

رویکرد ارزیابی نظام های بازیابی اطلاعات پس زمینه و چشم انداز پیش رو

نام نشریه:	فصلنامه کتابداری و اطلاع رسانی www.isc.gov.ir (این نشریه در (نمابه می شود)
شماره نشریه:	شماره دوم، جلد 15 _ 58
پدیدآور:	محمد رضا کیانی
مترجم:	

رویکردهای ارزیابی نظامهای بازیابی اطلاعات: پس زمینه و چشم انداز پیش رو

محمد رضا کیانی^[1]

چکیده

این نوشتار با نگاهی گذرا به سیر تکاملی نظامهای بازیابی اطلاعات و ارزیابی آنها، انواع ارزیابی در این نظامها را از هم تفکیک و به طور مفصل تر درباره رویکردهای موجود در ارزیابی عملکرد نظامهای بازیابی، بحث کرده است. با نقد هر کدام، مدلی ترکیبی با رویکردی نظری - تصمیمی از «وانگ و فورگیونه» را معرفی می کند و در پایان نتایج و پیشنهادهای خود را ارائه می دهد.

کلیدواژه ها: ارزیابی نظامها، بازیابی اطلاعات، رویکردهای ارزیابی.

مقدمه

نظامهای بازیابی اطلاعات^[2] به بازنمون، ذخیره سازی و بازیابی اطلاعات و دانش مرتبط با نیازها و سؤالات خاص کاربران می پردازند. کاربرد در قالب یک پرسش، با واژه هایی از زبان طبیعی نیاز خود را مطرح می کند. این پرسش به وسیله نظام و بر اساس یک الگوریتم مشخص با مدارک نمایه سازی شده مقایسه و مدارکی که بیشترین مشابهت را با پرسش کاربر دارند، عرضه می شوند. در نهایت، این کاربر است که میزان ربط آنها را با نیاز خود مورد سنجش قرار می دهد (مندل^[3]، 2008). بازیابی اطلاعات، فرایندی دارای ابهام است. این ابهام به مبهم بودن عناصر دخیل در آن یعنی زبان طبیعی، نیاز اطلاعاتی، ربط و ... باز می گردد.

پس زمینه تاریخی نظامهای بازیابی رایانه ای، به گسترش روشهای نمایه سازی خودکار متون در دهه 60 میلادی باز می گردد؛ روشی که همچنان به صورت گسترده توسط ارائه دهندگان اطلاعات و همچنین در خدمات اینترنتی به کار می رود. در دهه 70، مدل های بازیابی بولی و فضای برداری و بعدها مدل های بازیابی احتمالاتی گسترش یافتند. لکن پذیرش تجاری آنها به عنوان مدل های انطباق جزئی (در مقابل مدل های انطباق کامل اولیه) در موتورهای کاوش اینترنتی که نتایج را به صورت فهرستی رتبه بندی شده ارائه

می‌دهند، تا دهه 90 به تعویق افتاد.

امروزه نظام‌های بازیابی اطلاعات نقش مهمی را در تمرکز اطلاعاتی^[4] و توسعه دانش بازی می‌کنند (وانگ و فورگیونه^[5]، 2006) چنان‌که مدیریت دانش به منظور دسترس‌پذیر کردن حجم قابل توجه اطلاعات و مجموعه‌های گسترده علمی متنی، بدون این نظام‌ها قابل تصور نیست. در زمان ما، استفاده از وب و موتورهای کاوش جزئی از زندگی روزمره تلقی می‌شود که با روشها، ابزار و رویکردهای نظام‌های بازیابی اطلاعات عجین شده‌اند. در این راستا، همچون بسیاری از مفاهیم و فرایندهای دیگر، انسان دست به ارزیابی می‌زند و درباره کیفیت کار اظهار نظر می‌کند.

ارزیابی نظام‌های بازیابی اطلاعات

زمانی که صحبت از ارزیابی می‌شود، منظور داوری در مورد کیفیت و کارآمدی است. به اعتقاد «ساراسویک»^[6] (1995) ارزیابی نفوذ عمده‌ای در تحقیقات، توسعه و کاربردهای مربوط به بازیابی اطلاعات دارد. به نظر «بائزا - بیتس و ریبرو - نتو»^[7] (1385) در ارزیابی نظام‌های بازیابی اطلاعات، به دو جنبه کلی تحلیل کارکردی^[8] و ارزیابی عملکرد^[9] نظام پرداخته می‌شود. در بخش تحلیل کارکردی نظام، موضوعاتی مانند آزمون تک تک کارکردها با معیارهای اثربخشی (هدفها)، هزینه - کارآیی (هزینه‌های نهایی) و هزینه - سودمندی (سود ناشی از خدمات) و همچنین تحلیل خطا، مورد توجه قرار می‌گیرد.

عملکرد نظام از نظر فضا و زمان و معیارهای دیگری که مهم‌ترین آنها ربط^[10] است، ارزیابی می‌شود. در متون حوزه ما، زمانی که بحث ارزیابی نظام‌های بازیابی اطلاعات مطرح می‌شود، بیشتر منظور مورد اخیر یعنی ارزیابی عملکرد بازیابی نظام از نظر ربط است.

هدف از توضیح این قسمت این بود که باید بین ارزیابی کل نظام بازیابی، ارزیابی عملکرد آن و ارزیابی بازیابی نظام از نظر ربط، تفاوت قایل شد؛ هرچند در بعضی از سنجش‌های ارائه شده، علاوه بر ربط، ارزیابی عملکرد به طور کلی، یعنی با توجه به جنبه‌های زمان و فضا، مورد توجه بوده است.

هدف از ارزیابی کل نظام، آمادگی برای پاسخگویی به منابع تأمین بودجه و توجیه عملکرد نظام و اثبات ارزشمندی آن، اطمینان از تداوم خدمات مناسب و اثربخشی (با مقایسه سطح مورد نظر عملکرد و سطح کنونی)، حل مشکلات نظام (بررسی نقاط قوت و ضعف) و به دست آوردن اطلاعات برای طرح‌ریزی عملکرد آتی و یا طراحی و پیاده‌سازی نظام‌های جدید است (پائو، 1379).

در ارزیابی عملکرد یک نظام، هرچه زمان پاسخ کوتاه‌تر و فضای مورد استفاده [ی داده‌ها] کوچکتر باشد، نظام در وضعیت بهتری است که مربوط می‌شود به عملکرد ساختارهای نمایه‌سازی، تعامل با سیستم عامل، تأخیرهای کانالهای ارتباطی و پیامدهای لایه‌های نرم‌افزاری عمل کننده در پس زمینه (بائزا- بیتس و ریبرو- نتو، 1385).

بنابراین، یک نظام خوب، نظامی است که نیازهای کاربر را برآورده سازد. برآوردن این

نیازها، به عملکرد خوب نظام در ابعاد مختلف بستگی دارد. کیفیت پاسخها در ارتباط با نیاز اطلاعاتی، سرعت پاسخگویی و محیط رابط کاربر نظام، مهم‌ترین ابعاد هستند که در بین این سه، ارزیابی مورد اول از همه دشوارتر است.

رویکردهای ارزیابی عملکرد ارزیابی نظام (ربط)

الف) رویکرد نظام مدار

ارزیابی‌ها اغلب در ارزیابی اطلاعات از رویکرد نظام مدارانه^[11] استفاده می‌کنند که آن هم بیشتر بر تعیین الگوریتمهای جستجو با استفاده از سنج‌های آماری بویژه دو ضریب بسیار شناخته شده «بازیافت و دقت»^[12] متمرکز است (سالتون، 1992). نمونه‌های این روش‌شناسی ارزیابی که به «پارادایم کرنفیلد»^[13] معروف است (کلوردون، 1997) پروژه‌های کرنفیلد (کلوردون، میلز، و کین، 1966)، «اسمارت»^[14] (سالتون، 1971، 1989)، «استیرز»^[15] (بلرو میرون، 1990، 1985)، و «ترک»^[16] (از سال 1992 توسط موسسه استاندارد و فناوری) است.

این مطالعات اگرچه به صورت طرح تجربی استاندارد انجام شده‌اند، چشم‌انداز آنها به ارزیابی، کامل نیست. دو اشکال اساسی به این رویکرد گرفته‌اند. نخست، از ربط موضوعی به عنوان مبنا یا تنها معیار ارزیابی ارزیابی استفاده می‌شود. سنج‌های آماری عملکرد مانند بازیافت و دقت، به زمینه متن کار ندارند و مستقل از متن هستند و منعکس کننده فرایند تعاملی ارزیابی نیستند و تنها نوعی داور مقدماتی در مورد بازنمون دوجبهی ربط (مربوط یا نامربوط) می‌دهند (بورلند، 2000، 2003). از طرفی، تنها برون‌داد را مورد توجه قرار می‌دهند و راهبردی یکجایه (که در ادامه توضیح داده خواهد شد) دارند و به فرایند تعاملی ارزیابی توجهی ندارند. از سوی دیگر، مدارک را به مربوط و نامربوط تقسیم می‌کنند. هرچند جامعیت و مانعیت باز تعریف شده‌اند (ککالاین و جارولین^[17]، 2002)، همچنان این سنج‌های ربط محور از انعکاس طبیعت پویا و چند بعدی ربط ناتوان است (ساراسویک، 1996).

دوم، بُعد کاربری از فرایند ارزیابی حذف شده است، در حالی که ربط مفهومی شناختی است که معنای آن به درک کاربران از اطلاعات ارزیابی شده و موقعیتهای نیاز اطلاعاتی خودشان ارتباط عمده دارد (بورلند، 2003). این رویکرد تنها به درجه‌ای که بازنمون پرسش (عبارت جستجو)^[18] با متن عینیت‌های اطلاعاتی ارزیابی شده مطابقت دارد، می‌پردازد. به بیان دیگر، مطالعات ارزیابی در این حوزه، تعریف نیاز اطلاعاتی توسط سیستم را به جای تعریف‌های کاربران پذیرفته‌اند. کاربر در این رویکرد یک انتزاع^[19] است نه کاربر واقعی (مندل، 2008).

ب) رویکرد کاربر مدار

با تغییر مسیر پارادایم به سمت جنبه‌های رفتاری و شناختی در ارزیابی اطلاعات، حجم عمده‌ای از مطالعات بر ارزیابی رضایت، عملکرد و استفاده کاربر نهایی از نظامهای ارزیابی اطلاعات تمرکز کردند (الیس، 1996؛ رابرتسون و هنکاک بولیو، 1992؛ چمبر، 1994). مطالعات کاربرمدار^[20] به کاربران و سطوح استفاده آنها می‌پردازند. به دلیل دشواری

ارزیابی جنبه‌های فردی جستجوها و ارزیابی قضاوت کاربر از سودمندی نتایج، کار در این رویکرد بسیار دشوارتر از رویکرد نظام مدار است. در این رویکرد، تمرکز بیشتر بر فرایند است تا خروجی‌ها.

دسته‌بندی از نظر راهبرد جستجو

شاید مناسب باشد قبل از تقسیم مطالعات مربوط به ارزیابی عملکرد به نظام مدار و کاربر مدار، آنها را از نظر توجه به ماهیت درخواست پرسش، دسته‌بندی کنیم. در این تقسیم‌بندی، راهبرد جستجو مورد توجه قرار می‌گیرد. دو نوع راهبرد جستجوی ذکر شده، عبارتند از: راهبرد یکجایه^[21] و راهبرد تعاملی^[22].

در نوع اول، کاربر پرسش خود را ارائه و پاسخی را دریافت می‌کند. اما در راهبرد تعاملی، کاربر نیاز اطلاعاتی خود را از طریق مجموعه‌ای از مراحل تعامل با نظام، مشخص می‌کند.

در راهبرد یکجایه، وجه غالب کیفیت مجموعه پاسخ ارائه شده است، که خود می‌تواند نظام مدار یا کاربرمدار باشد. مطالعاتی که با این راهبرد صورت گرفته، ابتدا در محیط آزمایشگاهی (مانند ترك) و سپس در محیط واقعی انجام شدند. معیارهای ارزیابی^[23] راهبرد یکجایه در رویکرد نظام‌مدار عبارتند از: جامعیت و مانعیت، سنجه ارزیابی E، و میانگین توافقی F. این معیارها در رویکرد کاربرمدار عبارتند از: دستاورد جامعیت^[24]، جامعیت نسبی^[25]، نسبت تازگی^[26] و نسبت پوشش^[27].

در راهبرد تعاملی، بیشتر از کیفیت مجموعه پاسخ ارائه شده، به فرایند تعامل و اجزای آن مانند تلاش کاربر، ویژگی‌های طراحی محیط رابط^[28]، راهنمایی‌های ارائه شده از سوی نظام و طول مدت تعامل توجه می‌شود. بنابراین، باید بین راهبرد تعاملی با رویکرد کاربرمدار یکجایه، تفاوت قایل شد.

منظور از تعامل، ارتباط بین اطلاع‌جویان انسانی و نظام‌های ارزیابی اطلاعات است. وقتی کاربران به یک نظام ارزیابی اطلاعات وارد می‌شوند، اغلب از نحوه دستیابی خود به اطلاعات، درک مبهمی دارند. بنابراین، تعامل میان سیستم و کاربر در مورد درک و بیان نیازهای اطلاعاتی، تدوین پرسشها، انتخاب از میان مآخذ موجود، درک نتایج کاوش و پیگیری روند پیشرفت کاوش کاربران، می‌تواند انجام شود. رفتار اطلاعاتی^[29] و بازخورد ربط^[30] در این بخش با نظام‌های ارزیابی پیوند می‌خورند. بر اساس این نوع نگاه، ارزیابی با آزمایش اطلاعاتی واقعی کاربر نهایی و فرایندهای ارزیابی آغاز می‌شود (جارولین و اینگورسن^[31]، 2004). مدل اطلاعاتی^[32] (میوه چینی) بیتس و سایر مدل‌های تشریح و بازنمایی اعمال و فعالیت‌های شناختی کاربر در خلال فرایند ارزیابی اطلاعات (به عنوان مثال بلکین، 1980؛ اینگورسن، 1996؛ کولثا، 1991) از این دسته‌اند. در ارزیابی نظام‌های تعاملی، به طراحی محیط رابط، بازیافت و دقت (در بسیاری از محیط‌های تعاملی، کاربران فقط به اسناد معدود با ربط و با مانعیت بالا نیاز دارند)، زمان لازم برای یادگیری کار با نظام، زمان لازم برای دستیابی به هدفها در وظایف معیار، نرخ خطا، و تداوم استفاده از محیط رابط در طی زمان ... توجه می‌شود.

در مدل استاندارد، فرایند دسترسی به اطلاعات، این واقعیت مورد توجه قرار نمی‌گیرد که بسیاری از کاربران دوست ندارند با سیاهه‌های طولانی و به هم ریخته از نتایج بازیابی (که به طور مستقیم به نیاز اطلاعاتی‌شان نمی‌پردازد) مواجه شوند. همچنین، حاوی این فرض اساسی است که نیاز اطلاعاتی کاربر، ایستاست و فرایند اطلاع‌جویی یک فرایند اصلاح‌پیاپی است، تا جایی که فقط اسناد مربوط به نیازهای اطلاعاتی اصلی، بازیابی شوند. تعاملی که در هنگام مرور واژه‌های پیشنهادشده به عنوان نتیجه بازخورد ربط، مرور ساختارهای اصطلاحنامه‌ای، یا مشاهده‌ی نمایه‌ی موضوعی گردآوردهای اسناد به وسیله‌ی کاربر رخ می‌دهد، در مدل استاندارد چندان مورد توجه قرار نمی‌گیرد.

چنان‌که ملاحظه می‌شود، یکی از مسائل مهم در این نوع نظامها، بازخورد ربط است. بازخورد ربط در شکل اصلی آن به یک چرخه‌ی تعامل اطلاق می‌شود که در آن، کاربر مجموعه‌ی کوچکی از اسناد را که به نظر می‌رسد با پرسش ربط دارند، انتخاب می‌کند. سپس، نظام از ویژگیهای برگرفته شده از اسناد انتخاب شده مرتب، برای بازیابی در پرسش اصلی استفاده می‌کند. بعد، این پرسش بازیابی شده اجرا و مجموعه‌ی جدیدی از اسناد برگردانده می‌شود. اسناد مجموعه‌ی اصلی ممکن است در سیاهه‌ی یافته‌های جدید نیز ظاهر شوند، هرچند ممکن است رتبه‌ی متفاوتی داشته باشند. از جمله روشهای اعمال بازخورد ربط می‌توان به محیط رابط برای بازخورد استاندارد ربط (جعبه‌های علامتگذاری، اسناد دیگر شبیه به این، بازخورد تصاعدي)، قرار دادن اطلاعات با ربط در پس‌زمینه (کنترل‌پیشرفت و رفتار کاربران در دوره‌های تعاملی طولانی مدت)، قضاوت ربط گروهی (استفاده از قضاوت‌های ربط گروهی از کاربران مختلف، برابری بندی یا درجه بندی اطلاعات عمومی)، و بازخورد شهبند اشاره کرد (در این روش، نظام به جای اتکا بر کاربر برای انتخاب اسناد با ربط برتر، فرض می‌کند که اسناد رتبه بندی شده صورتی با ربط هستند، و از این اسناد برای طولانی کردن پرسش با یک الگوریتم رتبه بندی بازخورد ربط، استفاده می‌کند) (بائزا - بیتس و ریبرو - نتو، جلد دوم، 1389).

بررسی یک مدل ترکیبی

در پژوهش مورد بررسی (وانگ و فورگیونه، 2006) سعی شده است معیاری جامع و چند وجهی برای ارزیابی نظامهای بازیابی اطلاعات عرضه شود. مبنای این پژوهش بر فرایند تحلیل سلسله مراتبی^[33] در بافتار نظامهای بازیابی قرار گرفته است. بحث نو این پژوهش، تمرکز بر جنبه‌ی کاربری و کاربرد نظریه‌های تصمیم‌گیری^[34] در حوزه‌ی بازیابی اطلاعات است.

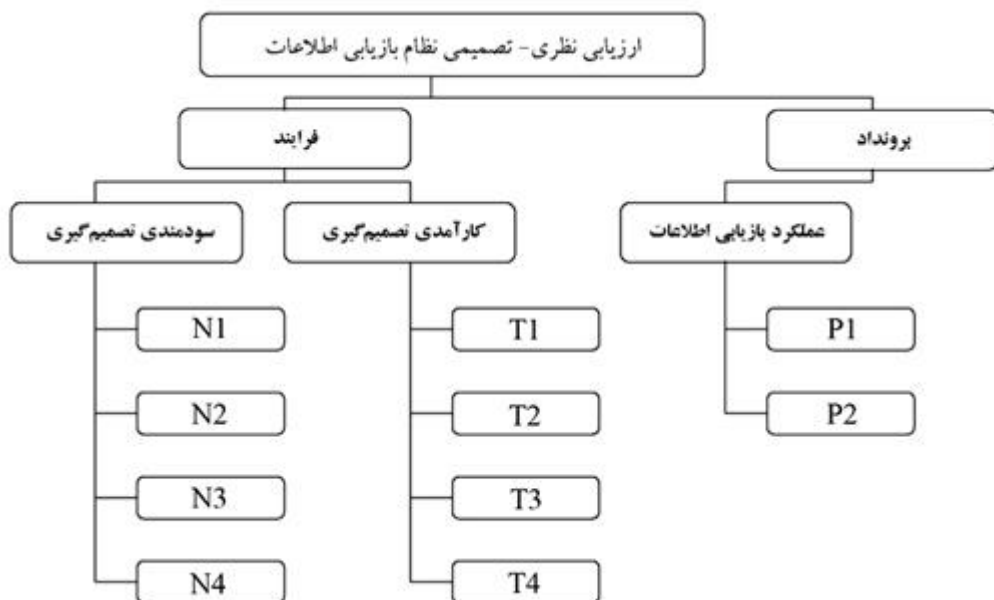
این پژوهش فرایند جستجوی اطلاعات^[35] از منظر کاربر را در 6 مرحله (شروع، گزینش، شناسایی، تدوین عبارت جستجو^[36]، گردآوری و ارائه) مطابق مدل «کولثاو» (1991) در نظر گرفته و آن را با فرایند تصمیم‌گیری مطابقت داده است. وی معتقد است تفکر در مورد جستجوی اطلاعات به عنوان فرایند ساخت تصمیم به طور خاص در زمان ارزیابی نظامهای بازیابی اطلاعات، اهمیت ویژه‌ای دارد و از این راه می‌توان مشکل مرتبط با رویکردهای متفرق ارزیابی نظامها را هم در مورد برون داد و هم فرایند حل کرد. این تطابق در جدول 1 تشریح شده است.

جدول 1. مراحل تصمیم‌گیری در فرایند جستجوی اطلاعات (مبتنی بر مدل وانگ و فورگینه، 2006)

مراحل فرایند جستجوی اطلاعات	مراحل تصمیم‌گیری در فرایند جستجوی اطلاعات	مراحل عام تصمیم‌گیری
شروع به کار	بازشناسی مسئله	بازشناسی مسئله
انتخاب موضوع	ارزیابی جایگزینهای موضوع عام تشخیص يك موضوع عام	تشخیص عینیت‌ها و برقراری معیارها
شناسایی اولیه	جمع‌آوری اطلاعات مرتبط با موضوع عام	جمع‌آوری داده و تولید
	بررسی ارتباط بین موضوعات جایگزین و موضوع عام	جایگزینها
تمرکز بر فرمول‌بندی	برقراری معیارها (تمرکز)	تعیین روابط بین معیارها و جایگزینها
گردآوری اطلاعات	ارزیابی جایگزینهای مقاله	ارزیابی جایگزینها و توسعه اعمال توصیه شده در بیشترین اشتراك با معیارهای تصمیم
انجام جستجو	اتخاذ انتخابهای نهایی	توسعه يك نقشه پیاده‌سازی و به‌کاربردن آن در عمل

بنابراین، مبانی این پژوهش را می‌توان به صورت زیر دسته‌بندی کرد. نخست، مطالعه‌ای است در مورد مراحل تصمیم‌گیری درگیر در فرایند جستجوی اطلاعات توسط کاربران و نقش نظریه‌های تصمیم‌گیری در ارزیابی نظامهای بازیابی اطلاعات. دوم، به تشریح مدلی نظری - تصمیمی که سنج‌های رویکردهای برون‌داده‌ای و فرایندی را ترکیب می‌کند، برای ارزیابی نظامهای بازیابی اطلاعات می‌پردازد. کاربرد معیار چند وجهی مبتنی بر روش تصمیم‌گیری که فرایند سلسله مراتبی تحلیل نامیده می‌شود، ارائه مدل ارزیابی و سپس کاربرد آن در حوزه خاص کاربرد بازیابی اطلاعات، قدمهای بعدی این پژوهش است. مدل ارائه شده مبتنی بر این رویکرد، مطابق شکل 1 است.

به اعتقاد محققان طراح، این مدل سنج‌های مختلف ارزیابی مرتبط به زمینه جستجوی اطلاعات را به شکل مجتمع به هم مرتبط می‌سازد. به دلیل اینکه يك نظام بازیابی اطلاعات بندرت در تمامی مراحل و رویه‌های ساخت تصمیم به کاربر یاری می‌رساند، تنها سنج‌های مرتبط، در ارزیابی باید در نظر گرفته شوند. بعضی سنج‌ها به فرایند و بعضی به برون‌داد می‌پردازند که به تفکیک بیان شده‌اند.



شکل 1. مدل چند معیاره ارزیابی نظام بازیابی اطلاعات (وانگ وفورگیونه، 2006)

که در آن حروف ذکر شده (سنجه‌های خاص ارزیابی) نمایانگر موارد زیر هستند:

- T1 زمان صرف شده برای بازشناسی مسئله
- N1 تعداد جایگزینهای موضوعی عام
- T2 زمان صرف شده برای تشخیص موضوع عام
- N2 تعداد مقاله‌های تولید شده برای موضوع عام
- T3 زمان صرف شده برای ایجاد معیار
- T4 زمان صرف شده برای تشخیص مقاله مرتبط
- N3 تعداد مقاله‌های مرتبط تشخیص داده شده
- N4 تعداد مقاله‌های اضافی تشخیص داده شده در واریسی دوباره
- P1 سودمندی نظام
- P2 یادگیری کاربر

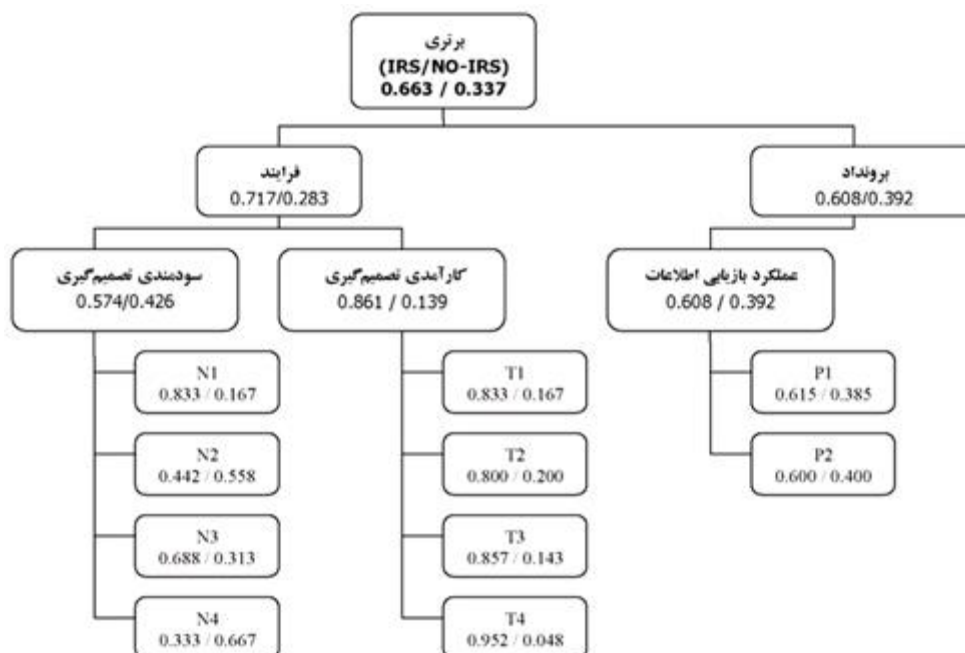
تشریح فرایند سلسله مراتبی تحلیل، به عنوان یکی از پر کاربردترین روشهای مورد استفاده در پرداختن به مسائل دارای معیارهای چند وجهی تصمیم‌گیری^[37] (که ارزیابی نظامهای بازیابی اطلاعات نیز از این دسته مسائل است) در بخش بعدی مورد توجه قرار گرفته و در نهایت، مدل ارائه شده را در یک نظام بازیابی اطلاعات (شبکه پیشگیری از

خودکشی^[38] به کار برده است و نتایج را در دو حالت (با استفاده یا بدون استفاده) از این نظام، مقایسه کرده است. با به کارگیری فرایند سلسله مراتبی تحلیل، این دو حالت در معیارهای فرایند و برون داد می‌توانند مقایسه شوند. لازم به ذکر است، برای تولید داده‌های مورد نیاز فرایند سلسله مراتبی تحلیل، از یک سناریوی شبیه‌سازی شده استفاده شده است. نتایج در جدول 2 ارائه شده‌اند.

جدول 2. ارزش سنج‌های ارزیابی در موارد با استفاده یا بدون استفاده از نظام ارزیابی اطلاعات

بدون استفاده از نظام ارزیابی اطلاعات	با استفاده از نظام ارزیابی اطلاعات	سنج‌های ارزیابی
150 ثانیه	30 ثانیه	T1
4	20	N1
60 ثانیه	15 ثانیه	T2
29	23	N2
360 ثانیه	60 ثانیه	T3
300 ثانیه	15 ثانیه	T4
5	11	N3
2	1	N4
5	8	P1
6	7	P2

ارزش‌های عددی نتایج به دست آمده در جدول 2، پس از تغییر به وزن‌های مساوی به منظور سنجش برتری هرکدام از دو روش، در شکل 2 آورده شده است. اگرچه شاید برخی معیارها مهم‌تر از بقیه باشند، به دلیل پرهیز از سوگیری اوزان یکسان در نظر گرفته شده‌اند.



نتایج ارائه شده در شکل 2، نمایانگر معناداری نسبی تفاوت میان دو جایگزین در بهتر شدن ساخت تصمیم کاربر در خلال جستجوی اطلاعات است.

همچنین، نشان می‌دهد که سودمندی تصمیم‌گیری به میزان قابل توجهی با استفاده کاربر از نظام بازیابی اطلاعات، در مقایسه با زمانی که از آن استفاده نمی‌کند، بالا می‌رود (0/574 در مقابل 0/426) و کارآمدی تصمیم‌گیری به میزان بسیار زیاد با نظام بازیابی بهتر می‌شود (0/861 در مقابل 0/139). همچنین، عملکرد بازیابی اطلاعات نیز به طور معناداری با نظام بازیابی اطلاعات افزایش می‌یابد (0/608 در مقابل 0/392). حالت استفاده از این نظام بازیابی اطلاعات، برتری معناداری بر حالت استفاده نکردن از این نظام در هر دو بخش معیارهای فرایندی و بروندادی نشان می‌دهد و بیشترین برتری در فرایند است. روی هم رفته، امتیاز به دست آمده 0/663 برای نظام بازیابی اطلاعات در مقایسه با امتیاز 0/337 برای مورد استفاده نکردن از این نظام، نشان می‌دهد نظام شبکه پیشگیری از خودکشی (PSN) قدرت تصمیم‌گیری کاربر را در بازیافت اطلاعات مرتبط در مورد پیشگیری از خودکشی، افزایش می‌دهد.

محقق سپس به بحث و نتیجه‌گیری می‌پردازد. وی معتقد است این روش برای بازیافت اطلاعات حقوقی و کتابخانه‌ای نیز قابل استفاده است که می‌تواند در پژوهش‌هایی مورد سنجش قرار گیرد (وانگ و فورگیونه، 2006).

نتیجه و پیشنهاد

چنان‌که گفته شد، تحقیقات ارزیابی در زمینه بازیابی اطلاعات بر یک جنبه (رویکرد نظام مدار) عمدتاً متمرکز شده و در عین حال از هم گسیخته است. مطالعات ارزیابی در

رویکردهای مختلف از هم جدا افتاده‌اند. به اعتقاد «ساراسویک» (1995) این جدا افتادگی در سطوح ارزیابی می‌تواند به عنوان کمبود بنیادی در تمامی ارزیابی‌های بازیابی اطلاعات در نظر گرفته شود. نظام‌های بازیابی به عنوان فعالیتی اجتماعی، با هدف پیوند و ممکن ساختن تعاملات میان تهیه‌کنندگان/ نویسندگان اطلاعات و کاربران/ خوانندگان اطلاعات، ایجاد می‌شوند (ساراسویک، 1995). این فعالیت، ماهیتی پیچیده، شناختی و ذهنی دارد (بورلند، 2003). بنابراین، ارزیابی یک نظام نه تنها در زمینه فنی و سیستمی، بلکه در زمینه‌های کاربردی باید سنجش شود. علاوه بر این، در ارزیابی باید ترکیبی از معیارهای بروندادی و فرایندی، برای انعکاس ماهیت پویا و چند بعدی ربط استفاده شود. چالش عمده در راه تلاش برای ارزیابی بازیابی اطلاعات، کاربرد مقیاس‌های متنوع به روشی متحد شده و جامع است (ساراسویک، 1995) که مسیر پژوهش‌های آینده را روشن می‌کند. به اعتقاد برخی از محققان، چالش آینده بر تشریح ابعاد چندگانه و اجزای نسل جدید نظام‌های بازیابی اطلاعات از طریق آزمایش آنها با رویکردهای ارزیابانه متنوع، متمرکز است (رابرتسون و هنکاک بولیو، 1992). در این نوشته، پژوهشی مبتنی بر این رویکرد به عنوان نمونه معرفی شد که می‌تواند مسیر را روشن‌تر کند.

منابع

- پائو، میراندا لی (1378). **مفاهیم بازیابی اطلاعات**. ترجمه اسدالله آزاد و رحمت‌الله فتاحی. مشهد: دانشگاه فردوسی مشهد.

- بائزا - بیتس، ریکاردو؛ ریرو - نتو، برتیه (1385). **قلمروهای نو در بازیابی اطلاعات**. جلد اول. ترجمه علی حسین قاسمی با همکاری سیروس آزادی و علی جوامع. تهران: چاپار؛ دبیرش.

- _____ ، (1389). **قلمروهای نو در بازیابی اطلاعات**. جلد دوم. ترجمه علی حسین قاسمی با همکاری سیروس آزادی و علی جوامع. تهران: چاپار؛ دبیرش.

- Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information retrieval. **Canadian Journal of Information Science**, 5, 133-143.
- Blair, D. C., & Maron, M. E. (1990). Full-text information retrieval: further analysis and clarification. **Information Processing & Management**, 26(2), 437- 447.
- Blair, D. C., & Maron, M. E. (1985). An evaluation of retrieval effectiveness for a full-text document retrieval system. **Communications of the ACM**, 28, 289-299.
- Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. **Journal of Documentation**, 56(1), 71-90.
- Borlund, P. (2003). The concept of relevance in IR. **Journal of the American Society for Information Science and Technology**, 54(10), 913-925.
- Cleverdon, Cyril (1997). The Cranfield Tests on Index Language Devices. In: Sparck-Jones, Karen; Willett, Peter (Eds.): **Readings in Information Retrieval**. Morgan Kaufman, 47-59.
- Cleverdon, C.W., Mills, J., & Keen, E.M. (1966). **An inquiry in testing of information retrieval systems**. (2 vols.). Cranfield, U.K.:Aslib Cranfield Research Project, College of Aeronautics.
- Ellis, D. (1996). The dilemma of measurement in information retrieval research. **Journal of the American Society for Information Science**, 47(1), 23-36.
- Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. **Journal of Documentation**, 52(1), 3-50.
- Jarvelin, K., & Ingwersen, P. (2004). Information seeking research needs extension toward tasks and technology. **Information Research**, 10(1), paper no. 212. Available from <http://InformationR.net/ir/10-1/paper212.html>.
- Kekalainen, J., & Jarvelin, K. (2002). Evaluating information retrieval systems under the challenges of

- interaction and multidimensional dynamic relevance. In: Bruce, H., Fidel, R., Ingwersen, P., Vakkari, P. (Eds.), **Proceedings of the 4th CoLIS Conference** (pp. 253–270). Greenwood Village, CP: Libraries Unlimited.
- Kuhlthau, C. (1991). Inside the Search Process: Information Seeking from the User's Perspective. **Journal of the American Society for Information Science**, 42(5), 361–371.
 - Mandl, T. (2008). Recent Developments in the Evaluation of Information Retrieval Systems: Moving Towards Diversity and Practical Relevance. **Informatica**, 32, 27-38.
 - Robertson, S. E., & Hancock-Beaulieu, M. M. (1992). On the evaluation of IR systems. **Information Processing & Management**, 28(4), 457– 466.
 - Salton, G. (1992). The state of retrieval system evaluation. **Information Processing & Management**, 28(4), 441– 449.
 - Salton, G. (1989). Automatic text processing: the transformation, analysis and retrieval of information by **computer**. Reading, MA: Addison-Wesley.
 - Salton, G. (1971). The SMART retrieval systems: experiments in automatic document processing. Englewood Cliffs, NJ: Prentice-Hall.
 - Saracevic, T. (1995). Evaluation of Evaluation in Information Retrieval. **Proceedings of the Eighteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval**, 138–146.
 - Saracevic, T. (1996). Relevance reconsidered. In Ingwersen, P. and Pors, N.O. (Eds.), **Information Science: Integration in Perspective: Proceedings of the 2nd International Conference on Conceptions of Library and Information Science (CoLIS 2)**, 201–218.
 - Schamber, L. (1994). Relevance and information behavior. **Annual review of information science and technology**, 29, 3– 48.

Wang, Y.D., & Forgionne, G. (2006). A decision-theoretic approach to the evaluation of information retrieval systems. *Information Processing & Management*, 42, 863-874.

[1]. عضو هیئت علمی دانشگاه بیرجند و دانشجوی دکتری کتابداری و اطلاع‌رسانی دانشگاه فردوسی مشهد (kiani.mreza@gmail.com).

- [2]. Information Retrieval System (IRS).
- [3]. Mandl, T.
- [4]. Information focusing.
- [5]. Wang, Y.D., & Forgionne, G.
- [6]. Saracevic, T.
- [7]. Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier.
- [8]. Functional analysis.
- [9]. Performance.
- [10]. Relevance.
- [11]. System oriented approach.
- [12]. Recall & Precision.
- [13]. Cranfield paradigm.
- [14]. Smart.
- [15]. Stairs.
- [16]. TREC (Text REtrieval Conference).
- [17]. Kekalainen, J., & Jarvelin, K.
- [18]. Query.
- [19]. Abstraction.
- [20]. User oriented.
- [21]. Batch mode.
- [22]. Interactive.

[23]. برای توضیح بیشتر به کتاب «قلمروهای نو در بازیابی اطلاعات» مراجعه شود.

[24]. Recall effort.

- [25]. Relative recall.
- [26]. Novelty ratio.
- [27]. Coverage ratio.
- [28]. User interface.
- [29]. Information-seeking behavior.
- [30]. Relevance feedback.
- [31]. Jarvelin, K., & Ingwersen, P.
- [32]. Berry-picking.
- [33]. Analytic Hierarchy Process(AHP).
- [34]. Decision making.
- [35]. Information search process(ISP).
- [36]. Formulation.
- [37]. Multiple criteria decision making(MCDM).
- [38]. Preventing Suicide Network(PSN).