

IBM® SPSS® Collaboration and Deployment Services 8

Každý si uvědomuje, že využití analýz v rozhodovacím procesu může výrazně zlepšit výsledky. Aby bylo možno z těchto analýz získat maximum, je potřeba zajistit infrastrukturu pro sdílení a zajištění spolehlivosti a kvality klíčových procesů. Takovou infrastrukturu poskytuje platforma IBM SPSS Collaboration & Deployment Services, platforma pro efektivní sdílení analytických aktiv. Integrace analytických nástrojů i výsledků analýz do klíčových obchodních aplikací vyžaduje spolehlivé, škálovatelné a zabezpečené prostředí. Jednotlivci i celá oddělení zde mohou publikovat výsledky svých analýz a jednoduchou formou je distribuovat napříč organizací. Výrazně se zjednodušuje spolupráce mezi různými typy uživatelů. Mezi těmi, kteří data analyzují a těmi, kteří získané znalosti a predikční modely využívají v obchodních procesech. Uložená aktiva mohou být použita pro automatizaci a Real Time Scoring.

Hlavní přínosy

- Přehledná práce analytiků na společných projektech
- Uložení, správa a ochrana analytických aktiv na jednom místě
- Auditovatelnost změn a zásahů
- Možnost Offline nebo Real Time Scoring
- Integrace prediktivních analýz do obchodních procesů

O softwaru

Software IBM SPSS Collaboration & Deployment Services je navržen tak, aby organizaci zajistil spolupráci v analytických týmech a návaznost při realizaci projektů. Jeho další úlohou je správa analytických aktiv, ať už se jedná o syntaxe, predikční modely či celé sekvence datových manipulací a transformací. Software napomáhá efektivnímu předávání analytických výsledků do provozních procesů a zajišťuje automatizaci analytických procesů a implementaci výsledků. Rovněž výrazně podporuje rozhodovací proces v nejrůznějších provozních oblastech i celkovou efektivitu analytických činností a rozhodování.

Sdílení a správa analytických aktiv

Platforma primárně slouží jako úložiště analytických aktiv, tzv. repository. Všechna další využití se odvíjí od této funkce. Na rozdíl od klasických souborových úložišť a úložišť dokumentů, představuje toto úložiště aktivní analytickou platformu pro automatizaci úloh a Real Time Scoring.

Samozřejmostí jsou klasické funkce a zabezpečení přístupu. Základní funkcí je verzování a tzv. štítkování verzí (label). Uloženému obsahu lze přiřadit tzv. topics a klíčová slova pro lepší vyhledávání. Označení verze pomocí labelu umožní jednodušší správu produkčních a testovacích verzí. Analytická aktiva mohou být uspořádána ve virtuálních složkách. Pro ně mohou být definovány různé úrovně přístupu pro různé uživatele či skupiny. Řízení přístupu je možné na úroveň uživatele, skupiny, složky, případně konkrétního uloženého objektu. Vhodné typy analytických aktiv jsou syntaxe a proudy softwarových nástrojů IBM SPSS, které lze dále automatizovat nebo použít pro Real Time Scoring.

Podporované jsou i grafické formáty (jpg, png, tiff aj.) a dokumenty klasických formátů (MS Excel, MS Word, PDF aj.). Využit lze i dávkové soubory operačního systému či skripty programovacího jazyka Python.

Analytické úložiště podporuje celou řadu uznávaných standardů a postupů jako například klastrování aplikačních serverů, virtualizaci, single sign-on přihlašování a nabízí další aplikační a programovatelné rozhraní. Významnými rozšířeními je i integrace systémů zasílajících a přijímajících zprávy (messaging) nebo šifrování přenášených dat pro splnění přísných bezpečnostních požadavků některých organizací.

Pro přístup k uloženým objektům slouží dva typy klientů. Tenký klient Portal User přistupuje přes klasické rozhraní webového protokolu http. Tato licence je poskytována v neomezeném počtu – stačí, aby manager nastavil přístupová práva. Tlustými klienty jsou administrátorská aplikace Deployment Manager nebo analytické nástroje IBM SPSS Statistics a IBM SPSS Modeler. V těchto nástrojích je funkcionalita pro načtení a zápis z/do úložiště již interně zabudována. Úložiště slouží rovněž pro uložení výstupů automatizovaných analytických úloh. Typicky se zde ukládají některé výstupní tabulky a grafy nebo celé výstupní soubory. Technicky je úložiště realizováno pomocí vybraného databázového serveru. Pro jeho zálohování, případnou obnovu, se pak používají nativní databázové prostředky.

Možnost sledování verzí dává managementu možnost náhledu do analytických procesů a jistotu, že tyto procesy budou efektivně auditovány pro splnění všech obchodních, případně i zákonných požadavků, které jsou na ně kladeny.

Analytici i koncoví uživatelé mohou své znalosti sdílet s ostatními po celou dobu vývojového, vyhodnocovacího i implementačního cyklu. Systém pro všechny slouží jako dokumentace postupů, ale i změn v souborech a procesech. Automatická kontrola verzí dokumentů zajišťuje používání aktuálních a správných verzí souborů, proměnných, syntaxí, modelů, výstupních dokumentů a zpráv. Výstupy není nutné předávat formou mailů a příloh a nestane se tak, že by organizační kolovaly různé verze neaktuálních informací.

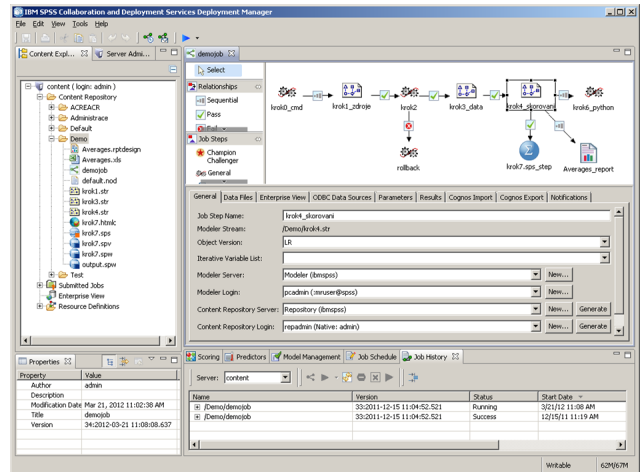
Procesní automatizace

IBM SPSS Collaboration & Deployment Services poskytuje analytikům prostředí a vizuální rozhraní pro tvorbu strukturovaných a dokumentovaných procesů i komplexních analytických dávkových úloh, tzv. jobů. Tyto procesy mohou být v určité časové posloupnosti a mohou si mezi sebou předávat různé parametry. Procesy obsahují logická pravidla a definované podmínky pro zpracování různých situačních scénářů.

Automatizace procesů zvyšuje informační hodnotu celého systému. Aktivní analytické úložiště spouští a řídí vybrané joby na vyžádání uživatele, na základě načasování v konkrétní den/hodinu, při definované změně systému, která zpracování dávky inicializuje, či jako reakci na zprávu z jiného procesu. Tato vlastnost rovněž dovoluje monitorovat průběh zpracování dávek a informovat lidi nebo jiné návazné procesy o úspěšném nebo neúspěšném dokončení každého jednotlivého kroku dávky.

Automatizace je podmíněna existencí aplikačních serverů. Samotné úložiště výpočty nerealizuje, pouze předává konkrétní analytická aktiva ke zpracování nativní serverové aplikaci. Po dokončení zpracování obdrží od serverové aplikace zprávu o tom, jak dílčí úloha proběhla,

vyhodnotí jak dále postupovat a pokračuje ve zpracování úlohy podle její definice (zpracuje další dílčí úlohu, ukončí zpracování aj.).



Rozhraní aplikace Deployment Manager pro správu obsahu a definici úloh pro automatizaci. Je vidět virtuální adresářová struktura úložiště, podrobná definice úlohy a všech jejích dílčích částí, které jsou vykonávány různými aplikačními servery v různé časové a podmíněné návaznosti. Rovněž je vidět historie spouštění úlohy.

Analytické procesy jsou stále komplexnější a složitější. Cílem každé organizace je nastavit zpracování procesů tak, aby se každý krok vždy zpracovával jednotným, snadno auditovatelným způsobem a aby jeho výstupem byly konzistentní a spolehlivé výsledky. Zajištění aktuálnosti analytických modelů a jejich kvality je pro rozhodování v organizaci naprosto zásadní.

Předdefinované modely evaluačních procesů usnadňují práci analytikům a jsou zárukou spolehlivosti výsledků a trvalé relevantnosti a přesnosti modelů. Celková spolehlivost posiluje efektivní řízení všech prací managementu i analytického prostředí.

Skórování v reálném čase

Automatizované spouštění úloh se většinou používá pro Offline skórování. Typicky se přes noc, týdně, měsíčně, provede skórování všech entit (zákazníků, zařízení aj.). K dané entitě se pak přiřadí ve vybrané časové frekvenci aktuální charakteristika, předpověď, vypočtená například predikčním modelem. V některých případech je ale Offline skórování nedostatečné, hlavně v situacích, kdy předem přesně nevíme, kdy má být požadovaná charakteristika známá, nebo když některé informace potřebné k výpočtu musí být získány v interaktivním procesu s klientem či zařízením. Pro tyto účely umožňuje úložiště skórování v reálném čase.

Real Time Scoring usnadňuje implementaci analytických procesů do praxe a je často využíván v klíčových rozhodovacích procesech. Tato funkcionality umožňuje data analyzovat, zpracovat a zobrazit výsledky v reálném čase. Real Time Scoring pracuje se vstupními daty z různých datových zdrojů i různé povahy. Může jít o data historická, transakční i data získaná během interakce s klientem, například data vyplněná v dialogu webové stránky nebo na přepážce.

Skórování se používá na různých místech, často jako integrovaná součást většího procesu. Integrace do externích procesů je jednoduchá pomocí programovatelného rozhraní API. Znamená to minimální nutnost zásahu do existujících aplikací. Proces funguje tak, že aplikace, do které je Real Time Scoring integrováno, předá skórovací službě požadavek, nejčastěji

formou webové služby, a po vrácení a zpracování výsledku pokračuje ve svém běhu.

Shrnutí

IBM SPSS Collaboration & Deployment Services je efektivní platforma pro sdílení a automatizaci zpracování analytických aktiv. Umožňuje přístup různým uživatelům na různých úrovních. Jednoduše lze propojit například s existující doménou a využít ji pro autentizaci uživatelů. Spojením této platformy a výkonných aplikačních serverů dostáváme k dispozici nástroj pro automatizaci a řízení produkčních úloh. Evidují se změny uložených objektů, eviduje se historie spouštění automatizovaných úloh. Posílají se emailové notifikace o změnách uloženého obsahu a výsledcích zpracování. Skórovací engine nástroje pro Real Time Scoring lze plně integrovat do provozních procesů. Případná změna použitých skórovacích scénářů je velice jednoduchá. Smyslem celého řešení je, využít vysoce sofistikovaný IT nástroj s integrací do operačních a databázových systémů tak, aby si pak koneční uživatelé vystačili s předdefinovanými nastaveními a minimalizovali tak počet požadavků na IT.

