



بانک دانش

موضوع: فرایند کاوی (Process Mining)

آبان ماه ۱۴۰۰
گرد آورنده: ملیحه بزرگوار



فرایند کاوی

فرایند کاوی همانند پلی بین داده کاوی، هوش تجاری و مدیریت فرایندهای کسب و کار قرار دارد. فرایند کاوی، تکنیکی برای بازسازی فرایندهای کسب و کار از رویدادهای ثبت شده در سیستم های اطلاعاتی است.

هدف فرایند کاوی استخراج اطلاعات غیربدیهی و کاربردی مربوط به فرایندها از تاریخچه رخداد است.



تاریخچه رخداد (نگاره رویداد)

تاریخچه رخداد(نگاره رویداد) ابتدای کار فرایندکاوی است.

یک تاریخچه رخداد از انواع داده‌های مشخصی تشکیل شده است که معمولاً از منابع مختلف استخراج شده و به شکل واحد ترکیب می‌شوند.

هر رخداد در تاریخچه به یک نمونه فرایند مربوط است و می‌تواند با یک فعالیت یا وظیفه ارتباط داشته باشد.

یک تاریخچه رخداد مناسب، اطلاعاتی در مورد نمونه‌های فرایندهایی که فعالیت به آنها تعلق دارد و پاسخ این سوال که چه فعالیتی در چه زمانی و توسط چه کسی اجرا شده است را شامل می‌شود.



انواع فرایندکاوی

در متداولترین نوع دسته بندی، فرایندکاوی به سه نوع **اکتشاف**، **انطباق** و **بهبود** تقسیم می شود.

✓ **اکتشاف:** اولین و متداولترین نوع فرایندکاوی است. هدف الگوریتم های کشف فرایند، استخراج اطلاعات و رسم مدل فرایند با کمک تاریخچه رخداد بدون استفاده از اطلاعات و مدل های پیش بینی است. یکی از ویژگی های مطلوب تکنیک اکتشاف، قابلیت کشف مجدد است که میزان توانایی یک تکنیک برای پیدا کردن فرایندهای حقیقی که بخشی از مشاهدات رخداد را تولید کرده اند را محاسبه می کند.

✓ **انطباق:** دومین نوع فرایندکاوی است. در این حالت از تاریخچه رخداد استفاده می شود تا واقعیتهای که از فرایندها در تاریخچه ثبت شده است با مدلی که از پیش موجود است مقایسه گردد و تطابق آن ها بررسی شود. با کمک این روش، انحرافات، مکان آن ها و میزان آن ها شناسایی می شوند.

✓ **بهبود:** سومین نوع فرایندکاوی است. برای این حالت فرض می شود که یک مدل پیش بینی از قبل موجود است و هدف، بهبود یا گسترش این مدل با توجه به داده های ثبت شده در تاریخچه رخداد در راستای اهداف و عملکرد مجموعه مربوط است.



انواع فرایندکاوی - ادامه

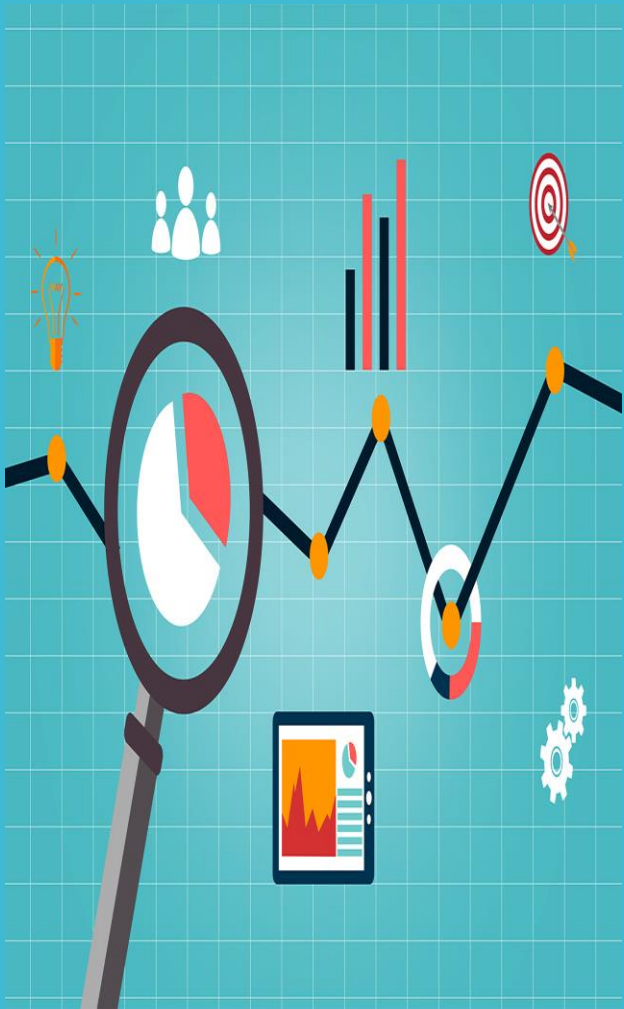
دومین نوع دسته بندی بر اساس دیدگاه‌ها صورت می‌گیرد که شامل دیدگاه‌های کنترل-جریان، سازمانی، نمونه فرایند و زمان می‌باشد.

✓ دیدگاه کنترل - جریان: کار اصلی در این دیدگاه توجه به جریان کار و ترتیب اجرای وظایف است. هدف یافتن مطلوب همه ی الگوهای ممکن است. برای بیان نتایج آن از زبان‌ها و نشان گذاری هایی همانند شبکه پتری استفاده می شود.

✓ دیدگاه سازمانی: در دیدگاه سازمانی توجه روی اطلاعات منابع است که باید با بررسی تاریخچه به دست آید. چه اجراکنندگانی درگیر بوده اند و چه ارتباطی با هم دارند.

✓ دیدگاه نمونه فرایند: این دیدگاه روی ویژگی های هر نمونه فرایند متمرکز می شود. نمونه های فرایند به کمک مسیر آن‌ها در فرایند یا آغاز گرشان مشخص می شوند.

✓ دیدگاه زمان: توجه و نگرانی دیدگاه زمان، زمان بندی و فرکانس وقوع رخدادها است.



متدولوژی های فرایندکاوی

متدولوژی PDM

- این متدولوژی به گونه‌ای طراحی شده است که در اسرع وقت دیدی کلی از فرایندها فراهم می‌کند.
- اما این روش تعداد کمی از تکنیک‌های فرایندکاوی را پشتیبانی می‌کند و برای تحلیل از دامنه دانش استفاده نمی‌شود. به همین دلیل برای پروژه‌های پیچیده و بزرگ مناسب نیست



متدولوژی های فرایند کاوی

متدولوژی Life-cycle

این روش جنبه های مختلف و تعداد قابل توجهی از تکنیک های فرایند کاوی را پوشش می دهد. بر اساس این متدولوژی، هر پروژه فرایند کاوی از چهار مرحله تشکیل شده است. شکل ذیل مراحل مختلف پروژه ی فرایند کاوی را نشان می دهد.

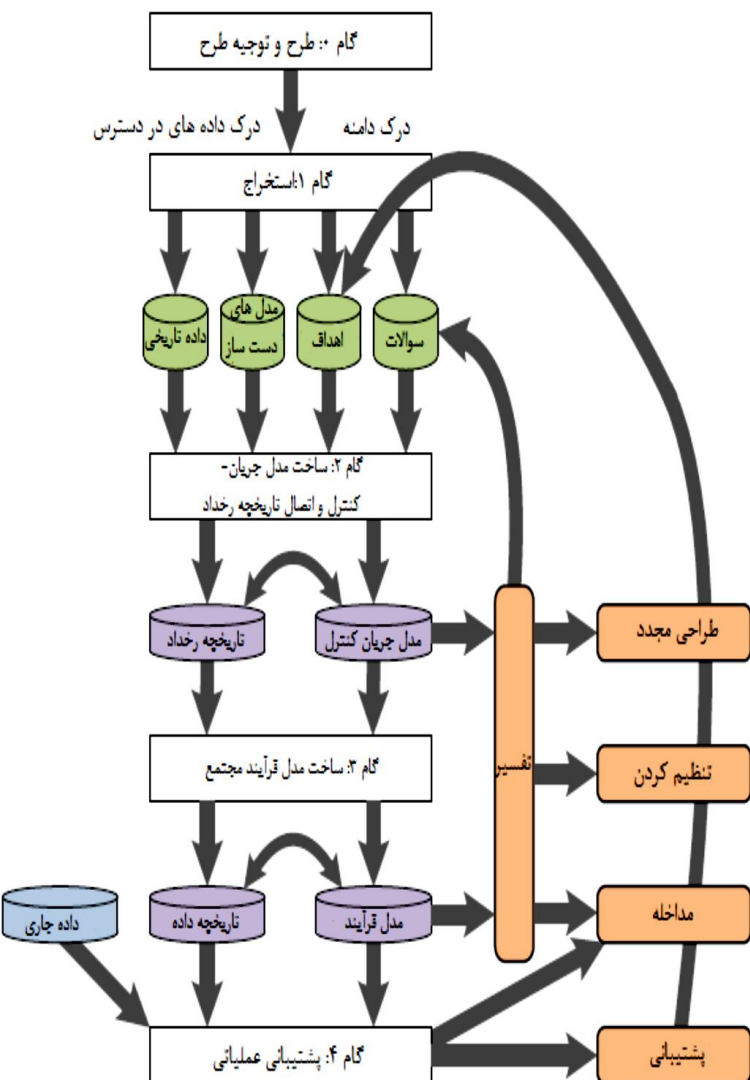
گام ۰ : طرح ریزی به همراه توجیه طرح

گام ۱ : مدل ها، داده ی رخداد، اهداف و سوالات باید با کمک مدیریت و متخصصین حوزه ی مورد بررسی از سیستم استخراج شوند.

گام ۲ : مدل جریان کنترل ساخته می شود و به تاریخچه رخداد مرتبط می شود. در این گام از تکنیک های خودکار کشف فرایند می توان استفاده کرد. با کمک مدلی که در این مرحله بدست می آید، برخی از سوالاتی که در گام قبل طراحی شده اند احتمالاً قابل پاسخگویی هستند. در این صورت امکان طراحی مجدد و اقدامات اصلاحی همانند قرار دادن فیلتر روی تاریخچه رخداد فراهم می شود.

گام ۳ : مدل جریان کنترل با توجه به دیدگاه های دیگر توسعه پیدا می کند، البته پس از اینکه ساختار ارتباطات فرایندها کاملاً مشخص گردد. ارتباطی که بین تاریخچه رخداد و مدل در گام ۲ برقرار شده است، برای گسترش مدل در این گام استفاده می شود.

گام ۴ : مجموعه فعالیت های پشتیبانی عملیاتی را شامل می شود



چالش های فرایند کاوی

- پیدا کردن، ادغام و پاک سازی داده رخداد: جمع آوری داده های مرتبط با رخدادها معمولاً با مشکلاتی رو به رو است. داده ها ممکن است بین منابع مختلف توزیع شده باشند. اگر هر منبع از روش های مخصوص به خودش برای آدرس دهی داده ها استفاده کرده باشد، یکپارچه سازی همه داده ها کنار هم دشوار خواهد بود. از طرفی داده رخداد ممکن است ناقص باشد و رخدادها صریحاً به نمونه های فرایندها اشاره نکنند یا شامل مقادیر پرت باشند.

- تاریخچه رخدادهای پیچیده با ویژگی های گوناگون: برخی از تاریخچه های رخداد می توانند بسیار بزرگ یا بسیار کوچک باشند که در هر دو حالت مدیریت و کار با آن ها دشوار خواهد بود. از طرف دیگر ممکن است داراری ویژگی های متفاوتی باشند. تاریخچه های رخداد، شامل نمونه ای از رفتارها هستند و نباید آن ها را کامل در نظر گرفت. ممکن است حالتی که تاکنون اتفاق نیفتاده و در تاریخچه ثبت نشده است، در آینده رخ دهد. با توجه به این نکته چالش بزرگتر زمانی است که اندازه ی تاریخچه کوچک و تنوع آن بالا است.

- تغییر مفاهیم: در مواردی این امکان وجود دارد که به هر دلیلی فرایندها زمانی که در حال بررسی و تحلیل هستند، دچار تغییراتی شوند. این تغییرات در مدل ها و نتایجی که فرایند کاوی به دنبال آن ها است تأثیر گذارند. به همین دلیل ضروری است که کشف و پیگیری شوند.

- کاوش وظایف پنهان: یکی از فرضیات پایه ای که برای فرایند کاوی در نظر گرفته می شود، ثبت هر رخداد در تاریخچه است. واضح است که برای وظایفی که در تاریخچه ثبت نشده اند، نمی توان اطلاعاتی به دست آورد. اما با توجه به برخی از زبان های خاص، امکان بررسی وظایفی که به عنوان وظایف پنهان شناخته می شوند، فراهم می شود. وظایف پنهان، اعمالی است که با توجه به رفتار وظایف دیگر از مدل فرایند به دست می آیند. اگر فرایندها پیچیده باشند، پیدا کردن و اضافه کردن وظایف پنهان به مدل فرایند دشوار خواهد بود.



چالش های فرایندکاوی – ادامه

- کاوش وظایف تکراری: این مشکل به زمانی اشاره می کند که مدل فرایند شامل دو گره است، که هر دو یک وظیفه را نشان می دهند. قابل ذکر است که این مشکل با وظایف پنهان در ارتباط است. می توان بسیاری از فرایندهایی که شامل وظایف پنهان هستند ولی وظیفه تکراری ندارند را به معادلی که شامل وظیفه تکراری است ولی وظیفه پنهان ندارد، تبدیل کرد.
- کاوش حلقه ها: برای هر فرایندی این احتمال وجود دارد که یک وظیفه را بیش از دوبار اجرا کند، در این صورت معمولاً یک حلقه در مدل شکل گرفته است. کشف حلقه در مدل می تواند پیچیده و غیربدیهی باشد و معمولاً با مشکل وظایف تکراری ارتباط دارد.
- برخورد نقصان: یکی از مشکلاتی که در برخی از تاریخچه ها وجود دارد، ناقص بودن آنها است. اگر تاریخچه فاقد اطلاعات کافی برای استخراج فرایندها باشد، ناقص شمرده می شود. مدل فرایندی که از چنین تاریخچه ای استخراج می شود، اشتباه خواهد بود.
- جمع آوری داده از منابع ناهمگون: سیستم های اطلاعاتی که امروزه استفاده می شوند، بسیار پیچیده و معمولاً متشکل از چندین جزء و برنامه کاربردی هستند. هر یک از برنامه ها معمولاً تنها بخشی از اطلاعات را تولید می کنند به همین دلیل برای جمع آوری داده های مورد نیاز برای فرایندکاوی، باید اطلاعات از چندین منبع مختلف و گاهی ناهمگون استخراج شوند. یکی از راهکارهایی که برای حل این مشکل معرفی شده، استفاده از انبار داده برای استخراج اطلاعات از تاریخچه ها است.
- نمایش نتایج: ارایه نتایج حاصل از فرایندکاوی برای افراد به نحوی که بتوانند نسبت به فرایندها بینشی مناسب به دست آورند، یکی دیگر از دشواری های فرایندکاوی است. اطلاعات مدیریتی غیربدیهی باید به گونه ای ارایه شوند که به راحتی قابل فهم باشند.





منابع

۱. دولو، فاطمه و سید رئوف خیامی. (۱۳۹۴). معرفی مفاهیم فرایندکاوی و بررسی یک نمونه از کاربردهای آن. سومین کنفرانس بین‌المللی پژوهش‌های کاربردی در مهندسی کامپیوتر و فن‌آوری اطلاعات
۲. قالیبافان. شکوفه، کاهانی، محسن و بهشید بهکمال. (۱۳۹۵). تحلیل نگاره رویداد با هدف فرایندکاوی: چالش‌ها و راه‌کارها. هشتمین کنفرانس بین‌المللی فناوری اطلاعات و دانش