مروری بر کاربرد نظام‌های خبره و هوشمند در بازیابی اطلاعات

کیوان کوشا

چکیده

با وجود مطالعات نظری نسبتاً زیادی که در زمینهٔ استفاده از نظام‌های خبره و هوشمند در بازیابی اطلاعات صورت گرفته، به‌طور نسبی نظام‌های بازیابی اندکی را می‌توان مشاهده کرد که در آنها از قابلیت‌های نظام‌های خبره و هوش مصنوعی استفاده شده باشد. به تعبیر دیگر، در اغلب این‌گونه نظام‌ها، استفاده از الگوهای سنتی بازیابی اطلاعات نظیر منطق بولی قابل مشاهده است. بررسی متون مرتبط نشان می‌دهد که اغلب پیشرفت‌ها در این زمینه با مفاهیم پردازش هوشمند متن، واسط‌ها و عامل‌های هوشمند مرتبط بوده است. این مقاله سعی دارد به‌طور نظری میان اهداف مورد انتظار از نظام‌های خبره و هوشمند و فرایند بازیابی اطلاعات ارتباط برقرار کند. سرانجام، با رویکردی کاربرمدار این موضوع مورد تأکید قرار می‌گیرد که ماهیت برخی فرایندها در بازیابی اطلاعات مرتبط با نیاز اطلاعاتی واقعی کاربر به گونه‌ای پیچیده است که به دشواری می‌توان نظام‌های خبره و هوش مصنوعی را به‌طور کامل جایگزین آن کرد.

کلیدواژه‌ها: نظام خبره. نظام هوشمند. بازیابی اطلاعات. ربط. هوش مصنوعی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مروری بر کاربرد نظام‌های خبره و هوشمند در بازیابی اطلاعات

کیوان کوشا[1]

مقدمه

مقالات و متون متعددی در مورد مفهوم و کاربردهای نظام‌های خبره[2] و هوش مصنوعی[3] نوشته شده است. با این حال، جست‌وجو در ویرایش پیوسته (۲۰۰۴) *پایگاه چکیده مقالات کتابداری و اطلاع‌رسانی (لیزا)*[4] نشان داد که به‌طور نسبی مقالات اندکی در خصوص کاربردهای عینی نظام‌های خبره و هوشمند در بازیابی اطلاعات نوشته شده است. هدف مقالهٔ حاضر پاسخگویی به این سؤال است که آیا نظام‌های خبره و هوش مصنوعی (هوشمند) را می‌توان در نظام‌های بازیابی اطلاعات مورد استفاده قرار داد و اگر چنین است کاربرد عینی آنها در این‌گونه نظام‌ها چگونه است. مروری بر تحقیقات انجام شده در زمینهٔ کاربرد نظام‌های خبره و هوشمند نشان می‌دهد که ابهام‌هایی در خصوص کاربردهای عینی این‌گونه نظام‌ها در بازیابی اطلاعات وجود دارد. یکی از دلایل این امر انتشار متون و مقالاتی است که بیشتر بر جنبه‌های نظری تاکید داشته‌اند. این موضوع توسط لنکستر[5] و وارنر[6] در کتاب "کاربرد فن‌آوری‌های هوشمند در خدمات کتابداری و اطلاع‌رسانی" برجسته شده است (لنکستر و وارنر، ۲۰۰۱، ص ۱). در واقع، این ابهام وجود دارد که آیا نظام‌های مورد بحث در این‌گونه

۱. دانشجوی دکتری علوم‌کتابداری و اطلاع‌رسانی دانشگاه تهران و عضو هیأت علمی مرکز اطلاعات و مدارک علمی کشاورزی
kkoosha@ut.ac.ir
www.koosha.tripod.com

2. Expert systems
3. Artifical intelligence
4. Library and Information Science Abstract (LISA)
5. Lancaster
6. Warner

متون و مقالات، به‌طور عینی به مرحلهٔ بهره‌برداری رسیده‌اند و یا در پایگاه‌های اطلاعاتی که حاوی انواع مختلفی از اطلاعات بر روی دیسک فشرده یا شبکه‌های پیوسته هستند تنها ساختار و عملکردهای آنها به‌طور نظری مورد بحث قرار گرفته در قالب پیش نمون[1] ارائه شده‌اند. جدول ۱ دیدگاهی کلان نسبت به مقالات مرتبط منتشر شده در این زمینه را که در پایگاه *لیزا* نمایه شده‌اند به‌دست می‌دهد.

جدول ۱. مقاله‌های مرتبط نمایه شده در پایگاه پیوستهٔ چکیده مقالات کتابداری و اطلاع‌رسانی (لیزا: ۲۰۰۴)

	کلیدواژه(های) مورد جست‌وجو	فیلد مورد جست‌وجو	تعداد مقالات
۱	(Expert System* OR AI OR Intelligen*) AND Retrieval	توصیفگر	۲۵۵
۲	(Expert System* OR AI OR Intelligen*) AND Retrieval	عنوان مقالات	۲۸
۳	(Expert System* OR AI OR Intelligen*) AND Retrieval AND Application*	عنوان مقالات	۳
۴	(Expert System* OR AI OR Intelligen*) AND Retrieval AND Application*	توصیفگر	۱

دو رویکرد را می‌توان در خصوص کاربرد نظام‌های خبره و هوشمند در متون مختلف مشاهده کرد. تعدادی از مؤلفان عقیده دارند که رایانه می‌تواند به‌طور مجازی تقریباً هر کاری را که انسان انجام می‌دهد شبیه‌سازی کند. نمونهٔ بارز این ادعا را می‌توان در طراحی برنامه‌هایی که قادرند در سطح بسیار بالایی شطرنج بازی کنند مشاهده کرد. برخی دیگر از مؤلفان، این ادعا را تا حدود زیادی اغراق آمیز می‌دانند و مدّعی هستند که رایانه هرگز نمی‌تواند اعمالی را انجام دهد که واقعا هوشمندانه باشد. باید دانست که فارغ از بحث‌های جنجالی میان این دو طرز تفکر، در حال حاضر نظام‌های خبره و

1. Prototype

هوشمند چه کاربردهای عینی را توانسته‌اند در بازیابی اطلاعات ایفا کنند و آیا ماهیت بازیابی اطلاعات به گونه‌ای است که بتوان از نظام‌های خبره و هوشمند برای بهبود عملکرد آن، یعنی همان بازیابی "بیشترین اطلاعات مرتبط با نیاز اطلاعاتی کاربر"، در یک پایگاه اطلاعاتی بهره جست. در اینکه رایانه می‌تواند برخی امور (نظیر ذخیره و بازیابی حجم بسیار زیادی از اطلاعات یا انجام تجزیه و تحلیل‌های آماری پیچیده) را بهتر از انسان انجام دهد شکی نیست، اما در مواردی که نیاز به تصمیم‌گیری مبتنی بر تجزیه و تحلیل نیاز اطلاعاتی کاربر و مطابقت آن با محتوای یک مدرک مطرح است آیا نظام‌های خبره و هوشمند می‌توانند بهبود اساسی را در فرایند تشخیص این‌گونه مدارک با نیاز اطلاعاتی کاربران نهایی ایفا نمایند. تشخیص ربط خود از طریق کاربر یا متخصصان موضوعی با پیچیدگی‌های فراوانی روبه‌رو است. به‌طور مثال، نیازهای اطلاعاتی و رویکردهای موضوعی افراد در رویارویی با یک مدرک ممکن است متفاوت باشد. همچنین، حتی نیازی واحد از سوی یک فرد نیز ممکن است در زمان‌های مختلف تغییر پیدا کند (نشاط، ۱۳۸۲، ص ۳۹-۴۰).

سؤالاتی از این قبیل در حوزهٔ بازیابی اطلاعات از دیدگاه هرمنوتیک، که بر قلمرو شناختی تأکید داشته و آن را فرایندی تفسیری قلمداد کرده است، مورد بحث قرار گرفته است (نشاط، ۱۳۸۲). بنابراین، این سؤال بنیادی مطرح است که آیا نظام‌های خبره و هوشمند را نیز می‌توان در قلمرو شناختی، یعنی تطابق زبان جست‌وجو با نمادهای معرف محتوای مدارک از یک سو و نیاز اطلاعاتی استفاده‌کنندگان از سوی دیگر به‌کار برد. به عبارت دیگر، هر سه عنصر اصلی در بازیابی اطلاعات یعنی زبان جست‌وجو، توصیفگرهای معرف محتوای مدارک در یک پایگاه اطلاعاتی، و نیاز اطلاعاتی ابراز شدهٔ استفاده‌کننده تحت تأثیر دیدگاه هرمنوتیک هستند.

البته، رویکرد دیگری را نیز می‌توان مشاهده کرد که در آن تأکیدی بر نیازهای اطلاعاتی کاربران نمی‌شود. به‌طور مثال، در طرح‌های گرانفیلد[1] یک و دو که تحقیقات بازیابی اطلاعات به صورت تجربی و آزمایشگاهی دنبال شده است، به نیاز اطلاعاتی کاربران توجه نشده و تنها بهبود کارآیی نظام‌های بازیابی اطلاعات از نقطه‌نظر تطابق میان زبان جست‌وجو و نمادهای معرف محتوای مدارک مورد آزمون قرار گرفته است. چنین رویکردی را می‌توان در متون مرتبط با حوزهٔ رایانه بیشتر مشاهده کرد. برای اینکه بتوان تا حدودی به

1. Granfield

سؤال مطرح شده در این مقاله پاسخ داد، می‌توان هدف نظام‌های بازیابی اطلاعات را از یک سو، و امکانات و قابلیت‌های نظام‌های خبره و هوشمند را از سوی دیگر مورد بحث قرار داد.

نظام خبره و هوشمند چیست؟

الف. نظام هوش مصنوعی (هوشمند). در اغلب متون، این مفهوم دلالت بر نظامی دارد که قادر است برخی اعمال هوشمندانه انسان را شبیه‌سازی کند. فنلی[1] تعریف شفافی از نظام‌های هوش مصنوعی ارائه داده است: "برنامه‌های رایانه‌ای که به‌منظور شبیه‌سازی قدرت استدلال‌گرایی[2] انسان طراحی شده‌اند و می‌توانند از اشتباهات خود درس بگیرند و قادرند به صورت سریع و خبره اعمالی را انجام دهند که انجام آنها نیاز به تخصص انسان دارد" (فنلی، ۱۹۹۲). کاسی[3] در کتاب خود تحت عنوان "ماهیت هوش مصنوعی" این مفهوم را حوزه‌ای کلان از علم به‌شمار آورده است که ممکن است در اذهان مردم معانی مختلفی را (نظیر روبات‌های مصنوعی) تداعی کرده باشد. هوش مصنوعی قرار است وظایفی را انجام دهد که نیاز به هوش انسان دارد (کاسی، ۱۹۹۸). در این وادی، این پرسش مطرح است که آیا وظایفی نظیر محاسبات و تجزیه و تحلیل‌های پیچیدهٔ ریاضی که برنامه‌های رایانه‌ای قادرند آنها را با سرعت، دقت، و صحت بیشتری نسبت به انسان انجام دهند (نظیر برنامه‌های تجزیه و تحلیل آماری) می‌توان مصادیقی از یک نظام هوشمند به‌شمار آورد. اگرچه رایانه چنین اعمالی را می‌تواند با دقت و صحت بیشتری نسبت به انسان انجام دهد، اما در نقطه مقابل، کارهای بسیار ساده‌ای نیز وجود دارند که انسان قادر است آنها را به سهولت انجام دهد. به‌طور مثال، تشخیص چهرهٔ افراد برای انسان کاری بسیار معمولی محسوب می‌شود، ولی برای نظام‌های رایانه‌ای فرایندی بسیار پیچیده، دشوار، و چه‌بسا ناممکن محسوب می‌گردد. نظام‌های هوشمند قصد دارند به اموری بپردازند که برای انجام آنها نیاز به سطح بالایی از استدلال‌گرایی و تصمیم‌گیری است.

فورد[4] چنین تعریفی از نظام هوشمند ارائه داده است: "نظام هوشمند نظامی است که حداقل برخی قواعد دانش اکتشافی[5] ــ نه لزوماً در سطح خبره ــ استفاده کند یا بتواند قضاوت‌ها یا تصمیم‌گیری‌هایی را با وجود نامعلوم بودن یا ناقص بودن شواهد موجود انجام دهد" (فورد، ۱۹۹۱، ص ۸).

1. Fenly
2. Reasoning
3. Cawsey
4. Ford
5. Heuristic knowledge

با بررسی متون مرتبط دو هدف اصلی را می‌توان در ایجاد نظام‌های هوش مصنوعی به منظور شبیه‌سازی اعمال هوشمندانه انسان شناسایی کرد:

- درک و بررسی بهتر جنبه‌های مختلف هوش انسان و امکان شبیه‌سازی آن توسط ماشین
- طراحی برنامه‌های کاربردی هوشمند برای انجام وظایفی که نیاز به هوش انسان دارد

طبق نظرات مک‌کارتی و هیس[1]، مشکلی که حوزهٔ هوش مصنوعی با آن روبه‌رو است، با جنبه‌های معرف شناختی و اکتشافی مسائل در ارتباط است. برخی اعمالی را که نظام‌های خبره انجام می‌دهند به هیچ وجه نمی‌توان هوشمند دانست. در واقع، نظام‌های هوش مصنوعی باید از درجه بالایی از استدلال‌گرایی و تصمیم‌گیری در شرایط و موقعیت‌های مختلف برخوردار باشند. استفاده از منطق اکتشافی اساس کار نظام‌های هوشمند به شمار می‌آید. بنابراین، حتی برنامه‌هایی که از سطح بسیار بالایی از تحلیل‌های آماری و ریاضی استفاده می‌کنند نمی‌توان نظامی هوشمند تلقی کرد.

ب. نظام خبره. فورد نظام خبره را این‌گونه تعریف کرده است: "برنامه‌های رایانه‌ای همراه با دانش، اطلاعات، و پایگاه‌های اطلاعاتی که در ارتباطی پویا با یکدیگر سعی دارند به حل مسائل و فرایندهای تصمیم‌گیری در مقیاس نسبی محدودتری نظیر آنچه انسان انجام می‌دهد بپردازند". ویژگی‌های خاص نظام خبره طبق نظر فورد (۱۹۹۱، ص ۸) عبارتند از:

- توانایی رویارویی با داده‌هایی که ممکن است در آنها عدم قطعیت یا ناکامل بودن دیده شود (به‌طور مثال، سیستم بتواند پاسخی نظیر "نمی‌دانم" یا "ممکن است" یا "احتمالاً" را ارائه دهد).
- توانایی دسترسی به نتیجه‌گیری‌هایی با عدم قطعیت و در جایی که لازم است ارائه چندین رهیافت یا راه‌حل برای مسئله‌ای مشخص (هنگامی که جواب قطعی در مورد یک مسئله وجود ندارد صورت گیرد).
- توانایی توضیح اینکه چرا سؤالی مشخص پرسیده می‌شود و چگونه سیستم می‌تواند برای آن پاسخ مناسب و مشخصی پیدا کند.

فورد همچنین در کتاب خود به خوبی به تشریح وجوه تمایز نظام خبره با دیگر

1. McCarthy and Hayes

نظام‌های رایانه‌ای الگوریتمی پرداخته است:

آنچه به‌طور نسبی در ذخیره و بازیابی اطلاعات موضوع جدیدی تلقی می‌شود قابلیت ذخیره و بازیابی نه تنها داده‌های عددی و الگوریتمی[1] بلکه داده‌های نمادین و اکتشافی[2] است. با توسعهٔ نرم‌افزارها و سخت‌افزارها، انجام فرایندهای مبتنی بر دانش اکتشافی با سهولت بیشتر و هزینهٔ کم‌تری قابل انجام است. دانش نمادین (در نقطه مقابل دانش عددی) چیزی است که ما برای تصمیم‌گیری‌ها و قضاوت‌های خود از آن استفاده می‌کنیم، و این دانش اغلب از طریق دسترسی به نتایجی از شواهد موجود به‌دست می‌آید.

اسپانگلر و می[3] (۲۰۰۰) بر این باورند که:

نظام خبره، ابزاری نرم‌افزاری است که مبتنی بر فن‌آوری هوش مصنوعی توسعه یافته است. در ابتدا، هدف از طراحی نظام‌های خبره، شبیه‌سازی و چه‌بسا جایگزینی توانایی استدلال‌گرایی انسان در رویارویی با مسائل مختلف بوده است. اما امروزه، عملکردهای معمولی‌تری از نظام‌های خبره به منظور کمک به انجام برخی امور نظیر تشخیص بیماری‌ها، کنترل فرایندها در چرخهٔ تولید محصولات، تجزیه و تحلیل‌های مالی، تفسیر داده‌ها، و مانند آن انتظار می رود.

ارتباط نظری نظام‌های خبره و هوشمند با فرایند بازیابی اطلاعات

نکاتی که در تعاریف فوق آمده به اجمال مورد اشاره قرار می‌گیرد و در مورد ارتباط منطقی هریک از آنها با فرایند بازیابی اطلاعات بحث می‌شود:

الف. شبیه‌سازی قدرت استدلال و تصمیم‌گیری انسان. این پرسش را می‌توان مطرح کرد که در نظام بازیابی اطلاعات چه فرایندی باید توسط ماشین شبیه‌سازی شود که قبلاً وابسته به قدرت استدلال انسان بوده است. شاید یکی از مهم‌ترین نقش‌های انسان در مرحلهٔ بازیابی اطلاعات، انتخاب کلیدواژه‌های مرتبط با موضوع مورد نظر و ترکیب آنها از طریق عملگرهای مختلف از یک سو و تشخیص میزان ربط مدارک بازیابی شده با موضوع مورد نظر از سوی دیگر باشد. اینکه چنین اعمالی نیاز به استدلال دارد یا خیر خود قابل بحث است. بدون شک، حداقل حوزهٔ اطلاع‌رسانی به دنبال چیزی بیش از تطابق صرف کلیدواژه‌های مورد جست‌وجو (نیاز اطلاعاتی بیان شده نه واقعی) و

1. Numeric and algorithmic data
2. Symbolic and heuristic data
3. Spangler and May

نمادهای موضوعی در مدارک است. با این همه، تحقیقات زیادی به صورت تجربی و آزمایشگاهی تحت عنوان طرح‌های گرانفیلد یک و دو صورت گرفته است که در اغلب آنها چنین رویکردی مورد تأکید بوده است. چه‌بسا، فرایند فوق نیاز زیادی به قدرت استدلال انسان نداشته باشد و ماشین بتواند از طریق نظام‌های فرعی نظیر اصطلاحنامه‌ها و پرونده‌های واژگان فرایند کاوش را انجام دهد. اگر هدف پاسخ به نیاز اطلاعاتی واقعی کاربران باشد، آنگاه قدرت استدلال انسان در تشخیص میزان ربط کلیدواژه‌های جست‌وجو از یک سو و مدارک بازیابی شده از سوی دیگر اهمیت پیدا می‌کند. بنابراین، حداقل به‌طور نظری اگر نظامی به دنبال پاسخ‌گویی به چنین نیازی از طریق از میان برداشتن یا کم‌رنگ کردن نقش انسان (اغلب متخصصان موضوعی و کاوش) باشد نظام‌های خبره و هوشمند می‌توانند نقش اساسی ایفا نمایند.

ب. انتقال تخصص و یادگیری از طریق بازخوردها. در اغلب نظام‌های خبره و هوشمند به عنصری اساسی نیاز است و آن انتقال تخصص انسان[1] به شکلی قابل ارائه به یک نظام است. در واقع، تولید قواعد یا ترکیب آن با شبکه‌های معنایی یا قالب‌های مختلف، باید توسط شخصی متخصص به نظام انتقال یابد. یک متخصص ممکن است دانش‌های مورد نیاز را از طریق شیوه‌های مختلف نظیر مصاحبه، مشاهده، یا تجربه به‌دست آورده و وارد نظام کرده باشد. اما گردآوری دانش در حوزهٔ بازیابی اطلاعات توسط متخصصان، به‌منظور به رمز در آوردن آنها برای رایانه، کاری دشوار است؛ زیرا بسیاری از دانش‌های به‌دست آمده در این حوزه متکی بر پیش‌فرض‌هایی است که ممکن است به‌سادگی مورد تردید قرار گیرند. به‌طور مثال، مفهوم ربط و نیاز اطلاعاتی از جمله پیش فرض‌هایی است که چه‌بسا در بسیاری از تحقیقات انجام شده در حوزهٔ بازیابی اطلاعات به‌طور واقعی به آنها نگاه نشده است. از این رو، دانش‌های به‌دست آمده مبتنی بر این‌گونه تحقیقات نیز خود با ابهام‌هایی روبه‌رو است.

پورس[2] (۲۰۰۰) در مقالهٔ خود تحت عنوان "بازیابی اطلاعات: الگوهای تجربی و تجزیه و تحلیل‌های آماری" روایی برخی تحقیقات انجام شده در حوزهٔ بازیابی اطلاعات را مورد انتقاد قرار می‌دهد. به‌طور مثال، وی اشاره دارد که طبق اصول آماری، تعداد کلیدواژه‌های مورد کاوش، تعداد رکوردهای بازیابی شده، تعداد افرادی که به

1. Expertise
2. Pors

تشخیص مرتبط بودن یا نبودن مدارک بازیابی شده می‌پردازند، و تعداد قضاوت‌های صورت گرفته در خصوص ربط در تحقیقات بازیابی اطلاعات باید منطبق با قوانین نمونه‌گیری باشد تا بتوان نتایج به‌دست آمده را به کل یک پایگاه اطلاعاتی تعمیم داد؛ حال آنکه نمونه‌گیری از کلیدواژه‌هایی که ممکن است استفاده‌کنندگان برای کاوش خود انتخاب کنند تقریباً غیرممکن است. از این رو، این سؤال مطرح است که دانش به‌دست آمده از حوزهٔ بازیابی اطلاعات مبتنی بر چنین تحقیقاتی تا چه حد قابل انتقال به یک نظام خبره یا هوشمند است. این موضوع با نظر مک‌کارتی و هیس که قبلاً به آن اشاره شد ارتباط نزدیک دارد؛ زیرا مشکلی که حوزهٔ هوش مصنوعی با آن روبه‌رو است با جنبه‌های معرفت‌شناسی و اکتشافی مسائل در ارتباط است.

هرگاه پذیرفته باشیم که دانش مورد نیاز در حوزهٔ بازیابی اطلاعات برای انتقال به نظام خبره و هوشمند برای تعیین و اجرای وظایف یا قواعد، خود با مشکلات بنیادی روبه‌رو است، این سؤال مطرح است که آیا به شیوه‌های دیگری می‌توان به طراحی و توسعه این‌گونه نظام‌ها پرداخت.

برخی متخصصان طراحی نظام‌های خبره در بازیابی اطلاعات، بر قابلیت یادگیری ماشین تأکید دارند. به‌طور مثال، در فرایند بازیابی اطلاعات، رایانه قادر خواهد بود از بازخورهای به‌دست آمده از کاربران در خصوص مرتبط بودن یا نبودن اسناد بازیابی شده، برای تصمیم‌گیری‌های بعدی خود استفاده کند. یکی از رایج‌ترین فنون اعمال یادگیری توسط رایانه در نرم‌افزارهای نظام‌های خبره و هوشمند، تولید خودکار قواعد[1] است. از طریق چنین نرم‌افزارهایی می‌توان نظامی دانش‌مدار به‌طور خودکار تولید کرد. در واقع، این نرم‌افزارها به صراحت از دستورات "اگر"، "پس"[2] برای ارائه قواعد استفاده نمی‌کنند، بلکه به‌طور خودکار به تولید قواعد مورد نیاز می‌پردازند. اما در این میان، طراح سیستم باید نمونه‌های مورد نظر اولیه را برای برنامه مشخص کند. به این ترتیب، این نرم‌افزارها توانایی شناسایی قواعد در داده‌ها را دارند. معنی ضمنی این نکته آن است که برنامه می‌تواند از اطلاعات مربوط به مثال‌ها و نمونه‌های ارائه شده، قواعد مورد نظر را استخراج کند و دانشی را ارائه دهد که قبلا خالق آن طراح برنامه نیست. در واقع، چنین نظام‌هایی قادرند در موارد خاصی دانش جدیدی را تولید کنند که پیش از آن برای انسان قابل دسترس نبوده است. در بسیاری از موارد، کاربران پس از کاوش اولیه به اصلاح و پالایش کاوش خود می‌پردازند و به این ترتیب ممکن است به اطلاعات مرتبط و مورد

1. Rule induction
2. IF, THEN

نیاز خود دست یابند. حال اگر نظامی بتواند حداقل به‌طور نظری از اشتباهات و بازخوردهای کاربران برای اصلاح یا تغییر برخی عملکردها در فرایند بازیابی اطلاعات بیاموزد (به‌طور مثال، کلیدواژه‌های مرتبط را با املای صحیح وارد کند و آنها را با دیگر کلیدواژه‌های مترادف در اصطلاحنامه یا سرعنوان موضوعی، یا بازخوردهای دیگر کاربران ترکیب نماید؛ و یا اینکه بر اساس سطح دانش کاوش کاربر از واسطهای جست‌وجوی مختلفی استفاده نماید، و ...)، می‌توان سطحی از خبرگی را در آن یافت.

ج. انجام اعمالی که نیاز به تخصص انسان دارد. نظام‌های خبره برنامه‌های رایانه‌ای هستند که قادرند وظایف معمولی افرادی را که دارای نوعی مهارت یا تخصص ویژه[1] هستند انجام دهند. هدف از طراحی نظام‌های خبره این است که بتوان برخی مهارت‌های خاص را در دسترس عموم قرار داد. مهارت کاوش در پایگاه‌های اطلاعاتی مختلف با واسطها و فرمان‌های کاوش مختلف، اغلب خاص متخصصان کاوش در کتابخانه‌ها و مراکز اطلاع‌رسانی است. حال اگر نظامی بتواند محیطی تعاملی برای جایگزینی چنین مهارت‌هایی را پدید آورد، می‌توان آن را نظام بازیابی خبره به‌شمار آورد. به‌طور مثال، حداقل به‌طور نظری، نظامی که قادر باشد از طریق مصاحبه مرجع تا حدودی میان نیاز بیان شده و نیاز واقعی کاربر ارتباط منطقی برقرار کند و خود بهترین راهبرد کاوش را در پیش بگیرد می‌توان آن را خبره تلقی کرد. البته، این سؤال مطرح است که هنگامی که انجام چنین فرایندهایی با استفاده از انسان (متخصص کاوش) با مشکلات و پیچیدگی‌های زیادی روبه‌رو است، ماشین چگونه خواهد توانست بهبودی در آن ایجاد کند. به عبارت دیگر، در نظام‌های خبرهٔ دانش ـ مدار باید از قبل دانش و مهارتی را در نظام گنجاند و از آن خواست در شرایط مختلف و مورد نیاز از آن استفاده کند. تجربه و مهارت کسب شده توسط متخصصان کاوش (نظیر فهرست‌نویسی و نمایه‌سازی) موضوعی است که به سادگی آن را نمی‌توان از طریق نظام خبره ارائه داد. با این همه، تحقیقات در این عرصه می‌کوشند تا هرچه بیشتر بر اساس نیاز بیان شدهٔ کاربر به نیاز واقعی او پی ببرند: مسئله‌ای که قرن‌ها بشر نتوانسته است آن را به‌طور کامل حل کند.

در مثال زیر، نمونهٔ فرضی یک پایگاه دانش نشان داده شده است تا از طریق آن بتوان کاربران را با استفاده از پرسش و پاسخ به پایگاه اطلاعاتی مرتبط با نیاز آنها

1. Expertise

هدایت کرد:

اگر موضوع مورد کاوش "اسلام" باشد

و نوع منابع مرجع مورد کاوش مقالات مجلات فارسی باشد

و متن کامل مقالات مورد نیاز باشد

سپس توصیه این است که پرس‌وجوی کاربر در "پایگاه تمام متن نمایه" صورت گیرد

با این همه، با استفاده از شبکهٔ معنایی، به‌طور نسبی، ارائه نظامی که بتواند به‌طور خودکار کاربران را در کاوش‌های خرد یا کلان کمک کند آسان است. این فرایند را می‌توان از طریق گنجاندن قواعدی که از یک شبکهٔ معنایی برای ارائهٔ نتیجه‌گیری دربارهٔ ارتباط میان اصطلاحات استفاده می‌کند انجام داد. یکی از نظام‌هایی که از ساختار شبکهٔ معنایی در بازیابی اطلاعات استفاده می‌کند MenUSE[1] نام دارد. این برنامه، امکان کاوش پیوستهٔ مؤثری را به کاربران، بدون آنکه نیاز به آموزش یا تجربهٔ زیادی در امر کاوش اطلاعات داشته باشند، ارائه می‌دهد. اساس برنامهٔ MenUSE را سرعنوان موضوعی پزشکی تشکیل داده است (پولیت[2]، ۱۹۹۸). هدف از توسعهٔ این برنامه حذف انجام اعمالی است که نیاز به تخصص انسان ــ یعنی کاوش توسط متخصصان موضوعی با تجربه ــ دارد.

د. قضاوت با وجود نامعلوم بودن یا ناقص بودن شواهد موجود. چه‌بسا، مهم‌ترین قضاوتی که در فرایند کاوش باید صورت گیرد نیاز واقعی استفاده‌کننده است، در حالی که ابعاد آن نامعلوم و اغلب ممکن است اطلاعات به‌دست آمده دربارهٔ آن ناقص باشد. این موضوع مورد تأکید قرار گرفته است که باید میان نیازهای بیان شده از سوی کاربران نهایی و نیازهای واقعی آنها تفاوت قائل شد. یکی از مهم‌ترین شیوه‌های پی بردن به نیاز واقعی کاربران مصاحبه مرجع است. بنابراین، حداقل به‌صورت نظری، اگر نظامی قادر باشد مکانیزمی برای مصاحبه با کاربر قبل از انجام کاوش و نیز پس از مرحله بازیابی مدارک انجام دهد، می‌توان سطحی از خبرگی را در آن مشاهده کرد. اینکه نظامی از طریق پایگاه دانش خود بتواند به هر سؤالی که از سوی کاربر مطرح می‌شود پاسخ دهد و آن را به صورت منطقی ردگیری کند، اگرچه از طریق دستورات "اگر"، "پس" میسّر است، احتمالا ایجاد چنین پایگاهی هزینهٔ بسیار زیادی دربر خواهد داشت. تجربه متخصصان موضوعی و کتابداران مرجع در طول سالیان متمادی در طرح سؤالات پیش‌بینی نشده و

1. Menu-based User Search Engine
2. Pollit

یا حدس زدن نیاز اطلاعاتی آنها از روی رفتار و حتی ظاهر کاربران چیزی است که به سادگی آن را نمی‌توان در قالب نظامی ماشینی اجرا کرد. به‌طور مثال، آیا یک نظام خبره در پاسخ به این سؤال مطرح شده از سوی یک دانش‌آموز و دکترای فیزیک در باب نظریهٔ نسبیت به‌طور یکسان برخورد خواهد کرد. توانایی توضیح اینکه چرا سؤالی مشخص پرسیده می‌شود، چه کسی با چه ویژگی‌هایی آن را مطرح می‌کند، نیاز واقعی فرد برای دسترسی به اطلاعات مورد نظر چیست، و در نهایت چگونه نظام می‌تواند برای چنین سؤالاتی پاسخ مناسب و مشخصی پیدا کند، تقریباً به‌طور عملی ناممکن است. این نکته که نظام‌های خبره و هوشمند چگونه خواهند توانست مسائل هرمنوتیکی و ذهنی مرتبط با تشخیص ربط و نیاز واقعی کاربران را در فرایند بازیابی اطلاعات حل کنند، چالش‌های اساسی بر سر راه تحقیقات جاری نظام‌های خبره و هوشمند محسوب می‌گردد.

هـ. توانایی دسترسی به نتیجه‌گیری با عدم‌قطعیت: ارائه راه‌حل برای مسائل. توانایی هر نظام در رویارویی با مسائل و ارائه راه‌حل‌هایی با عدم قطعیت از استنتاج‌های به‌دست آمده، یکی از مهم‌ترین ویژگی‌های هر نظام خبره به شمار می‌آید. در زندگی روزمره و نیز تصمیم‌گیرهای تخصصی و علمی، اغلب ممکن است نتیجه‌گیرهای ما در مورد مسائل مختلف دارای قطعیت نباشد و با عبارت‌هایی نظیر شاید، احتمال دارد، چه بسا ، ممکن است و غیره همراه باشد. آیا نظام هوشمند یا خبره می‌تواند از یک سو درجه بالایی از استدلال‌گرایی، تصمیم‌گیری و هوشمندی را در فرایند تطابق پرسش‌های کاربران با نمادهای موضوعی مدارک موجود در یک پایگاه اطلاعاتی و از سوی دیگر، با نیاز اطلاعاتی واقعی آنها که احتمال دارد در پرسش‌ها منعکس نباشد (یعنی کلیدواژه‌های مورد جست‌وجو) برقرار سازد؟ شاید رتبه‌بندی نتایج کاوش و وزن دادن به کلیدواژه‌ها را بتوان نوعی نتیجه‌گیری با عدم‌قطعیت به‌شمار آورد. در واقع، برخی معتقدند که با نمایش میزان ربط احتمالی پرسش‌های کاربر و مدارک بازیابی شده (اغلب از طریق نشان دادن درصد ربط) می‌توان نوعی نتیجه‌گیری با عدم‌قطعیت را در فرایند بازیابی اطلاعات ابراز کرد. هنگامی که یک مدرک با درجهٔ ربط ۱۰ درصد بازیابی می‌شود، نظام بازیابی قطعیت کامل دربارهٔ ربط کامل آن با نیاز کاربر ارائه نداده است. اما خود تشیخص اینکه چه مدارکی (چه از لحاظ رتبه‌بندی نتایج کاوش و چه از لحاظ

درجه ربط به صورت درصد) با نیاز کاربر ربط دارند، مجدداً در چرخه‌ای از ابهام قرار دارد که مشابه آن را می‌توان در سنجش ربط در نظام‌های بازیابی اطلاعات مشاهده کرد.

و. استفاده از دانش اکتشافی به‌جای منطق الگوریتمی. در بسیاری از تعاریف ارائه شده از نظام‌های خبره و هوشمند بر این نکته تأکید شده است که آنها از دانش اکتشافی (دانشی که به‌نظر می‌رسد بر مبنای تجارب عملی، درست عمل می‌کنند) به‌جای منطق الگوریتمی استفاده می‌کنند. این برنامه‌ها، اعمالی را فراتر از پردازش داده‌های حرفی و رقمی انجام می‌دهند. دانش اکتشافی قواعدی تجربی هستند که اغلب مورد استفادهٔ متخصصان قرار می‌گیرند. این‌گونه قواعد، با قواعد الگوریتمی متفاوتند و مسیرهای کوتاهی را برای یافتن راه حل مسائل مختلف ارائه می‌دهند، و نسبت به قواعد الگوریتمی، به نیرو و زمان کمتری برای پردازش نیاز دارند. قواعد اکتشافی را می‌توان در شرایطی که قواعد الگوریتمی ناشناخته یا نامناسب هستند به‌کار برد. نمونه‌ای فرضی از یک منطق اکتشافی که می‌توان آن را در یک نظام بازیابی اطلاعات مورد استفاده قرار داد در مثال زیر آمده است:

سند الف دربارهٔ پرندگان است
اصطلاح طوطی، در زیر مجموعه اصطلاح کلان پرندگان قرار دارد
بنابراین سند الف دربارهٔ طوطی نیز هست

مثال فوق را می‌توان در فرایند بازیابی اطلاعات برای افزایش احتمال دسترسی کاربر به منابع مرتبط دیگر باز نیاز اطلاعاتی او، مورد استفاده قرار داد. به این ترتیب، کاربری که به دنبال اطلاعاتی در مورد "طوطی" است، می‌تواند به مدارک مرتبط کلان‌تر دیگر که موضوع پرندگان و احتمالاً طوطی را نیز دربر دارد دست یابد.

در اغلب نظام‌های رایانه‌ای غیرخبره، تصمیم‌گیری‌ها بر اساس قواعد ریاضی و الگوریتمی اتخاذ می‌شوند (نظیر برنامه‌های تجزیه و تحلیل آماری و ریاضی). به عبارت دیگر، توانایی استدلال‌گرایی این‌گونه نظام‌ها فقط در ارتباط با مسئله‌ای از پیش تعریف شده قابل اجراست. منطق الگوریتمی در نظام بازیابی اطلاعات را می‌توان مشابه تطابق عبارت جست‌وجو و نمادهای موجود در یک سند یا تجزیه و تحلیل فراوانی تکرار کلیدواژه‌ها در فیلدهای مختلف دانست که اغلب میزان ربط نیاز بیان شدهٔ

کاربر با مدارک موجود در یک پایگاه اطلاعاتی بر مبنای آنها سنجیده می‌شود. در مقابل، منطق اکتشافی قرار است از ظرفیت‌های استدلال‌گرایی مبتنی بر یک پایگاه دانش‌مدار برای حل مسائلی که قبلاً به‌طور تجربی برای آنها راه‌حلی به‌دست آمده استفاده نماید. حال این سؤال مطرح است که الگوی مشابه منطق اکتشافی در چه شرایطی در یک نظام بازیابی اطلاعات قابل اجراست. به تعبیر دیگر، چه موقع می‌توان گامی فراتر از تطابق صرف کلیدواژه‌های جست‌وجو با نمادهای گنجانده شده در یک مدرک برداشت (که البته خود این موضوع هیچ‌گاه هدف نهایی حوزهٔ اطلاع‌رسانی نبوده است). شاید بتوان این‌طور استنباط کرد که بازیابی مدارک مرتبط با نیاز اطلاعاتی واقعی کاربران مبتنی بر نیازهای بیان شده چیزی است که تنها از طریق به‌کارگیری منطق اکتشافی در نظام‌های خبره و هوشمند قابل اجراست. همان‌گونه که اشاره شد، منطق اکتشافی چیزی جز قواعدی که به‌طور تجربی از مهارت‌ها و تجارب کسب شده توسط انسان به‌دست آمده نیست. بنابراین، حداقل به‌طور نظری، ایجاد پایگاه دانش‌مدار در یک نظام خبره فقط با به‌کارگیری دانش تجربی که خود بر مطالعات صورت گرفته دربارهٔ کاربران (نظیر رفتار اطلاع‌یابی، تعامل کاربر با واسط جست‌وجو، نیاز اطلاعاتی و مانند آن) استوار باشد دست‌یافتنی است. از این رو، تا زمانی که نتوان از طریق تحقیقات گوناگون دانش مورد نیاز برای ایجاد چنین محیطی را فراهم ساخت، ایجاد و توسعه یک نظام بازیابی مبتنی بر منطق اکتشافی نیز دور از انتظار خواهد بود.

کاربردهای عینی نظام‌های خبره و هوشمند در بازیابی اطلاعات

لنکستر عقیده دارد که، در طول سال‌های متمادی تحقیقات در حوزهٔ بازیابی اطلاعات توسط عدهٔ معدودی از محققان صورت می‌گرفت و تحقیقات آنها آثار بسیار اندکی بر محصولات تجاری داشته است. به‌طور مثال، محصولات عرضه‌کنندگان خدمات و پایگاه‌های اطلاعاتی در جهان نظیر دیالوگ سال‌ها از رویکردهای منطق بولی در بازیابی اطلاعات برای مقابلهٔ عبارات جست‌وجو و مدارک در پایگاه‌های خود استفاده می‌کردند. هم‌اکنون، بسیاری از عرضه‌کنندگان این محصولات از دستاوردها و نتایج تحقیقات در عرصهٔ بازیابی اطلاعات استفاده نمی‌کنند. سال‌هاست که دریافته‌اند استفاده از منطق بولی برای کاربران معمولی مشکلات بسیاری را در فرایند بازیابی اطلاعات به‌وجود می‌آورد. این در حالی است که در نتایج برخی تحقیقات، استفاده از

نظام تطابق عبارت‌های جست‌وجو[1]، رتبه‌بندی نتایج کاوش، وزن دادن به واژگان جست‌وجو، و مانند آن تأکید شده است (لنکستر و وارنر، ۲۰۰۱، ص ۳۵).

عملکردی که در فرایند بیشتر نظام‌های بازیابی اطلاعات دیده می‌شود، تطابق بخشی از یک متن (اغلب کلمات یا عبارت جست‌وجو) با مدارک موجود در یک پایگاه اطلاعاتی است (که خود می‌تواند اطلاعات کتاب‌شناختی، تمام متن یا چندرسانه‌ای را دربر داشته باشند). در طول بیش از چهار دهه گذشته، فنون مختلفی برای بهبود این عملکرد به ظاهر ساده پیشنهاد شده و به‌کار رفته است تا بتوان ابزارهای مؤثرتر و کارآمدتری را برای دسترسی به اطلاعات مرتبط و مورد نیاز کاربران توسعه داد. نتیجهٔ چنین تحقیقاتی، توسعهٔ نظام‌های بازیابی اطلاعات پیچیده‌ای است که اغلب از تجزیه و تحلیل‌های آماری یا فنون پردازش زبان طبیعی استفاده می‌کنند. اما، به هر حال، به این‌گونه نظام‌ها نمی‌توان عنوان نظام بازیابی "هوشمند" اطلاق کرد؛ زیرا همان‌گونه که اشاره شد در یک نظام خبره و هوشمند مناسب‌تر آن است که به‌جای منطق‌های الگوریتمی و ریاضی از منطق‌های سمبولیک و اکتشافی استفاده شود. لنکستر و وارنر عقیده دارند که مطالعات با ارزشی در زمینهٔ نظام‌های بازیابی اطلاعات صورت گرفته است که برخی از آنها منجر به طراحی نظام‌های بازیابی اطلاعات بسیار پیچیده‌ای نیز گردیده است. در بعضی از این مطالعات، به خوبی از شیوه‌های پردازش آماری و نیز زبان طبیعی برای مطابقت دادن عبارت‌های جست‌وجو با رکوردهای موجود در پایگاه اطلاعاتی استفاده شده است که نمونهٔ بارز آنها را می‌توان در تحقیقات سالتون و مک گیل[2] (۱۹۸۳)، فریکز و بیزا- یتز[3] (۱۹۹۲)، و کورف‌هیج[4] (۱۹۹۷) مشاهده کرد. با این همه، پردازش‌هایی که در این‌گونه نظام‌ها صورت می‌گیرند به هیچ وجه هوشمند نیستند (لنکستر و وارنر، ۲۰۰۱، ص ۳۶).

دانشگاه ماساچوست، با تشکیل مرکزی تحت عنوان بازیابی هوشمند اطلاعات (سیر)[5] به تحقیق و مطالعه در عرصه طراحی نظام‌های بازیابی اطلاعات به‌صورت هوشمند و خبره پرداخته است. این مرکز در مقالات و کتاب‌های مختلف به‌عنوان مهم‌ترین مرکز تحقیقاتی در این زمینه مورد اشاره قرار گرفته است. با مراجعه به سایت این مرکز می‌توان اولویت‌های تحقیقاتی زیر را مورد شناسایی قرار داد که آگاهی از آنها در زمینهٔ بحث حاضر از اهمیت بسزایی برخوردار است (سیر، ۲۰۰۴).

- **دریافت بازخورد نسبت به چگونگی ربط**[6]: فرایندی که در آن کاربر نهایی اقلام

1. Matching phrase
2. Salton and McGill
3. Frakes and Baeza-Yates
4. Korfhage
5. Center for Intelligent Information Retrieval (CIIR)
6. Relevance feedback

اطلاعاتی مرتبط را در مرحلهٔ اول بازیابی اطلاعات مورد شناسایی قرار می‌دهد و سپس نظام خبره به‌طور خودکار عبارت کاوش جدیدی را بر اساس بازخوردهای استفاده‌کننده مورد کاوش و بازیابی قرار می‌دهد.

- **بازیابی اطلاعات چندرسانه‌ای**[1]: دسترسی به تصاویر و قطعات صوتی ویدئویی در پایگاه‌های اطلاعاتی بدون دربر داشتن اطلاعات توصیفی نوشتاری دربارهٔ آنها.
- **بازیابی اطلاعات مؤثرتر**[2]: بهبود جامعیت و مانعیت در بازیابی اطلاعات.
- **توسعهٔ واژگان جست‌وجو**[3]: توسعهٔ واژگان جست‌وجو فراتر از اصطلاحاتی که توسط کاربر نهایی مورد کاوش قرار گرفته است.
- **بازیابی کارآمد اطلاعات**[4]: توسعهٔ الگوریتم‌های کاوش به‌طوری که بتواند به‌طور همزمان[5] و در نظام‌های تعاملی عمل کند.
- **بازیابی توزیعی**[6]: ارائه فهرست رتبه‌بندی شده از نتایج کاوش مبتنی بر بازیابی اطلاعات از چند پایگاه اطلاعاتی.

یکی از نرم‌افزارهایی که توسط مرکز بازیابی هوشمند اطلاعات طراحی شده و از شیوه‌های مختلفی به منظور بازیابی اطلاعات مرتبط با نیاز اطلاعاتی کاربران استفاده کرده است INQUERY نام دارد. شاید این برنامه را بتوان یکی از مهم‌ترین برنامه‌های توسعه یافته به‌شمار آورد که از قابلیت‌های نظام‌های هوشمند و خبره به‌طور عینی و کاربردی در فرایند بازیابی اطلاعات استفاده‌کرده است. این برنامه مدل‌های زبان طبیعی، منطق بولی، نزدیک‌یابی[7]، بازخوردهای استفاده‌کننده در خصوص مرتبط بودن یا نامرتبط بودن اطلاعات بازیابی شده، و اصلاح نتایج کاوش را مورد استفاده قرار می‌دهد؛ و رکوردهای بازیابی شده را بر اساس احتمال مرتبط بودن با نیاز اطلاعاتی کاربر رتبه‌بندی می‌کند. این برنامه قادر است به‌طور خودکار عبارت‌های جست‌وجو را بر اساس بازخوردهای کاوشگران اصلاح کند، تا به این ترتیب بتوان رکوردهای مرتبط دیگر را نیز بازیابی کرد. اسناد طولانی هم بر مبنای محتوای درونی آنها و هم بر اساس مرتبط‌ترین عبارت[8] بازیابی می‌شوند (این کویری[9]، ۲۰۰۴). بررسی ویژگی‌ها و قابلیت‌های این برنامه از این لحاظ اهمیت دارد که ظاهراً سعی شده است نتایج بسیاری از تحقیقات نظری این مرکز و سایر تحقیقات در عرصهٔ به‌کارگیری نظام‌های خبره و هوشمند در بازیابی مؤثر و کارآمد اطلاعات در این نرم‌افزار به‌طور کاربردی پیاده‌سازی شود.

1. Multimedia retrieval
2. More effective retrieval
3. Vocabulary expansion
4. Efficient information retrieval
5. Real time
6. Distributed IR
7. Proximity
8. Best passage
9. INQUERY

پردازش هوشمند متن

از دهۀ ۱۹۵۰، استفاده از رایانه برای ذخیره‌سازی متن به‌صورت الکترونیکی آغاز شد. قابلیت فزایندۀ ذخیره‌سازی داده‌ها توسط رایانه، سبب گردید که در حال حاضر حجم فراوانی از اطلاعات در سراسر جهان در قالب الکترونیکی ذخیره شود. حال این سؤال مطرح است که آیا می‌توان از این حجم وسیع اطلاعات نوشتاری، که در قالب الکترونیکی ذخیره شده است، برای پاسخگویی به نیازهای اطلاعاتی کاربران نهایی بهره جست یا باید پردازش‌هایی بر روی آنها انجام شود. طرح چنین پرسشی سبب شد که یکی از کاربردهای نظام‌های پردازش متن در حوزۀ کتابداری و اطلاع‌رسانی جلوه‌گر شود. هدف از طراحی چنین نظام‌هایی فراتر از ارجاع کاربر به یک یا چند منبع مستقل است که احتمالاً جواب سؤال مورد نظر وی در آنها یافت می‌شود. در واقع، این‌گونه نظام‌ها قصد دارند تا حد ممکن با پردازش منابع مختلف و یکپارچه کردن بخش‌هایی مرتبط از آنها به نیاز کاربر پاسخی مناسب دهند.

رویکردهای جاری در خصوص طراحی نظام‌های پردازش متن را می‌توان به‌گونه‌ای "هوشمند" تلقی کرد. زیرا رایانه قادر است به تفسیر معانی یک کلمه یا جمله بپردازد و مبتنی بر آن پاسخی مناسب را ارائه دهد. به عبارت دیگر، از طریق تجزیه و تحلیل‌های نحوی می‌توان نقش یک کلمه در جمله (اسم، فعل و مانند آن) و نیز عناصر ساختاری مرتبط با آن را تشخیص داد. در حال حاضر، نظام‌های پردازش هوشمند متن به صورت تجربی یا کاربردی در زمینه‌های زیر مورد توجه قرار گرفته‌اند:

الف. مقوله‌بندی متن[1]. این شیوه، دلالت بر فرایند طبقه‌بندی اطلاعات نوشتاری در قالب تقسیمات موضوعی از پیش تعریف شده دارد. برای نمونه، نظام CONSTRUE در خبرگزاری رویتر قادر است اخبار و گزارش‌ها را به‌طور خودکار در قالب ۶۷۴ تقسیم‌بندی موضوعی سازماندهی کند (هیز[2]، ۱۹۹۲). طرح در حال اجرای "او سی الی سی[3]" نیز تحت عنوان *اسکورپیون* در نظر دارد به رده‌بندی خودکار صفحات یا سایت‌های وب با استفاده از طرح رده‌بندی دهدهی دیویی بپردازد و متن صفحات وب را با رده‌های دیویی مبتنی بر منطق الگوریتمی تطبیق دهد (تامپسون، شفر، ویزین ـ گوتز[4]، ۱۹۹۷). چنین طرح‌هایی با هدف طبقه‌بندی هوشمندانۀ متن و ارائه آن بر اساس نیاز اطلاعاتی کاربران دنبال می‌شوند.

1. Text categorization
2. Hayes
3. OCLC
4. Thompson, Shafer and Vizine-Goetz, 1997

ب. استخراج متن[1]. در این شیوه که گامی فراتر از طبقه‌بندی متن به‌شمار می‌آید، بخشی(هایی) از یک سند حاوی اطلاعات نوشتاری (به‌طور مثال، یک پاراگراف) در مورد یک موضوع مشخص مورد شناسایی قرار می‌گیرد و استخراج می‌شود. یکی از کاربردهای عینی این شیوه را می‌توان در مطالعه‌ای در زمینهٔ استخراج اطلاعات کتاب‌شناختی از فهرست مآخذ متن پروانه‌های ثبت اختراعات مشاهده کرد (لاوسون[2]، و دیگران ۱۹۹۶). نظام نمایه‌سازی استنادی خود محور CiteSeer نیز با هدف استخراج اطلاعات استنادی و فهرست مآخذ مقالات در محیط وب و ارائهٔ نوعی پایگاه نمایه استنادی صورت گرفته است. این برنامه، بر اساس قواعد منطق اکتشافی و الگوریتمی، اطلاعات استنادی را از متن مقالات در قالب PDF یا Postscript استخراج و سازماندهی می‌کند.

ج. برقراری پیوند میان متون[3]. در این شیوه، از تجزیه و تحلیل‌های آماری یا نحوی به‌منظور ایجاد ارتباط مفهومی میان دو بخش از یک متن یا متون مختلف که با یکدیگر مشابهت موضوعی دارند استفاده می‌شود. از این شیوه می‌توان برای ایجاد پیوندهای فرامتنی[4] به‌طور خودکار میان اسناد مختلف استفاده کرد. سالتون و باکلی[5] (۱۹۹۲) در مطالعات خود به این موضوع پرداخته‌اند.

د. نظام توسعه متن[6]. این شیوه را می‌توان نمونهٔ توسعه یافتهٔ نظام ایجاد پیوند میان متن به‌شمار آورد. به‌طوری که از طریق این رویکرد می‌توان ارتباط بخش‌هایی از متن را در متون و اسناد مختلف به صورت ترکیبی نمایش داد. تلاش‌هایی برای ارائهٔ مدلی رایانه‌ای به منظور یکپارچه‌سازی متون مرتبط و هم موضوع از منابع اطلاعاتی نوشتاری مختلف انجام شده است (چن[7]، ۱۹۹۳).

هـ. خلاصه کردن متن[8]. این شیوه مشابه روش‌هایی است که در فرایند چکیده‌نویسی خودکار به‌کار می‌رود و هدف از آن انتخاب جملاتی از متن است که بتواند به‌عنوان خلاصه‌ای گویا و مناسب از محتوای سند عمل کند. در این شیوه، جملات کلیدی از طریق تجزیه و تحلیل‌های آماری یا زبان‌شناختی از متن استخراج می‌شوند. در حال حاضر، از روش‌های پیشرفته‌تری در خلاصه کردن

1. Text extraction
2. Lawson
3. Text linkage
4. Hypertext links
5. Salton and Buckely
6. Text augmentation
7. Chen
8. Text summarization

متن استفاده می‌شود. در شیوه‌های سنتی جملات کلیدی در متن یک سند مورد شناسایی و عیناً استخراج می‌شدند. اما هم اینک نظام‌های خودکار خلاصه‌سازی متن قادرند که چند جمله را در هم ادغام کرده و جملهٔ جدیدی را با بار اطلاعاتی مناسب‌تر به‌وجود آورند.

و. تولید متن[1]. تولید متن، به برنامه‌ها و ابزارهایی اطلاق می‌شود که از طریق آن می‌توان به‌طور خودکار به تولید متون پرداخت. به‌طور مثال، این برنامه‌ها می‌توانند به افراد برای تهیهٔ انواع گزارش‌ها کمک نمایند و قادرند ضمن بررسی املای صحیح کلمات، ساختار نحوی و دستوری جمله را نیز مورد بررسی قرار دهند و در صورت وجود اشتباه نمونهٔ صحیح یک عبارت را پیشنهاد کنند. با این وجود، نمی‌توان به چنین نظام‌هایی لفظ هوشمند منتسب کرد و آنها اغلب تنها در سطح نظام‌های خبره توسعه یافته‌اند. برنامهٔ واژه‌پرداز MS Word XP در مقیاس کوچکی قادر است برخی از عملکردهای تولید متن را از طریق پیشنهاد املای صحیح کلمات و نیز ساختار دستوری یک جمله یا عبارت به اجرا در آورد.

ز. ترجمهٔ ماشینی[2]. یکی از پیشرفته‌ترین کاربردهای نظام‌های پردازش متن که از اوایل دههٔ ۱۹۵۰ به آن توجه شده ترجمهٔ ماشینی است. با وجود پیشرفت‌های بسیاری که در این حوزه صورت گرفته است، طراحی نوعی نظام ترجمهٔ تمام خودکار که قادر باشد کلیهٔ متون در موضوعات مختلف را بدون نیاز به انسان ترجمه کند تقریباً ناممکن است. در حال حاضر، برخی از نظام‌های ترجمه توانسته‌اند با ۹۰ درصد دقت و صحت در امر ترجمه به اجرا درآیند. یکی از قدیمی‌ترین و موفق‌ترین نظام‌های ترجمهٔ متن، سیستران[3] نام دارد که نمونهٔ تجاری و ناقص آن بر روی برخی موتورهای کاوش وب نظیر AltaVista (www.av.com) نیز قابل دسترس است؛ و قادر است یک متن (یا صفحات وب) را از انگلیسی به چند زبان دیگر و بالعکس ترجمه کند (سیستران، ۲۰۰۳). طبق بررسی‌های به‌عمل آمده، هنگامی که متنی حاوی ۲۰۰ هزار صفحه متن از روسی به انگلیسی توسط این برنامه ترجمه می‌شود، فقط ۵ درصد خطا را می‌توان در ترجمه متن اصلی مشاهده کرد که میزان بالایی از دقت در کار ترجمه را نشان می‌دهد (هاچینز و سومرز[4]، ۱۹۹۲).

1. Text generation
2. Machine translation
3. Systrann
4. Hutchins and Somers

کاربردهای دیگر در بازیابی اطلاعات

الف. عامل‌های هوشمند[1]. عامل‌های هوشمند نرم‌افزارهایی هستند که بدون دخالت انسان قادر به بازیابی داده‌ها به صورت هوشمند از چند منبع یا پایگاه اطلاعاتی هستند. لنکستر و وارنر اصطلاح عامل‌های هوشمند را به عنوان ابزارهایی که قادر به انجام وظایف مختلفی در زمینه بازیابی اطلاعات در اینترنت و دیگر پایگاه‌های اطلاعاتی هستند تعریف کرده‌اند (لنکستر و وارنر، ۲۰۰۱، ص ۵۹). برخی موتورهای کاوش نیز مدعی‌اند که در رتبه‌بندی نتایج کاوش خود از عامل‌های هوشمند استفاده می‌کنند.

ب. واسط‌های هوشمند[2]. طراحی واسط‌های هوشمند در نظام اطلاعاتی این امکان را به‌وجود می‌آورد که به‌وسیلهٔ طیف گسترده‌تری از کاربران مورد استفاده قرار گیرد. چنین نظام‌هایی اغلب برای سهولت بازیابی اطلاعات طراحی می‌شوند و قادرند خود را مطابق با مهارت و دانش جست‌وجوی استفاده‌کنندگان تطبیق دهند. به عبارت دیگر، کاربران نیاز کمتری به داشتن مهارت و دانش تخصصی در زمینهٔ جست‌وجوی پیشرفته و بازیابی اطلاعات مرتبط دارند. امکان برقراری ارتباط با کاربر از طریق واسط جست‌وجوی ساده، انعطاف‌پذیری برای تطبیق با مهارت‌های جست‌وجوی اطلاعات کاربران، امکان تغییر واسط جست‌وجو بنا بر نیاز کاربران، کمک به کاربر در انتخاب پایگاه اطلاعاتی، و تسهیل فرایند سفارش رکوردهای بازیابی شده از مهم‌ترین ویژگی‌های واسط‌های هوشمند به‌شمار می‌آیند.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

نتیجه‌گیری

اگرچه تلاش‌های بسیار زیادی در ارتباط با استفاده از کاربردهای نظام‌های خبره و هوش مصنوعی در بازیابی اطلاعات شده است، بررسی بحث‌های مختلف در متون مرتبط نشان می‌دهد که حجم بسیار زیادی از این تحقیقات به مطالعات نظری یا طراحی نظام‌های خبره به صورت تجربی و پیش‌نمون منجر شده است. لنکستر و وارنر، یکی از مهم‌ترین دلایل عدم توسعهٔ نظام‌های خبره و هوش مصنوعی در فعالیت‌های کتابداری و اطلاع‌رسانی را، که بازیابی اطلاعات نیز شامل آن می‌شود، پرهزینه بودن طراحی این نظام و نیز پیچیدگی آنها دانسته‌اند. به نظر آنها تصور اینکه طراحی نرم‌افزارهای خبره و هوشمند برای انجام برخی فعالیت‌های کتابداری و اطلاع‌رسانی هزینهٔ اندکی دربر دارد،

1. Intelligent agents
2. Intelligent interfaces

بسیاری از مدیران کتابخانه‌ها و سایر محققان را گمراه کرده است. برخلاف تصور موجود، برای ایجاد یک پایگاه دانش مرتبط با فعالیت‌های کتابداری و اطلاع‌رسانی باید هزینه و زمان نسبتاً زیادی صرف کرد. به علاوه، روزآمدسازی مستمر این‌گونه نظام‌ها خود بسیار هزینه بر هستند. این دو محقق در کتاب خود بر این نکته تأکید دارند که نظام‌های خبره و هوشمندی که تاکنون در زمینهٔ فعالیت‌های کتابداری توسعه یافته‌اند، در مقایسه با نظام‌های توسعه یافته در سایر حوزه‌ها، "اسباب‌بازیی" بیش نیستند. به عبارت دیگر، برای طراحی نظام‌های خبره و هوشمند در فعالیت‌های کتابداری باید هزینه و زمان زیادی را صرف کرد و این امکان یا در بسیاری از طرح‌های پژوهشی مهیّا نبوده است و یا اینکه پس از آغاز با مشکلاتی مواجه شده است (لنکستر و وارنر، ۲۰۰۱، ص ۱۰۷).

نکتهٔ بسیار مهم دیگری که چه‌بسا کاربرد نظام‌های خبره و هوشمند را در بازیابی اطلاعات تحت تأثیر قرار داده است مربوط به ماهیت این‌گونه نظام‌هاست. تشخیص ربط خود مسئله‌ای است که هنوز بحث‌ها و نقطه‌نظرات مختلفی دربارهٔ آن، حتی در نظام‌های سنتی، وجود دارد. این موضوع مورد قبول همگان است که کلیهٔ تحقیقات انجام شده و در دست انجام در زمینهٔ طراحی و توسعهٔ نظام‌های بازیابی اطلاعات، با هدف تسهیل و بهبود فرایند دسترسی به مدارک مرتبط منطبق با نیاز اطلاعاتی کاربر است. حتی اگر یک نظام هوشمند یا خبره در بهترین شرایط خود بتواند مدارک مرتبط را بنابر منطق‌های الگوریتمی و اکتشافی یا ترکیبی از هر دوی آنها بازیابی کند، لزوماً به معنای آن نخواهد بود که با نیاز اطلاعاتی کاربر منطبق است. ذهن‌گرایانه بودن مفهوم ربط و نیاز اطلاعاتی باعث شده است که با وجود تحقیقات گسترده در این زمینه هنوز مشکل رفع نیازهای واقعی کاربران به عنوان یکی از مسائل مهم حوزهٔ کتابداری و اطلاع‌رسانی باقی مانده باشد. طرح چنین مسائلی می‌تواند نقش نظام‌های خبره و هوشمند را به عنوان واسطی کارآمد میان نیاز اطلاعاتی کاربر و مدارک مرتبط در یک پایگاه اطلاعاتی با ابهام‌هایی روبه‌رو سازد و بر نقش آنها به عنوان "ابزاری کمکی" برای بازیابی اطلاعات تأکید بیشتری نماید. شاید کاربران و نیازهای اطلاعاتی واقعی آنها تنها عنصر کلیدی در طراحی نظام‌های بازیابی خبره و هوشمند به شمار آیند که باید دربارهٔ جنبه‌های مختلف آن مطالعات بیشتری صورت گیرد. شاید به همین دلیل است که مهم‌ترین اولویت پژوهشی مرکز بازیابی هوشمند اطلاعات بر این موضوع، یعنی

تحقیق در زمینهٔ "دریافت بازخورد نسبت به چگونگی ربط"، متمرکز شده است.

مآخذ

نشاط، نرگس (۱۳۸۲). "هرمنوتیک و بازیابی اطلاعات". اطلاع‌شناسی، س. اول، ۲، زمستان، ص ۳۳-۴۶.

Cawsey, A. (1998). *The essence of artificial intelligence*. London: Prentice Hall.

Chen, Z. (1993). "Let documents talk to each other: a computer model for connection of short documents. *Journal of Documentation*, 49, pp. 44-54.

CIIR (Center for Intelligent Information Retrieval). (2004). [online]. Available a: http://ciir.cs.umass.edu/ (8 Jan. 2004)

Fenly, C. (1992). "Technical services processes as models for assessing expert system suitability and benefits". in: *Artificial Intelligence and Expert Systems: Will they Change the Library?*; ed by F.W. Lancaster and L.C. Smith, pp. 50-66. Urbana-Champaign: University of Illinois, Graduate School of Library and Information Science.

Ford, N (1991). *Expert sysrtems and artificial intelligence: an information manager's guide*. London: Library Association Publishing.

Frakes, W. B. and Baeza-Yates, R (1992). Information retrieval: data structures & algorithms. Englewood Cliffs, NJ, Prentice Hall.

Hayes, P.J. (1992). "Intelligent high-volume text processing using shallow, domain-specific techniques". in: *Text-based intelligent systems*, ed. By P. S. Jacobs, pp. 227-241. Hillsdale, NJ: Lawrence Erlbaum.

Hutchins, W. J. and Somers, H. L. (1992). *An introduction to machine translation*. New York: Academic Press.

INQUERY Search Software (2004). [Online]. Available: http://ciir.cs.umass.edu/inquerypage.html (9 Jan. 2004)

Korfhage, R. R. (1997). "Information storage and retrieval". New York :Wiley.

Lancaster, F.W. and Warner, A. (2001). *Intelligent technologies in library and information service application.* New Jersey: Information Today.

Lawson, M. et al. (1996). "Automatic extraction of citation from the text of english-language patents- An example of template mining". *Journal of Information Science*, 22, pp. 423-436.

Pollit, A.S. (1993). *MenUSE for medicine: End user browsing and searching of the MEDLINE Database using the MeSH Thesaurus. RIAO 88: user oriented content-based text and image handing*, Vol. 1, Cambridge, MA, pp. 547-573.

Pors, N. O. (2000). "Information retrieval, experienmental models and statistical analysis". *Journal of Documentation*, 56(1): 55-70.

Salton G. and Buckely, C. (1992). "Automatic text structuring experiments". in: *Text-based intelligent systems*, ed. By P. S. Jacobs, PP. 199-210. Hillsdale, NJ: Lawrence Erlbaum.

Salton, G. and McGill, M. J. (1983). *The introduction to modern information retrieval.* New York: McGraw-Hill.

Spangler, W.E. and May, J.H. (2000). "Expert and knowledge-based systems". *Encyclopedia of Library and Information Science*. Edited by Allen Kent, Vol. 67.

Systran. [online] Available: http://www.systransoft.com. [Accesse 8 Sep. 2003].

Thompson, R.; Shafer, K. and Vizine-Goetz, D. (2003). "Evaluating Dewey concepts as a knowledge base for automatic subject assignment". [online]. Available: http://orc.rsch.oclc.org:6109/eval_dc.html. [13 Sep. 2003].