

# **DataOps: Teollisen data- arkkitehtuurinne puuttuva palanen**

**Kirjoittaja:** John Harrington, HighByte  
**Suomentaja ja editoija:** Sakari Aulanko,  
Novotek

# Tiivistelmä

DataOps on uusi menetelmä ohjelmistoratkaisujen piirissä. Se vastaa data-analytiikan ja -arkkitehtuurin tarpeisiin teollisuusyrityksissä, jotka ovat omaksumassa teollisuus 4.0:a, digitaalista transformaatiota ja älykästä tuotantoa. DataOps-menetelmien avulla pyritään lyhentämään analytiikan julkaisusykliä, sekä helpottamaan ylläpitoa. Käytännössä se tapahtuu esimerkiksi kontekstualisoimalla ja standardoimalla tietoa mahdollisemman lähellä syntypaikkaa, sekä varmistamalla turvallinen ja eheä datavirta sitä käyttäviin applikaatioihin, jotka voivat pyöriä niin laitoksella kuin datakeskuksissa tai pilvessä.

HighByte on kehittänyt Intelligence Hub -ratkaisun vastatakseen teollisen ympäristön DataOps-tarpeisiin – tilanteisiin, joissa data täytyy kerätä ja järjestää teollisista järjestelmistä ja valjastaa liiketoiminnan käyttöön koko organisaatiossa ja ekosysteemissä.




# Ongelma: Käyttökelvottoman datan tulva

Teollisuus 4.0:ssa, digitaalisessa transformaatiossa ja älykkäässä tuotannossa hyödynnetään informaatiota tuottamaan automatisoituja päätöksiä ja tarjoamaan päätöksiä tekevien tahojen käsiin hyödyllistä tietoa silloin, kun tietoa tarvitaan. Muutoksen ja tietopohjaisen päätöksenteon saavuttamiseksi on tarjolla useita teknologioita, kuten IIoT-alustat, edistyneet analytiikkatyökalut, lisätyt ja virtuaaliset visualisoinnit, mobiilialustat, pilvipohjaiset tietojärjestelmät, pienet ja edulliset sensorit sekä verkkoteknologiat.

Valittuja teknologioita yhdistelemällä on saatu tieto liikumaan tarpeen mukaisesti, mutta itse tiedon hyödyntäminen on osoittautunut käytännössä enemmän aikaa ja työtä vaativaksi toiminnaksi, kuin alussa on oletettu.

Teollisuus 4.0:n teknologioita omaksuessa usein ajatellaan, että API:en avulla on helppo yhdistää teollinen data,

sen analytiikka ja visualisointi ja hyödyntää dataa nopeasti. Kerätessä tietoa useista järjestelmistä ja laitteista usein käytännössä kuitenkin huomataan, että data on epäjohdonmukaista, ja että datavirroista puuttuu konteksti, joka selittäisi, mistä virrasta on kyse, mistä data on peräisin, mitä toleranssia siltä odotettiin, tai missä yksikössä mitaustulos oli. Automaatio- ja ohjausjärjestelmien data vastaa harvoin sellaisenaan liiketoiminnan tarpeisiin. Dataa on valtavasti tarjolla, mutta käyttötapaukset vaatisivat vain murto-osan sen tietosisällöstä, ja mielellään muotoiltuna selkeästi ja objektipohjaisesti.



Kun uutta teknologiaa on otettu käyttöön ja siihen yhdistytään, datan kontekstuaalisuus ja saavutettavuus ovat osoittautuneet kuluttavammiksi ja työläemmiksi tehtäviksi kuin odotettiin.

Nykyisin dataa tarvitaan lähellä laitteita omissa datakeskuksissa ja joissain tapauksissa pilvipohjaisissa järjestelmissä. Data-arkkitehtuurin haasteiden ratkaisemiseksi ja datan kontekstuaalisuuden ja standardisoinnin haasteiden ratkaisemiseksi on kehittynyt uusi sovellusratkaisujen kategoria, joka voi olla avainasemassa teollisuus 4.0:aa omaksumille yrityksille.

Tämä kategoria tunnetaan nimellä DataOps, ja kun ratkaisu on rakennettu juuri teollisen datan käyttöön, kutsutaan sitä teolliseksi DataOps:ksi

# Johdatus teolliseen DataOpsiin

Teollinen data-arkkitehtuuri kehittyi ennen teollisuus 4.0:aa vuosien varrella kerroksittaiseksi lähestymiseksi, jossa lopputulos määriteltiin Purdue-mallissa tai ISA-95:n kautta. Tässä mallissa data virtaa sensoreista ohjausjärjestelmiin, niistä SCADA/HMI:n ja sieltä MES:iin ja lopulta ERP:iin. Tiedon määrä väheni dramaattisesti tiedon siirtyessä tasolta seuraavalle. Datan resoluutio myös pienentyi siihen pisteeseen, että monissa yrityksissä ERP- ja jopa MES-järjestelmien dataa päivitettiin manuaalisesti, eikä niitä välttämättä ollut edes yhdistetty tehtaan laitteisiin.

Jokaiseen kerroksen yhdyskohtaan kehitettiin joukko datan viestintäprotokollia. Suurin osa niistä kuitenkin oli sovel-luskohtaisia ja yhtiöillä oli tapana kehittää omia protokollia liittyäkseen laitteistoihinsa ja ohjelmistoihinsa. Nämä protokollat olivat yksilöllisiä ratkaisuja yhden kerroksen liittämiseen seuraavaan, eikä niitä käytetty uudelleen toisten kerrosten rajoilla.

OPC on eräs teollisuuden omaksumis-ta protokollista, joka kehitettiin tiedon siirtämiseksi laitteistokerrokselta SCADA-kerrokseen. Laiteval-

## Teollisuus 4.0

Teollisen alan liiketoiminta muuttuu dramaattisesti. Muutos on niin merkittävä, että sitä kutsutaan neljänneksi teolliseksi vallankumoukseksi.

Ensimmäinen teollinen vallankumous kattoi 1700- ja 1800-luvut, kun tuuli- ja vesimyllyjen voima valjastettiin. Seuraava teollinen vallankumous piti sisällään tehtaiden sähköistymisen ja moottorikäyttöiset laitteet 1800-luvun lopulla ja 1900-luvun alussa. Kolmas teollinen vallankumous määrittyi ohjaujärjestelmien ja moottorien myötä 1900-luvun puolivälissä, yhä jatkuen. Neljäs aalto aalto on kyber-fyysinen vallankumous, joka tarjoaa reaaliaikaista palautetta ja päätöksentekoa laitteiden valvonnan, sensorien, historiatiedon, liiketoimintajärjestelmien ja analytiikkajärjestelmien välillä.

Jokaisessa teollisessa vallankumouksessa on tapahtunut fundamentaalisia muutoksia siinä, millaisia prosesseja käytetään, miten tuotteita valmistetaan, miten asiakkaita palvellaan, ja mitkä yritykset lopulta saavuttavat valta-aseman.

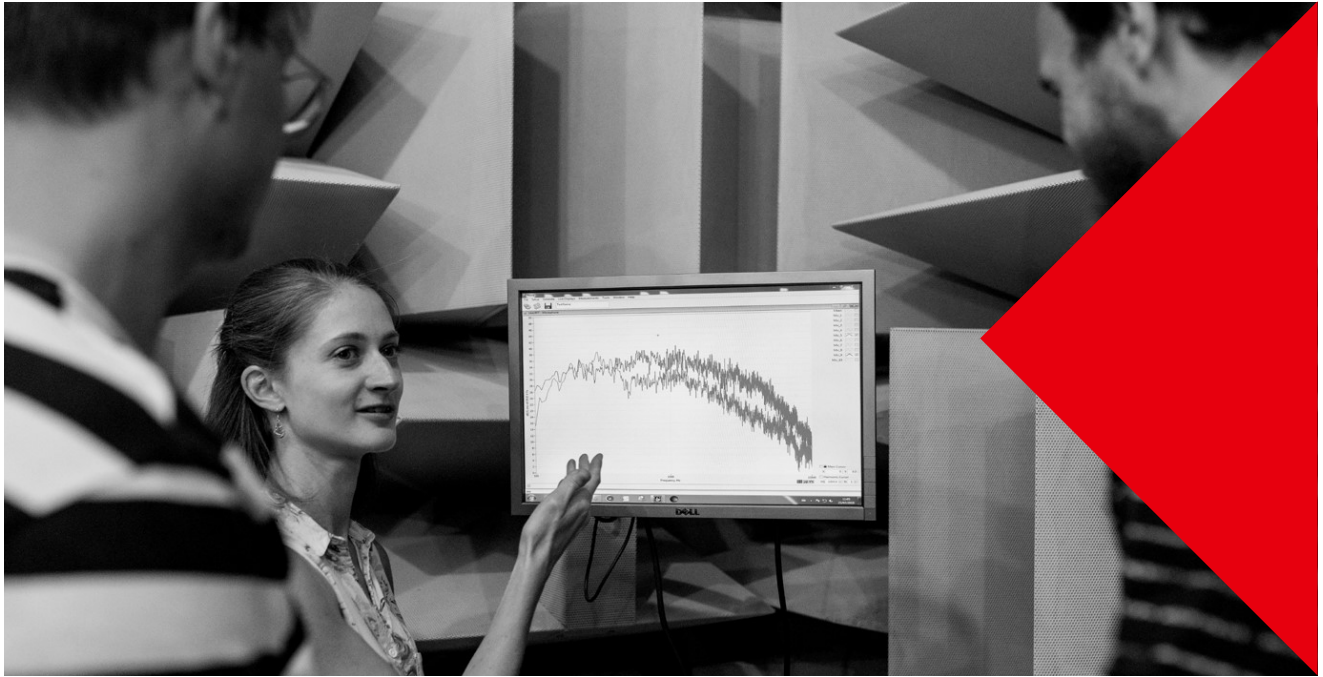
vonnan ja sovelluserroksen väliin kehitettiin OPC-palvelimet muuntamaan tietoa sovelluskohtaisesta valvonnasta ja OPC:stä.

Datan prosessointi järjestelmäkerrosten läpi on ollut aikaisemmin riittävä tapa toimia, sillä liikutettavan datan määrä on ollut rajallista, ja suuri osa datasta on luotu edellisen tason järjestelmässä. Teollisuus 4.0:n, digitaalisen transformaation ja älykkään tuotannon kohdalla tilanne ei ole enää sama. Esimerkiksi sensoridataa (taso 0) ei välttämättä tarvita ohjaukseen, eli sitä ei hyödynnetä PLC-tasolla (taso 1) tai SCADA-järjestelmässä (taso 2) – mutta sitä voidaan hyvin taas tarvita tason 3 huolto- tai kunnossapitojärjestelmässä esimerkiksi ennakoivaan kunnossapitoon.



Digitaalisessa transformaatiossa on kyse tiedon ja teknologioiden tehokkaasta hyödyntämisestä liiketoiminnan perustana. Tuotantoyrityksessä se tarkoittaa tuotannollisen ja operatiivisen tiedon jakamista läpi organisaation sellaisessa muodossa, joka on ymmärrettävä ja täytettävä liiketoiminnan laatuvaatimukset. Jokaisessa organisaatiossa on omassa toiminnassa kyettävä hyödyntämään tietoa mahdollisimman suoraviivaisesti ja tehokkaasti. Esimerkiksi datapisteen vastaanottaminen nimellä F8:4 ja arvolla 52.2 ei kerro huoltoinsinöörille, että kyseessä on leimauslaitteen hydraulikkaöljyn lämpötila, yksikkö on celsius, ja että mikäli lämpötila nousee yli 180 asteen, siitä aiheutuu hyvin todennäköisesti komponenttivikaantuminen.

Teollisuudessa datavirtojen optimoinnin ja hallinnan haasteet, sekä kontekstualisoinnin ja standardisoinnin puutteet ovat johtaneet DataOps:n lähestymiseen ja sitä myötä joukkoon uusia ohjelmistoteknologioita, jotka tukevat kyseistä toimintamallia. Käsiteltävä aikasarja- ja tapahtumapohjainen tieto voi olla hyvin erilaista ja erimuodoissa, ja siitä puuttuu usein konteksti. Järjestelmiä voi olla monilta vuosikymmeniltä ja toimittajilta, ja silti tietoturva ja tiedon laadulliset tekijät tulee olla kunnossa. On sanomattakin selvää, että teollisissa järjestelmissä DataOps vaatii omanlaisen osaamisen ja teknologiaportfolion.



# Teollinen DataOps – Onnistumisen kulmakivet

Seuraavaksi kuvataan viisi olennaista fokusaluetta, joita tarvitaan teollisen tiedon hyödyntämiseen läpi organisaation, mahdollisimman lyhyeen analytiikan julkaisusykliin ja ylläpitokustannusten optimointiin.

## 1. Kuvaa, normalisoi ja kontekstualisoi data

Teolliset tietojärjestelmät on luotu hallitsemaan moottoreita, venttiilejä, liukuhihnoja, työkoneita ja -laitteita sekä muita vastaavia komponentteja. Data tulee tyypillisesti ohjauslogiikoista, laitteiden valvomoista, etäkäyttöpäätteistä tai älykkäistä sensoreista. Ei ole tavatonta, että tehtaalla on satoja ohjelmoitavia logiikoita ja useita valvomoita. Kalusto ja valvomot on usein ostettu eri aikoihin ja eri toimittajilta, ja tehtaiden kasvaessa tarpeet ovat muuttuneet ja ratkaisut kehittyneet.

Ohjausjärjestelmien tietorakenteet ja kommunikaatiot ratkaisut on suunniteltu toimimaan tehokkaasti teollisessa ympäristössä. Datapisteet kuvataan usein pelkistetyksi positiokohtaisesti, ilman informaation täsmällisempää mallintamista tai kontekstualisointia.

Tätä tietoa tulee pystyä kuitenkin analysoimaan kautta laitteistojen, prosessien ja tuotteiden. Satojen laitteiden ja niiden kymmenien tuhansien datapisteiden käsittelyä varten DataOps-ratkaisujen tulee mahdollistaa tiedon mal-

linnus mieluusti niin, että tieto kuvaa itsensä (esimerkkinä Sparkplug B). Kun mallit korreloivat laitteiden, prosessien ja tuotteiden kanssa, on tieto helppo ottaa vastaan sitä hyödyntävissä palveluissa ja sovelluksissa.

## 2. Mahdollista liitynnät tuotannon ja IT:n järjestelmiin

Teolliset laitteet ja järjestelmät sekä IT-järjestelmät kommunikoivat luonnostaan eri tavoilla. Teolliset laitteet ja järjestelmät käyttävät usein sovelluskohtaisia protokollia, jotka soveltuvat hyvin aikasarjaisen ja tapahtumapohjaisen tiedon jatkuva-aikaiseen käsittelyyn. IT-järjestelmät taas käyttävät erityyppisiä rajapintaratkaisuja (muun muassa REST:iä ja SOAP:ia) ja integraatioita. Teollisuuden ja IT-järjestelmien välisessä kommunikaatiossa hyödynnetään usein API-rajapintoja ja/tai pub-sub-protokollia (esimerkkinä MQTT tai AMQP). Esimerkiksi MQTT tarjoaa joustavan ja kevyen pub-sub-kommunikaation, joka soveltuu useisiin käyttötarkoituksiin ja tietoturva vaatimuksiin.

Teollisuuden DataOps-ratkaisun on kyettävä integroitumaan ja toimimaan saumattomasti operatiivisten laitteiden ja datalähteiden sekä liiketoimintasovelluksien välissä. Ratkaisun on myös sopeuduttava tämän päivän IT-käytäntöihin.

### 3. Hallitse jakamista ja tietovirtaa

Tietovirtoja tulee hallita järjestelmistä, joista tietovirrat voi tunnistaa, muokata, ottaa käyttöön, ja taas poistaa käytöstä. Laitteiden ja tietovirtojen selkeä muutostenhallinta on tärkeää, jotta datan laatu säilyy ja tietovirtoja on helppo vakiinnuttaa, kun muutoksia tapahtuu. Myös tietoturvan kannalta on tärkeää pystyä hallitsemaan tietovirtoja. Monet ulkopuoliset toimittajat haluavat laitedataa tarjotakseen parempaa palvelua. Tiimi, joka operoi laitetta, haluaa voida kontrolloida, mikä data virtaa, millä frekvenssillä ja millä ehdoilla se siirtyy. He haluavat voida myös estää datan virran, jos toimittaja ei enää tarvitse sitä. Siksi tiedon virran hallinta on välttämätön osa teollista DataOps-ratkaisua.

### 3. Varmista, että ratkaisut täyttävät skaalautuvuuden ja tietoturvan vaatimukset

Teollinen data eroaa tyypillisestä liiketoiminnallisesta datasta, jota useimmat IT-järjestelmät käsittelevät. Teollinen data on usein aikasarja-pohjaista ja tulee sadoista tai tuhansista eri laitteista. Tämä data täytyy kaapata, kontekstualisoida ja toimittaa juuri tilanteen vaatimalla resoluutiolla analysoitavaksi tai visualisoitavaksi. Teollinen data hyödynnetään yleensä millisekunnissa luomisensa jälkeen, tietovirtatyyppisesti. Tähän tarpeeseen nähden esimerkiksi erämuotoinen tiedonsiirto (ETL eli extract, transform, load) ei välttämättä ole se optimaalisin.

Teollisen DataOps-ratkaisun tulee siis kyetä käsittelemään tietovirtoja tehokkaasti, sekä kuvata ne mahdollisimman lähellä synty- ja käyttöpaikkaa. Usein tarvitaan myös tiedon yhdistelyä ja esikäsittelyä ennen tiedon toimittamista eteenpäin.

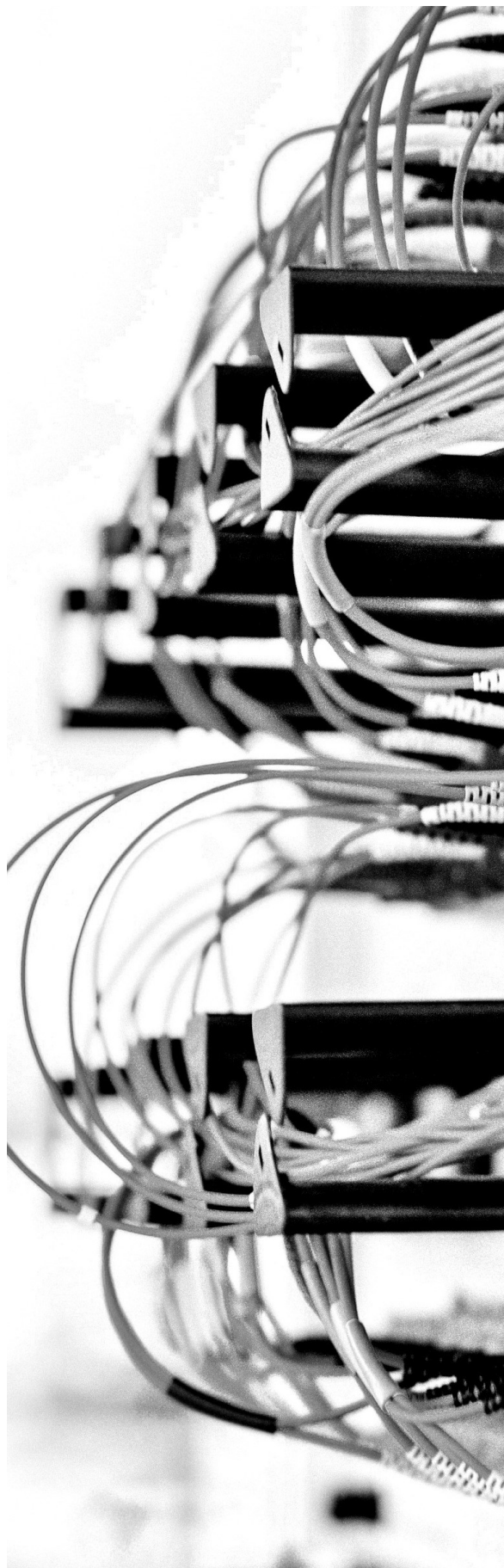
On tärkeä huomioida, että teollinen data sisältää tuotantolaitoksen luottamuksellista tietoa, joka täytyy pystyä toimittamaan turvallisesti sitä hyödyntäville sovelluksille.

#### Hyödynnä Edgeä

Analytiikka- ja visualisointiohjelmistosta riippuen dataa voidaan haluta prosessoida joko lähellä laitteita, laitoksilla, datakeskuksissa tai pilvessä. Teollisen DataOps-ratkaisun on hyvä kyetä toimimaan lähellä laitetta ja kommunikoidaan eri sovelluksille niiden tarvitsemalla datafrekvenssillä ja ehdoilla, sekä jakamaan malleja ympäri tehdasta ja yritystä samalla mahdollistaen datan standardoinnin ja normalisoinnin.



Kun tieto mallinnetaan lähellä sen syntypaikkaa, saadaan tietomalli hyödynnettyä mahdollisimman monessa palvelussa ja organisaation tasossa. Samalla mallin ylläpito voidaan toteuttaa käytännössä lähellä lattiatasoa, yhdessä paikassa. Teollisuudessa tietolähteiden muutokset syntyvät aina laittiatasolta, joten sen sijaan että muutos tehtäisiin useisiin kunnossapito ja tuotannonohjausjärjestelmiin, voidaan muutos tehdä yhteen paikkaan soveltuvan DataOps-ratkaisun avulla. Tällöin työn määrä optimoituu, ja samalla varmistetaan, että muutoksesta ei synny haittaa muihin tietoa hyödyntäviin palveluihin.



# HighByte Intelligence Hub

HighByte:n Intelligence Hub on esimerkki käytännön DataOps-ratkaisusta teollisuusyrityksille. Se on markkinoiden ainoa ratkaisu, joka yhdistää kontekstualisoidut ja standardoidut datamallit teollisiin ja IT-järjestelmiin, hallitsee tietovirtoja turvallisesti, ja on kehitetty Edge-natiivilla lähestymistavalla.

Seuraavat Intelligence Hubin ominaisuudet mahdollistavat skaalautuvuuden useisiin käyttötapauksiin:

- **Datan mallinnus.** Mallinna ja kuvaat laitteet, tuotteet, prosessit ja järjestelmät tarpeisiisi parhaiten sopivalla tavalla. Kontekstualisoi data kuvaaman itsensä metadatalle, standardoimalla tietoalkiot ja normalisoidulla mittayksiköt. Mallit on helppo käyttää uudelleen ja jakaa hubeista toisiin.
- **Datan muutokset.** Intelligence Hubiin kuuluu tiedonkäsittelyn moottori (JavaScript), jonka avulla tietovirtoja voidaan esikäsitellä tehokkaasti laskusuoritusten avulla esimerkiksi tiedon skaalaamiseksi tai keskiarvoistamiseksi. Tällöin datan käytettävyys kohenee ja dataa tarvitsee siirtää vähemmän.
- **Datavirtojen liittäminen.** Luo yhteys ja datavirtaus raakadatalle tai mallinnetulle tiedolle tarpeen mukaisilla frekvensseillä tai ehdoilla. Hallitse tehokkaasti datavirtoja eri lähteiden ja sovellusten välillä – yhdessä paikassa.
- **Integraatiot.** Integroidu eri järjestelmiin OPC UA:n ja MQTT:n avulla. Hub tarjoaa yhteyksien ja datapisteiden konfiguraation ja hallinnan.
- **Tietoturva.** Hyödynnä tiedonsiirrossa OPC UA:n ja MQTT:n mahdollistamien tietoturvakäytänteitä. Pääkäyttäjän on helppo hallita tietovirtoja määrittelemällä ja seuraamalla sallittuja yhteyksiä.
- **Edge-natiivius.** Intelligence Hub on sovelluksena kevyt, ja toimii ratkaisuna niin kevyillä rauta-alustoilla (esim. Raspberry Pi), teollisilla kytkimillä, tai muilla Linux-, Windows 10 - ja Windows Server -alustoilla.

HighByte



# Yhteenveto

Teollisen DataOps-ratkaisun avulla tehostetaan tiedon operatiivista ja ylläpidollista hyödyntämistä, sekä mahdollistetaan tiedon tehokas hyödyntäminen läpi organisaation. Lähestymistapa mahdollistaa sujuvan muutostenhallinnan ja nopean analytiikan ja palvelujen julkaisusyklin uusien sovelluksia kehitettäessä ja käyttöönotettaessa, tai kun

vastataan tulevia muutoksia tietolähteisiin, arkkitehtuuriin tai ulkoisiin palveluihin. Tämä on hyvin tärkeä ja olennainen kilpailutekijä nopeasti digitalisoituvassa liiketoimintaympäristössä.

## Pääpointit

- Teollinen DataOps on olennainen osa teollisuus 4.0:n, digitaalisen transformaation ja älykkään tuotannon onnistumista.
- Datan yhdistely, kontekstualisointi ja standardointi on tarpeen, jotta teollista dataa voi hyödyntää laajassa mittakaavassa.
- Teollisessa ympäristössä turvallisuus on kriittistä – järjestelmissä tulee olla kyky hyödyntää tietoturva-protokollia, sekä kyky hallita käyttöoikeuksia ja datavirtoja järjestelmien välillä.



Novotek auttaa sinua tekemään tuotannostasi tehokkaampaa ja kannattavampaa. Toimitamme innovatiivisia IT- ja automaatoratkaisuja teollisuuteen. Toimimme Pohjoismaissa, Benelux-maissa, Sveitsissä ja Britteinsaarilla tarjoten nopeaa ja mitattavaa lisäarvoa hyödyntämällä markkinoiden johtavia valmisohjelmistoja. Ratkaisujemme pääpainopiste on valmistavan teollisuuden tuotannossa.

HighByte on yhdysvaltalainen teollisten sovellusten kehittäjä, joka tuottaa ratkaisuja data-arkkitehtuurin ja tietoturvan tarpeisiin. HighByte uskoo, että kontekstuaalinen ja standardoitu data on välttämättömyys teollisuus 4.0:n haasteisiin vastaamiseksi, ja siksi he loivat HighByte Intelligence Hubin. Hubin avulla teollisuusyritykset pystyvät yhdistymään tuotantolaitteisiinsa ja hyödyntämään niiden dataa kirjoittamatta riviäkään koodia.