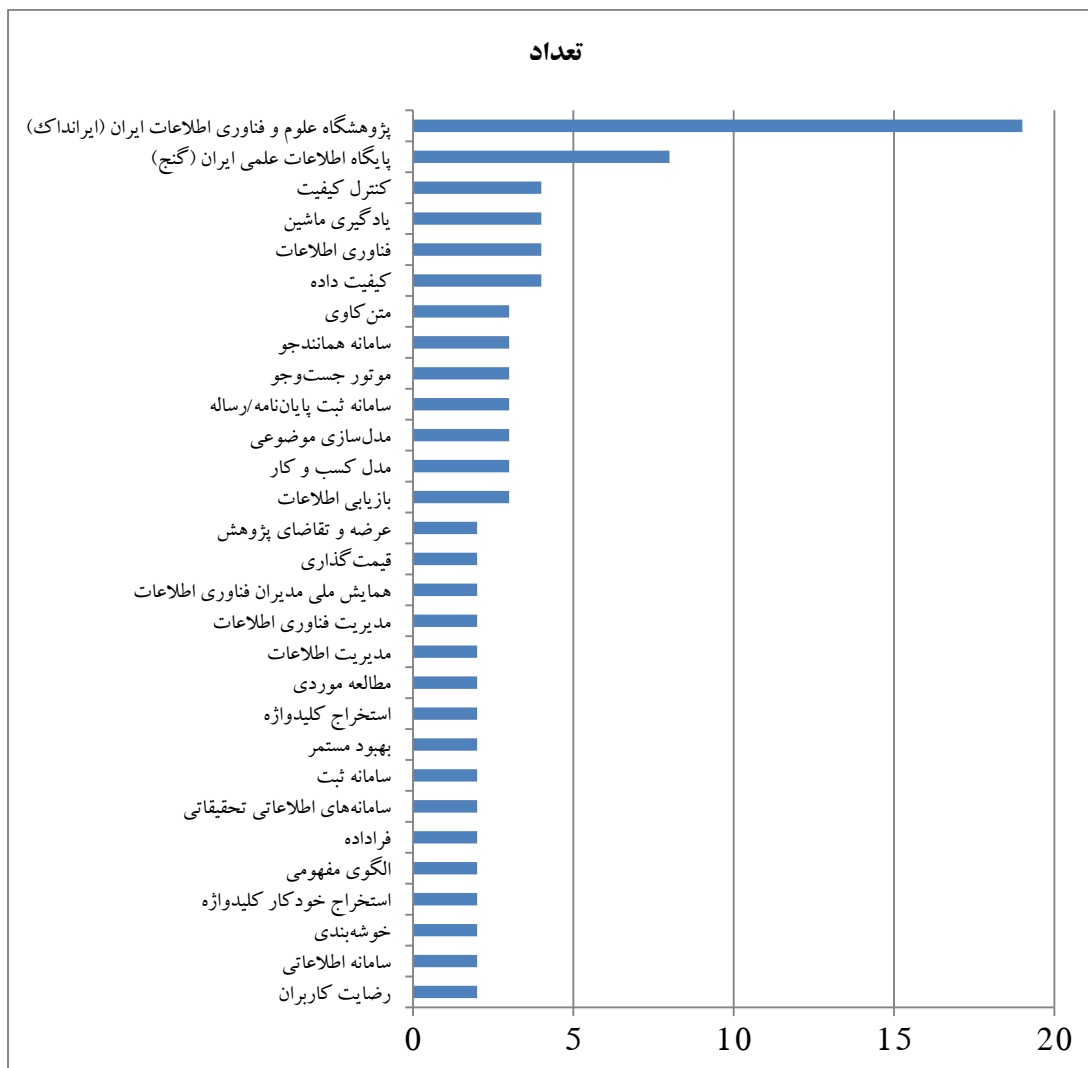
موضوعی موردنیاز جامعه است و در بسیاری از موارد جامعه دانشگاهی بجای «مسئله‌یابی» به دنبال «مسئله‌سازی» شده است. از سوی دیگر اولویت‌های پژوهشی که بیشتر وزارتخانه‌ها و سازمان‌ها مشخص کرده و به نحوی به اطلاع پژوهشگران می‌رسانند، صرف‌نظر از شکل و روش‌های ناکارآمد، کمتر به شکل واقعیانه و کاربردی مورد شناسایی قرار گرفته‌اند.

۴-۳-۱. بررسی میزان تطبیق

برخی اوقات با بررسی تقاضاهایی که از طرف جامعه خبره پژوهشی مطرح می‌شود حوزه‌های مختلف پژوهشی مورد تقاضا شناسایی می‌شود، مقایسه این حوزه‌ها با منابع پژوهشی تولیدشده و در دسترس، این امکان را فراهم می‌سازد تا کاستی‌های پژوهشی و نیز حوزه‌هایی که شامل پژوهش‌های بدون تقاضا هستند شناسایی شوند. سپس سیاست‌گذار با استفاده از منابع اطلاعاتی که خود در اختیار دارد و نیز در راستای سیاست‌های کلان پژوهشی، اسناد بالادستی و ... نیازهای پژوهشی را تعیین می‌کند و در ادامه با توجه به منابع و امکانات موجود و نیز معیارهای مختلف به اولویت‌دهی حوزه‌های پژوهشی می‌پردازد. در این پژوهش، تقاضاهای پژوهشی را با توجه به روندهای بین‌المللی مورد بررسی قرار می‌دهیم.

بدین منظور و برای تولید سبد واژگان هر پایان‌نامه از یکی از خصیصه‌های محتوایی (کلیدواژه‌های) آن استفاده شده است، پس در ابتدا برای این منظور سبد واژگان هر یک از خصیصه‌ها تولیدشده و درنهایت با ترکیب آن‌ها سبد واژگان هر طرح پژوهشی تشکیل خواهد شد.

از آنجا که کلیدواژه‌های موجود در مجموعه داده‌ها، توسط افراد خبره تولید کننده اثر و همچنین با نظارت شورای پژوهشی ایرانداک به هر یک از مستندات تخصیص داده شده است، فرض شده است که این کلیدواژه‌ها، نماینده مناسبی برای هر مستند هستند، ازاین‌رو تنها پیش‌پردازشی که روی آن‌ها اعمال شده است استانداردسازی متون کلیدواژه‌ها است. در این میان ۲۵۰ کلیدواژه استخراج شده است که به تفکیک و با فرکانس استفاده در جدول زیر فهرست شده است (تکرار بیش از ۱).



شکل ۲۹- تکرار کلیدواژگان در طرح‌های پژوهشی ایرانداک

در این پژوهش برای استخراج سبدها واژگان عنوان و چکیده به این شکل عمل شد که N-gram‌های (N=1,2,3,4) عنوان و کلیدواژه استخراج شد و اگر N-gram استخراج شده در مجموعه کلیدواژه‌های حوزه‌های استخراج شده قرار داشت به عنوان سبدها واژگان برای کلیدواژگان در نظر گرفته شد و در غیر این صورت از آن صرف‌نظر شد. پژوهش‌های پیشین نشان می‌دهد که در یک سند علمی احتمال وجود کلیدواژه‌ها در واژگان کلیدی بالاتر از عنوان آن و عنوان بیش از چکیده است و کلیدواژه‌های موجود در واژگان کلیدی نماینده نوآوری آن سند علمی است. از این رو در این طرح پژوهشی از کلیدواژه‌های ثبت شده در طرح‌های پژوهشی، به عنوان نماینده آن مستند استفاده می‌شود.

پس از آنکه سبدهای واژگان هر طرح پژوهشی استخراج شد، متناسب با هریک از سبدهای واژگان یک ماتریس بردار ویژگی برای طرح تولید شد و به این ترتیب سه ماتریس بردار ویژگی برای مجموعه

داده‌ها به دست آمد. همان‌طور که پیش‌تر نیز در این مورد بحث شد، هر یک از سطرهای این ماتریس‌ها نشان‌دهنده یک مستند و ستون‌های آن یکی از عبارت‌های حوزه‌های پژوهشی فصل قبلی است. در هر سطر، مقادیر هر سلول این ماتریس‌ها، برای ستون‌هایی که مقادیر آن‌ها در سبدها و واژگان مرتبط با آن مستند موجود نیست صفر است و برای عناصر موجود در سبدها و واژگان با استفاده از روش وزن‌دهی NG-TF2 مقدار گرفته است. کدهای مستندات استخراجی، با وزن دهی فوق در سه دسته کم متوسط و زیاد (با توجه به بازه امتیازات کسب شده از کلیدواژگان هر طرح پژوهشی) قرار می‌گیرند. در صورتی که میزان ارتباط یک طرح پژوهشی کم باشد، این سنجش را در گروه‌های پژوهشی دیگر مجدد مورد جستجو قرار می‌دهیم:

جدول ۱۴- میزان ارتباط طرح‌های پژوهشی ایرانداک با موضوعات استخراج شده

کد طرح	ارتباط اصلی با موضوع	میزان ارتباط	ارتباط دوم با موضوع	میزان ارتباط دوم
A01	G2	کم	B2	متوسط
A02	G2	متوسط		
A03	G4	متوسط		
A04	G2	متوسط		
A05	G2	متوسط		
A06	G2	متوسط		
A07	G3	متوسط		
A08	G8	متوسط		
A09	G3	متوسط		
A10	G4	متوسط		
A11	G4	متوسط		
A12	G4	متوسط		
A13	G1	کم	B2	زیاد
A14	G1	زیاد		
A15	G2	کم	B2	متوسط
A16	G4	کم	B5	متوسط
A17	G1	کم	B2	زیاد
A18	G1	زیاد		
A19	G4	متوسط		
B01	B5	کم	G2	متوسط
B02	B1	زیاد		
B03	B5	کم	G2	متوسط
B04	B3	زیاد		

کد طرح	ارتباط اصلی با موضوع	میزان ارتباط	ارتباط دوم با موضوع	میزان ارتباط دوم
B05	B1	زیاد		
B06	B1	متوسط		
B07	B2	متوسط		
B08	B1	متوسط		
C01	I6	متوسط		
C02	I2	زیاد		
C03	I3	متوسط		
C04	I3	زیاد		
C05	I6	زیاد		
C06	I5	زیاد		
C07	I3	زیاد		
C08	I4	زیاد		
C09	I6	زیاد		
C10	I2	متوسط		
C11	I2	زیاد		
C12	I5	متوسط		
C13	I3	متوسط		
C14	I1	زیاد		
C15	I6	کم	G2	متوسط
C16	I6	زیاد		
C17	I5	زیاد		
C18	I5	متوسط		
C19	I2	زیاد		
C20	I3	متوسط		
C21	I5	زیاد		
C22	I6	کم	G2	متوسط

اگر میزان ارتباطات را با توجه به دسته بندی کم، زیاد و متوسط امتیاز بندی کنیم، در سالهای مورد بررسی میزان تطبیق با گروه‌های پژوهشی میانگین بصورت زیر بوده است. احتمالاً دو سال میانی به دلیل شکل دهی گروه‌های پژوهشی و تطبیق مجدد پژوهشگران با موضوعات مختص به گروه ایشان بوده است:

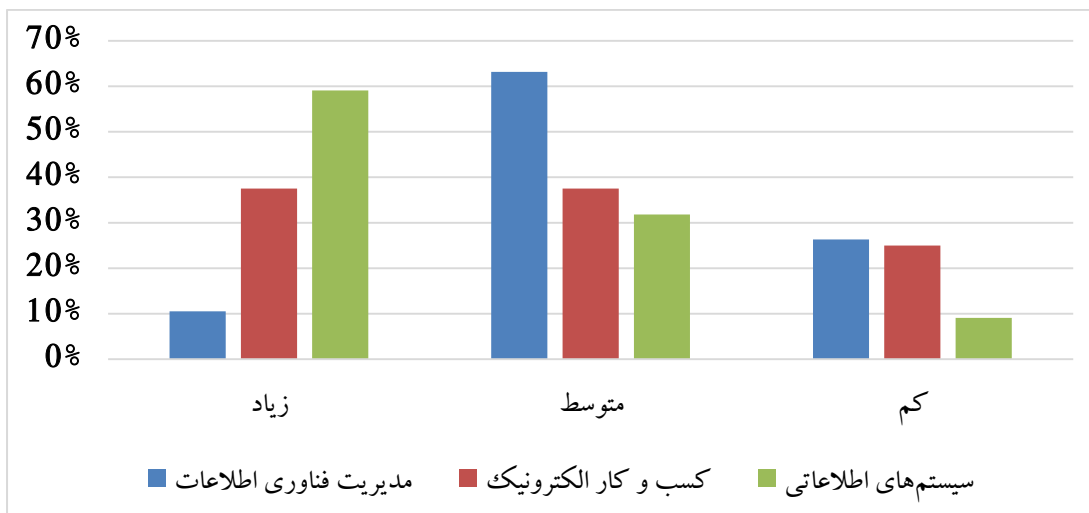


شکل ۳۰- میانگین ارتباط کلیدواژگان در طرح‌های پژوهشی با دسته‌بندی انجام شده

۴-۳-۲. مقایسه میزان تطبیق در هر گروه پژوهشی پژوهشکده فناوری اطلاعات

اولین مرحله از انجام یک مطالعه، شناسایی یک حوزه تحقیقاتی است. انتخاب یک منطقه بکر در زمینه تحقیقاتی در حالت‌های مختلفی روی خواهد داد. بسیاری از محققان ایده‌های روشنی در مورد مشکل تحقیقاتی که می‌خواهند دنبال کنند دارند. با این حال، محققان، به‌ویژه آنهایی که در مراحل اولیه حرفه خود هستند، زمانی که مجبور می‌شوند موضوع تحقیقاتی اصیل و نوآورانه را به ثمر برسانند، با مشکل مواجه می‌شوند. بهترین راه برای انجام این کار، شناسایی شکاف در تحقیقات موجود در این زمینه است. هنگامی که مقالات یا کتاب‌هایی را در مورد موضوعات مورد علاقه مطالعه می‌شود، ممکن است متوجه شد که حوزه‌هایی وجود دارند که زمینه قابل توجهی برای تحقیقات بیشتر دارند، اما محققان دیگر از آنها استفاده نکرده‌اند. شکاف پژوهشی یا شکاف ادبی به چنین حوزه‌های ناشناخته یا ناشناخته‌ای گفته می‌شود که زمینه برای تحقیقات بیشتر دارند. با توجه به مطالب مطرح شده در این فصل، شکاف پژوهشی مطرح شده در این مطالعه از طریق مقایسه پژوهش‌های بین‌المللی و دسته‌بندی آنها از طریق مدل‌سازی موضوعی (فصل قبل) و پژوهش‌های هر یک از سه گروه پژوهشی پژوهشکده انجام می‌پذیرد.

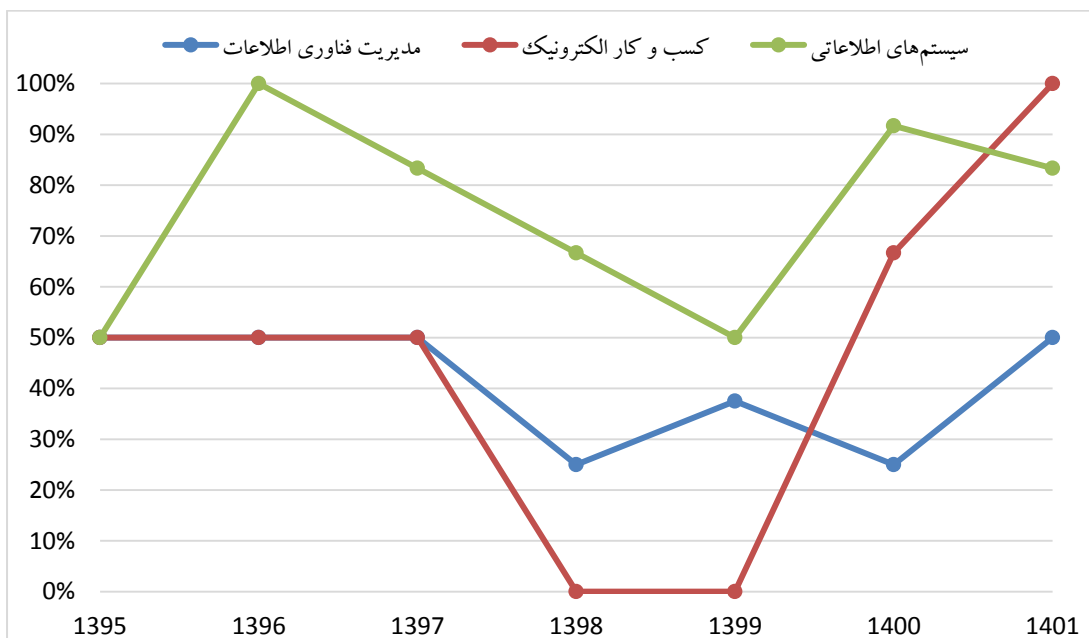
با توجه به مطالب فوق و جدول ۱۰، درصد تطبیق هر یک از مطالعات بر موضوعات مرتبط شناسایی شده در ادبیات موضوع، در شکل زیر نمایش داده شده است:



شکل ۳۱- تعداد ارتباط (کم، متوسط و زیاد) کلیدواژگان در هر گروه با دسته‌بندی انجام شده

همانگونه که در شکل فوق نمایش داده شده است، گروه پژوهشی سیستم‌های اطلاعاتی بیشترین درصد را در تطابق به موضوعات بین‌المللی داراست. پس از آن گروه کسب و کار الکترونیک در وضعیت مناسبتری قرار دارد. همچنین می‌توان ملاحظه نمود که گروه مدیریت فناوری اطلاعات در بیشتر مطالعات خود در تطبیق متوسط قرار دارد.

همچنین اگر بخواهیم این روند را در طول زمان‌های پژوهش بررسی کنیم به نمودار زیر می‌رسیم:



شکل ۳۲- روند میانگین سالانه ارتباط کلیدواژگان هر گروه با دسته‌بندی انجام شده

به نظر می‌رسد گروه کسب و کار الکترونیک در بازه زمانی تشکیل گروه مذکور و تخصیص اعضای جدید، در ابتدا در تطبیق بسیار کم وجود داشته و نهایتاً توانسته است در طی چند سال به تطبیق

حداکثری برسد. اما گروه پژوهشی مدیریت فناوری اطلاعات همواره در دامنه متوسط قرار داشته و هیچگاه مسیر رشد یا نزول خاصی را در آن نمی‌توان مشاهده نمود.

همانند تحلیل طرح‌های پژوهشی، در هر سطر از جدول زیر، برای عناصر موجود در سبد واژگان با استفاده از روش وزن‌دهی NG-TF2 مقدار گرفته است.

کدهای مقالات استخراجی، با وزن دهی فوق در سه دسته کم متوسط و زیاد (با توجه به بازه امتیازات کسب شده از کلیدواژگان هر مقاله) قرار می‌گیرند:

جدول ۱۵- میزان ارتباط مقالات پژوهشی ایرانداک با موضوعات استخراج شده

میزان ارتباط	ارتباط اصلی با موضوع	کد مقاله
کم	G4	D23
متوسط	G4	D24
کم	G4	D25
کم	B6	D26
کم	G6	D27
متوسط	G6	D28
کم	B5	D29
زیاد	B2	D30
کم	G6	D31
کم	G6	D32
کم	B2	D33
کم	I6	D34
کم	G4	D35
کم	G6	D36
کم	I6	D37
متوسط	I6	D38
کم	G4	D39
کم	B6	D40
کم	G6	D41
کم	B6	D42
کم	B6	D43
متوسط	B1	D44

میزان ارتباط	ارتباط اصلی با موضوع	کد مقاله
کم	B3	D01
متوسط	B4	D02
متوسط	G6	D03
کم	B5	D04
متوسط	G4	D05
متوسط	B5	D06
زیاد	G4	D07
زیاد	I6	D08
زیاد	I6	D09
زیاد	B2	D10
متوسط	G4	D11
متوسط	G4	D12
متوسط	G6	D13
زیاد	B2	D14
متوسط	G6	D15
کم	I6	D16
کم	B3	D17
کم	G6	D18
-	-	D19
کم	G6	D20
زیاد	B2	D21
متوسط	I6	D22

میزان ارتباط	ارتباط اصلی با موضوع	کد مقاله
کم	B3	E16
کم	G5	E17
متوسط	I5	F01
متوسط	I5	F02
متوسط	I5	F03
متوسط	I5	F04
متوسط	I5	F05
متوسط	I5	F06
کم	I3	F07
کم	I3	F08
متوسط	I2	F09
زیاد	I6	F10
متوسط	I6	F11
زیاد	I3	F12
متوسط	I3	F13
متوسط	I3	F14
زیاد	I6	F15
متوسط	B1	F16
متوسط	B1	F17
کم	I6	F18
کم	I3	F19
زیاد	I5	F20
متوسط	I1	F21
زیاد	B3	F22
کم	B1	F23
متوسط	B2	F24
کم	I6	F25
متوسط	B2	F26
متوسط	I6	F27

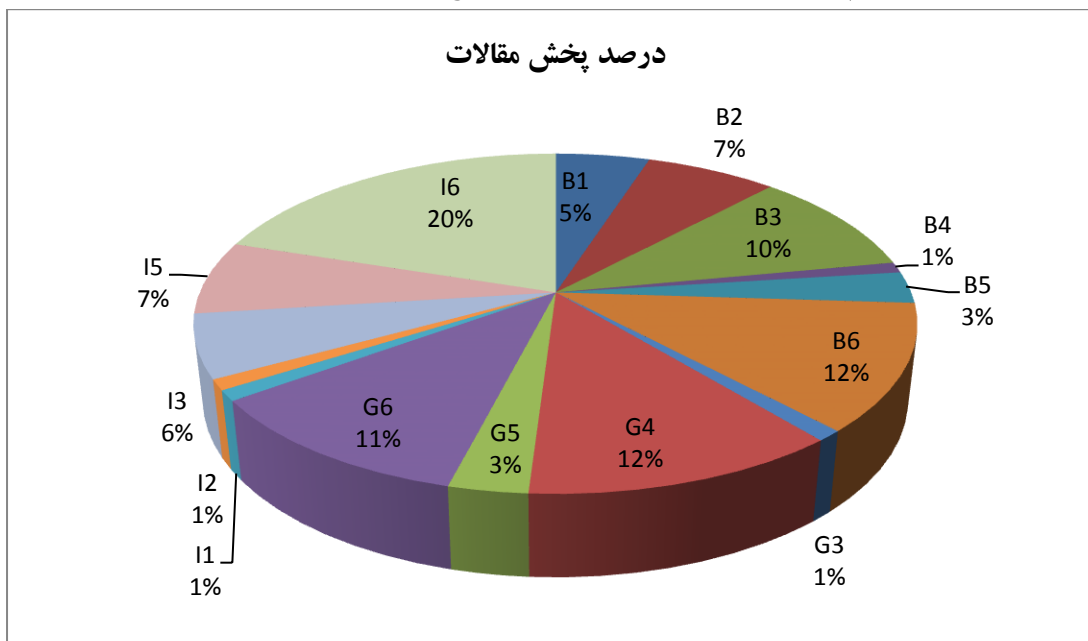
میزان ارتباط	ارتباط اصلی با موضوع	کد مقاله
متوسط	G4	D45
کم	I6	D46
کم	I6	D47
متوسط	G5	D48
کم	I6	D49
متوسط	G4	D50
کم	G3	D51
متوسط	I6	D52
کم	G4	D53
متوسط	B3	D54
متوسط	B3	D55
متوسط	I6	D56
کم	I6	D57
کم	B3	D58
متوسط	B3	E01
متوسط	B3	E02
کم	I6	E03
متوسط	B6	E04
متوسط	B6	E05
کم	B6	E06
کم	G5	E07
کم	B6	E08
کم	B6	E09
کم	B6	E10
کم	B6	E11
متوسط	B1	E12
کم	B6	E13
کم	B3	E14
-	-	E15

در جدول فوق ۱۰۰ مقاله پژوهشی استخراج و میزان ارتباط آنها به موضوعات استخراج شده مشخص شد. اگر بخواهیم نگاهی به نحوه پخش هر مقاله داشته باشیم به جدول زیر می‌رسیم:

جدول ۱۶- شمارش میزان ارتباط مقالات با هر موضوع

موضوع	ارتباط زیاد	ارتباط متوسط	ارتباط کم
B1		۴	۱
B2	۴	۲	۱
B3	۱	۴	۵
B4		۱	
B5		۱	۲
B6		۲	۱۰
G3			۱
G4	۱	۶	۵
G5		۱	۲
G6		۴	۷
I1		۱	
I2		۱	
I3	۱	۲	۳
I5	۱	۶	
I6	۴	۶	۱۰

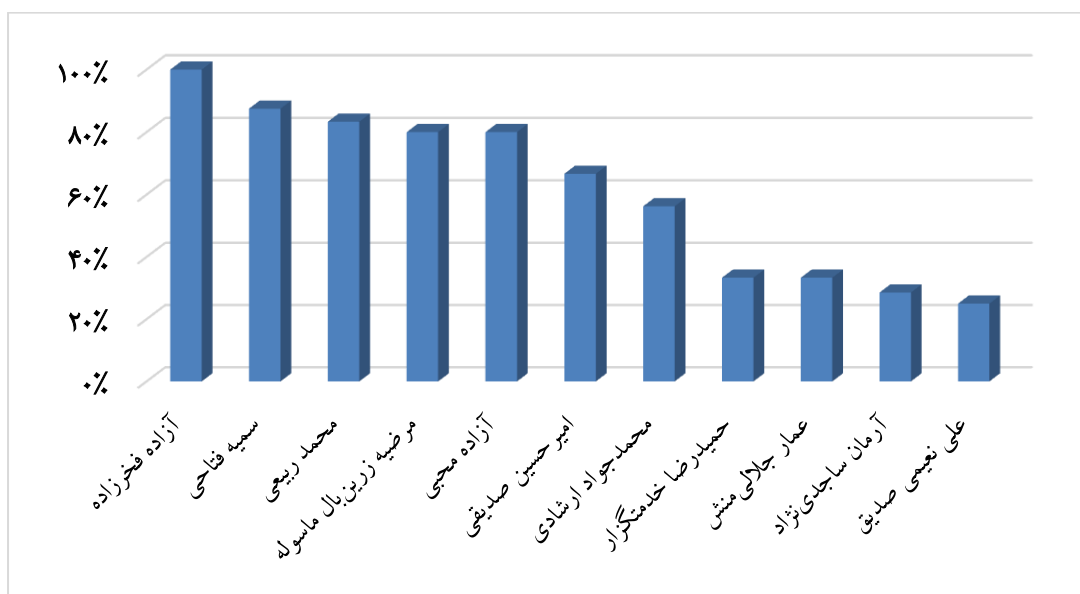
اگر تمامی ارتباطات (شامل زیاد، متوسط و کم) را در نظر بگیریم و میان ارتباطات‌های کم، متوسط و زیاد، تفاوتی قایل نشویم درصد پخش هر مقاله در هر موضوع به شکل زیر است:



شکل ۳۳- درصد پخش مقالات پژوهشده در هر گروه با دسته‌بندی انجام شده

۳-۳-۴. مقایسه میزان تطبیق در هر پژوهشگر

همکاران پژوهشی لازم است مشکلاتی را در زمینه تحقیقاتی خود که قبلاً به آنها پرداخته نشده است، بیابند. از این طریق نه تنها سرمایه و منابع خود را در پروژه مناسب سرمایه گذاری می کنند، بلکه شانس انتشار یافته‌های تحقیقاتی خود را نیز افزایش می‌یابد. میانگین تطبیق فعالیت‌های پژوهشی همکارانی که بیش از یک طرح پژوهشی را انجام داده‌اند به شرح زیر است:



شکل ۳۴- میزان تطبیق کلیدواژگان طرح‌های پژوهشی هر پژوهشگر با دسته‌بندی انجام شده

با توجه به بالاتر بودن تطبیق همکاران گروه سیستم‌های اطلاعاتی با موضوعات بین‌المللی، تفسیر شکل‌های بخش قبل بسیار نمایان می‌شود.

یافتن شکاف‌ها و ارائه موضوعات بدیع می‌تواند به بیش از یک دلیل دشوار باشد. در اینجا چالش‌هایی را فهرست می‌کنیم که ممکن است هنگام شناسایی شکاف‌های پژوهشی در هر حوزه با آن‌ها مواجه شد:

۱. تلاش برای برخورد با حجم عظیمی از اطلاعات: ممکن است سوالات بی پاسخ زیادی در زمینه مورد علاقه محقق وجود داشته باشد. بنابراین ممکن است با تعداد شکاف‌های تحقیقاتی که به آن برخورد می‌کنند غرق شده و در مورد اینکه باید روی کدام یک تمرکز کنند سردرگم شوند.
۲. دشواری جستجوی سازماندهی شده: ممکن است برای برخی از محققان در سازماندهی

اطلاعاتی که جمع‌آوری کرده‌اند دارای مشکل باشد. اگر مطالعات به درستی یادداشت و دسته‌بندی نشوند، به راحتی ایده‌ها از دست دادنی است.

۳. تردید در زیر سوال بردن هنجارهای تثبیت شده: برخی از محققین به اندازه کافی برای به چالش کشیدن دانش موجود در حوزه خود اطمینان ندارند و ممکن است در مورد سؤال کردن آنچه دیگران در کار خود ادعا کرده‌اند تردید کنند.

۴-۳-۴. مقایسه حوزه‌ای تناسب طرح‌ها

قدمت تولید علم را می‌توان از زمانی که بشر برای بقا و بهبود کیفیت زندگی خود در تلاش برای کشف واقعیت‌های جهان هستی برآمد متصور بود. ثبت و مستندسازی علم نیز قدمتی کهن دارد که در حیات طولانی خود شاهد تحولات بنیادینی بوده است. اختراع خط، اختراع کاغذ، اختراع ماشین چاپ و ... زمینه لازم برای تولید انبوه مستندات علمی را در سطح جهان فراهم ساخت. اختراع کامپیوتر و ظهور اینترنت بستر لازم برای نشر و عمومی‌سازی علم را در گسترده بسیار بزرگی پدید آورد و شیوه‌های نوینی از ارتباطات علمی را پایه‌گذاری کرد. این حجم زیاد از اطلاعات علمی سبب شد تا سیاست‌گذاران و متخصصین اطلاعات، شیوه‌ها و فنون جدیدی را برای سنجش، کنترل، مدیریت و برنامه‌ریزی علم و فناوری بکار گیرند. تحلیل استنادی، تکنیک‌های آماری، روش‌های مبتنی بر الگوریتم‌های پیچیده رایانشی، ابزارها و تکنیک‌های تحلیل داده و متن و ... تنها بخشی از این شیوه‌های نوین هستند که گاه به‌تنهایی و گاه به‌صورت ترکیبی موجب شده‌اند تا این حجم از اطلاعات و مستندات قابل رصد و ارزیابی باشند (اسدی ۱۳۹۳).

با وجود قدمت طولانی علم، دانش علم‌سنجی قدمت چندانی ندارد. به‌طور رسمی از نیمه دوم قرن بیستم علم‌سنجی به‌عنوان یک رشته علمی در آمریکا موردتوجه قرار گرفت. پس از پایان جنگ جهانی دوم روش پژوهش علمی مورد تأکید قرار گرفت و در فرایند انجام پژوهش از روش‌های آماری و تکنیک‌های مبتنی بر اعداد و ارقام بهره گرفته شد. به‌عنوان مثال رشته‌های روان‌سنجی و زیست‌سنجی به‌عنوان شاخه‌های کمیت‌پذیر در رشته‌های روانشناسی و زیست‌شناسی بنانهاده شدند؛ علم‌سنجی، کتاب‌سنجی و اطلاع‌سنجی نیز به‌عنوان حوزه‌های سنجش کمی در حوزه کتابداری و اطلاع‌رسانی و مدیریت پژوهش مطرح شدند (امیرحسینی ۱۳۷۱).

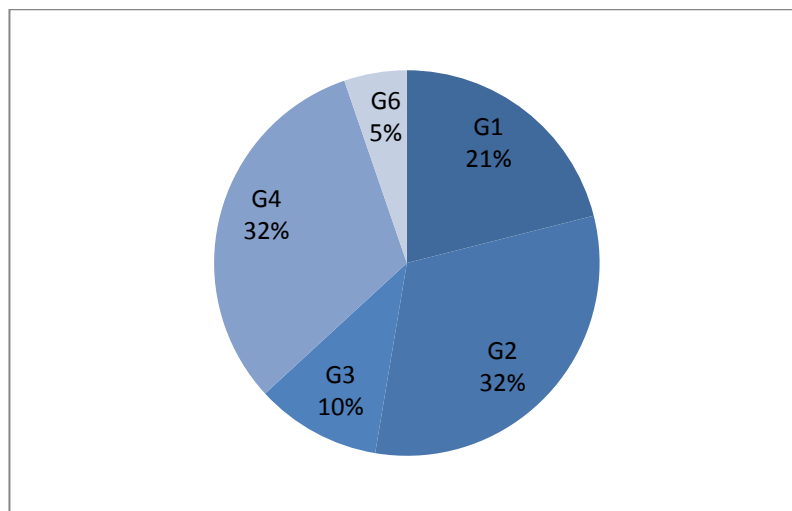
در این بخش قصد داریم میزان پوشش هر یک از مطالعات گروه‌های پژوهشی را با دسته‌بندی‌های تفکیک شده فصل قبل بررسی کنیم.

پژوهشکده فناوری اطلاعات در زمینه کاربرد این فناوری همچون منبعی کلیدی برای توانمندسازی استراتژیک و عملیاتی در طراحی، ساخت، و عرضه محصولات برای بیشینه‌سازی رضایت کاربران، بهره‌وری، و سودآوری کار می‌کند.

سامانه‌های مدیریت اطلاعات و اطلاعات مدیریت؛ تعامل انسان و اطلاعات؛ امنیت اطلاعات؛ پذیرش فناوری اطلاعات؛ عمومی‌سازی کاربرد فناوری اطلاعات؛ سامانه‌های ارزش‌افزا؛ آینده‌پژوهشی فناوری اطلاعات؛ پایش و سنجش نفوذ فناوری اطلاعات؛ سامانه‌های پایش و سنجش مکانی اطلاعات علم، فناوری، و نوآوری؛ هوش کسب‌وکار؛ و کاربست فناوری اطلاعات در زنجیره ارزش، شبکه ارزش، و کانون ارزش از زمینه‌های پژوهش پژوهشکده در چارچوب اساسنامه و برنامه استراتژیک ایرانداک و سندهای بالادست علم و فناوری هستند (ایرانداک، ۲۰۲۲ irandoc.ac.ir).

در ابتدا به بررسی گروه پژوهشی مدیریت فناوری اطلاعات می‌پردازیم. همانگونه که اشاره شد دسته‌های استخراج شده در گروه پژوهشی مدیریت فناوری اطلاعات بصورت زیر است:

- اینترنت اشیا (G1)
- رایانش ابری (G2)
- ارتباطات بی‌سیم (G3)
- یادگیری عمیق (G4)
- امنیت (G5)
- یادگیری ماشین (G6)



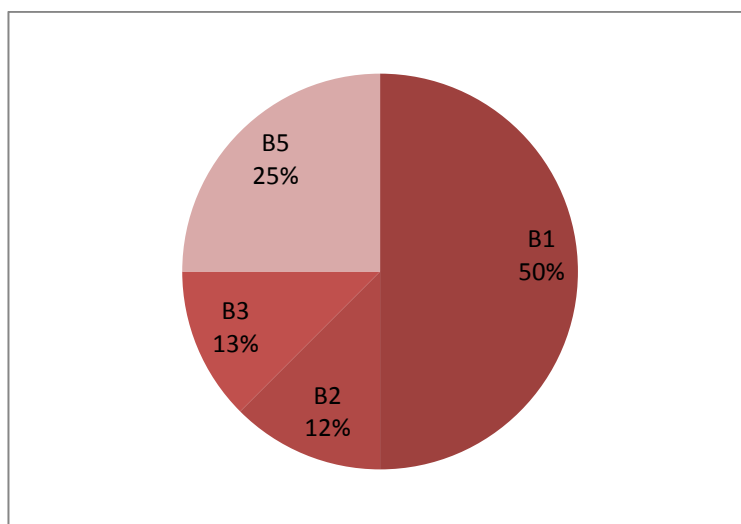
شکل ۳۵- ارتباط کلیدواژگان در طرح‌های پژوهشی گروه مدیریت فناوری اطلاعات با حوزه‌های دسته‌بندی

همانگونه که از شکل فوق مشخص است، تمرکز این گروه پژوهشی بر موضوعات ۳ و ۶ کمتر از دیگر موضوعات است.

موضوعات ارتباطات بی سیم و یادگیری ماشین در این گروه پژوهشی مورد مطالعه عمیق قرار نگرفته است. همچنین موضوع امنیت بدلیل عدم وجود تخصص در این زمینه کلاً مورد بررسی قرار نگرفته است. از نظر خبرگان که در این طرح پژوهشی مورد مطالعه قرار گرفته‌اند، ارتباطات بی سیم بعلت کمبود تجهیزات و همچنین تخصص کافی و یادگیری ماشین بعلت عدم تمرکز پژوهشگران این گروه بر این موضوع، مورد توجه قرار نگرفته است. همچنین در مباحث دیگر نیز ارتباط بصورت متوسط (و نه زیاد) مشاهده می شود که در حقیقت ناشی از عدم پرداختن طرح‌ها به کلیدواژه‌های اصلی موضوع پژوهشی استخراج شده است.

در گروه پژوهشی کسب و کار الکترونیک موضوعات مطرح شده در فصل قبل موارد زیر بوده‌اند:

- یادگیری برخط (B1)
- سلامت (B2)
- تجربه‌ی کاربری (B3)
- واقعیت مجازی (B4)
- شبکه‌های اجتماعی (B5)
- یادگیری عمیق (B6)



شکل ۳۶- ارتباط کلیدواژگان در طرح‌های پژوهشی گروه کسب و کار الکترونیک با حوزه‌های دسته‌بندی

همانگونه که در شکل فوق مشخص است بیشتر مطالعات در زمینه اول و انجام پذیرفته و موضوعات ۴ و ۶ مورد بررسی قرار نگرفته‌اند.

از نظر خبرگان شاید اتفاقات پاندمی اخیر و عزم ایرانداک بر تمرکز بر یادگیری الکترونیکی توانسته است تمرکز خوبی را در زمینه اول رقم بزند. موضوع چهارم به دلیل نوظهور بودن در این گروه مورد

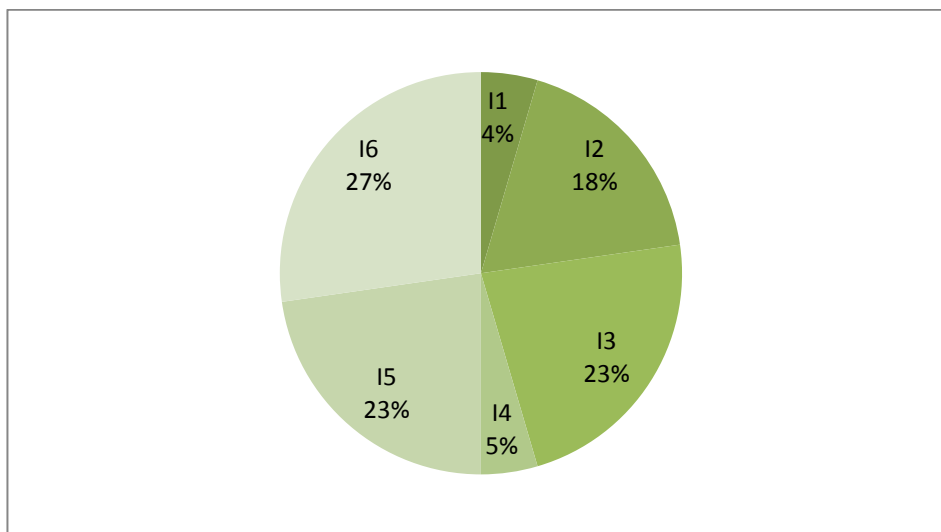
مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۱۳۰

توجه قرار نگرفته و مطالب مرتبط با موضوع ششم با توجه به عدم تخصص کافی در این گروه پژوهشی، مورد پژوهش قرار نگرفته است.

در گروه پژوهشی سیستم‌های اطلاعاتی تناسب بهتری در موضوعات مورد مطالعه به چشم می‌خورد. موضوعات بررسی شده و جمع‌بندی شده در ادبیات در فصل قبل به شکل زیر فهرست شده است:

- اینترنت اشیا (I1)
- هوش مصنوعی (I2)
- یادگیری ماشین (I3)
- امنیت (I4)
- سمانتیک (I5)
- سیستم‌های پیشنهاددهنده (I6)

در نمودار زیر می‌توان مشاهده نمود که موضوعات اول و چهارم می‌توانند بیشتر مورد توجه این گروه قرار گیرند.

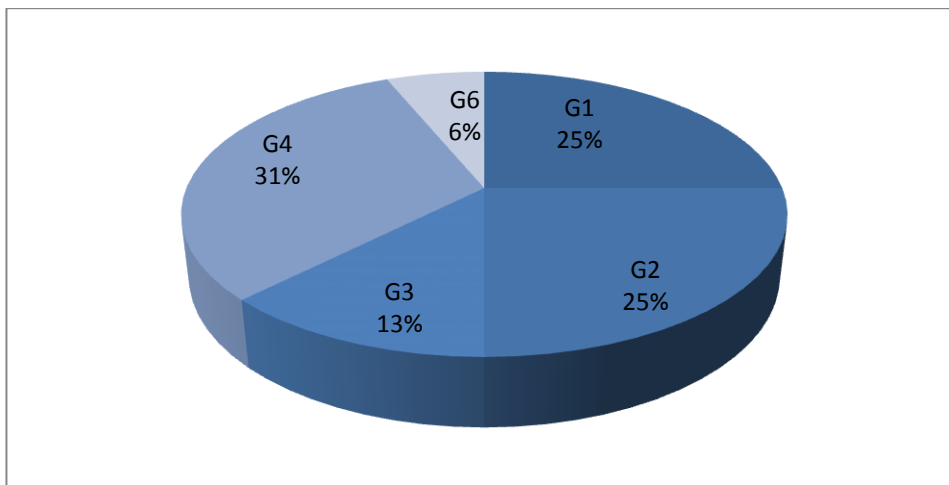


شکل ۳۷- ارتباط کلیدواژگان در طرح‌های پژوهشی گروه سیستم‌های اطلاعاتی با حوزه‌های دسته‌بندی

با توجه به شکل فوق در این گروه پژوهشی مورد اول و پنجم جای کار بیشتری را می‌طلبند. از نظر خبرگان عدم تقاضای بازار و سرمایه‌گذاری در زمینه اینترنت اشیا و همچنین موضوع امنیت به علت عدم وجود تخصص کافی در پژوهشگران این گروه قابل بهبود است. با وجود بررسی تعداد مطالعات در هر زمینه، میزان تناسب مطالعه (کم، متوسط و زیاد) نیز می‌تواند در چگونگی هزینه کرد در این موضوعات نشان‌دهنده و شفاف‌تر باشد.

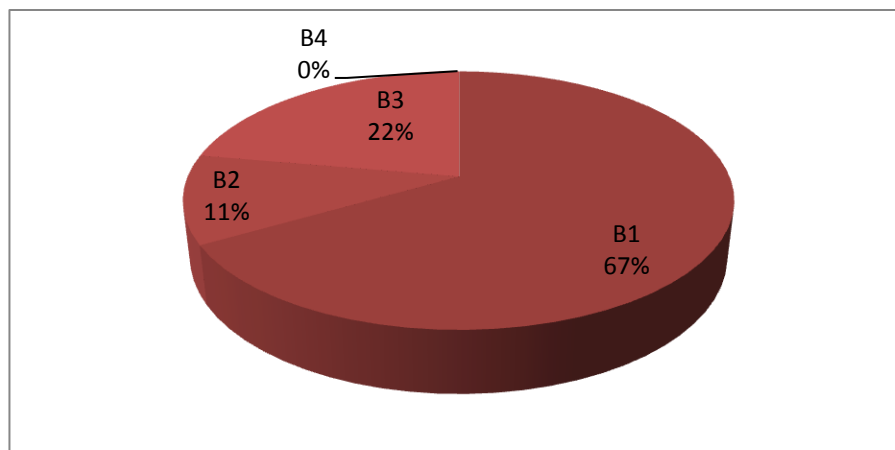
مقایسه تطبیقی حوزه‌های فناوری اطلاعات در گروه‌های پژوهشی ... □ ۱۳۱

با توجه به مساله فوق، ضرایب ۱ و ۲ و ۳ را به طرح‌های پژوهشی با میزان ارتباط کم، متوسط و زیاد اختصاص داده و درصد جمع موزون هر طرح را در نمودارهای زیر نمایش می‌دهیم. با توجه به شکل زیر به نظر می‌رسد موضوعات ۵ و ۶ دارای کیفیت و کمیت ناکافی در پژوهش‌های جاری گروه پژوهشی مدیریت فناوری اطلاعات هستند:



شکل ۳۸- ارتباط وزن‌دار کلیدواژگان در طرح‌های پژوهشی گروه مدیریت فناوری اطلاعات

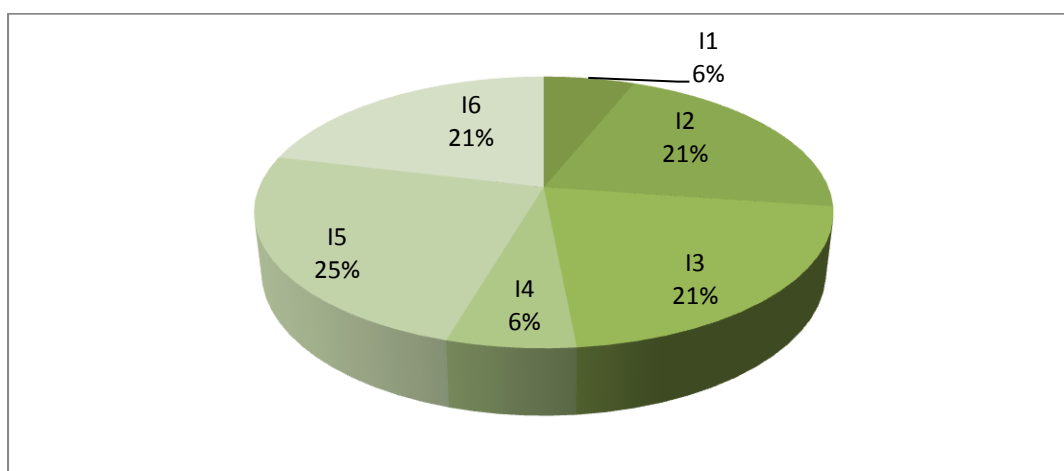
اما در گروه پژوهشی کسب و کار الکترونیک تمرکز عمده بر یادگیری الکترونیک و کیفیت همراهی نیز در این زمینه بیشتر است (شکل زیر):



شکل ۳۹- ارتباط وزن‌دار کلیدواژگان در طرح‌های پژوهشی گروه کسب و کار الکترونیک با حوزه‌های دسته‌بندی

بنابراین به نظر می‌رسد موضوعات دیگر و از جمله موضوعات واقعیت مجازی، شبکه‌های اجتماعی و یادگیری عمیق باید بسیار مورد توجه فعالیت‌ها قرار گیرد. حتی به نظر خبرگان جذب همکارانی با تخصص در آن زمینه‌ها در این گروه پژوهشی بسیار ضروریست.

در گروه پژوهشی سیستم‌های اطلاعاتی تناسب تعداد و کیفیت همراهی بسیار بهتری در مقایسه با دیگر گروه‌های پژوهشی به چشم می‌خورد (شکل زیر):



شکل ۴۰- ارتباط وزن دار کلیدواژگان در طرح‌های پژوهشی گروه سیستم‌های اطلاعاتی با حوزه‌های دسته‌بندی

همانگونه که اشاره شد، موضوعات امنیت و اینترنت اشیا می‌تواند محل تمرکز کمی بیشتری در این گروه پژوهشی باشد. با این حال گروه پژوهشی مذکور بهترین تناسب را با مطالعات بین‌المللی و دسته‌بندی آنها نسبت به دیگر گروه‌های پژوهشی داراست.

۴-۴. تحلیل تطابق پژوهش‌ها با اسناد بالادستی

اولویت‌بندی پژوهش، انتخاب طرح‌های پژوهشی بر مبنای یک سری از معیارهای کمی و کیفی است. اولویت‌بندی پژوهش‌های علمی، ذاتاً کار مشکلی است چراکه اهداف سازمانهای حامی و متقاضی پژوهش متفاوت است و شناسایی و درک این اهداف نیاز به آگاهی از حوزه‌های دانشی مختلف مانند ثبت اختراع، فروش یافته‌ها، کاربرد اطلاعات در حوزه‌های مختلف و ... دارد و از طرف دیگر نیازهای پژوهش و متعاقباً اولویت‌های پژوهشی به مرور زمان تغییر می‌کند و با توجه به سرعت تولید دانش در حوزه‌های مختلف، این تغییر در مدت زمان کوتاهی صورت می‌پذیرد (Mona and Vivian, 2013).

یکی از مباحث مهم در مقوله نیازسنجی پژوهشی، شناسایی و تعیین گروه‌های هدف در فرایند شناسایی و تعیین اولویت‌های تحقیقاتی است. به عبارتی پژوهش‌های مختلف درصدد حل مسائل و مشکلات مجموعه‌ای از مخاطبان هستند که ضرورتاً باید از پیش، شناسایی و مورد تجزیه و تحلیل قرار گیرند. مخاطبان پژوهش می‌توانند یکی یا ترکیبی از سه سطح زیر باشند:

- دریافت‌کنندگان خدمات پژوهشی نظیر کاربران، پژوهشگران، مصرف‌کنندگان، دانشجویان و ...
- افرادی که در ساختار سازمانی وظیفه سیاست‌گذاری، برنامه‌ریزی، پشتیبانی و ... در حوزه‌ی پژوهشی مدنظر را بر عهده دارند.

- منابعی نظیر برنامه‌ها، نظام‌ها، طرح‌ها، فناوری‌هایی که به صورت مستقیم یا غیرمستقیم از نتایج پژوهش‌ها استفاده خواهند کرد.

همانگونه که پیشتر ذکر شد رویکردهای متفاوتی در زمینه کشف حوزه‌های مورد نیاز برای مطالعه وجود دارد.

در رویکرد اداره‌گری/مدیریتی، ساختار سازمانی و جایگاه نهادهای حاکمیتی در تعیین نیازها و اولویت‌های پژوهشی حائز اهمیت است. در این روش هر یک از نهادهای بالادستی اولویت‌ها و نیازهای سطح کلان را مشخص می‌کند و سطوح پایین‌تر بر مبنای اولویت‌های کلان ارائه شده، لیستی از نیازها و اقدامات پژوهشی را تهیه می‌کنند.

رویکرد فوق در برخی موارد بافت و ساختار مورد استفاده در نیازسنجی و اولویت‌یابی پژوهش را مدنظر قرار داده‌اند و در برخی موارد از نظر اهداف نیازسنجی و اولویت‌یابی با یکدیگر متفاوت‌اند. به همین دلیل، در بسیاری از موارد استفاده از یک رویکرد به تنهایی نتایج موفقیت‌آمیزی نخواهد داشت و ترکیبی از رویکردها انتخاب می‌شود و تلاش می‌شود تا با ترکیب نگرش‌های بیان‌شده، نیازهای پژوهشی از ابعاد مختلف شناسایی شده و در اولویت‌بندی پژوهش‌ها، حوزه‌هایی مدنظر قرار گیرد که بیشترین میزان تأثیر را بر جامعه پژوهشی داشته و بهره‌وری حاصل از سرمایه‌گذاری پژوهشی را بیشینه می‌کنند.

با توجه به نکته فوق ذکر این مساله ضروریست که هر ساله فرم‌های متعددی برای ایرانداک توسط نهادهای بالادست ارسال می‌شود که بخشی از آنها با برنامه استراتژیک ایرانداک پیوند دارد. در ادامه به ارائه گزارش‌های درخواستی سه نهاد بالادست و نظارتی ایرانداک یعنی سازمان برنامه و بودجه در قالب بودجه ریزی مبتنی بر عملکرد، دیوان محاسبات در قالب گزارش پیاده سازی بودجه ریزی مبتنی بر عملکرد و هیئت امنا در قالب گزارش عملکرد ایرانداک در راستای سیاست‌های هیئت امنا پرداخته می‌شود. همچنین گزارش‌های زیر نیز سالانه از ایرانداک دریافت و مورد رصد قرار می‌گیرد.

- اولویت‌های عمومی و اختصاصی ریاست جمهوری
- سامانه مدیریت پژوهش و فناوری ایران (مپفا) وزارت عتف
- گزارش به تسما وزارت عتف
- گزارش به مرکز آمار
- گزارش به سمات ملی
- گزارش عملکرد سیاست‌های هیئت امنا
- گزارش به مؤسسه پژوهش و برنامه‌ریزی آموزش عالی
- گزارش به مرکز نظارت، ارزیابی و تضمین کیفیت وزارت عتف

- گزارش به «آی.اس.سی.»
- گزارش بودجه سالانه
- گزارش اقتصاد مقاومتی
- گزارش عملکرد برنامه و بودجه

در بودجه سالانه ایرانداک که بر پایه بودجه‌ریزی بر مبنای عملکرد تنظیم می‌شود فرم‌های زیادی وجود دارد که در آنها برخی شاخص‌ها پیوند نزدیکی با برنامه استراتژیک ایرانداک دارد چرا که بودجه‌ریزی عملیاتی یا مبتنی بر عملکرد از برنامه استراتژیک سازمان شروع می‌شود. موارد مدنظر در فرم‌های بودجه‌ریزی عبارتند از: اهداف، راهبردها، برنامه‌های اجرایی، سنجه عملکرد/هدف کمی/پیامدها یا نتیجه نهایی/خروجی (محصول/خدمت/طرح)، هزینه‌های پرسنلی، هزینه‌های غیر پرسنلی، قیمت تمام شده خدمت/محصول در دستگاه مجری، و پرداخت‌های انتقالی.

از سال ۹۸، دیوان محاسبات خواستار گزارش عملکرد ایرانداک در راستای پیاده‌سازی بودجه‌ریزی بر مبنای عملکرد شده و خواستار مدارک و مستندات در ۲۰ مورد شدند که در زیر موارد مرتبط با برنامه استراتژیک آمده است.

۱. برنامه رهردی و عملیاتی مستند و مدون
 ۲. درختواره اهداف و برنامه‌های سازمان
 ۳. ارتباط بین محصولات و خدمات سازمان با درختواره برنامه‌ها
 ۴. درختواره فعالیت‌ها و فرآیندهای سازمان
 ۵. تعریف ارتباط بین درختواره فعالیت‌ها و فرآیندها با محصولات
 ۶. احصاء زمان صرف شده برای فعالیت‌ها در هر واحد سازمانی
 ۷. احصاء خدمت و محصولات سازمان
 ۸. هزینه تمام شده فعالیت‌ها، محصولات و خدمات، برنامه‌ها و ... به تفکیک واحدهای مجری
 ۹. احصاء سنجه‌ها و شاخص‌های کلیدی عملکرد بر اساس اهداف سازمان
- هیئت امنا به عنوان تصویب‌کننده برنامه استراتژیک، افزون بر دریافت گزارش عملکرد سالانه برنامه استراتژیک ایرانداک، گزارشی نیز با عنوان عملکرد سیاست‌های هیئت امنا دریافت می‌کند که جدول زیر فرم گزارش است. همانگونه که از جدول مشاهده می‌شود موارد زرد شده در بندهای گزارش معیارها و الزاماتی برای تدوین برنامه استراتژیک ایرانداک نشان می‌دهد.
- در این قسمت موضوعات استراتژیک ایرانداک در اسناد بالادستی استخراج شده و تعداد مرجع شدن آنها در قالب سه پژوهشکده ایرانداک استخراج شده و توسط ریاست ایرانداک تایید شده است:

- شکوفایی علمی

- توسعه فناوری اطلاعات
- افزایش بهره‌وری نظام ملی نوآوری
- توسعه نوآوری
- ساماندهی نظام ملی نوآوری
- توسعه اخلاق حرفه‌ای در محیط‌های علمی و پژوهشی
- توسعه مهارت‌های تحقیقاتی اهالی علم و فناوری
- ایجاد، ساماندهی و تقویت نظام ملی اطلاع‌رسانی رایانه‌ای
- انسجام بخشی به نظام اطلاعات علمی و فناوری
- ثبت پیشنهادیه‌ها و پژوهش‌ها در سامانه ملی ثبت ایرنداگ
- تغییر ساختارها و سیستم‌های آموزش و پژوهش در جهت دستیابی به توسعه پایدار
- سامان‌دهی نظام مدیریت پژوهش در کشور
- تربیت نیروی متخصص در فناوری اطلاعات
- تقاضا محوری
- استقرار دولت الکترونیک
- بهینه سازی عملکرد در نظام آموزشی و تحقیقاتی
- ایجاد بانک‌های اطلاعاتی یکپارچه برای رساله‌ها، پایان‌نامه‌ها، طرح‌های پژوهشی و فناوری، مقالات، مجلات و کتب علمی، اختراعات و اکتشافات پژوهشگران
- آینده‌نگری در خصوص آثار تحولات فناوری اطلاعات در سطح ملی و جهانی
- ترویج فرهنگ احترام به حقوق مالکیت فکری
- بهینه‌سازی ساختار نظام آموزشی و تحقیقاتی کشور
- ارتقاء رتبه جهانی دانشگاه‌ها
- ساماندهی نظام ملی آمار و اطلاعات علمی
- ایجاد نظام پایگاه داده و اطلاعات پژوهش و فناوری کشور
- رشد تولید مقالات
- تثبیت جایگاه کشور در فناوری اطلاعات
- شفافیت اطلاعات
- مدیریت دانش
- دسترسی همگانی به اطلاعات
- استانداردسازی و اصلاح فرآیندهای تولید، ثبت، داوری و سنجش علم و فناوری
- توسعه تحقیقات کاربردی

- تکمیل زیرساخت‌ها و قوانین مالکیت فکری
 - حمایت از مالکیت فکری و معنوی
 - ارزیابی جامع عملکرد نظام ملی علوم، تحقیقات و فناوری
 - اجرای نظام اشتراک منابع میان مراکز آموزشی و پژوهشی در کشور
 - نیازسنجی پژوهشی
 - تسهیل و کارآمدسازی فرآیند عرضه و تقاضا و انتقال و انتشار علم و فناوری
 - گنجاندن دروس مرتبط با حمایت از مالکیت فکری در محتوای آموزشی
 - الزام مراکز علمی به اعمال نظارت بر رعایت حقوق مالکیت فکری
 - تحلیل جریان علمی کشور
 - توسعه پژوهش
 - تعیین استاندارد های بومی در حوزه علم و فناوری
 - تسهیل تبادل دانش
 - افزایش کارایی و اثربخشی تحقیقات
 - جلوگیری از انجام پژوهش‌های تکراری
 - مسئله محور کردن پژوهش
 - رعایت موازین دینی، فرهنگی و ارزشی در بکارگیری علم و فناوری
 - پیشگیری و مقابله با تقلب در تهیه آثار علمی
 - ثبت پژوهش‌ها در سامانه سمات
 - ساده‌سازی زبان علم برای عموم مردم
 - تقویت جریان معادل‌سازی واژگان علمی در تمام رشته‌ها
 - حمایت از توسعه علوم و فناوری‌های میان رشته‌ای
 - تدوین شاخص‌های سنجش بهره‌وری نظام علم و فناوری و پایش آنها
 - تدوین اطلس ظرفیت‌های علم و فناوری
 - ارتقاء جایگاه جهانی کشور در علم و فناوری
 - ساماندهی نظام اعتبارسنجی
- به منظور تحلیل فاصله بین عرضه و تقاضای موضوعات استخراج شده، لازم است تا میزان عرضه و تقاضای هر یک از موضوعات شفاف شود. برای این منظور موضوعات فوق را با دسته بندی اجماعی خبرگان به چند دسته زیر تقسیم می‌شوند:
- پژوهش و توسعه علم
 - نوآوری و بهره‌وری

- توسعه مهارتی
- پشتیبانی از سیاست‌گذاری
- اخلاق علم
- استانداردسازی

جدول زیر توسط خبرگان تشکیل شده است و دسته‌بندی‌های انجام شده در هر گروه پژوهشی که می‌تواند پوشش‌دهنده هر دسته راهبردی فوق باشد استخراج شده است. در این جدول سعی شده میزان موضوعات پوشش‌دهنده هر دسته راهبردی توسط دسته‌بندی مطالعات گروه‌های پژوهشی مشخص گردد.

جدول ۱۷- تطبیق دسته‌بندی‌ها و مطالعات پژوهشکده بر نیازهای بالادستی

نام دسته راهبردی	نام واژه ذکر شده در اسناد و گزارش‌های بالادست و تایید شده توسط ایرانداک	دسته‌بندی‌های مرتبط	درصد وزن‌دار
پژوهش و توسعه علم	شکوفایی علمی	رایانش ابری یادگیری عمیق یادگیری ماشین هوش مصنوعی	۳۲
	توسعه فناوری اطلاعات		
	آینده‌نگری در خصوص آثار تحولات فناوری اطلاعات در سطح ملی و جهانی		
	ارتقاء رتبه جهانی دانشگاه‌ها		
	رشد تولید مقالات		
	مدیریت دانش		
	نیازسنجی پژوهشی		
	تسهیل و کارآمدسازی فرآیند عرضه و تقاضا و انتقال و انتشار علم و فناوری		
	توسعه پژوهش		
	تسهیل تبادل دانش		
	مسئله محور کردن پژوهش		
	ثبت پژوهش‌ها در سامانه سمات		
	حمایت از توسعه علوم و فناوری‌های میان رشته‌ای		
نوآوری و بهره‌وری	ارتقاء جایگاه جهانی کشور در علم و فناوری	ارتباطات بی‌سیم سلامت تجربه‌ی کاربری واقعیت مجازی شبکه‌های	۱۵
	افزایش بهره‌وری نظام ملی نوآوری		
	توسعه نوآوری		
	ساماندهی نظام ملی نوآوری		
	ایجاد، ساماندهی و تقویت نظام ملی اطلاع‌رسانی رایانه‌ای		
	انسجام بخشی به نظام اطلاعات علمی و فناوری		

نام دسته	نام واژه ذکر شده در اسناد و گزارش‌های بالادست و تایید شده توسط ایرانداک	دسته‌بندی‌های مرتبط	درصد وزن‌دار
	تغییر ساختارها و سیستم‌های آموزش و پژوهش در جهت دستیابی به توسعه پایدار	اجتماعی اینترنت اشیا	
	سامان‌دهی نظام مدیریت پژوهش در کشور		
	استقرار دولت الکترونیک		
	بهینه سازی عملکرد در نظام آموزشی و تحقیقاتی		
	ساماندهی نظام ملی آمار و اطلاعات علمی		
	استانداردسازی و اصلاح فرآیندهای تولید، ثبت، داوری و سنجش علم و فناوری		
	اجرای نظام اشتراک منابع میان مراکز آموزشی و پژوهشی در کشور		
	افزایش کارایی و اثربخشی تحقیقات		
توسعه مهارتی	توسعه مهارت‌های تحقیقاتی اهالی علم و فناوری	یادگیری برخط سیستم‌های پیشنهاددهنده	۱۸
	تربیت نیروی متخصص در فناوری اطلاعات		
	تثبیت جایگاه کشور در فناوری اطلاعات		
	توسعه تحقیقات کاربردی		
	ساده‌سازی زبان علم برای عموم مردم		
پشتیبانی از سیاستگذاری	ثبت پیشنهادیه‌ها و پژوهش‌ها در سامانه ملی ثبت ایرانداک	امنیت سمانتک	۷
	تقاضا محوری		
	بهینه سازی ساختار نظام آموزشی و تحقیقاتی کشور		
	ایجاد نظام پایگاه داده و اطلاعات پژوهش و فناوری کشور		
	شفافیت اطلاعات		
	دسترسی همگانی به اطلاعات		
	تحلیل جریان علمی کشور		
	جلوگیری از انجام پژوهش‌های تکراری		
	پیشگیری و مقابله با تقلب در تهیه آثار علمی		
	تدوین اطلس ظرفیت‌های علم و فناوری		
اخلاق علم	توسعه اخلاق حرفه‌ای در محیط‌های علمی و پژوهشی		۱۰
	ترویج فرهنگ احترام به حقوق مالکیت فکری		
	تکمیل زیرساخت‌ها و قوانین مالکیت فکری		
	حمایت از مالکیت فکری و معنوی		

نام دسته	نام واژه ذکر شده در اسناد و گزارش‌های بالادست و تایید شده توسط ایرانداک	دسته‌بندی‌های مرتبط	درصد وزن‌دار
	گنجانیدن دروس مرتبط با حمایت از مالکیت فکری در محتوای آموزشی		
	الزام مراکز علمی به اعمال نظارت بر رعایت حقوق مالکیت فکری		
	رعایت موازین دینی، فرهنگی و ارزشی در بکارگیری علم و فناوری		
استانداردسازی	ایجاد بانک‌های اطلاعاتی یکپارچه برای رساله‌ها، پایان‌نامه‌ها، طرح‌های پژوهشی و فناوری، مقالات، مجلات و کتب علمی، اختراعات و اکتشافات پژوهشگران		۰
	ارزیابی جامع عملکرد نظام ملی علوم، تحقیقات و فناوری		
	تعیین استاندارد های بومی در حوزه علم و فناوری		
	تقویت جریان معادل‌سازی واژگان علمی در تمام رشته‌ها		
	تدوین شاخص‌های سنجش بهره‌وری نظام علم و فناوری و پایش آنها		
	ساماندهی نظام اعتبارسنجی		

با توجه به نقاط نظر خبرگان، مسایلی مانند شبکه می‌توانست در دسته توسعه علمی نیز قرار بگیرد اما تمرکز بیشتر مطالعات جاری دنیا بر کاربردی‌سازی آنها بوده است. با توجه به جدول فوق به نظر می‌رسد از دید اسناد بالادستی، پژوهشکده فناوری اطلاعات در زمینه استانداردسازی و فعالیت‌هایی مانند بانک‌های اطلاعاتی یکپارچه و ... نیاز به تمرکز بالایی دارد. زمینه‌هایی مثل اخلاق علم و پشتیبانی از سیاستگذاری طبق سیاست‌های داخلی می‌توانند توسط دیگر پژوهشکده‌های ایرانداک توسعه و فعالیت نمایند.

فصل پنجم

نتیجه گیری و تحقیقات آتی

۵-۱. مقدمه

در این طرح مفاهیم نظری کاستی پژوهش، نیازسنجی و اولویت‌یابی پژوهش مطرح و رویکردهای اصلی که در نیازسنجی و اولویت‌یابی پژوهش مطرح هستند معرفی شدند. پژوهش‌های پیشین بیشتر معطوف به تحلیل‌های آماری از اطلاعات کتابشناختی پژوهش‌های عرضه‌شده داشته و درنهایت با دسته‌بندی این حوزه‌ها به شناسایی نیازهای پژوهشی پرداخته شده است. به نظر می‌رسد در بخش قابل‌توجهی از پژوهش‌های پیشین، پس از تحلیل وضع موجود از نظرات تعدادی از خبرگان برای تعیین نیازها و اولویت‌های پژوهشی بهره‌برده شده است. این موارد باعث شده است تا شناسایی کاستی و نیازسنجی پژوهش، متکی به بخش عرضه پژوهش باشد و نیازهای جامعه پژوهشی کمتر موردتوجه قرار گیرد.

در بخش اولویت‌یابی پژوهش نیز، پژوهش‌های پیشین عمدتاً مبتنی بر روش‌های بالا به پایین بوده و ساختارهای تصمیم‌گیر و سیاست‌گذار پژوهشی مبتنی بر شهود و دانش مدیران رده‌بالا، نتیجه‌گیری کرده‌اند. این امر باعث شده است تا کمتر به موضوعات جدید و ناب (مطرح نشده در پژوهش‌های پیشین) پرداخته و مقایسه انجام شود.

هرچند که استفاده از روش‌های ماشینی و تکیه بر تحلیل حجم انبوه اطلاعات، اخیراً در فرایندهای تصمیم‌گیری و برنامه‌ریزی موردنظر قرار گرفته‌اند اما کاربرست این روش‌ها در منابع زبان فارسی بسیار اندک بوده و الگوریتم‌ها و ابزارهای توسعه داده‌شده در این روش‌ها نیز بیشتر مبتنی بر متون عمومی مانند روزنامه و وب‌سایت‌های خبری و ... است و کمتر به متون علمی و پژوهشی پرداخته شده است.

۵-۲. تطبیق طرح‌های پژوهشی با نیاز بالادستی

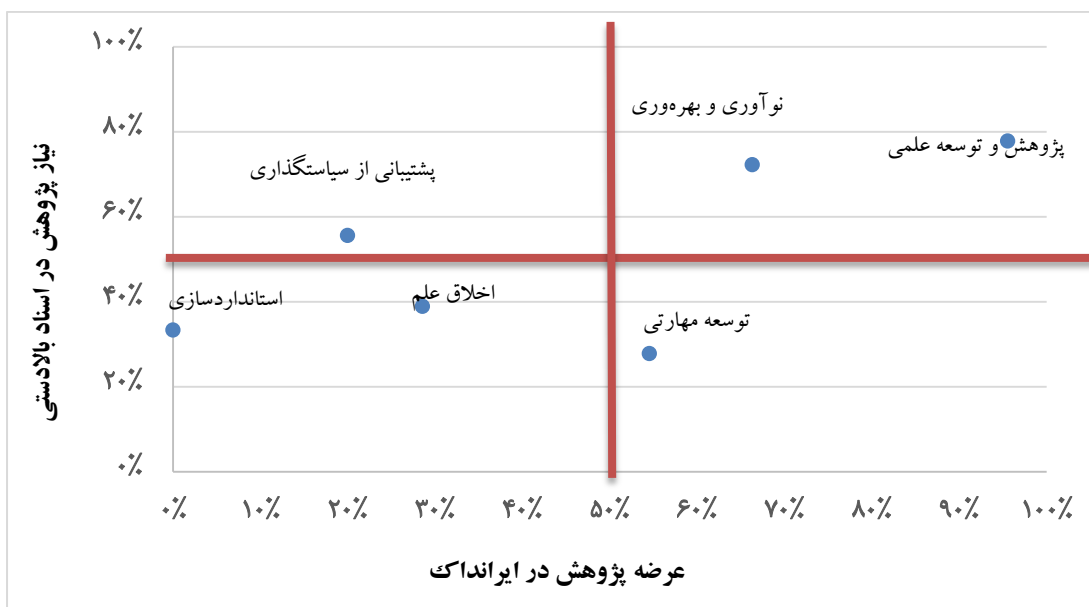
شناسایی کاستی‌های پژوهشی در حوزه‌های مختلف یکی از پیش‌نیازهای اساسی برای نیازسنجی و تعیین اولویت‌های پژوهشی است. رصد وضعیت پژوهش و شناسایی کاستی‌های پژوهش از منظر پژوهشگران که عهده‌دار عرضه و تقاضای پژوهش هستند مدنظر این پژوهش بوده است. روش‌های

گوناگونی برای شناسایی نیازمندی‌های پژوهشی در یک حوزه ارائه شده است که عمده آنها به جمع‌آوری و تحلیل نظرات خبرگان آن حوزه معطوف می‌شود. از آنجا که نظرسنجی از خبرگان دارای هزینه بسیار بالایی بوده و زمانبر است و نیز خبرگان بر اساس دانش و برداشت شخصی خود از محیط پیرامون، کاستی‌ها، نیازها و اولویت‌ها را مشخص می‌کنند که ممکن است با واقعیت‌های موجود در جامعه پژوهشی متفاوت باشد، نیاز به ارائه روشی است که با تحلیل رفتار واقعی پژوهشگران و رصد کنشگری آن‌ها در مواجهه با منابع پژوهشی مورد نظر بتواند کاستی‌های واقعی پژوهش را شناسایی کند. رفتار جویشگری پژوهشگران در پایگاه‌های اطلاعاتی قابل ذخیره شدن و تحلیل است، اما از آنجا که تعداد رکوردهای مربوط به این رفتار بسیار بالاست و همچنین تعداد منابع پژوهشی نیز عموماً از چند صد هزار رکورد بیشتر است، نیاز به ارائه روش‌هایی است که بصورت ماشینی و مستقل از خبرگان بتواند این فرایند را انجام دهد.

با این مقدمه با توجه به نرخ رشد جستجوی نتایج پژوهشی و با درنظر گرفتن رفتار کاربران و جستجوگران پایگاه اطلاعاتی، مسیر حرکت هر یک از دسته کلی موضوعات قابل پیش‌بینی است. انتظار می‌رود موضوعات داغ و موضوعات دارای کاستی پژوهش که نرخ رشد تقاضای بالایی دارند در آینده نزدیک شروعی برای پژوهش‌های جدید توسط جستجوگران موضوعات کنونی باشند. از سوی دیگر پیش‌بینی می‌شود موضوعات خاموش و نیز موضوعات سرد که نرخ رشد تقاضای پایینی داشته باشند در آینده نیز پژوهش‌های کمتری را دربرخواهند گرفت.

برای این منظور عرضه پژوهش در پژوهشگاه و نیاز بالادستی در هر حوزه پژوهشی (در شکل ۵-۱) مورد مقایسه قرار گرفت. بنابراین نیازهای پژوهشی به ترتیب زیر توسعه پیدا کرد:

- ۱- استخراج نیازهای کلان با تحلیل اسناد بالادستی
 - ۲- مشخص نمودن محل تمرکز ایرانداک با توجه ماموریت‌های آن توسط خبره
 - ۳- دسته‌بندی نیازهای محل تمرکز به ۶ دسته ذکر شده در شکل
 - ۴- شمارش تعداد طرح‌های پژوهشی متمرکز بر هر دسته
- سپس موضوعات بر مبنای موقعیت خود در دو محور عرضه و تقاضای پژوهش نگاشت شدند (بین ۰ و ۱). با تقسیم محور افقی عرضه به بخش «عرضه بالا» و «عرضه پایین» و نیز با تقسیم محور عمودی به دو بخش «تقاضای بالا» و «تقاضای پایین» ماتریسی شامل چهار بخش بدست خواهد آمد. مقدار ۰,۵ به عنوان مرز بین عرضه و تقاضای بالا و پایین تعیین شده است. شکل زیر این چهار بخش و موقعیت هر یک از موضوعات را نشان می‌دهد.



شکل ۴۱- نواحی تطبیق کلیدواژگان طرح‌های پژوهشی با نیاز بالادستی

با توجه به شکل فوق، ناحیه سمت راست و بالای این نمودار دربردارنده موضوعاتی است که دارای عرضه بالای پژوهشی در پژوهشکده بوده و همچنین مورد تقاضای نیازهای پژوهشی بالادستی هستند. موضوعات پژوهشی در زمینه رایانش ابری، یادگیری عمیق و یادگیری ماشین که در این ناحیه واقع شده‌اند و به عنوان «موضوعات داغ» در نظر گرفته شده‌اند.

ناحیه سمت راست و پایین شکل در بردارنده موضوعاتی است که اصطلاح «موضوعات خاموش»^{۲۶} برای آنها انتخاب شده است. این بخش شامل موضوعاتی است که دارای تعداد قابل توجهی منبع هستند و به میزان زیادی در آنها پژوهش صورت پذیرفته است، اما در حال حاضر با توجه به سیاست‌های بالادستی علاقه زیادی به استفاده از این منابع ندارند. به نظر می‌رسد برخی از پژوهش قدیمی و نیز حوزه‌های پژوهشی که تا حدی اشباع شده باشند در این بخش قرار می‌گیرند. موضوع یادگیری الکترونیک در این زمینه قرار گرفته است.

ناحیه سمت چپ و پایین، ناحیه‌ای است که دربردارنده موضوعاتی است که دارای تعداد اندک پژوهش انجام شده بوده و همچنین این موضوعات مورد استقبال اسناد بالادستی نیز قرار نگرفته‌اند. این ناحیه که دقیقاً نقطه مقابل پژوهش‌های داغ قرار گرفته است با عنوان «موضوعات سرد»^{۲۷} نامگذاری شده است. «اخلاق علم»، و «سمانتک» در این ناحیه قرار گرفته‌اند.

ناحیه سمت چپ و بالا، آخرین بخش از این ماتریس است. این ناحیه دربردارنده موضوعاتی است که اگر چه مورد توجه اسناد بالادستی قرار گرفته است و تقاضای پژوهشی در موضوعات آن بالاست

²⁶ Silent Topics

²⁷ Cold Topics

اما به اندازه کافی منابع پژوهشی برای آنها وجود ندارد. به عبارتی به ازای تقاضای بالای پژوهشی در این موضوعات، با منابع پژوهشی کافی مواجه نمی‌شوند. از این رو نام این ناحیه «موضوعات دارای کاستی»^{۲۸} گذاشته شده است که دربردارنده کاستی‌های پژوهش است. «سیستم‌های پیشنهاددهنده که در راستای نیاز سیاستگذاران باشد» مهم‌ترین موضوعاتی است که دارای کاستی پژوهشی هستند و همانطور که پیش‌تر گفته شد، فعالیت‌های زیادی باید در زمینه مذکور قرار گیرد.

۳-۵. تطبیق طرح‌های پژوهشی با روندهای پژوهشی

با توجه به نرخ رشد و افول پژوهش‌ها در حوزه فناوری اطلاعات و روندهای آن، بررسی میزان پژوهش‌های ایران‌دک با روندهای بین‌المللی بسیار پراهمیت است. بر این اساس در این بخش قصد داریم میزان تمرکز پژوهش‌ها در ایران‌دک را با کلیدواژه‌های پرتکرار بسنجیم. اما روند این کلیدواژه‌ها در مطالعات بین‌المللی در حال افت و صعود متفاوت است. بر این اساس از طرح پژوهشی بررسی روندهای بین‌المللی در حوزه فناوری اطلاعات (ساجدی‌نژاد ۱۳۹۹)، به جدول نرمال شده رشد و صعود پژوهش‌های پرتکرار استناد می‌کنیم.

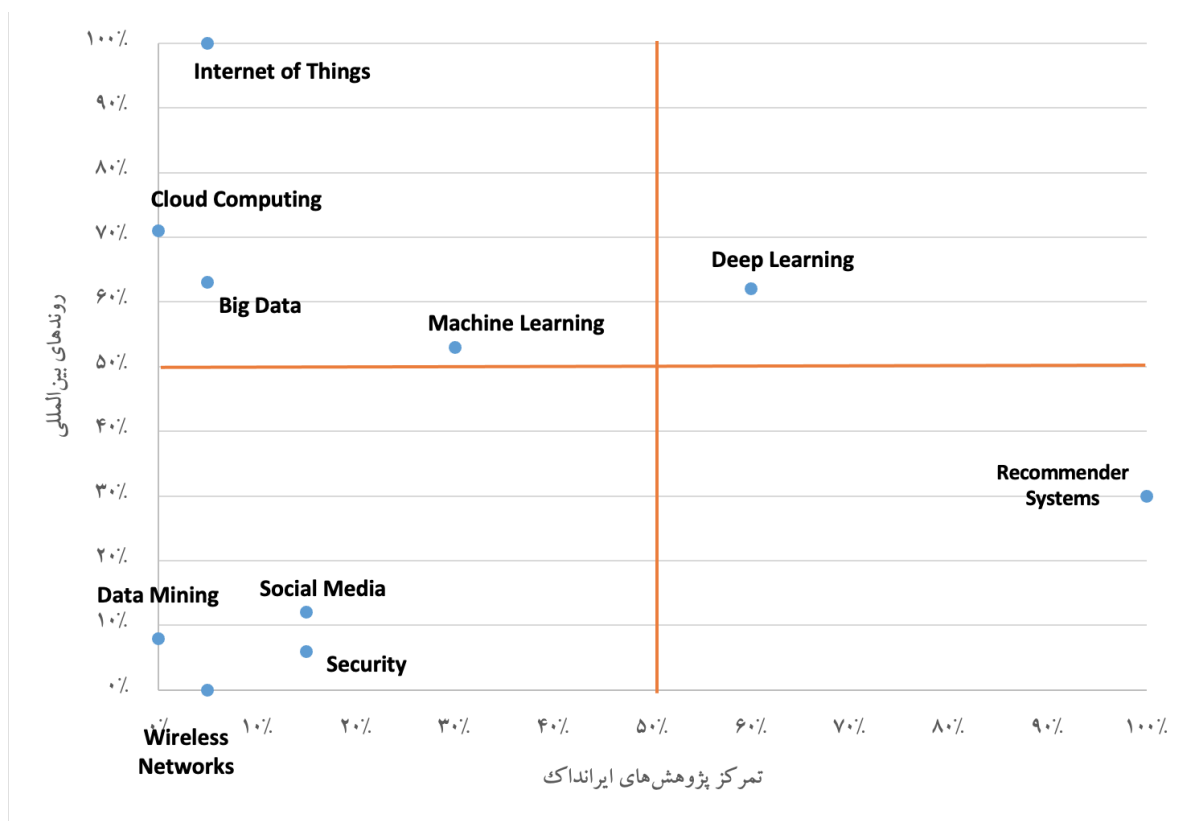
همانگونه که مشاهده می‌شود رشد این زمینه علمی بویژه در حوزه‌های اینترنت اشیا و یادگیری عمیق بسیار مشهود است و زمینه‌های علمی مانند اینترنت اشیا همچنان در حال ادامه است. تعداد کمتر مشاهدات علمی در سال ۲۰۱۹ به علت کامل نشدن تمام انتشارات (یا آنلاین نشدن آنها) است.

جدول ۱۸- مقایسه کلیدواژه‌های پرتکرار در دو بازه زمانی (ساجدی‌نژاد ۱۳۹۹)

کلیدواژه	درصد در بین کلیدواژگان پرتکرار در سال ۲۰۱۱	درصد در بین کلیدواژگان پرتکرار در سال ۲۰۱۹	رشد	درصد نرمال شده
Internet of Things	1.71	22.99	21.28	100%
Cloud Computing	7.85	19.54	11.69	71%
Big Data	0.00	9.20	9.20	63%
Deep Learning	0.14	9.20	9.05	62%
Machine Learning	4.28	10.34	6.07	53%
Recommender Systems	1.43	0.00	-1.43	30%
Social Media	8.56	1.15	-7.41	12%
Data Mining	9.70	1.15	-8.55	8%
Security	20.68	11.49	-9.19	6%
Wireless Networks	12.41	1.15	-11.26	0%

²⁸ Gap Topics

در جدول فوق، درصد‌های نرمال شده بالای ۰,۵، روندهای در حال رشد و درصد‌های زیر ۰,۵ در حال نزول هستند. با توجه به پخش پژوهش‌های ایرانداک (مقالات) در زمینه‌های فوق، نمودار زیر قابل ارائه است:

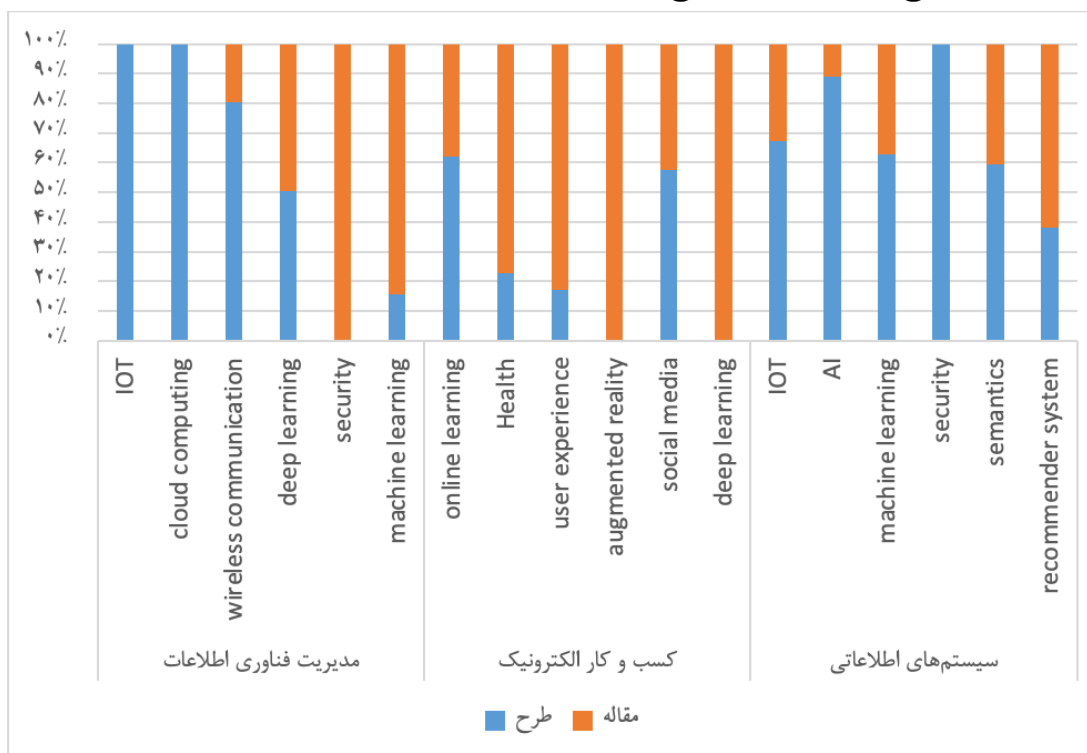


شکل ۴۲- نواحی تطبیق کلیدواژگان طرح‌های پژوهشی با روندهای جهانی

همانگونه که از شکل فوق برمی‌آید متأسفانه روند پژوهش‌های ایرانداک با روندهای بین‌المللی مطابقت زیادی ندارد. در پژوهش‌های ایرانداک تمرکز ویژه بر سیستم‌های پیشنهاددهنده شده که در روندهای بین‌المللی در حال افول است. شاید این تمرکز به دلیل استفاده از این سیستم‌ها در سامانه‌های داخلی ایرانداک باشد. همچنین حوزه‌های داده‌های بزرگ و اینترنت اشیا و محاسبات ابری که در محل رشد تمرکز پژوهش‌های بین‌المللی هستند، متأسفانه در وضعیت مطلوبی در ایرانداک به سر نمی‌برند. حوزه یادگیری ماشین نیز با وجود انجام فعالیت‌هایی همچنان نیاز به تمرکز بیشتر پژوهشگران در ایرانداک خواهند داشت.

۵-۴. مقایسه پخش طرح‌ها و مقالات در موضوعات

در این بخش سعی داریم میزان تمرکز در حوزه‌های پژوهشی در طرح‌ها و مقالات ایرانداک را مقایسه نماییم. با توجه به اینکه تعداد مقالات در بازه زمانی مشابه بیش از دو برابر طرح‌های پژوهشی است (احتمالا از هر طرح بصورت متوسط ۲ مقاله مستخرج می‌شود)، اعداد قابل مقایسه در هر دسته بصورت درصد تعداد به کل، نمایش داده خواهد شد. در نمودار زیر درصد تمرکز بر هر حوزه موضوعی استخراج شده به تفکیک طرح و مقاله آمده است:



شکل ۴۳- درصد تطبیق کلیدواژه‌گان طرح‌های پژوهشی با مقالات مستخرج

در شکل فوق چند نکته قابل اشاره و بررسی است:

- در حوزه‌های اینترنت اشیا، محاسبات ابری و امنیت (در گروه پژوهشی سیستم‌های اطلاعاتی)، طرح‌های پژوهشی انجام پذیرفته که هیچ‌یک از آنها به مقالات علمی منجر نشده است.
- در حوزه‌های امنیت (در گروه پژوهشی مدیریت فناوری اطلاعات)، واقعیت افزوده و یادگیری عمیق، پژوهشگران فعالیت‌هایی را به ثمر رسانده‌اند که در طرح‌های پژوهشی به چشم نمی‌خورد. احتمالا این مقالات ناشی از فعالیت‌های مشترک ایشان با دیگر پژوهشگران یا دانشجویان ایشان باشد. این مساله نشان می‌دهد ایجاد تمرکز و تشویق بر همکاری پژوهشی

یا فعالیت به عنوان اساتید راهنما و مشاور، بیش از فعالیت‌های درونی ایرانداک، می‌تواند پژوهش‌ها را به سمت روندهای جهانی سوق دهد.

- فعالیت‌های پژوهشی مانند ارتباطات و هوش مصنوعی حوزه‌هایی است که میزان بهره‌وری طرح به مقاله پایین بوده و احتمالاً پژوهشگران در این حوزه‌ها برای انتشار مقالات ناشی از طرح‌های خود دارای سختی‌هایی هستند.
- حوزه‌هایی مانند یادگیری ماشین، سلامت و تجربه کاربری با وجود طرح‌های اندک دارای خروجی مقالات زیادی است که احتمالاً ناشی از دانش و تجربه پژوهشگران این حوزه‌ها در انشار آثار علمی یا سخت‌گیری ایرانداک در تعریف پژوهش‌هایی متنوع در این حوزه‌هاست.
- باقی حوزه‌های موضوعی از روند متناسبی برخوردارند.

۵-۵. پیشنهاد برای مطالعات آتی

بخش تحقیقات آینده اغلب با بخش نتیجه‌گیری طرح‌های پژوهشی ترکیب می‌شود. این به این دلیل است که پیشنهادهای تحقیقاتی آینده به طور کلی ناشی از محدودیت‌های پژوهشی است که در فعالیت‌های پژوهشی شناسایی شده‌اند.

این پیشنهادها عموماً به یکی از دسته‌های زیر تعلق دارند:

- (۱) پیشنهاد بر اساس یک یافته خاص در پژوهش
- (۲) رسیدگی به یک نقص در پژوهش و بررسی (یا آزمایش) یک نظریه (چارچوب یا مدل)
- (۳) انجام پژوهش در یک زمینه، مکان و/یا فرهنگ جدید
- (۴) ارزیابی مجدد نتایج پژوهش
- (۵) بسط یک نظریه (چارچوب یا مدل)

قبل از اینکه در مورد هر یک از انواع پیشنهادهای تحقیقاتی آینده بحث کنیم، باید توضیح دهیم که این پیشنهادها از تجربه این پژوهش به دست آمده و لزوماً تمامی پیشنهادهای آتی را دربر نمی‌گیرد. پیشنهادهای پژوهش‌های آتی در زیر مورد بحث قرار گرفته شده است:

- در بعد پیشنهاد بر اساس یک یافته خاص در پژوهش، به نظر می‌رسد تمرکز پژوهش‌هایی در آینده باید کشف چرایی فاصله میان پژوهش‌های ایرانداک و پژوهش‌های بین‌المللی باشد. دلیل چرایی این فاصله می‌تواند بصورت اجماعی از خبرگان و پژوهشگران مرتبط استخراج شده و راهکارهایی برای مرتفع نمودن مشکلات پیشنهاد شود.

- در بعد رسیدگی به یک نقص در پژوهش، در این پژوهش به نظر رسید بعلت همپوشانی زیاد حوزه‌ها و موضوعات فناوری اطلاعات با یکدیگر، خوشه‌بندی آنها کار چالش برانگیزی بوده و مفاهیم و کلیدواژه‌های مشابه در موضوعات مختلف یافت شده‌اند. همانگونه که در پیشنهاد طرح آمده است، برای خوشه‌بندی در این طرح از روش مدل‌سازی موضوعی استفاده شده است اما به نظر می‌رسد برخی از روش‌های خوشه‌بندی ممکن است بهتر بتوانند مرز بین موضوعات را مشخص کنند. پیشنهاد می‌شود بکارگیری سایر روش‌های خوشه‌بندی به عنوان یکی از زمینه‌های پژوهش‌های آتی مورد توجه قرار گیرد.
- در بعد انجام پژوهش در زمینه متفاوت، موکدا توصیه می‌شود مشابه این پژوهش در دیگر پژوهشکده‌های ایرانداک تکرار شود تا فاصله‌ی احتمالی پژوهش‌های جاری و سطح جهانی در کل ایرانداک مورد بررسی قرار گیرد و تا آینده فاصله‌ی بزرگتری در این میان ایجاد نگردد. هرچند پژوهش‌های حوزه‌ی فناوری اطلاعات دارای سرعت بالاتری نسبت به دیگر حوزه‌های علمی است اما به نظر می‌رسد انجام چنین پژوهشی می‌تواند دستاوردهای مفیدی از آینده پژوهش‌های آن پژوهشکده‌ها نیز ارائه دهد.
- در بعد ارزیابی مجدد نتایج پژوهش، به نظر می‌رسد تکرار پژوهش‌هایی ازین دست در بازه‌های زمانی ۳ تا ۵ سال می‌تواند مسیر پژوهش‌های ایرانداک را در هر بازه زمانی نشان داده و از انحرافات دامنه‌دار و بلندمدت جلوگیری نماید.
- در نهایت در بعد بسط نظریه یا چارچوب، به نظر می‌رسد می‌توان نتایج کیفی مستخرج از داده‌های بدست آمده را در پژوهشی مستقل آزمون نمود. از جمله نتایج قابل آزمون می‌توان تاثیر کمبود امکانات سخت‌افزاری در کمبود پژوهش‌ها در یک زمینه علمی را بررسی کرد. همچنین می‌توان تاثیرپذیری روند پژوهش‌های ملی ویا ایرانداک را از اسناد و الزامات بالادستی مورد بررسی و آزمون قرار داد.

فهرست منابع

۱-۶. فهرست منابع فارسی

- اسدی، سعید. ۱۳۹۳. طرح پژوهشی سنجش و مصورسازی پراکندگی جغرافیایی ثروت علمی: مفاهیم و کاربردها. تهران: پژوهشگاه علوم و فناوری اطلاعات ایران.
- امیرحسینی، مازیار. ۱۳۷۱. کتابسنجی و اطلاع سنجی. فصلنامه کتاب. شماره ۳. ۱۸۳-۲۰۹.
- پژوهشگاه علوم و فناوری اطلاعات ایران. پژوهشگاه علوم و فناوری اطلاعات ایران» (ایرانداک). ۱۴۰۰. <https://irandoc.ac.ir/research-department>. (دستیابی در ۰۱/۰۱/۲۰۲۲).
- پژوهشگاه علوم و فناوری اطلاعات ایران. پژوهشگاه علوم و فناوری اطلاعات ایران» (ایرانداک). درباره. ۱۳۹۸. www.irandoc.ac.ir/about (دستیابی در ۰۹/۰۱/۲۰۲۰).
- پژوهشگاه علوم و فناوری اطلاعات ایران. پژوهشگاه علوم و فناوری اطلاعات ایران» (ایرانداک). ۱۳۹۵. <http://irandoc.ac.ir/about-us/about-us.html> (دستیابی در ۰۲/۰۱/۲۰۲۲).
- جلالی منش، عمار. ۱۳۹۹. طرح پژوهشی طراحی و تدوین منشور پژوهش، پژوهشکده‌ها، و گروه‌های پژوهشی پژوهشگاه علوم و فناوری اطلاعات ایران. تهران: ایرانداک.
- خاصه، علی‌اکبر، و فرامرز سهیلی. ترسیم چشمانداز پژوهش در علمسنجی و حوزه‌های سنجشی وابسته. فصلنامه علمی پژوهشی پردازش و مدیریت اطلاعات (۳۳)، ۱۳۹۷: ۹۴۱-۹۶۶.
- ساجدی نژاد، آرمان. طرح پژوهشی بررسی کارکردهای مدیریت اطلاعات علمی و فناوریانه در بافت دولت الکترونیکی و دولت همراه. تهران: پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، ۱۳۹۵.
- فتاحی، رحمت‌الله. حکایت استفاده کاربردی از پژوهشها در ایران. ۱۳۹۰. فصلنامه علوم و فناوری اطلاعات، ۴۶. ۷۷۷-۷۷۹.
- فتحی‌واجارگاه، کورش. ۱۳۸۹. طراحی مکانیزم اولویت‌یابی پژوهشی برای نظام آموزش و پرورش ایران (مورد آموزش و پرورش تهران). مطالعات تربیتی و روان‌شناسی، شماره ۱۱، ۱۸۵-۲۱۰.
- قنادی‌نژاد، فرزانه، غلامرضا حیدری، و رحیم چینی پرداز. تحلیل محتوای متون مربوط به اولویتهای پژوهشی در علم اطلاعات و دانششناسی. ۱۳۹۷. پژوهشنامه کتابداری و اطلاع رسانی، ۸، ۵۵-۷۴.
- مجبی، آزاده. طرح پژوهشی تدوین نقشه راه فناوری اطلاعات پژوهشگاه با تمرکز بر پیشرفتهای نوین در حوزه تجزیه و تحلیل اطلاعات. پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک). ۱۳۹۶.
- نامداریان، لیلا. طرح پژوهشی بررسی و تبیین جایگاه پژوهشگاه علوم و فناوری اطلاعات در نظام سیاستگذاری علم و فناوری کشور. پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، ۱۳۹۶.

- Al-khamayseh, Shadi, and Elaine Lawrence. "Towards Understanding Success Factors in Interactive Mobile Government." 2007.
- Banerjee, Rohit . "Initiating GEIT Using COBIT 5 at the Oman Ministry of Manpower." 2016.
- Blei, David M. "Probabilistic topic models." *Communications of the ACM* 55, 2012: 77-84.
- Blei, David, Y. Andrew, and M. I. Jordan. "Latent dirichlet allocation." *Journal of machine Learning research*, 3, 2003: 993-1022.
- Blei, David, Y. Andrew, and Michael Jordan. "Latent dirichlet allocation." *Journal of machine Learning research* 3, 2003: 993-1022.
- Chang, J., and D. M. Blei. "Relational topic models for document networks." *International conference on artificial intelligence and statistics*. 2009. 81-88.
- Choudhary, A. K., P. Oluikpe, J. A. Harding, and P. M. Carrillo. "The needs and benefits of Text Mining applications on Post-Project Reviews." *Computers in Industry*, vol. 60, 2009: 728-740.
- De Battisti, Francesca, Alfio Ferrara, and Silvia Salini. "A decade of research in statistics: a topic model approach." *Scientometrics* 103, 2015: 413-433.
- Furner, J. "Information science is neither." *Library Trends*, 63(3), 2015: 362-377.
- Gru, B., and K. Hornik. "Topicsmodels: An R package for fitting topic models." *Journal of Statistical Software*, 40(13), 2011: 1-30.
- Hall, David, Daniel Jurafsky, and Christopher Manning. "Studying the history of ideas using topic models." *Proceedings of the conference on empirical methods in natural language processing*. 2008. 363-371.
- Jalalimanesh, Amar, and Elaheh Homayoon Vala. "Promoting Research Collaboration Based on Data Mining Techniques in Library Information Systems." *IADIS Collaborative Technologies*, 2012: ۷۳-۸۲.
- Jens-Erik, M., O. Donald, and L. Given. "Looking for information: A survey of research on information seeking, needs, and behavior: ." *Emerald Group Publishing*, 2016.
- John W, Mohr, and Bogdanov Petko. "Introduction-Topic models: What they are and why they matter ." *Poetics* 41, 2013: 545-569.

- Mona, N., and W. Vivian. "Prioritization of systematic reviews leads prioritization of research gaps and needs." *Journal of clinical epidemiology*, vol. 66, 2013: 522.
- Robinson, K. A., I. J. Saldanha, and N. A. Mckoy. "Development of a framework to identify research gaps from systematic reviews." *Journal of clinical epidemiology*, 64, 2011: 1325-1330.
- Robinson, K. A., I. J. Saldanha, and N. A. Mckoy. "Development of a framework to identify research gaps from systematic reviews." *Journal of clinical epidemiology*, 64, 2011: 1325-1330.
- Wallach, H. M. "Topic modeling: beyond bag-of-words." *Proceedings of the 23rd international conference on Machine learning, Pittsburgh, Pennsylvania, USA*, 2006.
- Wallach, H. "Topic modeling: Beyond bag-of-words." *the 23rd International Conference on Machine Learning*. Pittsburgh, Pennsylvania, U.S., 2006. 977-984.
- Web of Science. *Keywords Plus® (Web of Science Core Collection and Current Contents Connect only)*. 01 29, 2020. http://images.webofknowledge.com/WOKRS534DR1/help/WOS/hp_advance_d_search.html (accessed 01 29, 2020).
- Weidong, Z., and W. Dai. "Research paper recommendation based on the knowledge gap." *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, 2015: 373-380.

پیوست

دسته‌بندی پایگاه WOS از انتشارات علوم

Acoustics	Emergency Medicine	Literature, American	Physics, Nuclear
Agricultural Economics & Policy	Endocrinology & Metabolism	Literature, British Isles	Physics, Particles & Fields
Agricultural Engineering	Energy & Fuels	Literature, German, Dutch, Scandinavian	Physiology
Agriculture, Dairy & Animal Science	Engineering, Aerospace	Literature, Romance	Plant Sciences
Agriculture, Multidisciplinary	Engineering, Biomedical	Literature, Slavic	Poetry
Agronomy	Engineering, Chemical	Logic	Political Science
Allergy	Engineering, Civil	Management	Polymer Science
Anatomy & Morphology	Engineering, Electrical & Electronic	Marine & Freshwater Biology	Primary Health Care
Andrology	Engineering, Environmental	Materials Science, Biomaterials	Psychiatry
Anesthesiology	Engineering, Geological	Materials Science, Ceramics	Psychology
Anthropology	Engineering, Industrial	Materials Science, Characterization & Testing	Psychology, Applied
Archaeology	Engineering, Manufacturing	Materials Science, Coatings & Films	Psychology, Biological
Architecture	Engineering, Marine	Materials Science, Composites	Psychology, Clinical
Area Studies	Engineering, Mechanical	Materials Science, Multidisciplinary	Psychology, Developmental
Art	Engineering, Multidisciplinary	Materials Science, Paper & Wood	Psychology, Educational
Asian Studies	Engineering, Ocean	Materials Science, Textiles	Psychology, Experimental
Astronomy & Astrophysics	Engineering, Petroleum	Mathematical & Computational Biology	Psychology, Mathematical
Audiology & Speech-Language Pathology	Entomology	Mathematics	Psychology, Multidisciplinary
Automation & Control Systems	Environmental Sciences	Mathematics, Applied	Psychology, Psychoanalysis
Behavioral Sciences	Environmental Studies	Mathematics, Interdisciplinary Applications	Psychology, Social
Biochemical Research Methods	Ergonomics	Mechanics	Public Administration
Biochemistry & Molecular Biology	Ethics	Medical Ethics	Public, Environmental & Occupational Health
Biodiversity Conservation	Ethnic Studies	Medical Informatics	Quantum Science & Technology

Biology	Evolutionary Biology	Medical Laboratory Technology	Radiology, Nuclear Medicine & Medical Imaging
Biophysics	Family Studies	Medicine, General & Internal	Regional & Urban Planning
Biotechnology & Applied Microbiology	Film, Radio, Television	Medicine, Legal	Rehabilitation
Business	Fisheries	Medicine, Research & Experimental	Religion
Business, Finance	Folklore	Medieval & Renaissance Studies	Remote Sensing
Cardiac & Cardiovascular Systems	Food Science & Technology	Metallurgy & Metallurgical Engineering	Reproductive Biology
Cell & Tissue Engineering	Forestry	Meteorology & Atmospheric Sciences	Respiratory System
Cell Biology	Gastroenterology & Hepatology	Microbiology	Rheumatology
Chemistry, Analytical	Genetics & Heredity	Microscopy	Robotics
Chemistry, Applied	Geochemistry & Geophysics	Mineralogy	Social Issues
Chemistry, Inorganic & Nuclear	Geography	Mining & Mineral Processing	Social Sciences, Biomedical
Chemistry, Medicinal	Geography, Physical	Multidisciplinary Sciences	Social Sciences, Interdisciplinary
Chemistry, Multidisciplinary	Geology	Music	Social Sciences, Mathematical Methods
Chemistry, Organic	Geosciences, Multidisciplinary	Mycology	Social Work
Chemistry, Physical	Geriatrics & Gerontology	Nanoscience & Nanotechnology	Sociology
Classics	Gerontology	Neuroimaging	Soil Science
Clinical Neurology	Green & Sustainable Science & Technology	Neurosciences	Spectroscopy
Communication	Health Care Sciences & Services	Nuclear Science & Technology	Sport Sciences
Computer Science, Artificial Intelligence	Health Policy & Services	Nursing	Statistics & Probability
Computer Science, Cybernetics	Hematology	Nutrition & Dietetics	Substance Abuse
Computer Science, Hardware & Architecture	History	Obstetrics & Gynecology	Surgery
Computer Science, Information Systems	History & Philosophy of Science	Oceanography	Telecommunications

Computer Science, Interdisciplinary Applications	History of Social Sciences	Oncology	Theater
Computer Science, Software Engineering	Horticulture	Operations Research & Management Science	Thermodynamics
Computer Science, Theory & Methods	Hospitality, Leisure, Sport & Tourism	Ophthalmology	Toxicology
Construction & Building Technology	Humanities, Multidisciplinary	Optics	Transplantation
Criminology & Penology	Imaging Science & Photographic Technology	Ornithology	Transportation
Critical Care Medicine	Immunology	Orthopedics	Transportation Science & Technology
Crystallography	Industrial Relations & Labor	Otorhinolaryngology	Tropical Medicine
Cultural Studies	Infectious Diseases	Paleontology	Urban Studies
Dance	Information Science & Library Science	Parasitology	Urology & Nephrology
Demography	Instruments & Instrumentation	Pathology	Veterinary Sciences
Dentistry, Oral Surgery & Medicine	Integrative & Complementary Medicine	Pediatrics	Virology
Dermatology	International Relations	Peripheral Vascular Disease	Water Resources
Development Studies	Language & Linguistics	Pharmacology & Pharmacy	Women's Studies
Developmental Biology	Law	Philosophy	Zoology
Ecology	Limnology	Physics, Applied	
Economics	Linguistics	Physics, Atomic, Molecular & Chemical	
Education & Educational Research	Literary Reviews	Physics, Condensed Matter	
Education, Scientific Disciplines	Literary Theory & Criticism	Physics, Fluids & Plasmas	
Education, Special	Literature	Physics, Mathematical	
Electrochemistry	Literature, African, Australian, Canadian	Physics, Multidisciplinary	

ABSTRACT

In many researches, the comparison of research trends in the field of information technology of the country and the global was carefully analysed. Also, emphasizing the trend of keywords, and trend of information technology in the world and the country was compared (Sajdinjad, 2019). So in this Research, different aspects of interest areas of each research group of Irandoc would be analysed. The aim is to examine the (possible) research gap of each research group in contrast with the trends of international research's (papers). Considering the rapid changes in the field of information technology (internationally), the period of the last 1 year will be monitored in this project. Assuming the intertwining of the researches of Irandoc research groups and their missions, creating a proper and systematic guide of suitable and appropriate research activities can lead irandoc to a greater advances in missions. There are many possible information technology subject areas in the research literatures and also there are many common keywords in each topics, but the importance of each keyword may be different in any topic. The reason for the mapping of international researches and research groups in this project is to extract research topics that have not been addressed in each research group in comparison with the international state. From the results of this project, it can be pointed out that in a significant part of the researches in Irandoc, they are relatively influenced by the upstream needs and the decision-making and research policy-making structures have made this impact on the research process. However, it seems that despite of the high demand for research in some subjects, not enough researches have been done, which may be due to the lack of sufficient credit for these researches. For example "recommender systems that are in line with the needs of policymakers" is one of the most important topics that requires serious focus in the research aspect of the research institute which is required in the upstream documents. In the review and comparison of Irandoc research trends with international trends, unfortunately, not much matching has been observed. In the researches of Irandoc, special focus has been placed on the researches that are declining in the international trends. On the other hand, the fields of Big Data and Internet of Things and Cloud Computing, which are the focus of international research, have been observed in an unfavorable situation in Irandoc.

Keywords: Topic modeling, IRANDOC, IT Management, e-Business, Information System



**Iranian Research Institute
for Information Science and Technology (Irandoc)**

Research project

**Comparative comparison of topics in the research
groups of the Information Technology Research
Institute of IRANDOC with international topics**

By:

Arman Sajedinejad

And

Mohammad Rabeie

September 2022