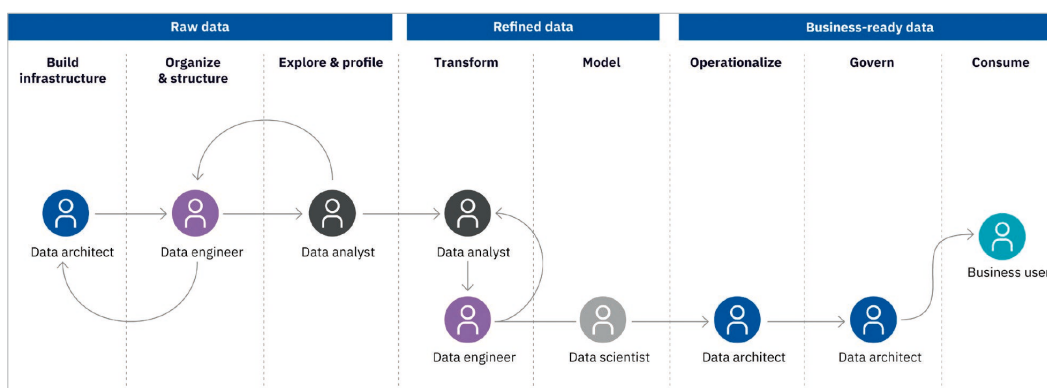


Razumijevanje podataka

U modernim informacijskim sustavima, neovisno o branši poslovanja, podaci kojima tvrtka raspolaže glavni su pokretač za inovacije i održavanje prednosti ispred konkurencije. Razumijevanje vlastitih digitalnih podataka kojima se raspolaže, ključni je faktor za shvaćanje poslovnih trendova i prilika, izradu analitike i statistike poslovanja te pomaže i ubrzava put prema AI (journey to AI)

Vedran Podubski, Senior Implementation Specialist, Megatrend Poslovna Rješenja



DataOps workflow po ulogama

Prema istraživanjima s kraja 2019. godine, 89% poduzeća ima problema s upravljanjem podacima (Data Management), gdje najveći dio problema čini veliko kašnjenje u dovršetku projekata za uvid u podatke (izvešća itd.) te samim time, nedostatak povjerenja u temeljne podatke koji se stvaraju zbog toga.

Razumijevanje poslovnih ciljeva organizacije odlučujuće je za razvijanje učinkovite podatkovne strategije u područjima analitike i umjetne inteligencije. Kako bi se uspješno zadovoljilo potrebe tržišta, točnije, klijenta, sam uspjeh ovisi o racionalizaciji podatkovnih operacija (data operations) s integriranom podatkovnom protočnom strukturom (data pipeline), koja pruža cjelovit i konzistentan prikaz poslovanja u bilo kojem trenutku.

Brže dobivanje rezultata iz više različitih izvora podatka poboljšava operativnu učinkovitost i djelotvornost te bolje donošenje odluka, neovisno o vrsti poslovnog odjela. Također, uz samu brzinu dobivanja rezultata iz podataka, važno je da su

dobiveni podaci pouzdani i spremni za poslovanje (business ready).

Organizacijama koje traže transformaciju unutar svojih podatkovnih operacija, tehnologija automatizacije procesa uvelike može pomoći i pružiti konkurentsku prednost. Ukupna metodologija, točnije, paradigma načina sakupljanja, obrade, racionalizacije i prikaza podataka, naziva se DataOps.

TO JE DATAOPS

Paradigma DataOps, točnije, metodologija, način je za organizaciju ljudi, procesa i tehnologije za brzu isporuku pouzdanih i visokokvalitetnih podataka svim njihovim korisnicima.

Sama praksa DataOpsa usredotočena je na omogućavanje suradnje u cijeloj organizaciji, kako bi se pokrenula agilnost, brzina isporuke i nove inicijative s podacima (agility, speed and data initiatives). Koristeći moć automatizacije, DataOps dizajniran je za rješavanje izazova povezanih s neučinkovitošću u pristupu, pripremi, integraciji i dostupnosti podataka.

Potencijalne koristi DataOpsa

uključuju značajna povećanja produktivnosti u isporuci informacija i podataka pojedincima te poboljšanje procesa radi postizanja učinkovitosti i optimizacije. Automatizirane podatkovne operacije, koje uključuju AI inicijative, vođene podacima (AI data - led initiative), mogu pomoći u postizanju sljedećih rezultata:

- Isporučivanje integriranih podataka spremnih za poslovanje koji pokreću analitiku i umjetnu inteligenciju
- Postizanje operativne učinkovitosti
- Omogućavanje privatnost

i usklađenost podataka (data privacy and compliance)

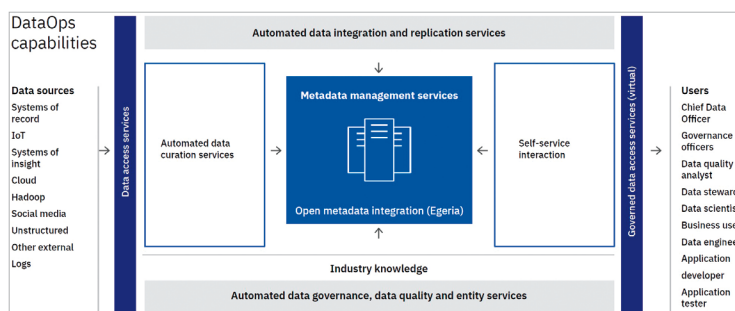
DATAOPS NIJE DEVOPS

U današnje vrijeme pojam DevOps-a već je vrlo raširen i donekle svima poznat te je velik broj organizacija već implementirao određene razine DevOps-a unutar svojih razvojnih disciplina. Zbog sličnosti konvenciji imenovanja metodologija, unatoč tome što obje tehnike u praksi služe za kreiranje najboljih operativnih praksi, svaka od ove dvije metodologije ima svoju jedinstvenu funkciju i mjesto unutar organizacije.

Primarna razlika između navedene dvije metodologije jest u osnovnom cilju, DevOps koristi se u razvoju i isporuci softvera, a DataOps, kao što je već i navedeno u prethodnom tekstu, služi za omogućavanje pouzdanih, visokokvalitetnih podataka, spremnih za sve oblike poslovanja i dostupnih za brzu upotrebu.

Daljnjom raščlambom razlika možemo navesti sljedeće karakteristike DataOpsa:

- Poticanje kontinuirane i brze inovacije posla, omogućujući samoposlužni pristup (self-service access) pouzdanim, visokokvalitetnim podacima za sve korisnike
- Omogućuje kontinuiranu isporuku podataka automatizacijom upravljanja podacima (data governance) i integracije, istodobno štiteći regulatorne probleme (regulatory concerns)
- Osigurava povratnu spregu za kontinuirano učenje od svih korisnika podataka praćenjem i



DataOps Arhitektura

optimizacijom protoka podataka (data pipeline)

- Ispravlja neusklađenost ljudi i ciljeva, njegujući bliže veze između podrške IT sustava, operative i poslovanja (operations and business)
- Ubrzava isporuku promjena i poboljšava kvalitetu isporuke uvođenjem automatizacije tijekom ciklusa isporuke podataka
- Poboljšava uvid u stvarnu vrijednost metapodataka i podataka, korištenjem dobivenih rezultata za ukupnu optimizaciju

DevOps, za razliku od prethodno navedenih karakteristika, ima sljedeće:

- Ubrzavanje kontinuiranog uvođenja ideja omogućavanjem suradničkog razvoja i testiranja u lancu razvoja softvera
- Omogućuje kontinuiranu isporuku inovacija automatiziranjem procesa isporuke softvera i uklanjanjem otpada, istodobno pomažući u rješavanju regulatornih problema
- Omogućuje povratnu spregu za kontinuirano učenje od klijenata nadgledanjem i optimizacijom softverskih inovacija
- Ispravlja neusklađenost ljudi i ciljeva, njegujući bliže veze između programera, operative i poslovanja
- Ubrzava i uklanja pogreške u isporuci promjena uvođenjem automatizacije tijekom razvojnog ciklusa
- Poboljšava uvid u stvarnu vrijednost aplikacija korištenjem povratnih informacija klijenata, u svrhu poticanja procesa optimizacije razvojnog sustava

LJUDI, PROCES I TEHNOLOGIJA

DataOps je orkestracija ljudi, procesa i tehnologije, a za održavanje predanosti prakse DataOpsa potrebna je duboka suradnja svih funkcija. Potreban je fokus na njegovanju praksi i procesa upravljanja podacima koji poboljšavaju brzinu i točnost analitike.

DataOps podržava visokoproduktivne timove s tehnologijom automatizacije, koji pomažu u postizanju povećanja učinkovitosti, kako u projektnim rezultatima, tako i u vremenu potrebnom za isporuku. S obzirom na to da više

poslovnih segmenata zahtijeva i želi upravljati podacima za postizanje kontekstualnih uvida, cilj DataOpsa je sljedeći:

- Povećanje kvalitete i brzine protoka podataka do organizacije
- Iskorištavanje predanosti vodstva za podršku i održavanje vizije na temelju podataka u cijelom poslu.

Ova vrsta transformacijske promjene započinje razumijevanjem istinskih ciljeva poslovanja:

- Kako podaci informiraju o odlukama i uslugama koje utječu na kupce?
- Kako podaci mogu pomoći u održavanju konkurentske prednosti na tržištu?
- Koji su financijski prioriteti u kojima nam podaci mogu pomoći u rješavanju?

Srž DataOpsa informacijska je arhitektura tvrtke. Znete li svoje podatke? Vjerujete li svojim podacima? Možete li brzo otkriti pogreške? Možete li postupno unositi promjene a da niste "razbili" cijeli protok podataka (data pipeline)? Kako bi se odgovorilo na ta pitanja, prvi korak je inventarizacija alata i praksi upravljanja podacima i integracije podataka.

Kada se razmišlja o alatima koji podržavaju DataOps praksu unutar tvrtke, potrebno je razmisliti kako automatizacija u ovih pet kritičnih područja može transformirati protok podataka (data pipeline):

1. Usluge čuvanja podataka (data curation services)
2. Upravljanje metapodacima (metadata management)
3. Upravljanje podacima (data governance)
4. Upravljanje glavnim podacima (master data management)
5. Samoposlužna interakcija (self-service interaction)

Pružanje podataka spremnih za poslovanje uključuje sve te aspekte, a svaka DataOps praksa mora uključivati holistički pristup koji uključuje svih 5 aspekata.

Organizacije koje su uspješno implementirale DataOps znaju kojim podacima imaju pristup, vjeruju značenju podataka i njihovoj kvaliteti, i koriste svoje podatke do maksimuma.

Podaci imaju vrijednost onda kada pouzdani podaci spremni za poslovanje pomažu u postizanju ra-

IBM DATAOPS

IBM DataOps pomaže u isporuci podataka spremnih za poslovanje, pružajući vodeću tehnologiju u industriji, koja radi s automatizacijom i omogućenom umjetnom inteligencijom (AI-enabled automation), infuzijskim upravljanjem (infused governance) i moćnim katalogom znanja, kako bi se operacionalizirali kontinuirani, visokokvalitetni podaci u cijelom poslu. Povećava učinkovitost, kvalitetu podataka, pronalaženje i ugrađuje upravljačka pravila kako bi se pružio protok podataka i samoposluživanje podataka pravim ljudima, u pravo vrijeme, iz bilo kojeg izvora.

IBM je identificirao šest faza u životnom ciklusu DataOpsa, plus nužna kulturna razmatranja, za uspješnu provedbu prakse DataOpsa. To se temelji, kao i interno usvajanje DataOpsa kao dio puta za transformaciju.

- **Razmišljanje.** (Think) Konceptualizacija, usavršavanje i određivanje prioriteta sposobnosti
- **Programiranje.** (Code) Generiranje, poboljšanje, optimizacija i ispitivanje značajki
- **Dostavljanje** (Deliver). Automatizirana proizvodnja i dostava značajki
- **Pokretanje.** (Run) Usluge, opcije i mogućnosti potrebne za pokretanje
- **Upravljanje.** (Manage) Stalno praćenje, podrška i povratak funkcionalnosti
- **Učenje.** (Learn) Kontinuirano učenje i povratne informacije na temelju rezultata eksperimenata

IBM nudi nove inovativne mogućnosti, koje uključuju ugrađeno strojno učenje (ML), AI automatizaciju, integrirano upravljanje (infused governance) i moćan katalog podataka (data catalog) za operative i kontinuirane visokokvalitetne podatke u cijelom poslovanju. Učinkovitost DataOpsa ovisi o ekstremnoj automatizaciji komponenata podatkovne tehnologije, koje se koriste za podatkovni tijek (data pipeline).

IBM Cloud Pak for Data, uključujući katalog znanja IBM Watson (WKC), može odgovoriti na te zahtjeve na učinkovit, robusan, automatiziran i ponovljiv način.

IBM Cloud Pak for Data Server može riješiti potrebu za kretanjem podataka, njihove objave i korištenja u podatkovnom tijeku, pomažući pritom u osiguranju kvalitete podataka i provođenju pravila. Uz učinkovito upravljanje kontrolom izvora, proces se može automatizirati i učinkovito izvršavati.

Ugrađeno strojno učenje (ML) u IBM Watson Knowledge Catalogu za IBM Cloud Pak for Data, dopunjuje postupak automatizacije i optimizira ga svakom iteracijom za robusan tijek podataka.

IBM Cloud™ DevOps Insights može pomoći u pružanju operativnog uvida u vizualizacije za tijek podataka. Pomaže u provođenju mjera sigurnosti i kvalitete koje se kontinuirano nadgledaju, otkriva sve neočekivane varijacije i generira operative statistike na temelju ekstremne automatizacije i prilagođene integracije s IBM Cloud Pak for Data.

Apache Airflow i NiFi mogu pomoći u dizajniranju tijeka rada i njegovoj orkestraciji.

Korištenje ekstremne automatizacije pomoću krajnjih točaka REST (REST endpoints), zajedno s parametrizacijom, može pomoći u dinamičnom odabiru određenih skupova podataka ili okoline, i promijeniti ponašanje bez utjecaja na izvorni kôd tijeka (pipeline) i prilagoditi se svakodnevnim potrebama stručnjaka za analitiku podataka. ◀

zličiti uvida, operative izvrsnosti, suradnje i konkurentske prednosti.

Ako želite smanjiti troškove učinkovitijom obradom podataka i analitikom s jedinstvenim setom proizvoda i kvalitetnijim utroškom resursa, isporučivati agilno i konti-

nuirano nove analitičke proizvode kroz DataOps, podići produktivnost, kooperaciju na jedinstvenoj integriranoj platformi, ili doznati više podataka, javite se našim stručnjacima na adresu poslovna_rjesenja@megatrend.com