اینفوکریستال‌ها و کاربرد آنها در بازیابی اطلاعات

محمدرضا فرهادپور | داریوش مطلبی

چکیده

هدف: بررسی کاربرد مصوسازی اطلاعات، به‌ویژه فن اینفوکریستال در فرآیند بازیابی اطلاعات، مصورسازی پرس‌وجوهای وزنی و فضاهای برداری است. مصورسازی اطلاعات به‌صورت عام و اینفوکریستال به‌صورت خاص و کاربردهای آنها در بازیابی اطلاعات، اینفوکریستال و مصوسازی پرس‌وجوهای وزنی، مصورسازی فضاهای برداری و تاثیر این روش در فهم راحت روابط موجود بین مفاهیم بسیار زیاد بازیابی شده، مورد بررسی قرار می‌گیرد.

روش/رویکرد پژوهش: این مطالعه با بررسی و تحلیل متون بازیابی اطلاعات و مصورسازی اطلاعات، تاثیر مصورسازی به‌ویژه فن اینفوکریستال در تحلیل نتایج جست‌وجو صورت گرفته است.

یافته‌ها: مصوسازی اطلاعات از توانایی دیداری انسان در تشخیص الگوها و روابط ساختاری اشیاء در بهبود نظام‌های بازیابی استفاده می‌کند و در علوم مختلفی از جمله زمین‌شناسی، جغرافیا و اخیراً کتابداری و اطلاع‌رسانی و ... کاربرد دارد. یکی از فنون مصوسازی اطلاعات، اینفوکریستال می‌باشد که می‌تواند فضای انتزاعی اطلاعات را به تصویر بکشد و تمامی روابط ممکن بین بی‌نهایت مفهوم را مجسم سازد.

نتیجه‌گیری: با توجه به اهمیت روزافزون نظام‌های بازیابی اطلاعات و کمبود منابع منتشر شده در حوزه مصورسازی، آشنایی با اینفوکریستال و کاربردهای آن در طراحی نظام‌های بازیابی اطلاعات، می‌تواند برای جامعه کتابداری و اطلاع‌رسانی ایران بسیار مفید باشد.

کلیدواژه‌ها

مصورسازی اطلاعات، بازیابی اطلاعات، اینفوکریستال

اینفوکریستال‌ها و کاربرد آنها در بازیابی اطلاعات

محمدرضا فرهادپور[1] | داریوش مطلبی[2]

دریافت: ۱۳۸۸/۹/۱ پذیرش: ۱۳۹۰/۲/۴

مقدمه

بازیابی اطلاعات یکی از بدیهی‌ترین فرآیندها در چرخه زیستی اطلاعات است. ازاین‌رو، بازیابی کارآمد اطلاعات، حتی فرآیندهای پیش از خود، نظیر سازماندهی و ذخیره‌سازی اطلاعات را تحت تأثیر قرار می‌دهد. به‌عبارت ساده‌تر، اطلاعات، هرچند دارای بار ارزشی بالایی است، در صورت عدم بازیابی، چیزی جز کالای منسوخ شده نخواهد بود. رشد فزاینده اطلاعات از یک‌طرف و ارزش تصمیم‌گیری مبتنی بر اطلاعات – جهت استفاده مناسب از فرصت‌ها در محیط رقابتی پیچیده– از طرف دیگر، بازیابی اطلاعات را به یکی از مهم‌ترین دغدغه‌های امروز حوزه اطلاع‌رسانی تبدیل کرده است. اگرچه توسعه ابزارهای ارتباطی از دغدغه انسان به لحاظ نبود یا فقدان اطلاعات تا حدودی کاسته است؛ مشکلات مربوط به بازیابی اطلاعات مرتبط، دسترسی به آن، ارزیابی ربط آن و مسائلی از این دست بیانگر وجود شکافی بین کارآیی نظام‌های بازیابی و انتظارات کاربران است. امروزه، پیچیدگی محیط اطلاعاتی افراد و سازمان‌ها، نیاز به داشتن اطلاعات درست برای تصمیم‌گیری مناسب را اجتناب‌ناپذیر کرده است. بررسی مطالعات مرتبط با حوزه گردآوری و استفاده از اطلاعات و یا تحلیل و پویش محیطی، نشان‌دهنده چالش‌های فراروی این حوزه است. از جمله این چالش‌ها توجه غالب مدیران به دسترس‌پذیری منابع اطلاعاتی و کم‌توجهی به کیفیت آنهاست. یکی از سؤال‌هایی که با مشاهده یافته‌های مطالعاتی مرتبط به ذهن می‌رسد، این

۱. دکتری کتابداری و اطلاع‌رسانی دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران
m.farhadpoor@gmail.com
۲. دانشجوی دکتری کتابداری و اطلاع‌رسانی دانشگاه آزاد اسلامی، واحد علوم و تحقیقات تهران
dariushmatlabi@yahoo.com

است که ناکارآمدی نظام‌های بازیابی یا نارضایتی از نتایج بازیابی چقدر مهم است؟

مطالعات مانینگ و همکارانش نشان داد که مردم ترجیح می‌دهند اطلاعات را، به‌جای نظام‌های بازیابی اطلاعات، از افراد دیگر دریافت کنند. اگر این مسئله را با چشم‌پوشی از تأکید مردم بر استفاده از عامل انسانی در فعالیت‌هایشان طی سال‌های اواخر دهه۹۰ مرور کنیم، ضرورت توجه به مسئله بازیابی روشن‌تر می‌شود (مانینگ، راگاوانو شوتز[3]، ۲۰۰۸، ص ۱۲).

رشد اطلاعات و همراهی فناوری‌های اطلاعاتی نوین، توانایی ذخیره و انتشار حجم عظیمی از مدارک اطلاعاتی را برای بشر فراهم کرد و به‌دنبال آن، امکان دسترسی بشر به اطلاعات نیز بیشتر شد؛ تا جایی که پدیده موسوم به انفجار اطلاعات از سال‌ها پیش به دغدغه فکری متخصصان حوزه‌های اطلاعات و ارتباطات تبدیل شد. همان‌گونه که رشد انفجارگونه اطلاعات، دسترسی گسترده بشر به اطلاعات را فراهم ساخت؛ بازیابی اطلاعات مرتبط به دغدغه‌ای مهم تبدیل شد و ناتوانی بشر در بازیابی، تفسیر، و تحلیل بهینه اطلاعات – به‌نحوی که کاربر بتواند به‌سرعت اطلاعات مرتبط و مشتمل بر نیازش را بیابد– به کانون پرچالشی بدل شد. رشد غیرقابل کنترل اطلاعات، وقتی که با فناوری وب پیوند خورد، گستره دسترسی بشر را وسعت بخشید و با دسترسی الکترونیکی به حجم عظیمی از اطلاعات بر اهمیت تمرکز بر حوزه بازیابی اطلاعات افزود. تحولی که فناوری وب درزمینه رشد کمّی اطلاعات و دسترسی به آن ایجاد کرد، این شبهه را به‌وجود آورد که بازیابی اطلاعات آغاز شده است. درحالی که، بازیابی اطلاعات در پاسخ به چالش‌های روزافزون دسترسی به اطلاعات در محیط‌های غیرالکترونیکی آغاز شد و همگام با رشد محتوا، به اشکال مختلف، به‌منظور ارائه دیدگاه اصولی برای جست‌وجوی محتوای آنها، رشد کرد. این حوزه، در ابتدا، برای انتشارات علمی و پیشینه‌های کتابخانه‌ای مطرح بود؛ ولی خیلی زود در سایر حوزه‌های موضوعی نظیر روزنامه‌نگاری، حقوق، و پزشکی نیز گسترش یافت؛ به‌طوری‌که بسیاری از پژوهش‌های بازیابی اطلاعات در این حوزه‌ها انجام گرفت و بسیاری از فعالیت‌های مستمر بازیابی اطلاعات به ایجاد دسترسی به اطلاعات فاقد ساختار[4] در شرکت‌های مختلف و حوزه‌های دولتی منتهی شد.

هدف اصلی تبدیل بازیابی اطلاعات به یکی از نقاط کانونی پژوهش‌های اطلاع‌رسانی را باید در تلاش برای مطالعه و درک فرآیندهای بازیابی اطلاعات با هدف طراحی، ساخت، و سنجش سیستم‌های بازیابی دانست تا بتواند بین اطلاعات تولید شده از سوی عوامل انسانی و نیاز کاربران ارتباط برقرار کند (اینگورسن، ۲۰۰۲، ص ٤۹).

مفهوم اصلی نظام‌های بازیابی اطلاعات در ارائه مدارک مرتبط با نیازهای کاربران نهفته است (بورلاند، ۲۰۰۳) و این نقطه را می‌توان چالش پژوهش‌های حوزه بازیابی اطلاعات

3. Manning , Raghavan & Schutze
4. Unstructured information

دانست، چرا که با گذشت زمان، پژوهش‌ها از تمرکز صرف بر ساختار نظام‌های بازیابی اطلاعات فراتر رفته و به گستره‌ای وسیع‌تر، از قبیل طراحی رابط کاربر، رفتار، و معیارهای داوری رابط کاربر و نظایر آن راه یافته است. درونمایه همه این تحقیقات حول محور کاربر پسندی بیشتر نظام‌های بازیابی اطلاعات و رفع مشکلات پیشِ رو و کنترل هرچه بهتر اطلاعات تولید شده است.

درحالی‌که، توانایی بشر، برای خواندن نتایج جست‌وجو در وب یا پایگاه‌های اطلاعاتی کتابشناختی همچنان سیر آرامی را سپری می‌کند؛ نظام‌های بازیابی اطلاعات قادر به بازیابی صدها هزار مدرک با سرعت و دقت بالایی هستند. مسئله نگران‌کننده اینجاست که ما تنها می‌توانیم تعداد محدودی از نتایج را مطالعه کنیم. در کنار پژوهش‌های مرتبط با بازیابی اطلاعات، ظهور فنون جدید مصورسازی و نظام‌های مبتنی بر آن، از قبیل ترسیم مقیاس جزء[5]، کوگار[6]، گوئیدو[7]، وب وایب[8]، اینفوکریستال، و نظیر آن نویدبخش چیرگی بر برخی مشکلات حوزه بازیابی اطلاعات است. در این شیوه‌ها، با هدایت محتوا و ربط آن به سوی فضاهای چندبُعدی، کاربر قادر به مرور مجموعه‌ای از مدارک و بازیابی منابع مورد نیاز خود می‌شود (مورس[9] و همکاران، ۲۰۰۲، ص. ۲). اینفوکریستال یکی از روش‌هایی است که کمک می‌کند روابط بین عنصرهای مختلف مورد جست‌وجو، در نظام‌های بازیابی اطلاعات، به راحت‌ترین شکل ممکن به‌تصویر کشیده شود. در این مقاله تلاش می‌شود ابعاد مختلف این روش مورد بررسی قرار گیرد. در ابتدای بحث برای آشنایی بیشتر، بازیابی اطلاعات و انواع آن را مورد بررسی قرار می‌دهیم.

بازیابی اطلاعات

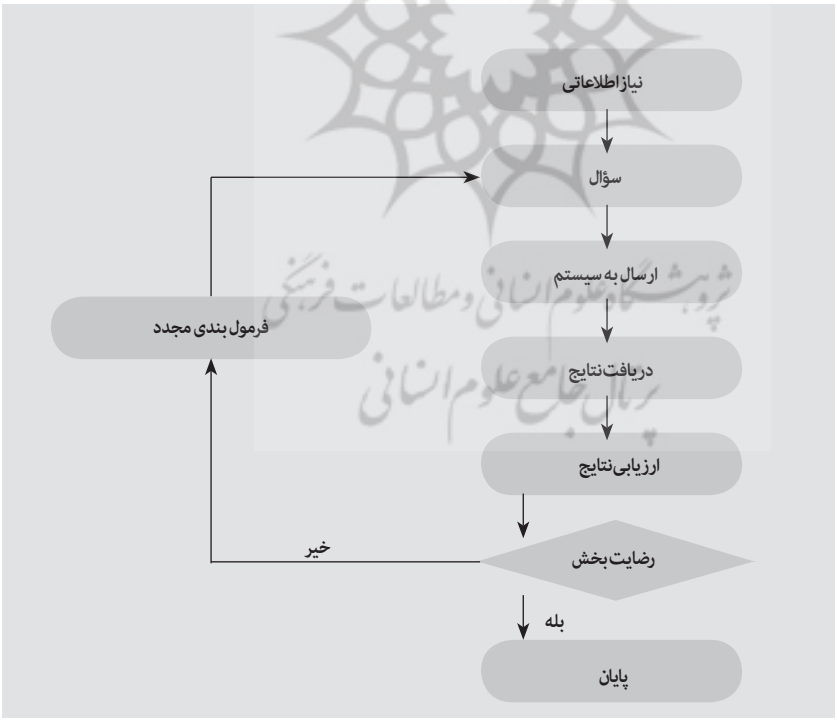
برای بازیابی اطلاعات تعاریف مختلفی ارائه شده است. والاگ[10] (۲۰۰٤، ص ۱۱)، بازیابی اطلاعات را فرآیند جایابی مدارک مرتبط با برخی معیارهای خاص کاربر، از مجموعه مدارک می‌داند. مانینگ و همکارانش بازیابی اطلاعات را یافتن مواد (معمولاً مدارک) از یک ماهیت فاقد ساختار (معمولاً متن) براساس نیاز اطلاعاتی و از درون یک مجموعه بزرگ (معمولاً ذخیره شده در رایانه) عنوان کردند (مانینگ، راگاوان و شوتز، ۲۰۰۸، ص ۱). با فراهم‌شدن امکان دسترسی به اینترنت و جست‌وجو در وب جهان‌گستر و رشد سریع پایگاه‌های اطلاعاتی، روزانه میلیون‌ها نفر با مسئله بازیابی اطلاعات و مسائل مرتبط با آن سرو کار دارند. در نتیجه، محیط متغیر وب تعریف بازیابی و مسائل مرتبط با آن را تحت تأثیر قرار می‌دهد.

درونمایه اکثر تعاریف مرتبط با بازیابی اطلاعات، بازیابی مدارک مرتبط براساس نیازهای کاربر و از درون یک ماهیت نظیر کتابخانه دیجیتال، اینترنت، و یاپایگاه‌های اطلاعاتی است.

5. Component scale drawing
6. Cougar
7. GUIDO
8. WebVIBE
9. Morse
10. Valvag

بنابراین، بازیابی اطلاعات فرآیندی است که در جهت دسترسی به اطلاعات و به‌منظور رفع نیازهای اطلاعاتی شکل می‌گیرد. از دیدگاه مارتی هرست[11]، فردی که در فرآیند اطلاع‌جویی قرار می‌گیرد، چند هدف را در ذهن خود دنبال می‌کند و از یک نظام جست‌وجو به‌عنوان ابزاری برای رسیدن به اهدافش استفاده می‌کند و مأموریت‌های دسترسی به اطلاعات برای رسیدن به این اهداف صورت می‌پذیرد. اودای[12]، بازنگری یک موضوع مشخص در طول زمان (مانند بازنگری فصلی فعالیت‌های پژوهشی رقبا)، پیگیری یک برنامه یا عادت جست‌وجو برای رسیدن به اهدافی خاص (روزآمد نگهداشتن خود درباره کسب و کار خوب)، و تشریح یک موضوع به سبکی نامشخص (تلاش برای شناخت یک حرفه ناآشنا) را سه نوع عمده مأموریت دسترسی به اطلاعات می‌داند (بایزا- یتس و ریبیرو- نیتو[13]، ۱۹۹۹).

فرآیندهای فراهم‌آوری، سازماندهی، و اشاعه اطلاعات برای پاسخگویی به نیازهای اطلاعات مردم و متخصصان شکل می‌گیرد و هر مراجعه‌ای ماحصل نیاز اطلاعاتی است که شخص را به طرح سؤال و استفاده از یک نظام اطلاعاتی ترغیب می‌کند. دسترسی به اطلاعات فرآیندی است که به شکل‌های مختلف مطرح شده است. مارتی هرست، فرآیند دسترسی به اطلاعات را در قالب شکل ۱ به شرح زیر ارائه می‌کند.



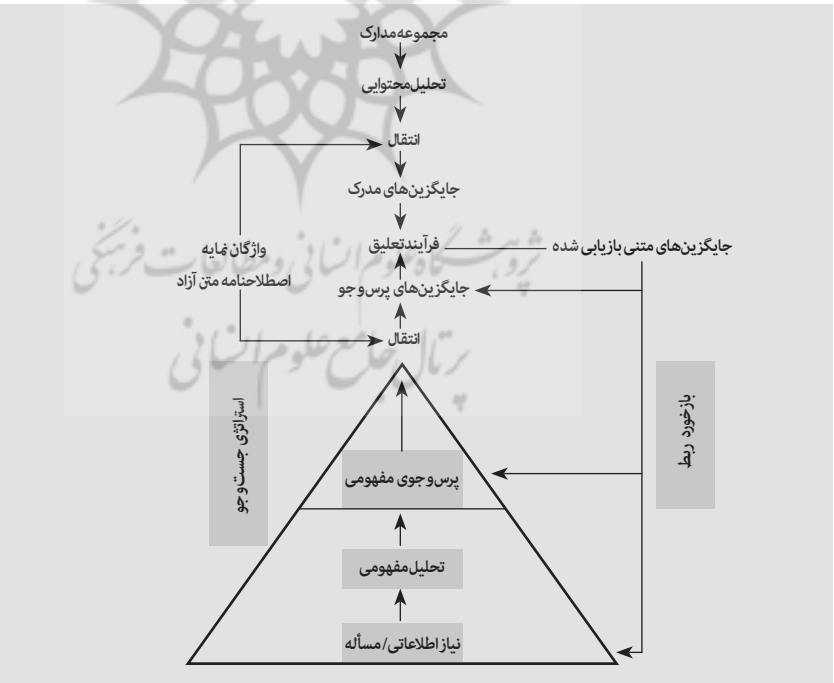
شکل ۱

نمودار ساده از مدل استاندارد فرآیندهای دسترسی به اطلاعات

11. Marti Hearst
12. Oday
13. Baeza-Yates & Ribeiro-Neto

شکل ۱ نشان می‌دهد که فرآیند جست‌وجو، از سوی کاربر، یک محدوده یادگیری تعاملی را برای او فراهم می‌کند. به‌عبارتی، کاربر اطلاعات را بررسی می‌کند، عناوین را در مجموعه نتایج می‌خواند، مدارک بازیابی شده را شخصاً فراخوانی می‌کند، فهرست موضوعات مرتبط با اصطلاحات پرس‌وجو را مرور می‌کند و فراپیوندها (در محیط وب) را دنبال می‌کند. این کارها در امتداد کارکرد یک نظام بازیابی صورت می‌گیرد؛ به‌طوری‌که، کاربر با ایجاد تعامل نزدیک با نظام، از بازخودهای دریافتی، برای آرایش مجدد جست‌وجو استفاده می‌کند. درنتیجه، نظام بازیابی اطلاعات به‌منظور پردازش درخواست‌های اطلاعاتی کاربران به‌کار گرفته می‌شود و انتظار می‌رود به درخواست‌های کاربران در قالب مجموعه‌ای از مدارک بازیابی شده پاسخ دهد.

بنابراین، در هر نظامی مجموعه‌ای از عوامل همچون درخواست‌های اطلاعاتی، پرس‌وجوها، و مدارک و فرآیندهایی نظیر تحلیل مفهومی، انتقال و تطبیق دخیل هستند که شامل روش‌هایی برای هدایت درخواست‌ها به مدارک می‌باشند. تعامل این مؤلفه‌ها را می‌توان در الگوی اصلاح شده لنکستر و وارنر (۱۹۹۳) در شکل ۲ مشاهده کرد که در آن نیازهای اطلاعاتی و مجموعه مدارک باید به شکل جایگزین‌هایی برگردانده شوند تا فرآیند تطبیق شکل گیرد.



شکل ۲

مدل اصلاح شده فرآیند بازیابی اطلاعات لنکستر و وارنر (۱۹۹۳)

برای بازیابی اطلاعات نیز مدل‌های متفاوتی ارائه شده‌اند که با توجه به شباهت‌ها و راهبردهای به‌کار گرفته‌شده، دانشمندان آنها را در قالب چند دسته مطرح کرده‌اند. از جمله بایزا- یتس و ریبرو- نیتو مدل‌های بازیابی اطلاعات را در سه دسته زیر خلاصه می‌کند (والاگ، ۲۰۰٤، ص، ۱۱).

الف) مدل‌های بولی: این مدل مدارک و پرس‌وجوها را در قالب مجموعه‌ای از اصطلاحات نمایه‌ای طرح‌ریزی می‌کند.

ب) مدل‌های برداری: این مدل، از بردارها و فضاهای دوبُعدی برای طرح‌ریزی مدارک و پرس‌وجوها استفاده می‌کند.

ج) مدل‌های احتمالی: از چارچوبی مبتنی بر نظریه احتمال استفاده می‌کنند.

اسپوئری[14] (۱۹۹۵)، دسته‌بندی دیگری به‌شرح زیر ارائه می‌کند:

الف) مدل بولی که اغلب از آن به‌عنوان مدل تطبیق عینی نام برده می‌شود،

ب) مدل آماری که مدل فضای برداری و بازیابی احتمالی را دربرمی‌گیرد، و

ج) مدل‌های زبانی و مبتنی بر دانش.

ارائه مدل‌های متفاوت، وجود نوعی احساس کمبود و نارضایتی از مدل‌های قبلی را مشهود می‌سازد و بازنمونی از توجه بیشتر به تکامل و کاربرپسندی است. اسپوئری (۱۹۹۵)، قلمرو بازیابی اطلاعات را درگیر سه چالش عمده می‌داند. نخست اینکه، امروز دیدگاه غالب درباره مدل بولی یا تطبیق عینی[15] این است که این مدل بیش از پیش نیازمند کاربرپسندی است؛ زیرا کاربران عادی استفاده از عملگرهای بولی و به‌کارگیری پرانتزها را در فرمول‌چینی پرس‌وجوهای بولی مؤثر، دشوار می‌دانند. همچنین، تعداد اندکی از نحوه استفاده کامل از قابلیت واقعی زبان پرس‌وجوی بولی، کافی دارند. دوم اینکه، دیدگاه‌های تطبیق جزئی[16] که بدواً برای استفاده آسان‌تر هستند، فهرست پی‌درپی از بهترین مدارک را به کاربر ارائه می‌کنند. این مسئله می‌تواند یک تأثیر بینایی کانونی[17] ایجاد کند، به این خاطر که فهرست رتبه‌بندی‌شده، نقش اصطلاح پرس‌وجو در رتبه‌بندی مدارک بازیابی‌شده را پنهان می‌کند و کاربران می‌توانند از این نوع بازخورد در چگونگی پیشبرد جست‌وجوهایشان استفاده کنند. سوم اینکه، آزمون‌های اخیر بازیابی نشان داده‌اند که مسئله دیدگاه‌های تطبیق عینی و جزیی به‌خاطر کمی همپوشانی مجموعه مدارک بازیابی‌شده مرتبط، مکمل یکدیگر، از این‌رو، یک اجماع کلی وجود دارد که ترکیب این دو دیدگاه برای افزایش اثربخشی بازیابی، لازم است.

مسائل مذکور از یک‌طرف و فقدان بازخورد بصری سبب احساس سردرگمی کاربر در هنگام جست‌وجو می‌شود، و در آغاز اعتماد و اثربخشی آنها را تحلیل می‌برد. برای رفع مشکلات فوق شیوه‌های مصورسازی اطلاعات با ارائه دیدگاه جامع از فضای اطلاعات

14. Spoerri
15. Exact matching
16. Partial matching
17. Tunnel vision

برای کاربران از فرآیند جست‌وجو پشتیبانی می‌کند. در ادامه، به‌تفصیل درباره مصورسازی اطلاعات صحبت خواهد شد.

مصورسازی اطلاعات

استفاده انسان از نمادهای تصویری برای انتقال مفاهیم و اندیشه‌ها، تاریخی بس طولانی دارد که علاوه‌بر تسهیل ارتباط، به ماندگاری و ثبت اندیشه بشری در قالب خطوط تصویرنگاری و اندیشه‌نگار کمک شایانی کرده است. بهره‌گیری از رویکرد دیداری برای تبیین اندیشه‌ها و احساسات انسان تحول قابل تأملی است. توانایی بشر در دریافت اطلاعات از دیده‌های پیرامونی با استفاده از حس بصری، منحصر به نمونه‌های ابتدایی خطوط تصویری نیست، بلکه تأکید انسان عصر حاضر در فراهم‌آوری اطلاعات نیز برمبنای این حس است و زمینه‌های پژوهشی از قبیل زبان تصویر[18]، سواد دیداری[19]، نمادشناسی و نظایر آن هریک از زاویه دید خاصی درصدد مطالعه و شناخت بیشتر این توانایی هستند.

درودی به نقل از مقاله «سواد دیداری در آموزش علم و فناوری» از خبرنامه بین‌المللی آموزش محیطی، فناوری، و علمی یونسکو زبان تصویر را شیوه‌هایی برای درک تصاویر در زندگی روزانه می‌داند که معمولاً فهم آسان‌تری داشته و نسبت به زبان شفاهی یا ریاضی فراگیرتر است. ینواین[20]، سواد دیداری را توانایی درک معنا در تصاویر می‌داند، و جیورجیس[21] و همکارانش سواد دیداری را توانایی برداشت ساختاریافته از تصاویر دیداری تعریف می‌کند (درودی، ۱۳۸۸، ص ۲۷۵ – ۲۷۹). از این‌رو، زبان تصویر و سواد دیداری را می‌توان از بنیان‌های پیشرو در حوزه مصورسازی اطلاعات برشمرد.

با گسترش شیوه‌های نوین بازیابی اطلاعات، دستیابی به حجم وسیعی از اطلاعات امکان‌پذیر شده است، ولی ریزش کاذب همچنان به‌عنوان یکی از دغدغه‌های این حوزه باقی است. ریزش کاذب به بازیابی نامربوط مدرک اشاره دارد وهرچه ریزش کاذب بیشتر شود نظام اطلاع‌رسانی ناکارآمدتر خواهد بود (پائو، ۱۳۷۹، ص ۱۰۸). از این رو، نیاز به روش‌هایی که توانایی جلوگیری از ریزش کاذب اطلاعات را داشته باشند، بیش از پیش احساس می‌شود. فنون مصورسازی اطلاعات می‌تواند روش‌های بهتری را برای دسترسی و درک فضاهای اطلاعاتی گسترده ارائه کند. به اعتقاد درودی (۱۳۸۸، ص۲۷٤)، معنا و مفهوم بالایی که در بطن تصاویر نهفته است می‌تواند به‌عنوان راهبردی مؤثر در درک بهینه اطلاعات، مورد استفاده قرار گیرد. به‌طوری‌که محققان زیراکس بر این عقیده‌اند که رابط‌های بصری که اطلاعات را به‌طور فزاینده‌ای در بازنمون‌های خلاصه و ساده‌تر کدگذاری می‌کنند، نقش محوری در مدیریت مؤثر فضاهای اطلاعاتی گسترده ایفا خواهند کرد (کارد و دیگران، ۱۹۹۱).

18. Language of vision
19. Visual literacy
20. Yenawine
21. Giorgis

مصورسازی اطلاعات، شیوه نمایش و ارائه دیداری اطلاعات به‌منظور درک بهتر و بهره‌گیری مناسب از اطلاعات است (درودی، ۱۳۸۸، ص۲۷٦)؛ و یا ابزاری است برای هدایت داده‌های پیچیده، به‌نحوی که کاربران به تفسیر بیشتر آن قادر باشند تا بارشناختی خود را با معنا بخشیدن به نمادها، اشکال، آرایش، و تشابهات بصری اعمال کنند (کوشمن، ۲۰۰۵، ص ۸۲٤).

با توجه به توانایی بالای انسان در اقتباس اطلاعات از محرک‌های بصری، و به روش‌های خاصی به الگوها و روابط بین آنها پی می‌برد، رؤیت رایانه‌ای[۲۲] در تلاش است چگونگی پردازش اطلاعات از سوی سیستم بصری انسان را تشخیص دهد و الزاماتی فراهم می‌کند تا درک سه‌بُعدی از ماهیت دوبُعدی را امکان‌پذیر سازد. اینفوکریستال نیز – با چنین رویکردی – به‌دنبال تجسم مجموعه داده‌ها به‌منظور درک بهینه ساختار زیرین داده‌هاست؛ به‌عبارت دیگر، در مصورسازی اطلاعات نمایش روابط ساختاری بین مدارک و محتوای آنها دنبال می‌شود، روابطی که درک آنها با استفاده از درخواست‌های بازیابی انفرادی بسیار مشکل می‌باشند.

توانمندسازی کاربران برای دسترسی به اطلاعات با ارائه روش‌هایی برای مصورسازی اطلاعات انتزاعی و فرمول‌بندی گرافیکی پرس‌وجوها، همواره از جمله دغدغه‌های متخصصان حوزه بازیابی اطلاعات بوده است. با توجه به نرخ تصاعدی رشد اطلاعات، لزوم به‌کارگیری روش‌های بازیابی بصری بیشتر در کانون توجه قرار می‌گیرد؛ روشی که استفاده از کاربران را به کاوش، دستکاری و ارتباط تعاملی با فضاهای بزرگ اطلاعاتی قادر می‌سازد و برای آنها چارچوبی فراهم می‌کند که اطلاعات بازیابی شده به روش‌های گوناگون و یا از منابع گوناگون را تجمیع و دستکاری کنند. روزنبلوم[۲۳] (۱۹۹٤)، براساس یافته‌های اخیر در حوزه مصورسازی، نشان می‌دهد که چگونه مجموعه‌های بزرگ داده‌ها می‌توانند همانند روش کاربردی انسان در تشخیص الگوهای زیرین ساختارها، مصورسازی شوند؛ روشی که در مقایسه با تحلیل مستقیم اعداد سریع‌تر است (اسپوئری، ۱۹۹۵، ص٦). چن[۲٤] (۲۰۰٦)، اضافه بار اطلاعات را یکی از معضلات رایج در رشد ظاهری اطلاعات قابل دسترس، در جوامع مدرن می‌داند که امکانات فیلترینگ و اشتراک اطلاعات برای حل آنها ضروری است و مصورسازی اطلاعات از توانایی بالقوه برای کمک به مردم در یافتن مؤثر و آگاهانه اطلاعات مورد نیازشان برخوردار است. زانگ و دنگ[۲۵] (۲۰۰۸)، براین باورند که برای کاهش هزینه فراهم‌آوری اطلاعات و ارتقای کیفیت بازیابی اطلاعات بهتر است که فناوری مصورسازی اطلاعات در بازیابی اطلاعات ادغام گردد تا کمک کند فضای پیچیده اطلاعات را بهتر درک کنند. مصورسازی اطلاعات، با کاربرد بازیابی اطلاعات، به‌دنبال مشخص کردن روابط ساختاری بین اسناد و بافت آنهاست، فرآیندی که تشخیص آن با استفاده از درخواست‌های بازیابی انفرادی به‌سختی

22. Computational vision
23. Rosenblum
24. Chen
25. Zhang &Deng

امکان‌پذیر است. چالش‌های بازیابی منابع اطلاعات موجود در شبکه جهانی وب به قوت خود باقی است و شیوه‌های جست‌وجو و بازیابی پیچیده اطلاعات سنتی، با مهارت همه کاربران سازگار نیست. شرمان[26] (۲۰۰۱)، بر این باور است که موتورهای کاوش سنتی– که پرس‌وجوی واژه را می‌پذیرد، تحلیل می‌کند، رابطه منطقی اعمال می‌کند، مدارک را می‌یوید تا واقعیت‌ها را یافته و هزاران میلیون نتیجه را بدون تحلیل و یا انتقال معنایی خاص نمایش می‌دهد– اغلب «ابزارهای تحلیلگران» هستند. این موتورها برای افرادی که از بخش چپ مغز (منطقی، تحلیلی، واقعی‌محور، و کمّی) خود بهره می‌گیرند، طراحی شده‌اند؛ در حالی که موتورهای کاوش بصری برای افرادی که از طرف راست مغز (بصری، ادراکی، رهیافتی، و یکپارچه) خود بهره می‌گیرند، طراحی شده‌اند. در واقع، استفاده از توانمندی ابزارهای مصورسازی اطلاعات، از راهکارهای مناسب برای مدیریت حجم وسیعی از اطلاعات تحت وب هستند که مشغله فکری بسیاری از محققان و آزمایشگاه‌های تحقیقاتی است.

نمونه‌هایی از این پژوهش‌ها را می‌توان در الگوی نینگ[27] (۲۰۰۸) با عنوان «اینفوویزمدل» یافت که نتیجه یک طرح تحقیقاتی انجام با حمایت مالی بنیاد ملی علوم طبیعی چین است، که به‌طرز موفقیت‌آمیزی ساختار الگوی مصورسازی را برای متن، صدا و تصویر ارائه نموده است و دارای ساختار ۵ مرحله‌ای (فراهم‌آوری اطلاعات، نمایه‌سازی اطلاعات، بازیابی اطلاعات، تولید داده‌های تصویری، و نمایش رابط بصری) است. نمونه اولیه آن پردازش بیش از ۱۰۰۰۰ هدف اطلاعاتی را فراهم می‌سازد.

منابع و پایگاه‌های اطلاعاتی آی.اس.آی. وب آو نالج[28] (Structure Search)، اسکوپوس[29] (Analytics)، ابسکو[30] (Visual Search)، آمازون[31] (AmazeType)، ویکی‌پدیا (SearchCrystal)، ومانند آن نمونه‌های دیگری هستند که هریک برحسب شرایط خود از توانمندی ابزارهای مصورسازی اطلاعات بهره گرفته‌اند.

مصورسازی اطلاعات و اینفوکریستال

مصورسازی اطلاعات دارای دو جنبه کاملاً مرتبط است: ۱) الگوسازی ساختاری[32] و ۲) بازنمون گرافیکی[33]. مفهوم الگوسازی ساختاری کشف، اقتباس، و تسهیل روابط اساسی بین اجزای سازنده یک شیء یا مفهوم است. این روابط، ساختاری را برای توصیف مجموعه‌ای از مدارک یا سایر مجموعه داده‌ها شکل می‌دهد.

از جمله پرسش‌هایی که الگوسازی ساختاری در پی پاسخ به آنهاست عبارت‌اند از: ساختار بنیادی یک شبکه پیچیده یا مجموعه‌ای از مدارک چیست؟ الگوهای ذهنی یک شهر یا باغ‌وحش در اذهان افراد مختلف چیست؟ ساختار ادبیات یک حوزه موضوعی چیست؟ و

26. Sherman
27. Ning
28. ISI Web of knowledge
29. SCOPUS
30. Ebsco
31. Amazon.com
32. Structural modeling
33. Graphical representation

نظایر آنها. در مقابل، هدف بازنمون گرافیکی، انتقال بازنمون اولیه یک ساختار به یک بازنمون گرافیکی است، به‌نحوی که آن ساختار بتواند به‌طور بصری آزموده شود و تعامل داشته باشد. به‌عنوان مثال، یک ساختار سلسله‌مراتبی می‌تواند به شکل یک درخت کاج (مخروطی)، یا نمودار هذلولی[34] نمایش داده شود.

اینفوکریستال، به‌عنوان یکی از روش‌های پیشرو در مصورسازی اطلاعات، با هدف استفاده از بازنمون بصری، برای کمک به کاربران در رویارویی با برخی از پیچیدگی‌های ذاتی بازیابی اطلاعات، و برای نخستین بار در ۱۹۹۵ توسط اسپوئری، در قالب پایان‌نامه دکتری آزمون و ارائه شد.

اینفوکریستال، یکی از ابزارهای مصورسازی اطلاعات است که می‌تواند فضای انتزاعی اطلاعات را به‌تصویر بکشاند (اسپوئری، ۱۹۹۵).

فضای اطلاعاتی، بخشی از سیستم‌های اطلاعاتی است که دارای اطلاعات بالقوه مرتبط با اهداف سیستم است که مطابق با تنظیمات سیستم ساختاریافته است (اینگورسن، ۲۰۰۲، ص، ۲۲۸).

اینفوکریستال، تمامی روابط ممکن بین بی‌نهایت مفهوم را مجسم می‌سازد. از این رو، کاربران می‌توانند به مفاهیم، وزن‌های ربطی‌داده و روابط موجود بین مفاهیم را به‌راحتی درک کنند. در این شیوه، امکان مشاهده پرس‌وجوهای بولی، وزنی، فضای برداری، مرور، و فیلترینگ نتایج به‌شکل انعطاف‌پذیر، پویا، و تعاملی فراهم است.

اینفوکریستال کارکردهای زیر را فراهم می‌سازد:

- کاربران می‌توانند یک فضای اطلاعاتی را از چند بُعد، به‌طور همزمان، با فرآیند مرور، کاوش کنند؛
- کاربران می‌توانند با انتزاع‌سازی (تجردآفرینی) اطلاعات را دستکاری کنند؛
- کاربر همانند [کار با] صفحات گسترده، قادر به طرح سؤال «چه می‌شود اگر» و مشاهده نتیجه آن، بدون نیاز به تغییر چهارچوب پرس‌وجو می‌باشد.
- در فرآیند جست‌وجو، کاربر، با استفاده از بازخورد بصری پویا، درباره نحوه اقدام بعدی، پشتیبانی می‌شود؛ و
- کاربر می‌تواند پرس‌وجوها را به‌طور گرافیکی فرمول‌بندی کند و به همین دلیل در استفاده از روش‌های مختلف، برای بازیابی اطلاعات، از قابلیت انعطاف برخوردار است. به‌عنوان مثال، کاربر می‌تواند بین رویکرد بولی و فضای برداری در حرکت باشد، و یا، از رویکرد بازیابی مبتنی‌بر کلیدواژه به رویکرد بازیابی تمام متن تغییر جهت دهد (اسپوئری، ۱۹۹۳، ص، ۱۱).

چنانچه پیش‌تر نیز اشاره شد، اینفوکریستال به‌دنبال تجسم‌بخشی به روابط بین مفاهیم

34. Hyperbolic

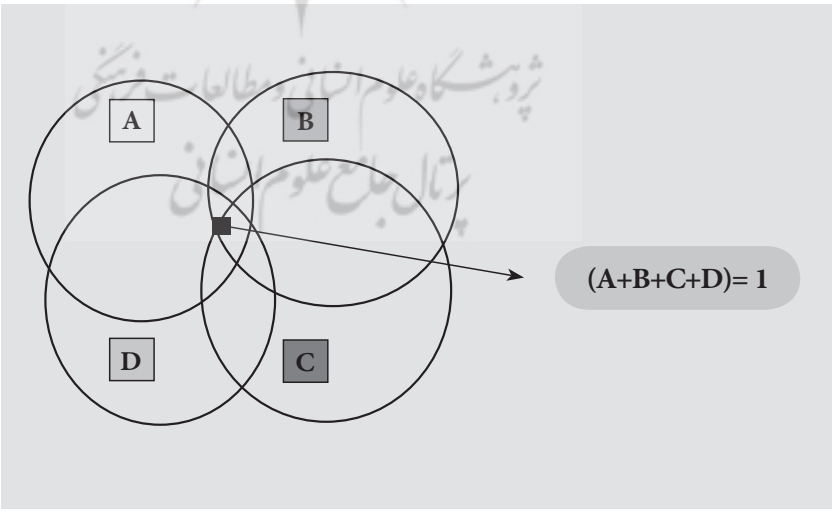
مختلف است. حال چگونگی تجسم‌بخشی همه ترکیب‌ها یا روابط بین معیارهای متعدد جست‌وجو، در یک نمایش دوبُعدی مسئله اصلی است. برای این منظور، یکی از راهکارها می‌تواند تعریف فضای چندبُعدی باشد که ابعاد آن برابر با تعداد معیارهایی است که باید مقایسه شوند. این رویکرد مشکل گشا نیست، زیرا باید بتواند دگرگونی خاصی وجود داشته باشد که فضای چندبُعدی را در نمایش دوبُعدی ترسیم کند (اسپوئری، ۱۹۹۳، ص ۱۵۱).

رویکرد دیگر، استفاده از نمودارهای ون برای نمایش مجموعه روابط با اشکال هندسی متداخل است. مشکل این شیوه نیز به محدودیت آنها در همه روابط ممکن برای یک مجموعه بیش از سه‌تایی مربوط می‌شود. به‌عنوان مثال، اگر به‌دنبال مدارکی باشیم که بتواند نیاز اطلاعاتی ما را در مورد «زبان‌های پرس‌وجوی دیداری را برای بازیابی اطلاعات–با توجه به نقش عوامل انسانی–» پوشش دهد، شاید یکی از ایده‌ها بیان نیاز در چارچوب مفاهیم و راهبرد زیر باشد:

بازیابی اطلاعات(A)- زبان پرس‌وجو(B)- عوامل انسانی (C)- دیداری یا گرافیکی (D)

بسیاری از نظام‌های بازیابی پیوسته برای ایجاد پرس‌وجو با ترکیب مفاهیم از عملگرهای بولی استفاده می‌کنند. در ترکیب فرمول با استفاده از عملگرهای بولی ناگفته پیداست که «و» به مانعیت و «یا» به جامعیت گرایش دارد.

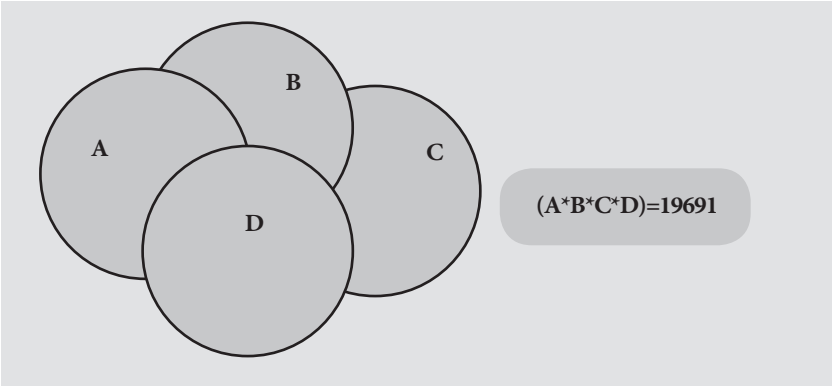
اسپوئری برای آزمون نتایج بین مفاهیم فوق و هدایت آن به پایگاه اطلاعاتی انسپک[۳۵] (۲–۱۹۹۱) از عملگرهای «و» و «یا» استفاده نمود که نتایج آن را می‌توان به‌صورت شماتیک در شکل‌های ۳ و ٤ مشاهده کرد.



شکل ۳

نمودار ون برای چهار مفهوم با عملگر AND

35. INSPEC



شکل ۴

نمودار ون برای چهار مفهوم با عملگر OR

برای تغییر وضعیت از مانعیت به جامعیت تنها جابه‌جایی عملگر «و» با «یا» کافی است، ولی ناگفته پیداست که برای کاربر به همان اندازه که یک مدرک بازیابی شده نمی‌تواند خیلی رضایتبخش باشد (به‌عبارتی وی به‌دنبال بیش از یک مدرک است) امکان مرور و بازبینی ۱۹۶۹۱ مدرک نیز برای وی فراهم نیست. لذا به‌دنبال بسط بازیابی مانع‌نگر و یا معقولانه ساختن جست‌وجوی جامع‌نگر خواهد بود. این شرایط با دگرگونی از نمودار متداخل ون به اینفوکریستال فراهم می‌گردد که در ادامه به تفصیل ذکر خواهد شد.

برای درک بهتر فرآیند تبدیل از نمودار ون به اینفوکریستال، روابط منتقل‌شده برای سه مفهوم فرضی در قالب اشکال هندسی به صورت شکل ۵ خواهد بود (اسپوئری، ۱۹۹٤).

شکل ۵

فرآیند تبدیل از نمودار ون به اینفوکریستال

شکل ۵ نشان می‌دهد که می‌توانیم نمودار ون را به نمایش تجسمی[۳۶] موسوم به اینفوکریستال تبدیل کنیم که همه روابط ممکن پرس‌وجوهای بولی را جداگانه در شکل عادی نشان دهد. آیکون‌های داخلی دارای بار معنایی بولی به این شرح هستند (اسپوئری، ۱۹۹٤):

1=(A and (not (B or C), 2=(A and C and (not B), 3=(A and B and C), 4=(A and B and (not C),5=(C and (not (A or B)), 6=(B and C and (notA), 7=(B and (not (A or C)).

36. Iconic display

آنچه در تشکیل اینفوکریستال (شکل ۵) اتفاق می‌افتد طی مراحل سه‌گانه زیر است:

۱. بسط: در این مرحله یک نمودار ون به اجزا و زیرمجموعه‌هایش شکسته می‌شود و اجزا از هم جدا می‌شوند.

۲. تغییر: در این مرحله، هریک از زیرمجموعه‌ها بسته به اینکه چه تعداد رابطه را بین چند معیار (مفهوم) بیان می‌کند، در اشکال هندسی با کدهای ریختواره مشخص نمایش داده می‌شود که به آن درجه زیرمجموعه[۳۷] گفته می‌شود. به‌عنوان مثال، این درجه برای نتایج مشترک سه مفهوم مثلثی خواهد بود.

۳. اینفوکریستال: شکل اینفوکریستال، که قالب نهایی است، تابعی از تعداد معیارها (مفاهیمی) است که در فرمول جست‌وجو چیده شده است. به‌عبارت دیگر، شکل اینفوکریستال برای یک عبارت جست‌وجوی مانع‌نگر دارای سه معیار (با عملگر AND بین مفاهیم) مثلثی و کد ریختواره آن نیز مثلثی خواهد بود. نتیجه بازیابی‌شده در قالب نماد اندرونی کانونی که می‌تواند نشان‌دهنده مدارک کاملاً مرتبط و نشان‌دهنده هر سه مفهوم باشد نیز به شکل مثلثی خواهد بود.

روابط عمده‌ای که در اینفوکریستال به کرات استفاده می‌شوند عبارت‌اند از:

• **کد ریختواره[۳۸]:** این کد برای نشان دادن تعداد معیارهای (مفاهیم) نمایش داده شده توسط اینفوکریستال به‌کار می‌رود. برای مثال، کد ریختواره برای یک مفهوم به‌صورت دایره‌ای، برای دو مفهوم به‌صورت مستطیلی، برای سه مفهوم به‌صورت مثلثی، برای چهار مفهوم (مثال عینی) به‌صورت مربع خواهد بود.

• **کد مجاورت[۳۹]:** این کد نشان‌دهنده نزدیکی یک نماد اندرونی به یک نماد معیار است و محتوای نماد اندرونی با نماد معیار مرتبط است.

• در اینفوکریستال فوق، نماد اندرونی با ریختواره مستطیلی که داخل آن عدد ٦ درج شده است، نشان‌دهنده سه نکته مهم است. نخست اینکه، تعداد معیار مشترک دو تاست، پس کد ریختواره برای نماد اندرونی مستطیلی است. دوم اینکه، در امتداد دو معیار «زبان پرس‌و جو» و «عامل انسانی» قرار گرفته است که به آن کد مجاورت گفته می‌شود؛ لذا نشان‌دهنده مدارکی است که زیر دو معیار مذکور نمایه شده‌اند؛ و سوم اینکه تنها ٦ مدرک بازیابی شده‌اند.

• **کدرتبه[۴۰]:** عبارت است از نمادهای دایره‌ای یک شکل که در چهار ضلع درونی اینفوکریستال نمونه مشخص است.

• رتبه یک نماد با معیارهای ادا شده برابر است و هرچقدر از کانون اینفوکریستال به سمت معیار حرکت کنیم، افزایش می‌یابد. به‌عنوان مثال، کد رتبه برای معیار «زبان پرس‌وجو» برابر با ٦١٢ است. کد رتبه برای چهار مفهوم در نماد کانون اینفوکریستال برابر با ۱ است.

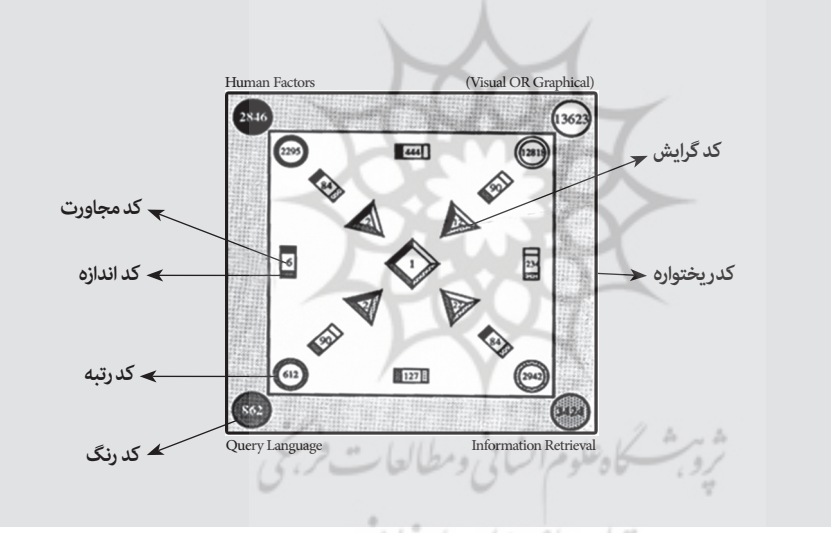
37. Rank of subset
38. Shape coding
39. Proximity coding
40. Rank coding

• **کد رنگ یا بافت**[41]: هر معیار در اینفوکریستال با یک رنگ نشان داده می‌شود. کد رنگ نماد معیار نشان داده شده به‌وسیله یک نماد اندرونی را می‌رساند.

• **کد گرایش**[42]: این کد به شیوه استقرار نمادها اشاره می‌کند و اضلاع هر نماد به سمت یک معیار است. به‌عنوان مثال، به نماد اندرونی مثلثی شکل در اینفوکریستال نمونه که عدد ۱۲ درون آن درج شده است، توجه کنید. اضلاع این نماد به طرف معیارهای «دیداری»، «بازیابی اطلاعات»، و «عامل انسانی» است.

• **کد اندازه یا درخشندگی و اشباع**[43]: این کد برای تجسم اطلاعات کمّی استفاده می‌شود. یعنی تعداد اقلام بازنمایی‌شده توسط یک نماد را می‌رساند.

• حال برای تشریح بیشتر کارآیی اینفوکریستال، نمایش نتیجه جست‌وجوی مثال قبلی«زبان‌های پرس‌و جوی دیداری برای بازیابی اطلاعات با توجه به نقش عوامل انسانی» شبیه شکل٦ خواهد بود.



شکل ۶

شکل اینفوکریستال برای چهار مفهوم «زبان پرس‌وجوی دیداری برای بازیابی اطلاعات با توجه به نقش عوامل انسانی»

در شکل٦ عدد درج‌شده بر روی هر آیکون تعداد مدارک بازیابی شده و وضعیت (استقرار) آیکون را نشان می‌دهد، (اسپوئری، ۱۹۹۳)؛ اشاره به این نکته ضروری است که چون تعداد معیار (مفاهیم) ٤ فقره است و لذا کد ریختواره مربعی است.

مجموع ۱۹٦۹۱ مدرک با محاسبه زیر قابل مشاهده است:

٦۱۲+۱۲۷+۲۹٤۲+۲۳٤+۱۲۸۱۸+٤٤٤+۲۲۹۵+٦+۸٤+۹۰+۲+۲+۱۲+۲۲+۱= ۱۹٦۹۱

اینفوکریستال فوق برای چهار مفهوم (معیار) مثال عینی، بیانگر کارآیی آن در نمایش زوایای پنهان رابطه بین مفاهیم است که اجازه استفاده از مهارت‌های مستدل بصری برای

41. Color or Texture coding
42. Orientation coding
43. Size or Brightness & Saturation coding

درک چگونگی ارتباط نمادهای اندرونی با نمادهای معیار را ه کاربر می‌دهد. به بیانی ساده، در جست‌وجو- با استفاده از عملگرهای بولی «و» و«یا»- مشاهده شد که یا تنها یک مدرک - که دربردارنده تمام معیارهای جست‌وجو بود- بازیابی شد؛ و یا ۱۹۶۹۱ مدرک بازیابی شد که دربردارنده همه معیارهای جست‌وجو بودند. پس، برای کاربر در هنگام جست‌وجو، معیارهای مختلف می‌توانند دارای ارزش‌های متفاوتی باشند. به‌عنوان مثال، در جست‌وجوی نمونه، اگر تأکید کاربر «بر زبان‌های پرس‌و جو و بازیابی اطلاعات» بیشتر از دو معیار دیگر باشد؛ برای بازیابی مدارک حاوی این دو معیار باید جست‌وجوهای دیگری فرمول‌بندی شود. در حالی که، اینفوکریستال فوق با انتخاب نمادی که عدد ۱۲۷ بر روی آن درج شده است، بدون تلاش مجدد، مدارک حاوی دو مفهوم معیار را بازیابی خواهد نمود و این همان امکان فیلترینگ اینفوکریستال است که با انتخاب نماد حاوی عدد ۱۲۷ معیارهای جست‌وجوی پیشین درنظر گرفته نمی‌شود و تنها ۱۲۷ مدرک مفاهیم معیار (زبان پرس‌وجو و بازیابی اطلاعات) نمایش داده می‌شود.

زبان پرس‌وجوی دیداری

نمایش بصری پرس‌وجوها و یا مجموعه داده‌های منتج، دیدگاهی بدیع در مشاهده روابط بین اقلام داده‌ای بازیابی شده را ارائه می‌کند. برعکسِ مصورسازی علمی، که به بیان اشیای دنیای واقعی می‌پردازد، مصورسازی در بازیابی اطلاعات در تلاش است تا مضامین غیرملموسی را که عموماً از اطلاعات نشأت گرفته‌اند، الگوسازی کند و نمود گرافیکی انتزاعی را برای نمایش مجموعه مدارک به‌کارگیرد. لذا یکی از ویژگی‌های رؤیایی اینفوکریستال در این است که نمادهای اندرونی، یک رابطه خاص بولی را میان معیارهای داده شده بیان کند و کاربر بتواند با استفاده از اینفوکریستال پرس‌وجوهای بولی را به‌صورت درهم‌کنشی -با دستکاری مستقیم رابط بصری- تعیین کند. این قابلیت، کاربر را از فرمول‌بندی مجدد و گرفتاری در استفاده از پرانتزها و عملگرهای بولی بی‌نیاز می‌کند. این کارکرد به یکی از شیوه‌های زیر ممکن است:

• با انتخاب یکی از نمادهای اندرونی. برای مثال، در نمونه ارائه‌شده، انتخاب نماد اندرونی با ریختواره مستطیل شکلی که عدد ٦ در آن درج شده است، همسان با فرمول‌بندی به شرح زیر خواهد بود:

(Query language AND Human Factors) = ٦

• با انتخاب نماد معیار امکان انتخاب مجموعه‌ای از روابط آن نماد معیار فراهم خواهد بود. به‌عنوان مثال، در نمونه فوق، یکی از نمادهای معیار زبان پرس‌وجو[44] می‌باشد که عدد ۸٦۲ بر روی آن درج شده است، که از این میزان ۲۵۰ مدرک با اشتراک ۳ معیار دیگر با

44. Query language

ترکیب‌های مختلف در اینفوکریستال قابل مشاهده است و ۶۱۲ مدرک دیگر نیز می‌تواند صرفاً در زیر مفهوم زبان پرس‌وجو یا اشتراک آن با معیارهای دیگر (غیر از ۳ مورد مندرج در اینفوکریستال نمونه) نمایه‌شده باشد (کد رتبه). با محاسبه زیر:

۲۵۰= ۲+۲۲+۶+۱+۲+۹۰+۱۲۷ تعداد مدارک موضوع اینفوکریستال نمونه

۶۱۲+۲۵۰= (کد رتبه) ۸۶۲

مجموعه مدارک موجود در پایگاه و نمایه شده زیر معیار زبان پرس‌وجو

با انتخاب نماد معیار زبان پرس‌وجو دسترسی به زیرمجموعه‌ای از نمادهای اندرونی امکان‌پذیر خواهد بود.

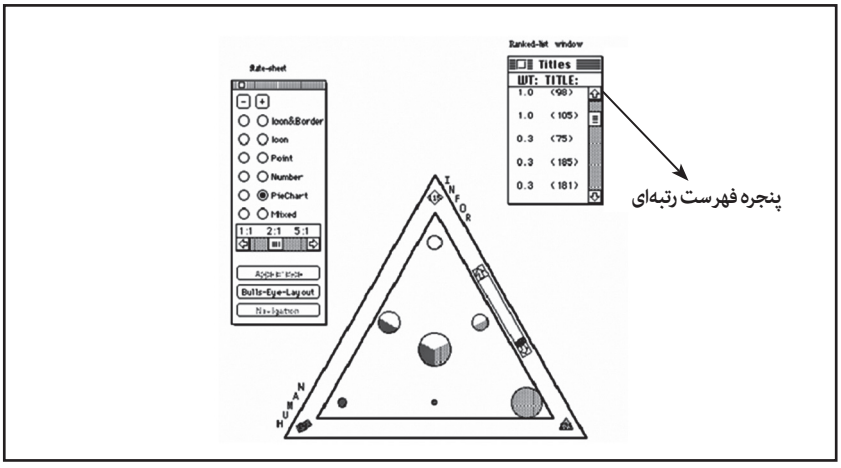
چنانچه پیش‌تر نیز گفته شد، بازیابی بولی تنها یکی از مدل‌های بازیابی است. اینفوکریستال قابلیت مصورسازی و بازنمون گرافیکی مدل‌های دیگر بازیابی را نیز دارد.

اینفوکریستال و مصورسازی پرس‌وجوهای وزنی[۴۵]

مدل بولی با تمام قابلیت‌هایی که دارد، مدلی آرمانی نیست. یکی از گلایه‌ها درباره این مدل ناتوانی آن در ارائه برونداد رتبه‌بندی شده، برای مدارک بازیابی شده است. به‌دلایل مختلفی، فرمول‌بندی پرس‌وجوهای وزنی برای کاربر حائز اهمیت است. نخست اینکه، در برخی شرایط معیارهای جست‌وجو به یک اندازه برای کاربر دارای اهمیت نیستند؛ دوم اینکه برای کاربر بدواً ارائه فهرستی از موارد دلخواه و الصاق وزن ربطی به آنها می‌تواند سهل‌تر باشد. این وزن‌ها می‌تواند به‌وسیله سیستم بازیابی در تولید پرس‌وجوهای بولی به‌کارگرفته شود و چگونگی انجام آن مسئله اساسی است که با اینفوکریستال قابل حل است.

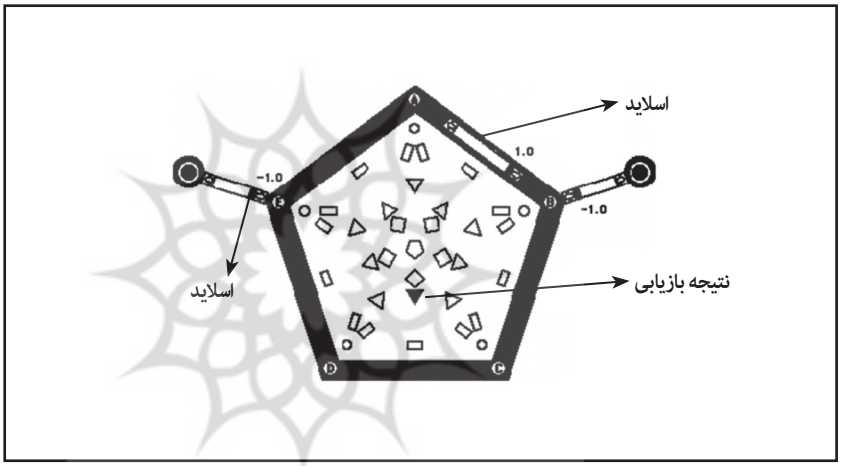
برای بسط بازنمون اینفوکریستال، به‌نحوی که قادر به فرمول‌بندی پرس‌وجوهای وزنی باشد، دو روش وجود دارد. نخست اینکه، کاربر به مفاهیم فهرست‌شده در طرح پرس‌وجو در میان دو کروشه و بعد از هرفقره وزن می‌دهد. دوم اینکه، با همراه کردن یک اسلاید با هر معیار اینفوکریستال، کاربر می‌تواند با تعامل با اسلاید وزنی، میزان اهمیت را انتخاب کند. این مقدار وزنی نباید بیشتر از یک باشد و کاربر می‌تواند بین ۱- تا ۱ را انتخاب کند. انتخاب مقدار وزنی ۱- برای یک معیار بیانگر علاقه کاربر به عدم مفهوم مورد نظر در نتایج است. به‌عبارت ساده‌تر، مقدار وزنی ۱- را می‌توان معادل عملگر منطقی «نه» دانست. شکل‌های ۷ و ۸ به‌خوبی گویای این مسئله هستند (اسپوئری، ۱۹۹۵).

45. Weighted queries



شکل ۷

انتخاب وزن با استفاده از پنجره فهرست رتبه‌ای که همزمان با نمایش اینفوکریستال ظاهر می‌شود



شکل ۸

نمایش پرس وجوهای وزنی با اینفوکریستال

در شکل ۷ اوزان با استفاده از پنجره فهرست رتبه‌ای امکان‌پذیر است، در حالی که شکل ۸ قابلیت اینفوکریستال در فرمول‌بندی پرس‌وجوهای وزنی با استفاده از اسلاید وزنی و اسلاید آستانه را نمایش می‌دهد. تنها برای معیارهای که وزن آنها برابر با ۱ نیستند در اینفوکریستال اوزان درج شده است (اسپوئری، ۱۹۹۵).

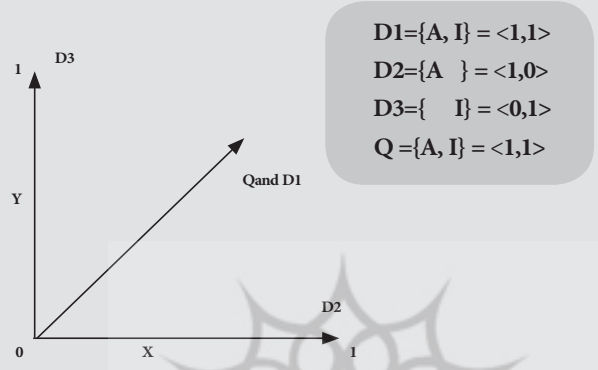
تعداد معیارهای این اینفوکریستال (شکل ۸) ۵ فقره که عبارت‌اند از: A, B, C, D و E، و لذا فرمول آنها می‌تواند به شرح زیر چیده شود:

(A[1.0] and(B[-1.0]) and C[1.0] and D[1.0] and (E[-1.0]))

نتیجه بازیابی‌شده با ریختواره مثلثی سیاه رنگ مشخص است که حاوی مدارکی است که معیارهای A، C، وD را نمایش می‌دهد.

مصورسازی فضاهای برداری

در فضاهای برداری، مدارک و پرس‌و جوها در فضای دوبُعدی ترسیم می‌شوند. رویکرد فضای برداری نمره ربطی را از وزن داده‌شده برای اصطلاحات نمایه که بازنمون مدارک و پرس‌وجوها هستند، محاسبه می‌کند (شکل ۹).



شکل ۹

فضای برداری برای سه مدرک که به مفاهیم A یا I مربوط می‌شوند و در امتداد محورهای X و Y با وزن ۱ نمایش داده شده‌اند.

اوزان ربطی چگونگی توصیف محتوای مدارک و پرس‌وجو به‌وسیله اصطلاحات نمایه را به‌خوبی نشان می‌دهند (اسپوئری، ۱۹۹۵). با دادن وزن از ۱- تا ۱ می‌توان میزان حضور یک نماد معیار را مشاهده نمود. همان‌گونه که در چنین حالتی مقدار ۱- نشان‌دهنده صفر مدرک است، مقدار ۱ نیز بیانگر انباشتگی[46] مدارک برای نماد معیار می‌باشد. لذا در این حالت مدارک دارای درجه ربط بالا در نزدیکی کانون مرکزی بازیابی می‌شوند. مدارک دارای درجه ربط پایین‌تر در موقعیتی دورتر از کانون قرار می‌گیرند (شکل۹).

شکل ۱۰

نمایش نتایج جست وجو با استفاده از رویکرد فضای برداری در قالب اینفوکریستال

46. Mass

شکل ۱۰ نمونه اینفوکریستال برای سه معیار با اوزان مشابه [۱] که نشان می‌دهد اینفوکریستال: الف) نشان‌دهنده مدارک مرتبط با هرسه معیار A وB وC با وزن [۱] می‌باشد که در کانون اینفوکریستال خوشه‌بندی شده است.

ب) نشان‌دهنده مدارک مرتبط با دو معیار A وB با وزن [۱] وC با وزن [۱-] می‌باشد که دورتر از کانون اینفوکریستال و مجاورت معیارهایA و B خوشه‌بندی شده است.

ج) نشان‌دهنده مدارک مرتبط با معیار A با وزن [۱] وB وC با وزن [۱-] می‌باشد که در نزدیکی معیار A خوشه‌بندی شده است.

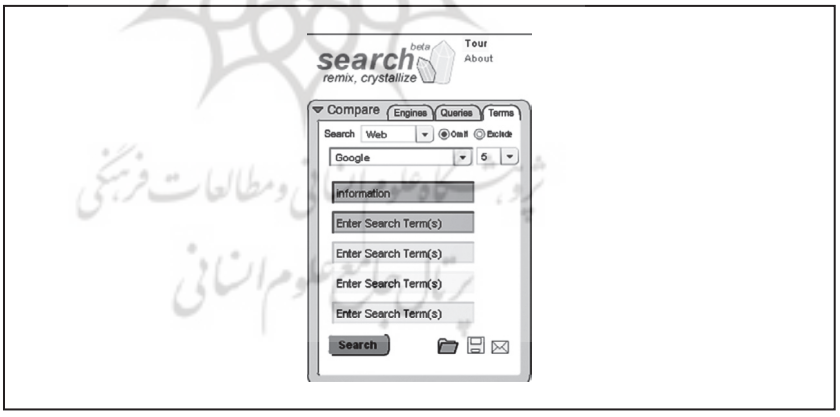
سخن آخر اینکه نمونه نرم‌افزار اینفوکریستال طی نتایج تحقیقات اسپوئری ارائه شده است که اطلاعات آنها از طریق وب‌سایت وی به نشانی زیر قابل بررسی است:

http://comminfo.rutgers.edu/~aspoerri/index.htm

قابلیت‌های فراوان اینفوکریستال به ارائه نمونه‌های دیگری نیز سبب شده است که SearchCrystal و MetaCrystal از آن جمله‌اند و نمونه آزمایشی SearchCrystal را می‌توان از طریق نشانی زیر مشاهده کرد:

http://comminfo.rutgers.edu/~aspoerri/searchCrystal

با استفاده از این نمونه می‌توان اصطلاحی را وارد نمود و میزان هدایت آنها را از طریق موتورهای جست‌وجوی مختلف به ویکی‌پدیا مشاهده نمود و یا برای چند اصطلاح از طریق یک موتور جست‌وجو همین کار را انجام داد (شکل ۱۱).



شکل ۱۱
نمونه رابط جست وجوی SearchCrystal

نتیجه‌گیری

رشد اطلاعات به لحاظ حجمی و محتوایی بر دغدغه‌های دسترسی به اطلاعات مرتبط افزوده است و این مسئله از مدت‌ها پیش بشر را وادار کرده است تا برای دستیابی به اطلاعات مفید به‌دنبال راهکار بازیابی باشد. هرقدر آهنگ رشد اطلاعات سریع‌تر شود بر پیچیدگی این مسئله

نیز افزوده می‌شود. شیوه‌های مختلفی که برای بازیابی ارائه شده‌اند، برخی از ابهامات این حوزه را برطرف کرده‌اند. ارائه شیوه‌های نوین نشان‌دهنده وجود نارضایتی از نظام‌های گذشته و سعی در کاربرپسند کردن بیش‌از پیش نظام‌های جدید است. در این بین، مصورسازی اطلاعات یکی از روش‌هاست که سعی دارد از توانایی دیداری انسان در تشخیص الگوها و روابط ساختاری اشیا در بهبود نظام‌های بازیابی استفاده نماید. یکی از فنون مصورسازی اطلاعات اینفوکریستال می‌باشد که می‌تواند فضای انتزاعی اطلاعات را به تصویر بکشاند.

اینفوکریستال، تمامی روابط ممکن بین بی‌نهایت مفهوم را مجسم می‌سازد و از این طریق، به کاربران کمک می‌کند تا مفاهیم وزن‌های ربطی داده و روابط موجود بین مفاهیم را به‌راحتی درک کنند. در این شیوه، امکان مشاهده پرس‌وجوهای بولی، وزنی و فضای برداری، مرور، و فیلترینگ نتایج به شکل انعطاف‌پذیر، پویا، و تعاملی برای کاربران فراهم است. با توجه به جدید بودن مبحث مربوط به این حوزه در متون کتابداری و اطلاع‌رسانی کشور در این مقاله سعی شده است تا این ابزار معرفی و قابلیت‌های آن مورد مداقه قرار گیرد.

اینفوکریستال، به‌عنوان یکی از فنون مصورسازی اطلاعات، بنا به اعتقاد اسپوئری (۱۹۹۵)، این قابلیت را دارد تا در حوزه‌های مختلفی از قبیل پایش اینترنت جهت یافتن منابع اطلاعاتی که از طریق اینترنت قابل دسترسی‌اند، برای کاوش پایگاه اطلاعاتی در پایگاه‌های اطلاعاتی رابطه‌ای، مدیریت سهام مالی، مدیریت منابع انسانی و کارهای گروهی، بازیابی اطلاعات هنگام ویرایش چندرسانه‌ای‌ها، فیلتر پست الکترونیکی، مرور فراپیوندها و ایجاد پیوند، مصورسازی آماری، شبکه‌های بولی، و شبکه‌های عصبی به‌کارگرفته شود. امید است که با شناخت قابلیت‌های اینفوکریستال شاهد پژوهش‌های جدید در آزمون و به‌کارگیری آن درحوزه‌هایی مذکور و حوزه‌های جدیدتر باشیم.

منابع

پائو، میراندا لی (۱۳۷۹). ***مفاهیم بازیابی اطلاعات***. ترجمه اسدالله آزاد و رحمت‌الله فتاحی. مشهد: دانشگاه فردوسی مشهد، مؤسسه چاپ و انتشارات.

درودی، فریبرز (۱۳۸۸). «نگارش تصویری با تصویرسازی مفاهیم: بهره‌گیری از سواد دیداری و زبان تصویر برای درک مؤثر اطلاعات». ***فصلنامه کتاب***، ۲۱(۱): ۲۷۳-۲۸۸.

Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier (1999). ***Modern information retrieval***. London: Addison-Wesley-Longman.

Borlund, P. (2003). "The concept of relevance in IR". ***Journal of the American Society for Information Science and Technology,*** 54(10): 913-925.

Card, Stuart K.; Robertson, George G.; Mackinlay Jock D. (1991). ***"The information visualizer, an information workspace"***. Proc. Retrieved November 2, 2009. from: http://www2.parc.com/istl/groups/uir/publications/items/UIR-1991-01-Card-CHI91-IV.pdf

Chen, Chaomei (2006). ***Information visualization, beyond the horizon***. 2nd ed., London: Springer-Verlag.

Ingwersen, Peter (2002). ***Information retrieval interaction***. London: Taylor Graham.

Koshman, Sherry (2005). "Testing user interaction with a prototype visualization-based information retrieval system". ***Journal of American Society for Information Science and Technology***, 56(8): 824-833.

Lancaster, F.; Warner, A. (1993). ***Information retrieval today***. Arlington, VA: Information Resources Press.

Manning, C.; Raghavan, P.; Schutze, H. (2008). ***Introduction to information retrieval***. Cambridge: Cambridge University Press.

Morse, E.; Lewis, M.; Olsen Kai, A. (2002). "Testing visual information retrieval methodologies case study: Comparative analysis of textual, Icon, Graphical, and "Spring" displays". ***Journal of American Society for Information Science and Technology***, 53(1): 28-40.

Ning, Zhou; et al. (2008). "A visualization model for information resource management". Proceeding of 12th International Conference Information visualization, pp.57-62. DOI: 10.11/IV.2008.105.

Sherman, Chris (2001). "Searching on the right side of the brain". Search Engine Watch. Retrieved 2 November 2007. from: http://searchenginewatch.com/showPage.html?page=2158651

Spoerri, Anselm (1993)." InfoCrystal: A visual tool for information retrieval". Proceedings of the IEEE Visualization Conference. San Jose, October 25-29 (Los Alamitos, CA: IEEE Computer Soc. Press), pp. 150 – 157.

_____ (1994). "Infocrystal: A visual tool for information retrieval & management". Conference Companion. CHI'94, (Boston, April 24-28), pp. 11-12.

_____(1995). "Infocrystal, A visual tool for information retrieval". Ph.D. Research and Thesis at MIT.

Valvag, Ottar Viken (2004). "Multiple evidence combination in information retrieval". Master of Technology thesis, Institutt for datateknikk og informasjonsvitenskap, NTNU.

Zhang, Fangfang; Deng, Shaoling (2008). "Studies on the visualization for web information retrieval".IEEE., pp.1-4.