

## «دیدهبانی فناوری» با ابزار «فناوری اطلاعات»

■ کیارش جهانپور

کارشناس ارشد مدیریت فناوری اطلاعات - مدیریت منابع اطلاعاتی  
مؤسسه آموزش عالی فارابی  
jkiarash@yahoo.com

تاریخ دریافت: ۱۳۹۰/۰۴/۲۸  
تاریخ پذیرش: ۱۳۹۰/۰۷/۳۰

### چکیده

اطلاعات موجود در اختراعات و مقالات منبع مناسبی از دانش مدون می‌باشند که نظارت بر روند تکامل این اختراعات «دیدهبانی فناوری» نامیده می‌شود. هدف دیدهبانی فناوری جمع‌آوری فرایند و یکپارچه‌سازی اطلاعات فنی‌ای است که برای بازیگران عرصه اقتصادی مفید می‌باشد. دیدهبانی فناوری هوشیاری و آگاهی را در تمامی سطوح علم و فناوری جهانی از طریق فنونی کارآمد حفظ می‌کند. فنونی قدرتمند مبتنی بر فناوری اطلاعات مانند متن‌کاوی، داده‌کاوی، فن‌کاوی و متن - داده‌کاوی که هم‌اکنون برای شناسایی و استخراج داده‌های مرتبط از متون علم و فناوری وجود دارند که خصوصاً در استنباط عقلایی از داده‌های غیرمتجانس و از هم گسیخته مفید می‌باشند. طی فرایند دیدهبانی فناوری، شاخص‌های مهم دیدهبانی شناسایی شده، اطلاعات لازم برای ردیابی آنها مشخص می‌شود و پس از تحلیل و پردازش این اطلاعات با ابزار فناوری اطلاعات نتایج به صورت گزارش توزیع و نشر می‌شوند و در فرایند تصمیم‌گیری که مقصد غایی این فرایند است استفاده می‌شوند. این مطالعه به شناخت هر چه بهتر و بیشتر ابعاد دیدهبانی فناوری در حوزه‌های مرتبط کمک نموده است. هدف از ارائه این مقاله معرفی مهمترین ابزارها، روش‌ها، فرایند و حوزه‌های مرتبط با دیدهبانی فناوری می‌باشد. در ارائه نتایج مباحث در این مقاله سعی شده است حوزه‌های مستعد برای پژوهش‌های آینده معرفی شوند.

### واژگان کلیدی

دیدهبانی فناوری، متن‌کاوی، پتنت (اختراع)، هوش تجاری، فن‌کاوی، داده‌کاوی.

### مقدمه

طی قرن‌های اخیر علم و فناوری به نیروی پیش‌برنده جوامع و موتور محرکه اقتصاد مبدل شده‌اند [۱]. به ویژه پس از انقلاب صنعتی، علم و فناوری نقش خود را به عنوان پیش‌ران اصلی موفقیت اقتصادی هر روز بیش از گذشته به اثبات رسانده است [۲].

نقشی که فناوری در شکل‌گیری زندگی ما بازی می‌کند و ظهور و بروزی که در یک اقتصاد رقابتی مبتنی بر دانش ایفای نماید بخشی از این واقعیت است [۳]. فناوری اصلی‌ترین عامل تولید ثروت است و می‌تواند مواردی چون ارتقای دانش، سرمایه فکری، استفاده مؤثر از منابع، حفظ و صیانت از منابع طبیعی و سایر عوامل مؤثر در ارتقای استانداردها و کیفیت زندگی نوع بشر را شامل شود [۴].

مهمترین تفاوت دنیای امروز با دیروز سرعت تحول فناوری عنوان شده است. دنیا در حال تغییر است و سرعت این تحولات روز به روز بیشتر می‌شود. به موازات گسترش علم و دانش، ابداعات و ابتکارات روند روزافزونی پیدا نموده و فناوری‌های جدیدی ظهور کرده‌اند.

بشر امروز در دریایی فراهم آمده از انبوه اطلاعات شناور است. با افزایش حجم داده‌ها و آگاهی‌های بشر، دسترسی مستقیم و مؤثر به اطلاعات نیازمند دستیابی به راه‌ها و راه‌کارهایی نوین برای مرور و مدیریت این حجم انبوه از اطلاعات متنی است [۱۱].

مدیران و پژوهشگران به موازات گسترش دامنه اطلاعات، به روش‌ها و ابزارهایی برای مرور سریع و بازیابی و مدیریت اطلاعات و نیز شناسایی به هنگام سمت و سوی تغییرات و روند گوناگونی آن نیاز دارند. تدابیر و تصمیم‌های متعدد و سرنوشت‌سازی در کشورها، دولت‌ها، سازمان‌ها و بنگاه‌های اقتصادی اتخاذ می‌شود که در گرو توان شناخت روند و سمت و سوی تغییرات علم و فناوری است [۱۲].

اولویت‌بندی اقدامات و برنامه‌ریزی از سطح کلان و ملی تا سطح مراکز تحقیق و توسعه و شناسایی فرصت‌های کسب و کار و سرمایه‌گذاری در آینده از این زمره است [۱۳]. در سازمان‌های امروزی توان واکنش در برابر تغییر فناوری یک

شاخص حیاتی به شمار می‌رود و بر این مبنا درک همه جانبه این امر ضرورتی انکارناپذیر یافته است [۴]. بر این اساس روش‌ها، شیوه‌ها و ابزارهای کمی و کیفی متفاوتی برای درک بهتر روند تغییرات فناوری و شناخت سمت و سوی آنها از فراز چارچوب‌های نظری گوناگون پیشنهاد شده یا به‌کار رفته است.

«دیدهبانی یا رصد فناوری»<sup>۱</sup> نیز از جمله این رویکردهاست. دیدهبانی فناوری بنابر تعاریف موجود عبارت است از روشی نظام‌مند برای گردآوری، تبدیل، تجزیه و تحلیل و اکتشاف اطلاعات مفید پژوهش و فناوری برای تصمیم‌گیری راهبردی در هر سازمان و یا بنگاه [۱۴].

دیدهبانی فناوری از هوشمندی تجاری<sup>۲</sup> یا رصد رقابت<sup>۳</sup> جدایی‌ناپذیر است؛ چنانکه این دو واژه اغلب با هم به کار برده می‌شوند و یا این که در بسیاری از مقالات مترادف یکدیگر در نظر گرفته شده‌اند [۱۵].

به هر حال در این پژوهش با توجه به اینکه اهداف فناورانه و علمی بیش از مقاصد تجاری صرف مدنظر می‌باشد از واژه «دیدهبانی فناوری» استفاده خواهد شد.

### مروری بر مبانی نظری و ادبیات موضوع

دیدهبانی فناوری در واقع تجمیع روش‌های مبتنی بر فناوری اطلاعات<sup>۴</sup> و انسانی<sup>۵</sup> است و در صدد فهم جایگاه علم و فناوری در حوزه‌های هدف در سطح جهانی است. از سوی دیگر تا زمانی که یک راه مهم انتشار اطلاعات<sup>۶</sup> علمی و فناوری، متون مربوط به آن دو باشد، تجزیه و تحلیل مستندات و متون فناوری یک جزء مهم و اساسی از فرایند دیدهبانی فناوری خواهد بود.

به علاوه آنکه نیازهای اساسی به وجود آمده در سازمان‌ها، بنگاه‌ها و مراکز تحقیق و توسعه به دسترسی سریع، آسان، مؤثر و مفید به اطلاعات مورد نیاز اهمیت و ضرورت دیده‌بانی فناوری را بیش از پیش مورد توجه قرار داده است. در حال حاضر روش‌های مستحکمی بر پایه فناوری و اطلاعات جهت شناسایی رابطه علم و فناوری و استخراج اطلاعات مورد نیاز و مؤثر از میان متون علمی و مستندات فناوری وجود دارد [۱۶].

در حوزه دیده‌بانی (رصد) فناوری، مطالعات متعددی انجام شده و به نظر می‌رسد با توجه به اهمیت نوآوری و ظرفیت ارتقای کیفیت محصول در شرکت‌ها و بنگاه‌های صنعتی و تجاری، نخستین مطالعات در این حوزه‌ها صورت پذیرفته است. تشخیص تحولات نوآوری در رقبا از جمله مهمترین و بنیادی‌ترین فعالیت‌ها در مدیریت هر سازمان به شمار می‌رود. پیش‌سازهای این طرز تلقی در گزارش پژوهش‌های مورین (۱۹۸۵) و کلین (۱۹۸۵) به چشم می‌خورد [۲۷].

دیدهبانی فناوری اخذ، تجزیه و تحلیل و طبقه‌بندی اطلاعات فنی برای رصد و رشد سازمان‌ها یا بنگاه‌های اقتصادی و تجاری است. دیدهبانی باید برای هر نوآوری فنی یا علمی، خواه به خلق فرصت بیانجامد، خواه تهدید، انجام شود [۱۴]. بدین دلیل در بستر اقتصاد رقابتی و مبتنی بر فناوری اطلاعات و ارتباطات، دیدهبانی فناوری جایگاه ویژه‌ای یافته و با توجه به ارزش راهبردی آن در کسب و کار و نیز همراهی و همبستگی آن با پژوهش‌های علمی مورد توجه قرار گرفته است [۲۷].

دیدهبانی فناوری یکی از اجزاء فرایند نظارت جامع<sup>۷</sup> است که خود ناشی از اثرات تقابل نیاز و فرصت است؛ نیاز به پیگیری تغییرات سریع در

علم و فناوری و فرصت به‌دست‌آمده از پیدایش روزافزون پایگاه داده‌های پیشرفته و مدون که دسترسی به آنها هر روزه با استفاده از شبکه‌های کامپیوتری و لوح‌های فشرده تسهیل می‌شود. امروزه اکثر اکتشافات از طریق پایگاه داده‌های کتاب‌شناسی قابل دسترسی هستند و راهبرد رقبا از طریق آزمون پتنت‌ها یا منابع مشابه قابل نمایش هستند. اطلاعات به یک مزیت راهبردی تبدیل شده است که می‌توان مجموعه دانش مفید و مرتبط از آن استخراج کرد [۳۵].

تمرکز جوامع، کشورها و بنگاه‌های اقتصادی رو به رشد «تحقیق، توسعه و نوآوری» است. مستندات و مدارک به ثبت رسیده از این سه مفهوم آرمانی بدین‌گونه برشمرده می‌شود: از مقالات علمی به ثبت اختراع و ابداع و از ثبت اختراع به محصول نوآورانه<sup>۸</sup>. این امر مسلم است که هرگونه نوآوری در عرصه فناوری باید به ثبت برسد و از سوی دیگر تعیین خطوط و مسیرهای اصلی فناوری و نوآوری، اطلاعات مؤثر و مفیدی از فرصت‌های موجود در بازار آینده و فضای کسب و کار پیش‌رو و نیز آینده فناوری در اختیار خواهد گذاشت.

فناوری بر اساس تعریف آکادمی ملی مهندسی فرایندی است که از طریق آن انسان در راستای تأمین نیازها و نیل به خواسته‌های خود به شکل‌دهی محیط و طبیعت می‌پردازد. بر این اساس فناوری شامل تمامی زیرساختارهای لازم برای طراحی، ساخت و تولید، عملیات و اجرا، تعمیر و بازسازی محصولات می‌شود. فناوری محصول علم است [۲].

بر اساس مطالعات انجام‌شده دیدهبانی فناوری یک روش مدیریت اطلاعات است و مستلزم درک کمی و به‌کارگیری روش‌های کمی ارزیابی نظیر

1. Technology Watch
2. Commercial Intelligence
3. Competition Watch

4. IT based
5. Human based
6. Dissemination

7. Overall Monitoring Process
8. Innovative Product

به آسانی و با سرعت به اطلاعات مورد نیاز خود دست پیدا کند.

### روش‌ها و ابزارهای دیده‌بانی فناوری

روش‌ها و ابزارهایی برای راهنمایی و کمک به دیدمبان فناوری در تحقیقات وجود دارد اما اساساً، امروزه رویکرد استخراج دانش<sup>۱</sup> از طریق خواندن متون حجیم<sup>۲</sup> در دسترس و متداول است [۳۵] که به متن‌کاوی یا متن-داده‌کاوی مشهور می‌باشد.

#### متن‌کاوی

عموماً به فرایند استخراج الگوهای غیردیده‌ی<sup>۳</sup> مورد نظر محقق یا دانش از اسناد متنی غیرساخت‌یافته اشاره دارد [۳۶]. در حقیقت متن‌کاوی یک شاخه کاربردی از دامنه ابزارهای داده‌کاوی است و حوزه‌ای از تحقیقات علمی را اخیراً به خود اختصاص داده است. فنون داده‌کاوی به گونه‌ای طراحی شده‌اند که بر اطلاعات ساختاریافته اعمال شده و تفسیر ارائه دهند و این در حالی است که بیشتر اطلاعات موجود، امروزه به صورت متون غیر ساخت‌یافته می‌باشند. به همین منظور متن‌کاوی براساس روش‌های داده‌کاوی جهت اعمال بر روی متون غیرساخت‌یافته طراحی شده است. متن‌کاوی خود از روش‌های شناخته‌شده‌ای از رشته‌های علمی نظیر داده‌کاوی، یادگیری ماشین، استدلال موردمحور<sup>۴</sup>، آمار و مدیریت دانش برای کمک به محقق جهت به دست آوردن درک و و شناخت حجم انبوه اطلاعات غیرساختار یافته یا نیمه‌ساختاریافته استفاده می‌برد [۳۷].

پژوهش‌های دیده‌بانی دارای دستاوردهای اکتشافی می‌باشند و روند تحقیق در آنها به طور خلاصه شامل گردآوری و بازیابی اطلاعات، پردازش

موسکارولا [۱۱]، یان گرکن [۱۲]، کامئوکا و همکاران [۲۴، ۲۳]، گروئن ولد [۲۵] و دیانا ماینارد و همکاران [۲۶] و بهرنگ قاسمی‌زاده [۳] اشاره نمود. تعدادی از پژوهش‌ها نیز بر به‌کارگیری و کاربست دیده‌بانی فناوری و روش‌های دیده‌بانی اعم از روش‌های مبتنی بر انسان و روش‌های بر پایه فناوری اطلاعات در عرصه مدیریت و برنامه‌ریزی علم و فناوری، بازرگانی، صنعت و کسب و کار تمرکز نموده‌اند که از آن میان پژوهش‌های قاسمی‌زاده [۳]، فورن لین [۲۰]، ویکتور کوالر [۲۷]، مارک ووگلرز [۲۸]، ماینارد و همکاران [۲۶]، کوستف [۲۱ و ۲۲]، گرکن [۴]، یون یون یانگ [۳۰]، ریتانو و همکاران [۳۱]، آه وی تان [۳۲]، آلان پورتر [۳۳] اشاره نمود. با این همه، پژوهشی با هدف دیده‌بانی و یا تعیین مسیرنمای فناوری کسب و کار به طور اخص گزارش نگردیده است. آن چه که از مطالعات گذشته و پیشینه پژوهش بر می‌آید آن است که اولاً دیده‌بانی فناوری بنابر تعریف اخذ، تجزیه و تحلیل، تفکیک و طبقه‌بندی و استفاده از اطلاعات فنی است که برای رشد و یا بقای شرکت‌ها و سازمان‌ها مؤثر می‌باشد (اسکورسا، ۲۰۰۱)، دوم اینکه دیده‌بانی باید در مورد هرگونه نوآوری علمی و فنی که ممکن است به خلق فرصت و ایجاد تهدید منجر شود، انجام شود، سوم این که دیده‌بانی فناوری فرایندی ممتد و مداوم است، چهارم دیده‌بانی فناوری حاصل تجمیع هر دو روش مبتنی بر انسان و فناوری اطلاعات است، پنجم روش‌های مبتنی بر فناوری اطلاعات که به طور اعم داده‌کاوی، متن- داده کاوی و در خصوص فناوری و مستندات به ثبت رسیده فناوری، فن‌کاوی یا Tech mining عنوان می‌شوند این توانایی را به بشر بخشیده‌اند که بتواند از میان حجم انبوه دریای داده‌ها و اطلاعات

کتاب‌سنجی<sup>۱</sup>، علم‌سنجی<sup>۲</sup>، رسانه‌سنجی<sup>۳</sup> و وب‌سنجی<sup>۴</sup> است. این روش‌ها به همراه روش‌هایی نظیر جامعه‌سنجی، روان‌سنجی و ... بسته به نوع تقاضا و فناوری مورد توجه استفاده خواهند شد. تاگو و سات کلیف<sup>۵</sup> در سال ۱۹۹۲ بر این اساس تعریفی از کتاب‌سنجی و علم‌سنجی مبتنی بر دورنمایی از مدیریت راهبردی و تحلیلی اطلاعات ارائه نموده‌اند. براساس این رویکرد کتاب‌سنجی به مطالعه جنبه‌ها و ابعاد کمی تولید، تفکیک و استفاده از اطلاعات به ثبت رسیده و توسعه مدل‌های ریاضی و سنجه‌هایی که ممکن است برای تعیین پیش‌آگهی و یا اخذ تصمیم حول یا در مورد برخی فرایندها به کار آید، اطلاق می‌گردد. کتاب‌سنجی در حقیقت سنجش متون مدارک و مستندات و سایر رسانه‌های ارتباطی است. علم‌سنجی نیز مطالعه ابعاد کمی علم است. علم‌سنجی بخشی از جامعه‌شناسی علم است و این روش توسط تاگو و سات کلیف برای توسعه سیاست‌گذاری علمی در سال ۱۹۹۲ به‌کار گرفته شده است. رویکرد کیفی در زمینه رابطه علم و فناوری در مطالعات الکانا و همکارانش در سال ۱۹۷۸ و دیگران سابقه‌دار است اما به نظر می‌رسد که در سال‌های قبل از دهه ۸۰ میلادی عمده مطالعات این حوزه با رویکرد کیفی از منظر جامعه‌شناسی علم بوده باشد.

[۳۳، ۳۲، ۳۱، ۱۱، ۱۶، ۴، ۲۷، ۲۶، ۲۱]

بسیاری از مطالعات و گزارش پژوهش‌ها بر نیاز به دیده‌بانی (رصد) و فن‌کاوی متمرکز شده‌اند، نظیر یون هسین تسنگ و همکاران [۱۸]، میشل فتوری و همکاران [۱۹]، فورن لین و همکاران (۲۰) و ... . تعدادی از پژوهش‌ها متوجه روش‌شناسی دیده‌بانی فناوری هستند که می‌توان به تحقیقات کوستف [۲۲، ۲۱، ۱۶]، پژوهش

1. Bibliometric  
2. Scientometric  
3. Mediametrics

4. Webmetric  
5. Tague & Sutcliffe  
6. Extracting Knowledge

7. Voluminous Texts  
8. non-trivial Patterns  
9. Case based reasoning



شکل ۱- مراحل متن کاوی

اطلاعات و یکپارچه سازی اطلاعات می باشد. در گام اول پژوهشگر یا دیدبان ضمن مشخص نمودن پایگاه داده های مورد نظر، اطلاعات استخراج شده از پایگاه داده ها را وارد نرم افزار واسطی که برای این امر طراحی شده است، می کند و پس از آنالیز اطلاعات مدارک فناوری، نتایج آماری اشکال مناسب و در قالب جداول و نمودارها با استفاده از روش های دیدهبانی فناوری خصوصاً متن کاوی و متن- داده کاوی<sup>۱</sup> عرضه خواهد شد. به طور خلاصه می توان مراحل متن کاوی را به صورت زیر شرح داد:

بازیابی اطلاعات<sup>۲</sup>: در این مرحله تمامی اطلاعات مربوط به موضوع مورد مطالعه جمع آوری می شود. در حقیقت یافتن اطلاعات مربوط به پرسش کاربر در میان انبوه اسناد موجود در این مرحله انجام می شود. در این مرحله اسناد بسیاری بازیابی شده که برای انسان مطالعه هریک از آنها مقدور نبوده است. در این مرحله اطلاعات نامربوط به پرسش کاربر نیز در سطح وسیعی در میان اطلاعات جمع آوری شده یافت می شود اما به طور کلی مرحله بازیابی اطلاعات دارای این هدف است که از میان مجموعه ای از مدارک و متون، آنهایی که به نیاز اطلاعاتی کاربر مرتبط ترند، انتخاب شوند.

استخراج و خلاصه سازی اطلاعات<sup>۳</sup>: در این مرحله اطلاعات و رویدادهایی که برای کاربر مهم و مورد علاقه است، استخراج می شود. می توان این مرحله را نوعی بازیابی اطلاعات دانست که براساس علاقه و نیاز کاربر مجدداً انجام می شود. یعنی اطلاعاتی که مرتبط با پرسش کاربر بوده در مرحله بازیابی جمع آوری شده و در این مرحله اطلاعات مرتبطی که مورد نظر کاربر است انتخاب می شود. در این مرحله دامنه انتخاب محدودتر

شده و از فناوری های پیشرفته ای نظیر NLP<sup>۴</sup> برای این گزینش استفاده می شود. داده کاوی<sup>۵</sup>: در این مرحله که اطلاعات مورد نظر کاربر جمع آوری شده است با استفاده از فنون داده کاوی ارتباط و همبستگی بین اطلاعاتی که از متون مستقل و متفاوتی جمع آوری شده اند ارائه می شود. طبقه بندی خلاصه سازی، تحلیل روند و دسته بندی کردن و... از جمله الگوریتم های متداولی هستند که برحسب نیاز کاربر و توانایی های نرم افزارها در این مرحله می توانند اجرا شوند. مراحل متن کاوی در شکل ۱ به صورت مختصر شرح داده شده است.

استفاده از نرم افزارهای رایانه ای در فرایند متن کاوی به نوعی قدرت اصلی این ابزار را در سرعت و دقت پردازش متون باعث شده است. از این رو شناخت نرم افزارهای رایانه ای متن کاوی، خصوصیات و توانایی های آنها لازم به نظر می رسد. ■ نرم افزارهای رایانه ای متن کاوی:

نرم افزارهای موجود در این حوزه به علت گستردگی، پیچیدگی و پیشرفت روزافزون الگوریتم های متن کاوی، دارای قابلیت های گوناگونی می باشند. قالب های گوناگونی از داده ها را به عنوان ورودی می پذیرند و خدماتی که به عنوان نتیجه و خروجی ارائه می دهند متفاوت است و به همین دلیل تقسیم بندی های گوناگونی می توان از آنها ارائه داد، لذا برای استفاده از این

1. Text-data mining  
2. Information Retrieval  
3. Information Extraction

4. Natural Language Processing  
5. Data Mining  
6. Open Source

7. Co-occurrence  
8. Federated

حضور گسترده ابزارهای متن‌کاوی در حوزه دیده‌بانی و به خصوص دیده‌بانی در حوزه‌های تجاری و کسب و کار باعث شده است که شرکت‌های مختلفی با تولید ابزارهای مناسب و قدرتمند برای برطرف کردن نیاز کسب و کارها به ابزار مناسب متن‌کاوی و داده‌کاوی به خصوص با رویکرد پیش‌بینی آینده و هوش کسب و کار روی آورند. در ادامه تعدادی از شناخته شده‌ترین ابزارهای متن‌کاوی که به صورت تجاری عرضه شده‌اند، معرفی می‌شوند:

**ClearForest**: این برنامه، خدماتی نظیر برچسب زدن<sup>۱</sup> پیشرفته متن، پلتفرمی پیشرفته برای خلاصه‌سازی متون، تحلیل متون و یک محیط توسعه یافته برای کاربر ارائه می‌دهد. از آنجایی که این برنامه قابلیت پذیرفتن قالب اطلاعاتی HTML را دارد در تحلیل متون و اخبار اینترنتی به طور گسترده‌ای مورد استفاده قرار می‌گیرد. از مزایای دیگر این برنامه این است که بر روی اطلاعات ساختاریافته و غیرساختاریافته می‌تواند به کار گرفته شود [۳۰].

#### IBM SPS (Statistics

: Package for Social Sciences)

این نرم‌افزار که البته بیشتر به علت قابلیت‌هایش در ارائه تحلیل‌های آماری به ویژه در پژوهش‌های علوم اجتماعی شهرت دارد، یکی از ابزارهای داده‌کاوی و متن‌کاوی نیز به شمار می‌رود.

این برنامه قابلیت مرتبط کردن مفاهیم<sup>۲</sup>، استخراج اطلاعات و دسته‌بندی آنها را داراست [۴۴]. بر اساس گزارش سالانه داده کاوش رکسر<sup>۳</sup>، این برنامه بیشترین رضایت‌مندی را در بین کاربران در سال‌های ۲۰۰۹ و ۲۰۱۰ داشته است [۴۵].

حجم گسترده‌ای از متون و تصاویر، افزونه‌های متنوع و متعدد که قابلیت کار روی زبان‌های مختلف (هندی، عربی و...) و کار روی پایگاه‌های صوتی و... را دارند، نام برد. این نرم‌افزار اسناد را در قالب‌های گوناگونی مانند RTF، XML، HTML، SGML و پست‌های الکترونیک پشتیبانی می‌کند.

#### Rapid Miner

این نرم‌افزار یکی از مؤثرترین نرم‌افزارهای متن‌باز در حوزه داده‌کاوی، متن‌کاوی و مدل‌سازی است که حاوی بیش از ۴۰۰۰ عملگر و ماژول داده‌کاوی می‌باشد. رپید ماینر عملگرهای انعطاف‌پذیر زیادی برای داده‌های ورودی و خروجی در قالب‌های مختلف مانند فایل‌های excel، SPSS و مجموعه داده‌های پایگاه داده‌های شناخته‌شده‌ای نظیر MySQL، Microsoft SQL Server، Sybase، Oracle، PostgreSQL، dBase در اختیار دارد. همچنین از قالب‌های پراکنده و غیرمتداولی نظیر Arff، standard datacsv، mySVM، SVMight و C4.5 پشتیبانی می‌کند. این برنامه نیز با زبان برنامه‌نویسی جاوا تهیه شده و قابلیت استفاده از افزونه‌های گوناگون طراحی شده برای اهداف مختلف را داراست [۴۰، ۴۱، ۴۲].

**Carrot<sup>2</sup>**: این برنامه کد باز نیز یک موتور دسته‌بندی نتایج جستجوها محسوب می‌شود. این برنامه می‌تواند به صورت خودکار مجموعه‌های کوچکی از اسناد و نتایج جستجوها را دسته‌بندی موضوعی نماید. لذا از این برنامه بیشتر برای دسته‌بندی اطلاعات استفاده می‌شود تا پیاده‌سازی الگوریتم‌های دیگر. این برنامه نیز در زبان برنامه‌نویسی جاوا توسعه یافته است [۴۳].

■ نرم‌افزارهای تجاری داده‌کاوی برای دیده‌بانی فناوری:

این ابزارها بر اساس حقوق مالکیت آنها قابلیت تهیه، هزینه و توسعه ابزارها برای محقق را بیشتر مدنظر قرار دهیم و سپس بعد از تفکیک ابزارها به دو دسته متن‌باز و تجاری قابلیت‌های فنی هر یک را بررسی کنیم.

#### ■ نرم‌افزارهای متن‌باز ابزار داده‌کاوی برای دیده‌بانی فناوری

به طور کلی نرم‌افزارهای متن‌باز، نرم‌افزارهایی هستند که کدهای سازنده آنها منتشر شده و در دسترس عموم قرار گرفته است و این اجازه را به هر شخص داده که این کدها را بدون پرداخت هیچگونه مبلغی به عنوان هزینه یا حق اختراع، نسخه‌برداری، تغییر یا توزیع مجدد کند [۳۹].

رایگان بودن و قابلیت تغییر در کدهای برنامه‌های متن‌باز از جمله مهمترین مزایای آنها محسوب می‌شود؛ زیرا تهیه این ابزارها را تسهیل می‌کند و از سوی دیگر به محقق این اجازه را می‌دهد که مطابق با نیاز تحقیق، چارچوب تئوریک مورد نظر و مورد مطالعه در ساختار ابزار دخالت کند و به عبارتی ابزار را برای اهداف تحقیق خویش بهینه کند. نمونه‌های زیر به عنوان کاربردی‌ترین نرم‌افزارهای متن‌باز متن‌کاوی می‌باشند:

**GATE (General Architecture for Text Engineering)**: این نرم‌افزار یک بسته نرم‌افزاری مبتنی بر زبان جاوا در حوزه متن‌کاوی است که توسط دانشگاه شفیلد انگلستان مدیریت می‌شود. از قابلیت‌های اساسی این نرم‌افزار می‌توان به استفاده از سیستم بسیار پیشرفته در فرایند خلاصه‌سازی اطلاعات (IE)، وجود گروه پشتیبانی متخصص و انجمن کاربران آنلاین پرتعداد، توانایی کار روی

1. Tagging

2. Concepts linking

3. Rexer's Annual Data Miner Survey



RefViz: این برنامه برای تحلیل متون و بصری سازی داده‌ها به کار می‌رود و در آن قابلیت‌های فراوانی برای تجزیه و تحلیل زبانی و آماری قرار داده شده است. لازم به ذکر است این برنامه صرفاً بر روی متون ساختاریافته کاربری دارد و بیشتر بر روی پایگاه داده‌های ISI در وب مانند Pub Med به کار گرفته می‌شود. این برنامه که یک نمای بصری خلاصه از حجم وسیعی از مقالات یا متون ارائه می‌دهد، از یک الگوریتم ریاضی برای دسته‌بندی کردن مقالات بر اساس مفاهیم پایه‌ای آنها بهره می‌برد [۴۶].

بدین‌گونه می‌توان ابزارهای معرفی شده در هر دودسته متن‌باز و تجاری را جزء شناخته‌شده‌ترین ابزارهای متن‌کاوی دانست. البته از آنجایی که در این زمینه ابزارهای بسیار زیادی با کارکردهای گوناگون وجود دارد و ارزیابی فنی ابزارها خود موضوعی تخصصی است، در این مقاله مجال بررسی بیشتر ابزارها نیست.

### فرایند دیده‌بانی فناوری

کاربرد اصلی دیده‌بانی فناوری، فراهم کردن اطلاعات فنی برای تصمیم‌گیری در بخش تولید یک شرکت می‌باشد. اگرچه فرایند دیده‌بانی بر فرایند تصمیم‌گیری تجاری نیز اعمال می‌شود. البته در مباحث تجاری در کنار نام دیده‌بانی به عباراتی نظیر دیده‌بانی تجاری<sup>۱</sup>، دیده‌بانی رقابت<sup>۲</sup> یا دیده‌بانی پیرامونی<sup>۳</sup> نیز برمی‌خوریم.

در کل دیده‌بانی فناوری به برنامه‌ریزی راهبردی<sup>۴</sup> شرکت بسیار مربوط است. برنامه‌ریزی راهبردی یک چارچوب کاری متمرکز بر فعالیت‌ها از طریق تعریف شاخص‌های حیاتی<sup>۵</sup> که باید دیده‌بانی شوند ارائه می‌دهد. اولین فاز از فرایند دیده‌بانی عبارت از قسمت‌های زیر است:

۱- تعیین و تحلیل نیازهای اطلاعاتی سازمان و تعریف شاخص‌های دیده‌بانی حیاتی؛ به طور کلی دو نوع نیاز اطلاعاتی خارجی برای یک شرکت متصور است:

الف) نیازهای اطلاعاتی فناورانه یعنی دانستن تغییرات فناورانه که بسیار سریع می‌تواند بر شرکت تأثیر بگذارد.

ب) نیازهای اطلاعاتی تجاری یعنی دانستن آنچه رقبا انجام می‌دهند و دانستن هر نوع تغییرات در محیط تجاری.

از طریق این دو دسته نیازهای نام‌برده، می‌توانیم شاخص‌های حیاتی دیده‌بانی (CWF)

را به عنوان شاخص‌های خارجی شرکت که به طور قابل ملاحظه‌ای بر قدرت رقابت شرکت تأثیر می‌گذارد، تعیین کنیم. باید به این مهم توجه داشت که هر فعالیت در هر بخش یک سازمان، CWFهای مخصوص به خود را دارد که عمدتاً به برنامه راهبردی تعریف شده در آن شرکت وابسته است. برای طراحی یک دیده‌بانی فناوری مناسب، CWFها باید به طور کاملاً تخصصی و ویژه هر بخش و فعالیت تعریف شوند.

۲- جستجو و فراهم آوردن اطلاعات لازم برای ردیابی CWF: بعد از اینکه CWFها را شناسایی کردیم باید منابع اطلاعاتی که برای ردیابی و شناسایی مستمر رویه‌ها مورد استفاده قرار می‌گیرند، شناسایی شوند [۱۵].

امروزه بسیاری از منابع اطلاعاتی به صورت آنلاین در دسترس هستند اما به هر حال باید وبسایت‌هایی را که مورد استفاده قرار می‌دهیم تجزیه و تحلیل کنیم و قابلیت اعتماد، اعتبار و نوبت‌های به‌روز شدن اطلاعات در این سایت‌ها را مورد ارزیابی قرار دهیم. منابعی مانند:

- اتحادیه اروپا؛ این پورتال اجازه دسترسی به

گستره وسیعی از اطلاعات مربوط به قدرت‌های مختلف اروپایی را به زبان‌های مختلف می‌دهد (اطلاعات مربوط به مجالس، کنگره‌ها، کمیسیون‌ها و ...)

- اداره‌های ثبت اختراعات؛ این ادارات پایگاه داده‌ای دارند که sp@acenet نامیده می‌شود.

این پایگاه‌ها دسترسی به اطلاعات مربوط به اختراعات ثبت شده در کل اروپا و جهان را میسر ساخته‌اند. گذشته از این، دسترسی رایگان به اختراعات ثبت شده در آمریکا و ژاپن نیز فراهم شده است. هر چند که متون مربوط به این اختراعات همیشه کامل نمی‌باشد.

- CORDIS: بکارگیری و ساختاردهی کلیه اطلاعات در مورد تحقیق و توسعه در اتحادیه اروپا در این پورتال انجام شده و در دسترس عموم قرار دارد. دسترسی به این سایت ساده و سریع بوده و مشاوره در زبان‌های مختلف و به صورت رایگان ارائه می‌شود.

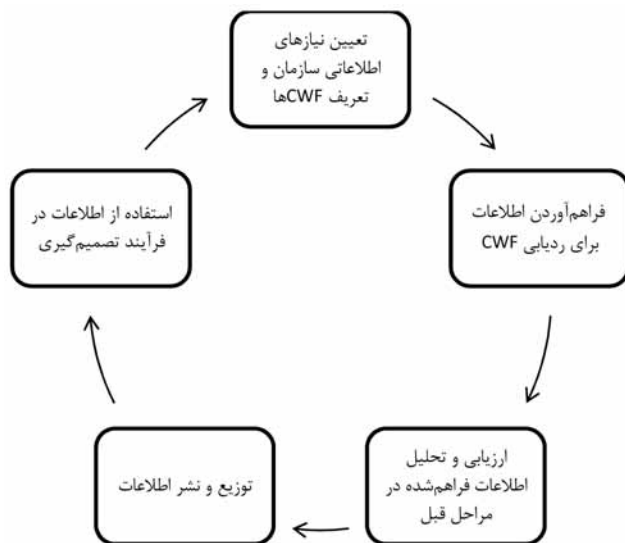
- کتاب‌شناسی: کتابخانه انگلستان و کتابخانه ملی اسپانیا دو مورد از شناخته شده‌ترین مؤسسات داده‌های کتاب‌شناسی و خلاصه‌سازی اسناد هستند.

گذشته از منابع نام برده که کاربرد بیشتری در جمع‌آوری اطلاعات مورد نیاز فرایند دیده‌بانی دارد، منابعی نظیر بولتن‌های اداری و سازمان‌های اداری نیز برای این منظور استفاده می‌شوند [۴۷].

۳- ارزیابی و تحلیل اطلاعات فراهم شده در مراحل قبلی: برای تصمیم‌گیری، صرفاً داشتن اطلاعات مناسب کافی نیست. این اطلاعات باید ارزیابی و تحلیل شوند تا آنچه مرتبط‌تر است انتخاب شود و سپس برای درک و پردازش ارائه شود. فراهم نمودن حجم عظیمی از داده‌ها در

1. Commercial Watch  
2. Competition Watch  
3. Surrounding Watch

4. Strategic Planning  
5. Critical Factors



شکل ۲- چرخه فرایند دیده‌بانی فناوری

پژوهشگران باشد. آنچنان که در ابزار متن‌کاوی تشریح شد، آخرین مرحله از پیاده‌سازی متن‌کاوی، داده‌کاوی است که محقق بنابر هدف و نیاز خویش می‌تواند خروجی یا خروجی‌های مورد نظر خویش را به دست آورد. مانند دسته‌بندی، تحلیل روند، طبقه‌بندی و ... این در حالی است که در حوزه‌های مختلف مدیریتی پیش‌بینی روند با استفاده از ابزارهای کلاسیک نظیر روش‌های آماری و ریاضی رواج داشته است اما استفاده از روش دیده‌بانی با ابزار متن‌کاوی در تحلیل روند در حوزه‌های مدیریتی به موارد معدودی محدود است. به عنوان مثال در پیش‌بینی روند قیمت‌ها که از موضوعات محوری در حوزه مدیریت مالی می‌باشد، استفاده از ابزارهای دیده‌بانی نظیر متن کاوی صرفاً به کارهایی نظیر میترمایر [۴۸] و فونگ [۴۹] محدود شده است و به نظر می‌رسد

طور اخص حاوی روش‌هایی برای جمع‌آوری، خلاصه‌سازی و دسته‌بندی اطلاعات متنی است و از سوی دیگر، دیده‌بانی فناوری بیشتر از اطلاعات متنی علمی و اسناد اختراعات (پتنتها) بهره‌میرد، لذا متن‌کاوی به عنوان یک ابزار کارآمد در این عرصه مورد استفاده قرار می‌گیرد.

همچنین با توجه به چرخه فرایند دیده‌بانی فناوری می‌توان به این مهم پی‌برد که «اطلاعات» در مرکز تمامی گام‌های دیده‌بانی و ابزارهای آن قرار دارد. این نتیجه بیانگر اهمیت ماهیت، اعتبار اطلاعات و منبع اطلاعاتی مورد استفاده در دیده‌بانی می‌باشد. این در حالی است که هیچ پژوهش درخوری پیرامون سنجش اعتبار منابع اطلاعاتی و روش‌های ارزیابی اعتبار این گونه منابع ارائه نشده و این نوع بررسی‌ها با توجه به کاربردشان در فرایند دیده‌بانی، می‌تواند حوزه تحقیقی مناسبی جهت فعالیت این دست از

صورتی که از پردازش آنها برای استفاده ناتوان باشیم، بی‌فایده است [۱۵]. از آنچه گفته شد می‌توان فهمید که در حقیقت در این مرحله است که ابزارهای اصلی دیده‌بانی فناوری نظیر داده‌کاوی، فن‌کاوی و متن‌کاوی بیشترین استفاده را دارند زیرا با حجم انبوهی از اطلاعات در این مرحله برخورد می‌کنیم و انتخاب اطلاعات مرتبطتر، درک و پردازش اطلاعات و تحلیل آنها همگی از توانایی‌های فنون متن‌کاوی و داده‌کاوی است.

### توزیع داخلی نتایج

**استفاده از اطلاعات در فرایند تصمیم‌گیری**  
این پنج مرحله به صورت پیوسته و ادواری اجرا می‌شوند و اغلب تصمیم‌های اتخاذ شده در مرحله نهایی دلالت بر وجود CWF جدید دارد که این امر موجب آغاز یک چرخه جدید می‌شود [۱۵]. در شکل ۲ فرایند دیده‌بانی فناوری، بنا بر آنچه گفته شد به صورت یک سیکل به صورت مختصر نشان داده شده است.

### نتیجه‌گیری

به‌طور خلاصه دیده‌بانی فرایندی تعریف‌گردید که طی آن روشی نظام‌مند برای گردآوری، تبدیل، تجزیه و تحلیل و اکتشاف اطلاعات مفید پژوهش و فناوری برای تصمیم‌گیری راهبردی در هر سازمان و یا بنگاه به اجرا در می‌آید. این فرایند به‌طور کلی به دو صورت مبتنی بر فناوری اطلاعات و یا مبتنی بر انسان، قابلیت اجرایی دارد که در این مقاله با توجه به موضوع که ابزار فناوری اطلاعات را مد نظر داشته به توضیح ابزارهای فناوری اطلاعات در این حوزه یعنی متن‌کاوی پرداختیم. از آنجایی که متن‌کاوی به

References

1. Lichtenhaler, E, (2004). Technological change and the technology intelligence process: a case study. *Journal of engineering and Technology management*, 21, 4.
2. Available from: [www.members.nae.edu/nae/techlithome.nsf/weblinks/KGRG-55A3ER?OpenDocument](http://www.members.nae.edu/nae/techlithome.nsf/weblinks/KGRG-55A3ER?OpenDocument) [Accessed 16 Jan 2011]
3. QasemiZadeh, B, (2010). Towards Technology Structure Mining from Text by Linguistics Analysis. *DERI - DIGITAL ENTERPRISE RESEARCH INSTITUTE*. October 2010, 2-8.12.
4. Gerken JM, Moehrle MG, and Walter L, (2010). Patents as an information source for product forecasting: Insights from a longitudinal study in the automotive industry. *The R&D Management Conferece 2010*. RADAMA Manchester, UK, 2010.3-9.
5. Kostoff, RN. (2003). Science and technology text mining: global technology watch. Office of Naval Research (ONR), [online], <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA415863> [Accessed 5 November 2010].
6. Mahmoudzadeh, L. Managing the future by coming technologies. Summary of book available from: [www.imi.ir/fa-ir/SummerayManagementBooks/techM.pdf](http://www.imi.ir/fa-ir/SummerayManagementBooks/techM.pdf) [20 November 2010].
7. The McKinsey Quarterly, 2007, eight business technology trends to watch, McKinsey Global Survey. Available from: [http://prism3.iprismglobal.net/Data/tsumon/iPGMarketing/Mckinsey\\_Trends.pdf](http://prism3.iprismglobal.net/Data/tsumon/iPGMarketing/Mckinsey_Trends.pdf) [20 November 2010].
8. Andriole, S. (2006). Business Technology Education in the early 21st Century: The Ongoing Quest for Relevance. *Journal of information technology education*. vol 5. p1-12
9. Schell, J. (2002). Trying IT to the business: using BTM to appropriately manage IT outsourcing. University of Maryland University College. Available from: [http://artandpenny.com/family/john/ITSM637/Schell\\_Business\\_Technology\\_](http://artandpenny.com/family/john/ITSM637/Schell_Business_Technology_)

افزایش توان، سرعت عمل و اعتبار دیده‌بانی مفید  
 فایده باشد و افق‌های جدیدی پیش روی  
 دیده‌بانان فناوری بگشاید.

Management.pdf[20 November 2010].

10. Hoque, F. (2001, May 2). *Aligning Business and Technology: The Principles of Business Technology Management (BTM)*. available from: [http://www.enamics.com/pdf/wp\\_btm.pdf](http://www.enamics.com/pdf/wp_btm.pdf) [24 December 2009]
11. Moscarola, J, Baulac, Y, Bolden, R.( 1998). *Technology Watch via Textual Data Analysis*. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.113.6548&rep=rep1&type=pdf> [20 November 2010]
12. Technology Futures Analysis Methods Working Group. (2004). *Technology futures analysis: Toward integration of the field and new methods*. *Technological Forecasting and Social Change*, 71, 3, 287-303.
13. Yoon, B. and Park, Y. (2005). A systematic approach for identifying technology opportunities: Keyword-based morphology analysis. *Technological Forecasting and Social Change*, 72, 2, 145-160.
14. Pere Escorsa Castells. De la vigilancia tecnológica a la inteligencia competitiva en las empresas, 2001a. available from: [http://www.uoc.es/web/esp/art/uoc/escorsa0202/escorsa0202\\_imp.html](http://www.uoc.es/web/esp/art/uoc/escorsa0202/escorsa0202_imp.html) [20 November 2010]
15. Rovira, C. (2008). *Technology Watch and Competitive intelligence for SEM-SEO*. Universitat pompeu fabra, HYPERTEXT.NET. Available from: <http://www.hipertext.net/english/pag1032.htm> [20 November 2010]
16. REPORT DOCUMENTATION PAGE. (2003). Available from: <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA458869> [20 November 2010]
17. Hwan Suh, J, Hoon Park, Ch, Hoon Park, S. (2010). Applying text and data mining techniques to forecasting the trend of petitions filed to e-People. *Expert Systems with Applications*. 37, 7255 -7268.
18. Zhang, CH. (2009) *Extracting chinese-english bilingual core*

به طور کلی تحلیل روند در حوزه‌های گوناگون  
 با ابزارهای مبتنی بر فناوری اطلاعات می‌تواند  
 بستری مناسب برای انجام پژوهش‌های جدید  
 باشد.

با توجه به نقش محوری اطلاعات در فرایند  
 دیده‌بانی و افزایش دسترسی به اطلاعات در اثر  
 استفاده از وب و فناوری‌های نوین می‌توان با  
 استفاده از دیده‌بانی و ابزارهای آن اطلاعات  
 بیشتری را در زمان کمتری تحلیل نموده و در  
 این حال تصمیماتی که با استفاده از این حجم  
 از اطلاعات مرتبط گرفته می‌شوند به واقعیت  
 رویدادها نزدیک‌تر بوده و مدیران را یاری‌رسان  
 خواهد بود. اما با این حال پژوهشی در زمینه  
 اندازه‌گیری، ارزیابی یا بررسی میزان موفقیت یا  
 عدم موفقیت تصمیماتی که بنا بر نتایج دیده‌بانی  
 اتخاذ شده‌اند، انجام نگرفته است. مسلماً این نوع  
 پژوهش‌ها و نتایج آنها برای تصمیم‌گیری مدیران  
 در استفاده یا عدم استفاده از دیده‌بانی در  
 سازمان‌ها و بنگاه‌ها نقش مؤثری ایفا خواهد نمود  
 و به نوعی ارزیابی اعتبار فرایند دیده‌بانی نیز تلقی  
 می‌شود و می‌تواند بازخورد مناسبی به طراحان  
 روش‌های دیده‌بانی ارائه دهد و تکامل فنون مورد  
 استفاده در این زمینه را موجب شود.

در بررسی فرایند دیده‌بانی فناوری به این نکته  
 اشاره شد که تعریف دقیق و مناسب CWFها  
 از مهمترین عوامل توفیق و اعتبار فرایند دیده‌بانی  
 می‌باشند. این در حالی است که در بخش‌ها،  
 سازمان‌ها و تحقیقات گوناگون این CWFها  
 می‌توانند به شکل‌های گوناگونی تعریف شوند.  
 لذا به نظر می‌رسد برای ارائه مناسب تعریف  
 CWFها در دیده‌بانی‌های مختلف نیاز به اصول،  
 چارچوب، روش‌ها یا چک‌لیست‌های استاندارد  
 داریم. لذا پژوهش در این زمینه نیز می‌تواند برای



- terminology from parallel classified corpora in special domain. IEEE Computer Society, 9, 271-274.
19. Fattori, M., Pedrazzi, G., Turra, R. (2003). Text mining applied to patent mapping: a practical business case. *World Patent Information*, 25, 335-342.
20. Lin, F., Wei, Ch., Lin, Y. (2008). DERIVING TECHNOLOGY ROADMAPS WITH TECH MINING TECHNIQUES. Available from: [www.pacis-net.org/file/2008/PACIS2008\\_Camera-Ready\\_Paper\\_255.pdf](http://www.pacis-net.org/file/2008/PACIS2008_Camera-Ready_Paper_255.pdf) [20 November 2010]
21. Kostoff, R., DeMarco, R. (2001). SCIENCE AND TECHNOLOGY TEXT MINING: ANALYTICAL CHEMISTRY. Information for defense community. Available from: <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA415945> [2 January 2011].
22. Kostoff, R. (2003). SCIENCE AND TECHNOLOGY TEXT MINING: MANAGEMENT DECISION AIDS. Available from: <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA415501> [2 January 2011]
23. Kameoka, A. (2003). Road-mapping for corporate strategy: A Japanese case study involving Delphi-scenario. Available from: [http://www.techmonitor.net/techmon/03jul\\_aug/tm/pdf/03jul\\_aug\\_sf4.pdf](http://www.techmonitor.net/techmon/03jul_aug/tm/pdf/03jul_aug_sf4.pdf) [20 November 2010]
24. Kameoka, A. and Y. Morikawa. A Forecast of a Business Product Development Process and the Innovation Strategy Planning by a Delphi-Scenario Writing Method. The 8th Technology Forecasting Symposium, the Proceedings. 19-24, 1979.
25. Groenveld, P. (1997). Roadmapping Integrates Business and Technology. *Research Technology Management*, September-October, 48-55.
26. Maynard, D et al. Ontology-based Information Extraction for Business Intelligence. Proceedings of the 6th international The semantic web. Lecture Notes in Computer Science. Available from: <http://www.springerlink.com/content/u38k622321825762> [20 november 2010]
27. Cavaller, V. (2007). LIMITS OF ANALYTICAL INSTRUMENTS IN TECHNOLOGY WATCH: APPLICATION OF A PRACTICAL CASE IN BIOTECHNOLOGY. Available from: [http://www.iadis.net/dl/final\\_uploads/200703L026.pdf](http://www.iadis.net/dl/final_uploads/200703L026.pdf) [20 novemebr 2010]
28. Veugelers, m., Bury, J., Viaene, S. (2010). Linking technology intelligence to open innovation. *Technological Forecasting & Social Change*. vol 77, issue 2, 335-343.
29. Cowan, R. et al. (2002). Software engineering technology watch. *Information Sciences*. Vol140, Issues 3-4, Pages 195-215.
30. Yang, YY, et al. (2008). Text mining and visualization tools - Impressions of emerging capabilities. *World Patent Information*. Vol30, Issue 4, Pages 280-293.
31. Reitano, A., Di Maio, F., Semirana, S. (2005). Data mining and Management Decisions. International conference on computer systems and technologies. Available from: <http://ecet.ecs.ru.acad.bg/cst05/Docs/cp/SII/II.7.pdf> [20 November 2010].
32. Tan, Ah. (1999). Text Mining: The state of the art and the challenges. In Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases. page, 65-70
33. Porter, A. (2007) Tech Mining to Drive Open Innovation. Proceedings of the First International Conference on Technology Innovation, Risk Management and Supply Chain Management (TIRMSCM 2007, Beijing), Universe Academic Press, Toronto, page 1-13
34. Yoon, B., Yongtae, P. (2004). A text-mining-based patent network: Analytical tool for high-technology trend. *The Journal of High Technology Management Research*. vol15, (1), pp.37-50
35. Jacquen, F., LARGERON, Ch. (2004). Discovering Unexpected Information for Technology Watch. pp 219-230. Springer-Verlag Berlin Heidelberg. vol.2004
36. Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996). From data mining to knowledge discovery: An Overview. In *Advances in Knowledge Discovery and Data Mining*. U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, eds., MIT Press, Cambridge, Mass., 1-36.
37. Text Mining and its Applications: Results of the NEMIS Launch Conference (Studies in Fuzziness and Soft Computing) by Spiros Sirmakessis (Feb 20, 2004).
38. Micah J. Crowsey, Amanda R. Ramstad, David H. Gutierrez, Gregory W. Paladino, and K. P. White, Jr., Member, IEEE, An Evaluation of Unstructured Text Mining Software
39. <http://www.lucidimagination.com/solutions/value-of-open-source>
40. <http://it.toolbox.com/wiki/index.php/RapidMiner>
41. <http://rapid-i.com/content/view/181/190/>
42. <http://rapid-i.com/content/view/181/196/>
43. <http://project.carrot2.org/>
44. An Evaluation of Unstructured Text Mining Software, Micah J. Crowsey, Amanda R. Ramstad, David H. Gutierrez, Gregory W. Paladino, and K. P. White, Jr., Member, IEEE
45. Karl Rexer, Heather Allen, & Paul Gearan (2010) 2010 Data Miner Survey Summary, presented at Predictive Analytics World, Oct. 2010.
46. <http://www.gtppc.org/gtppc/refviz.htm>
47. Report produced for the EC funded project, INNOREGIO: Business Intelligence Technology Watch: dissemination of innovation and knowledge management techniques, by LEIA Technological Development Center. MARCH 2000.
48. Marc-André Mittermayer, Forecasting Intraday Stock Price Trends with Text Mining Techniques, Proceedings of the 37th Hawaii International Conference on System Sciences - 2004
49. Y.-C. Phung, Text mining for stock movement predictions: a Malaysian perspective, available online: <http://library.witpress.com/pages/PaperInfo.asp?PaperID=14995> [9 April 2011]