Informatica

Driving Business Agility through DataOps: Governance, Automation, and Observability with Informatica

Jesse Garcia

Solution Architect – Healthcare Industry

Where data & AI come to LIFE

Your Speaker

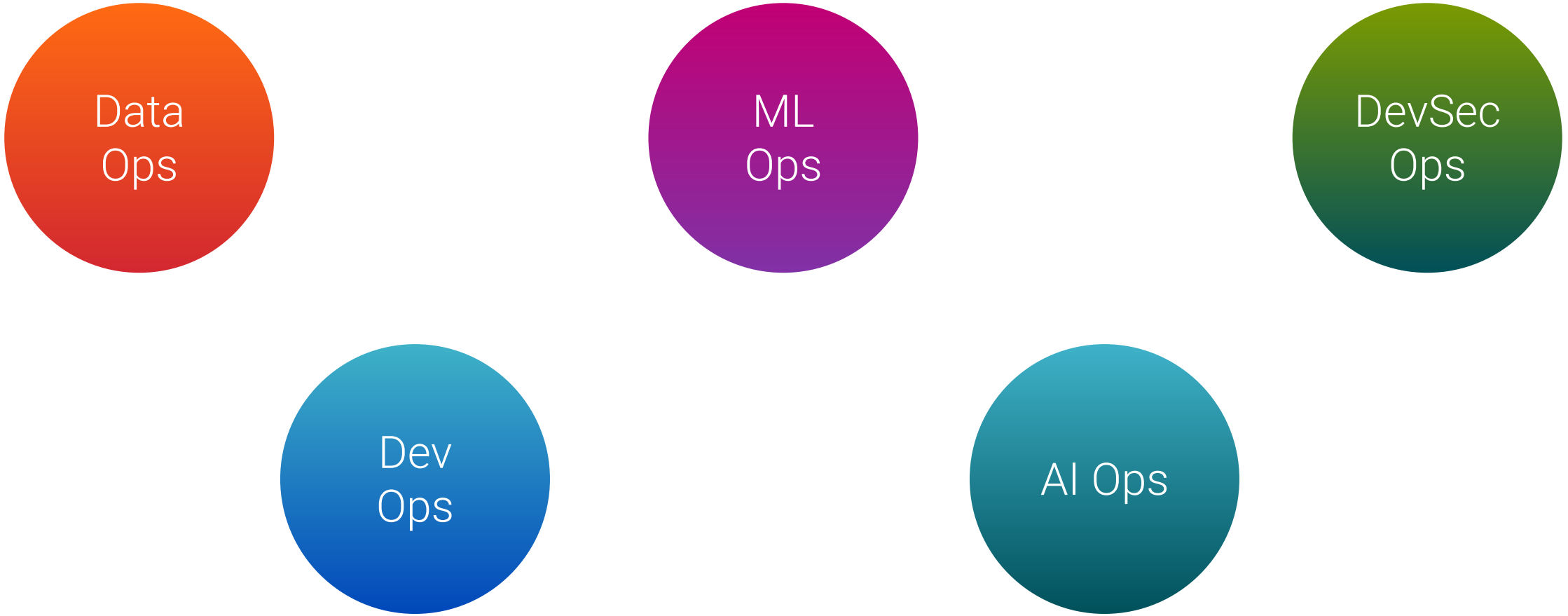
Jesse Garcia

Pre Sales Solution Architect at Informatica

Agenda

1. Overview of All the Ops
2. DataOps Capabilities and Building Blocks
3. DataOps + MLOps
4. Driving DataOps and MLOps Success with Informatica
5. Key Takeaways
6. Q&A

It’s an ”Ops” world … making sense of the terminology



What the Ops!?!?!

	Goal (Increase Agility Of)	People (Improve Collaboration For)	Process (Accelerate)	Technology (Enterprise Class)
DevOps	Combines software development and IT operations to **speed up delivery**	Development, QA, Operations	Plan → Code → Build → Test → Release → Deploy → Operate → Monitor	CI/CD for apps: SDLC, configuration management, incident tracking and collaboration tools
DataOps	**Optimizes** rapid data pipeline deployment	Data analysts, Data engineers, IT ops	Catalog → Govern → Integrate → Cleanse –> Orchestrate → Monitor → Collaboration & Governance→ Delivery	CI/CD for data: Data cataloging, data integration, data engineering, data quality, automation & orchestration, collaboration & data governance, and data delivery
MLOps	**Operationalizes** scalable, reproducible, **continuously monitored** ML deployments	Data scientists, Business owner, Data engineers, Data ops	Understand → Govern → Acquire → Develop → Deploy → Monitor	DevOps, DataOps, data science productivity, AI governance, model performance

What the Ops!?!?!

	Goal (Increase Agility Of)	People (Improve Collaboration For)	Process (Accelerate)	Technology (Enterprise Class)
AIOps	Leverages AI and ML to **automate**, **optimize**, and **enhance** IT operations	IT ops, Sec ops, Dev ops, Application Support Teams	Detect → Analyze → Automate → Respond → Optimize → Monitor	Observability platforms, log analytics, correlation engines, anomaly detection
InfoSecOps	Reduce risk of **sensitive data** security breach and regulatory non-compliance	Office of CISO, Security analysts	Identify → Map → Assess → Prioritize → Remediate → Track → Audit	Data access management, data masking, data encryption

Key Capabilities

DataOps: The Foundational Building Blocks

From Scattered Blocks to a Strong, Unified Structure

The Challenge: Fragmented Data and Disjointed Systems

Fragmented Data
Disconnected Systems
Inconsistent Quality
Limited Visibility
Data Chaos

Foundation: Data Integration
Sturdy Walls: Data Quality
Inspection: Data Governance & Data Observability
Open Doors: Data Democratization, Unified, Trusted Data

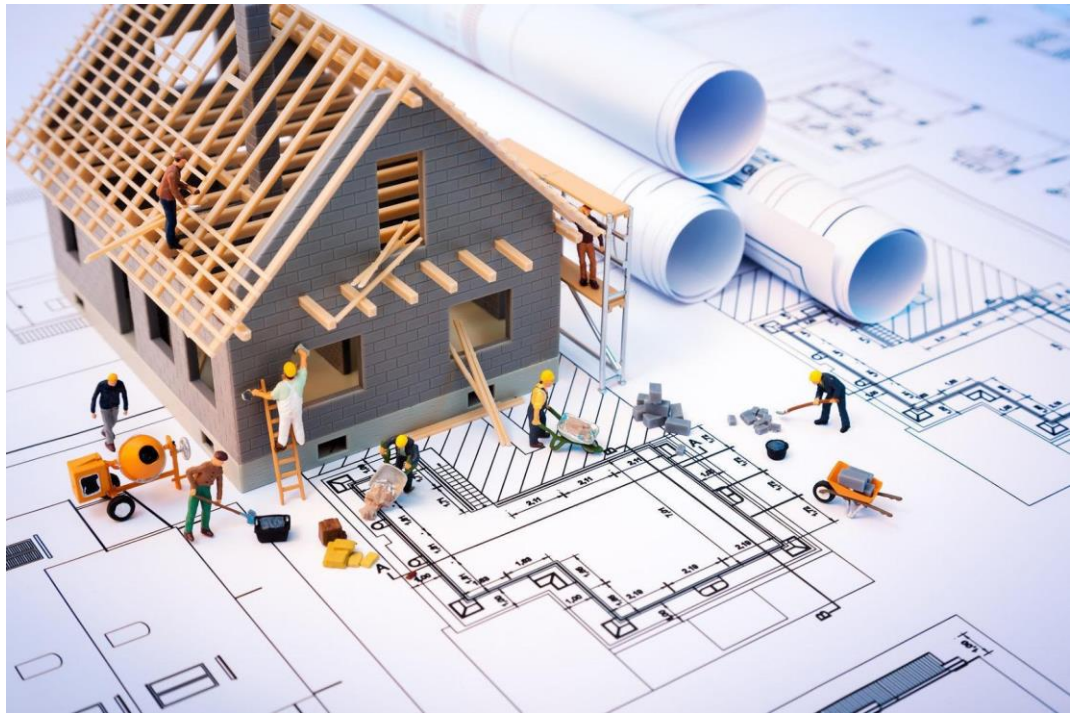
Creating a Structured, Trusted Data Environment

DataOps: Powering the Future of Data Management

Building the Foundation

Data Integration

- Automated data ingestion
- Batch and real-time integration
- Robust set of connectivity
- Scalable Data Pipelines
- Pipeline Orchestration and Automation
- Data Transformation
- Monitoring and Error Handling
- Secure Data Transfers



DataOps: Powering the Future of Data Management

Building Sturdy, Flawless Walls for your Data

Data Quality

- Data Profiling
- Data Validation
- Data Cleansing
- Standardization
- Data Enrichment
- Automated Quality Checks
- Root Cause Analysis
- Quality Metrics and Reporting
- Integration with Data Pipelines

DataOps: Powering the Future of Data Management

Inspecting the Rooms and Watching the House

Data Governance and Data Observability

- Policy Management
- Data Stewardship
- Data Cataloging
- Access Controls
- Audit Trails
- Regulatory Compliance
- Metadata Management
- Data Lineage
- Real-time Monitoring
- End-to-End Visibility
- Alerting Systems
- Anomaly Detection
- Root Cause Analysis
- Performance Metrics
- Data Freshness Tracking
- Dashboarding & Reporting

DataOps: Powering the Future of Data Management

Opening the Doors for Everyone to Access Insights

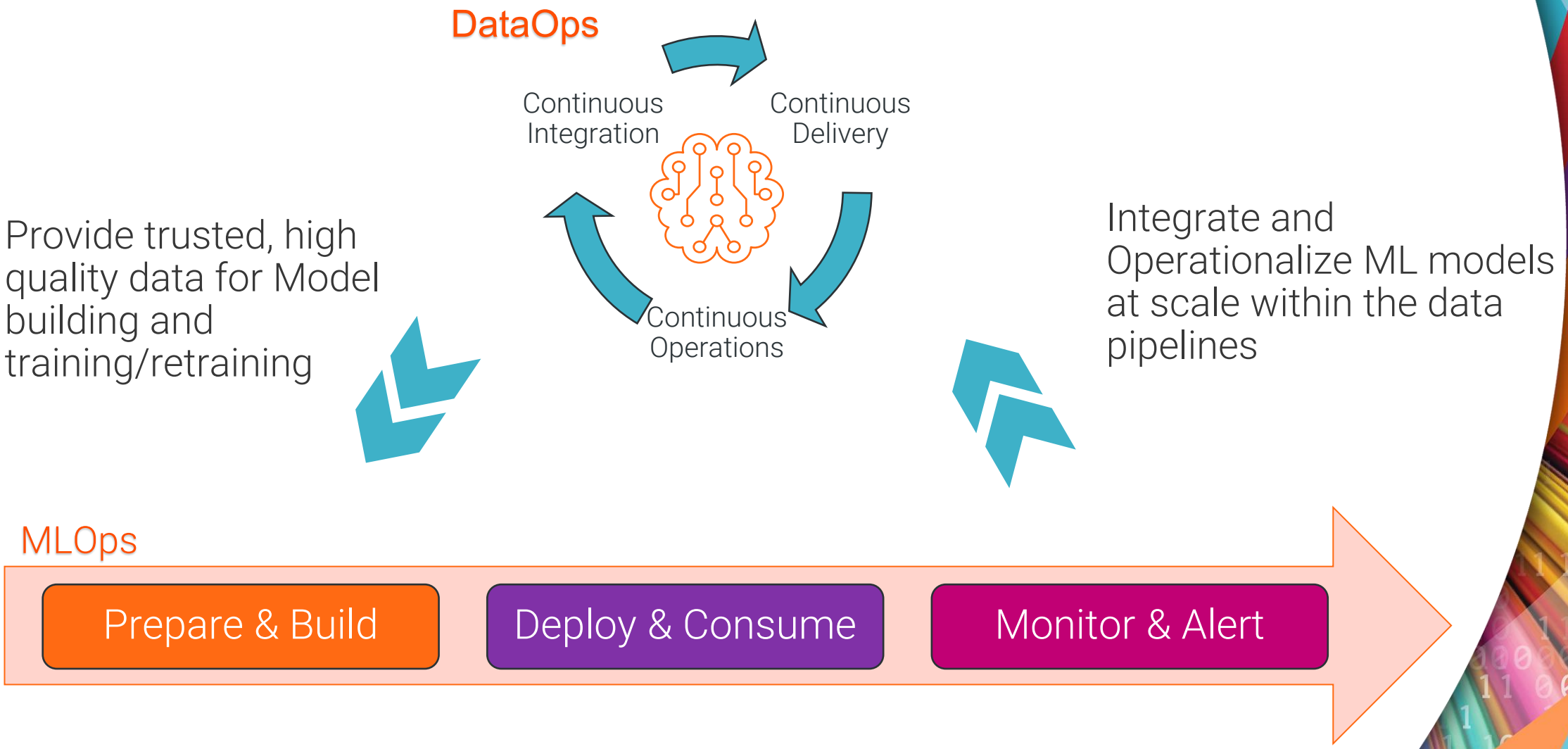
Data Delivery/Democratization

- Self-Service Access
- Data Catalogs and Marketplaces
- Role-Based Access Control
- Data APIs and Interfaces
- Data Transformation for Usability
- Data Sharing and Collaboration
- Automated Data Delivery
- Monitoring Usage and Adoption

Driving Intelligent Automation with DataOps and MLOps

DataOps + MLOps

DataOps goes hand in hand with MLOps



It’s an “Ops” World: Operationalize to Increase Agility

Technology

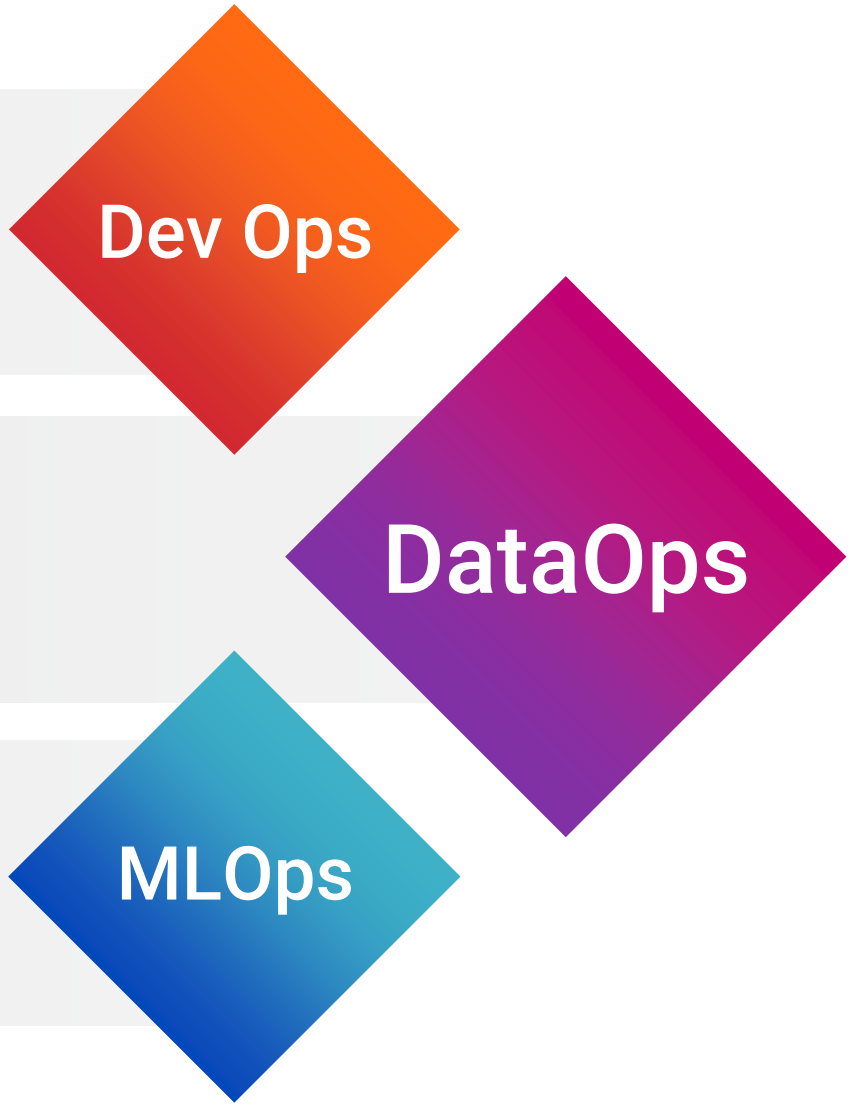
CI/CD for Apps: SDLC, configuration management, incident tracking and collaboration tools.

Technology

CI/CD for Data: Data cataloging, data integration, data engineering, data quality, data provisioning, data governance, collaboration and monitoring tools

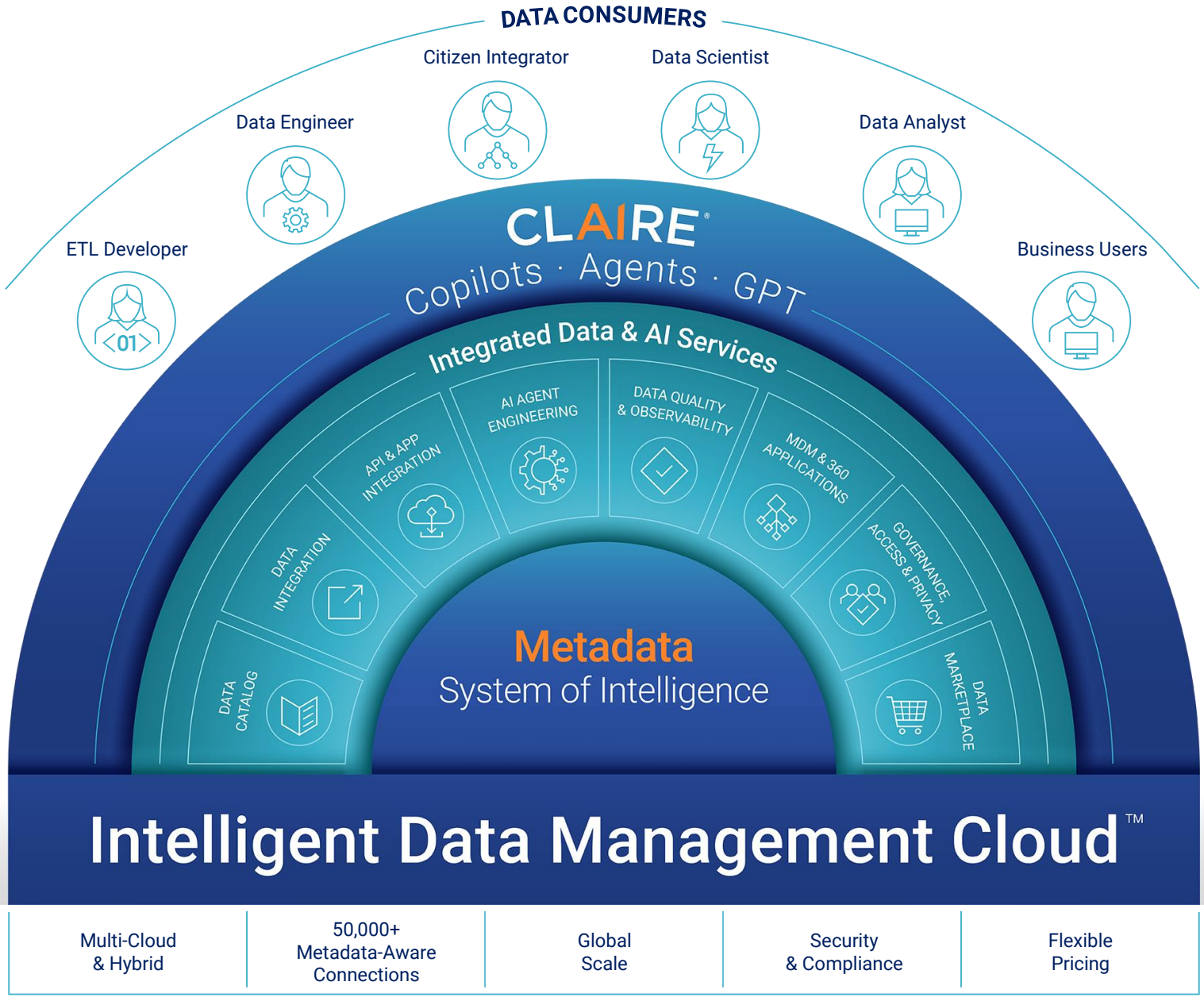
Technology

CI/CD for AI: DevOps, DataOps, data science productivity, AI governance, model performance.



Enabling Scalable DataOps with Informatica

Harnessing the Power of Intelligent Data Management



Unified Platform for Data Democratization

Cloud Data Marketplace

- Shared Business Context
- Ownership, Policy and Rules

Cloud Data Integration

- Replication, batch, real-time integration patterns
- Connectivity, scalable pipelines, orchestration and automation

Cloud Data Governance & Catalog

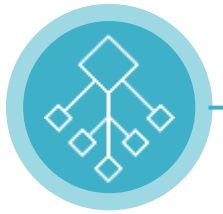
- Technical Metadata Inventory
- “Map of the Enterprise” Data

Cloud Data Quality

- Automated data quality checks
- Validation, cleansing, standardization

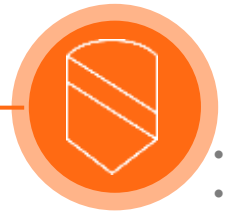
Data Observability

- Real-time monitoring
- End-to-end visibility



Master Data Management

- Ensure reliability of key data domains

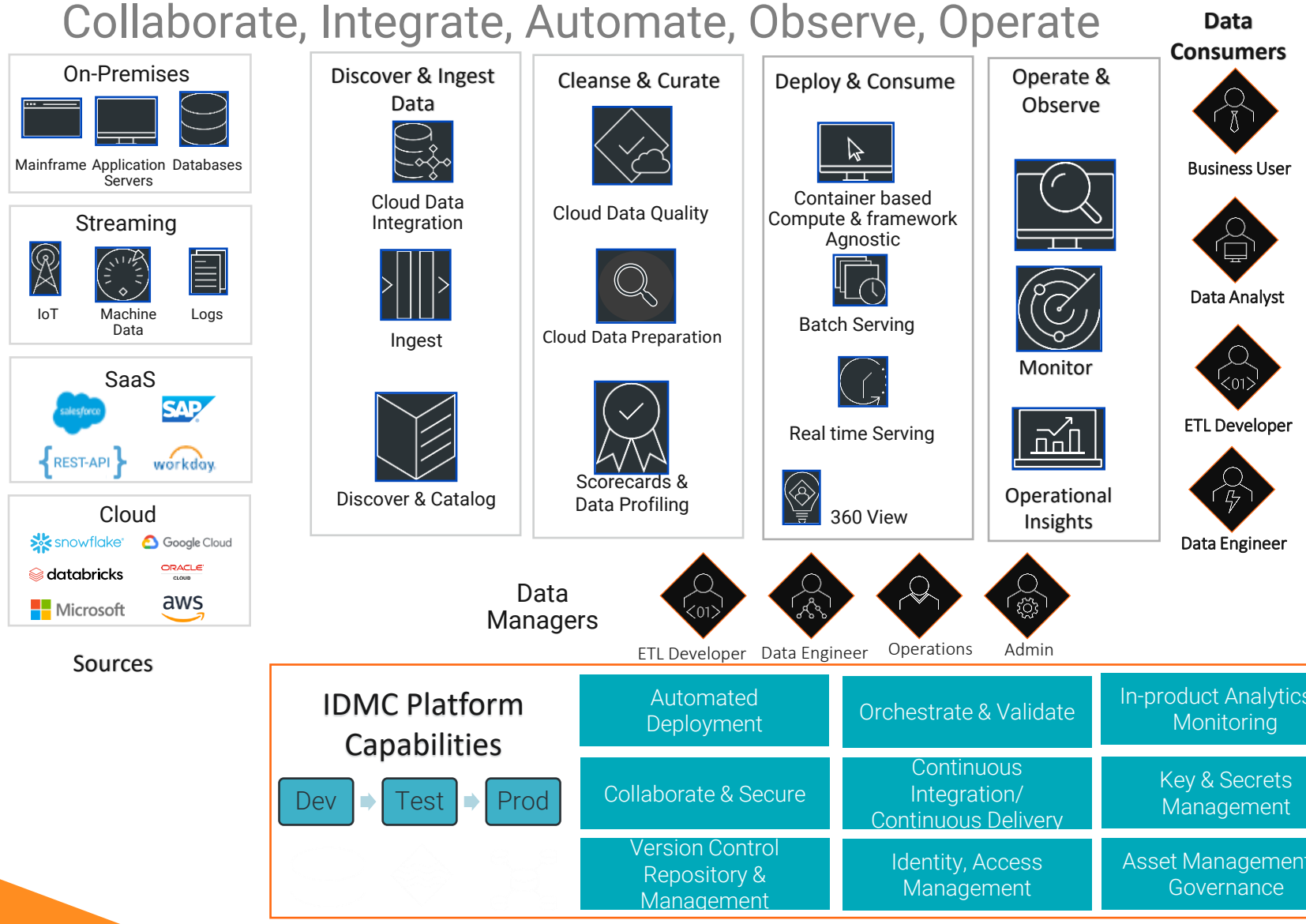


Cloud Data Access Management

- "Trust but Verify“ policies
- Monitor and manage data proliferation risk

- Where is the data I need?
- What’s in it?
- Is it any good?
- Can I use it?
- Should I use it?
- Who owns it?
- How do I get it?

It's an "Ops" World: Intelligent Data Management with IDMC Accelerates Outcomes



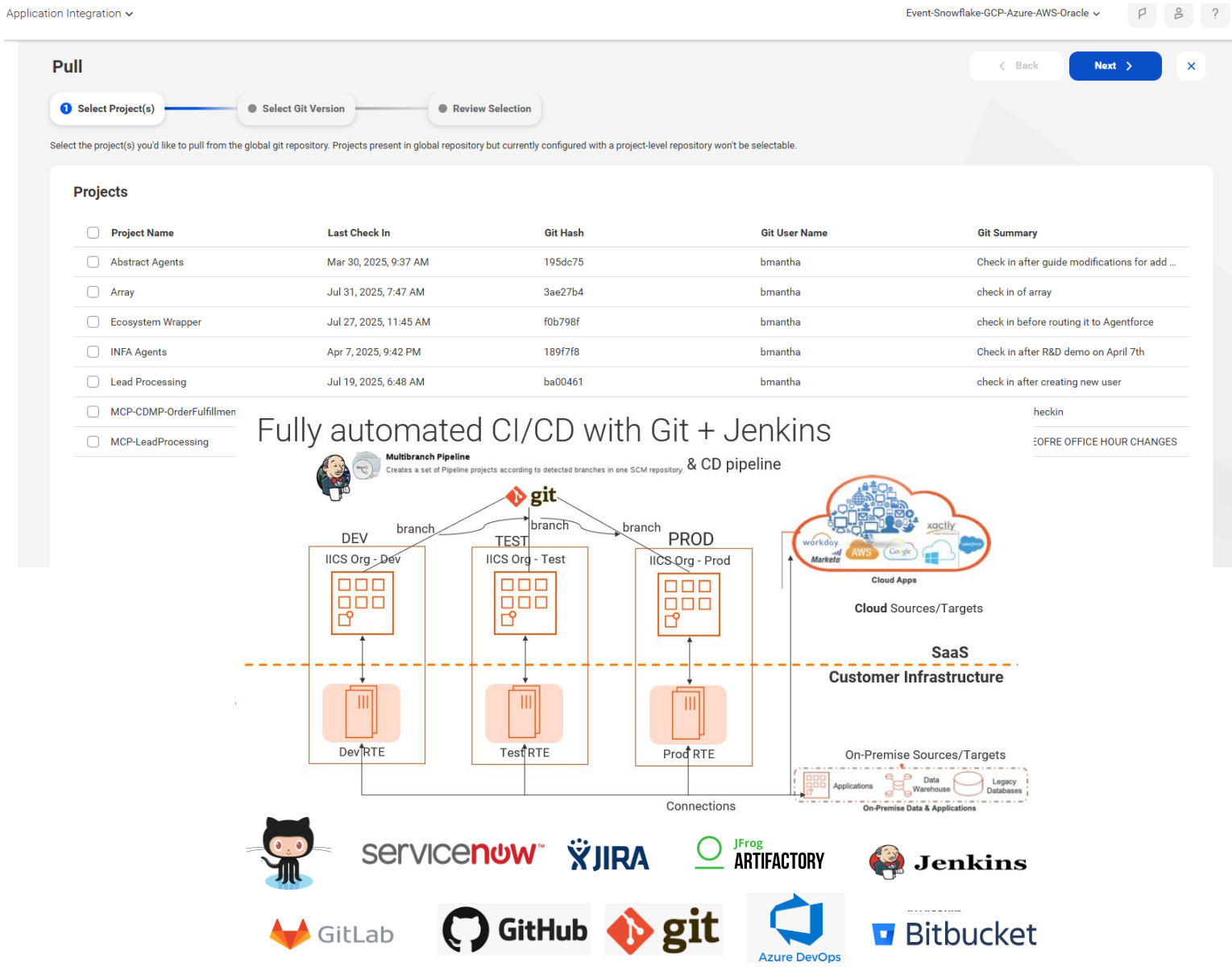
Enabling Robust DataOps Practice with

- Platform-wide Data Governance
- Built-in DevOps and CI/CD
- Automated Deployment
- Orchestration & Validation
- Infrastructure Performance visibility
- Operational Observability
- SLA Tracking
- Continuous Data Scoring & Cleansing

Collaboration

Collaborate

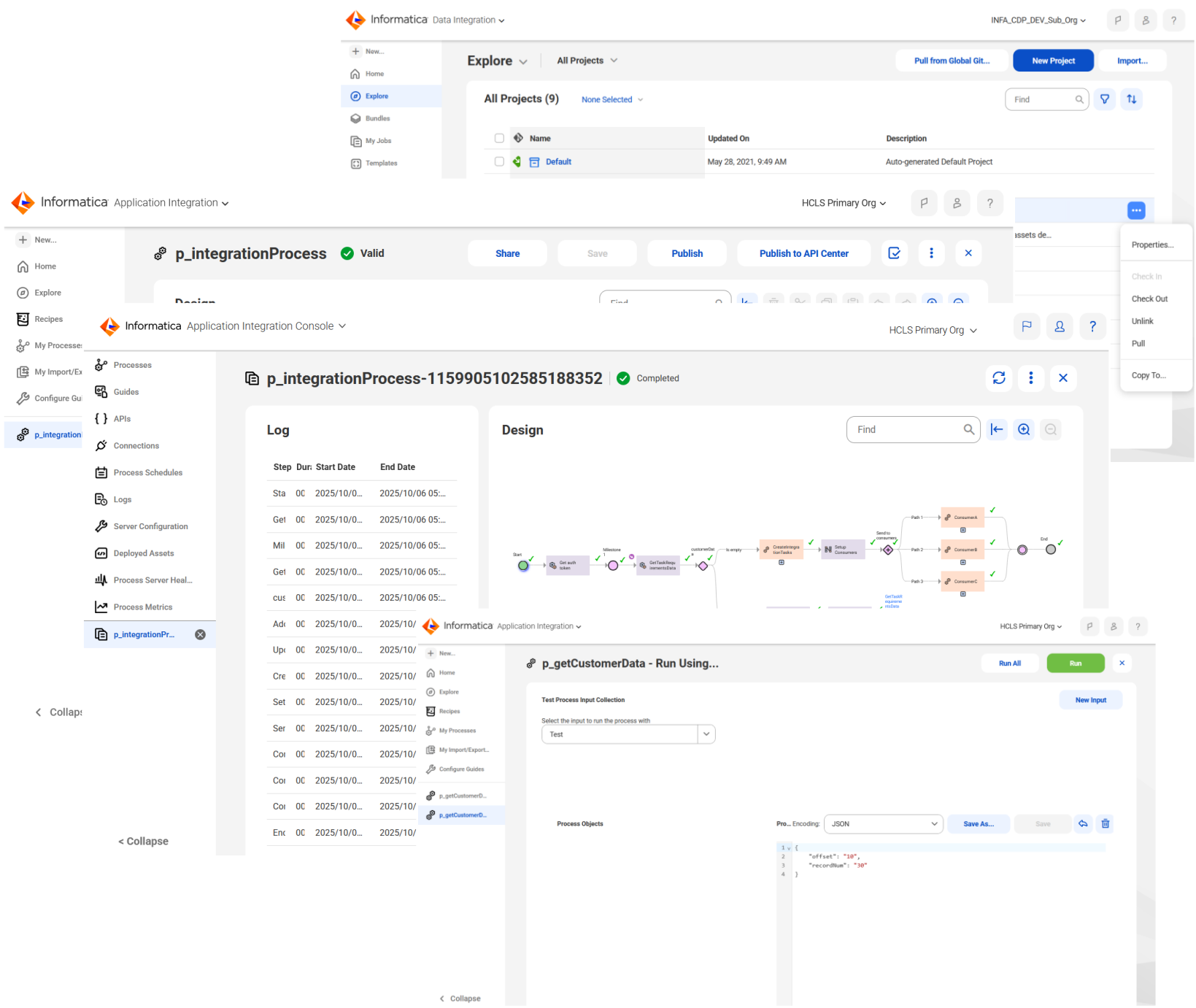
- Develop data pipelines in a way that promotes and automates collaboration and instant adaptation to change
- Collaboration with built in support for authorizations and approvals Workflow
- Democratized Data Integration
- Extensive support for Version control integration
- Infrastructure as code provisioning templates for runtime environment
- Connect data producers and sources with consumers



Integrate

Integrate

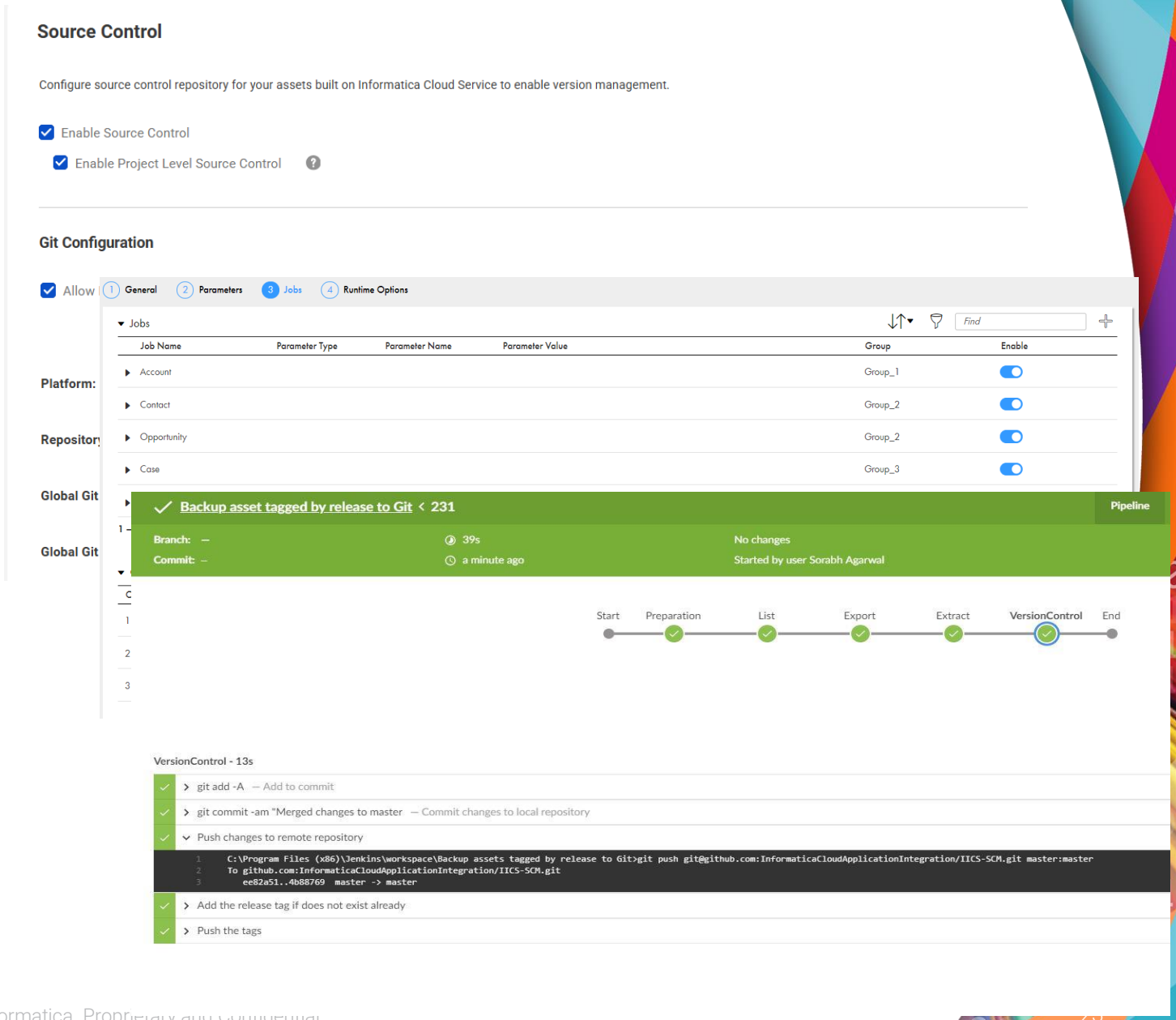
- Orchestrate complex pipelines and task flows
- Collaboration with built in support for authorizations and approvals Workflow
- Programmatically create, schedule, and monitor pipelines
- Provide test inputs for data pipelines for automated testing
- Extensive support for Version control integration



Automation

Fully Automated & Self-service data delivery

- Fully Automated Dev Ops –Code migration (CI/CD)
- Automated data pipeline with Testing & Validation
- Integrated Data Quality & Governance
- API based Automation - Different modes of Data delivery based on user persona & need
- Reusable Templates – Dynamic mappings
- Automated User role-group provisioning with SAML
- Infrastructure as code provisioning templates for runtime environment

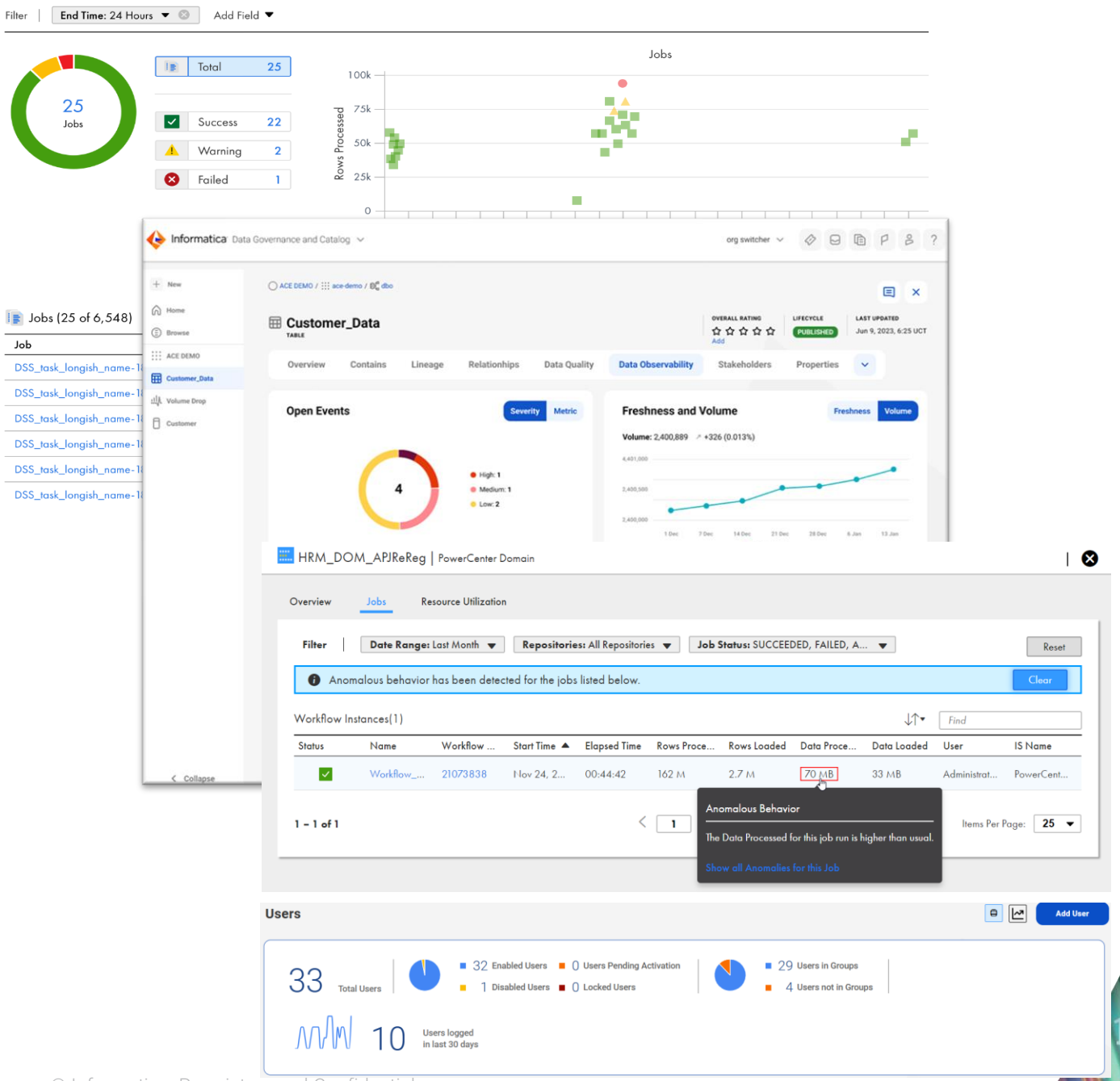


Informatica

Observability

Observe

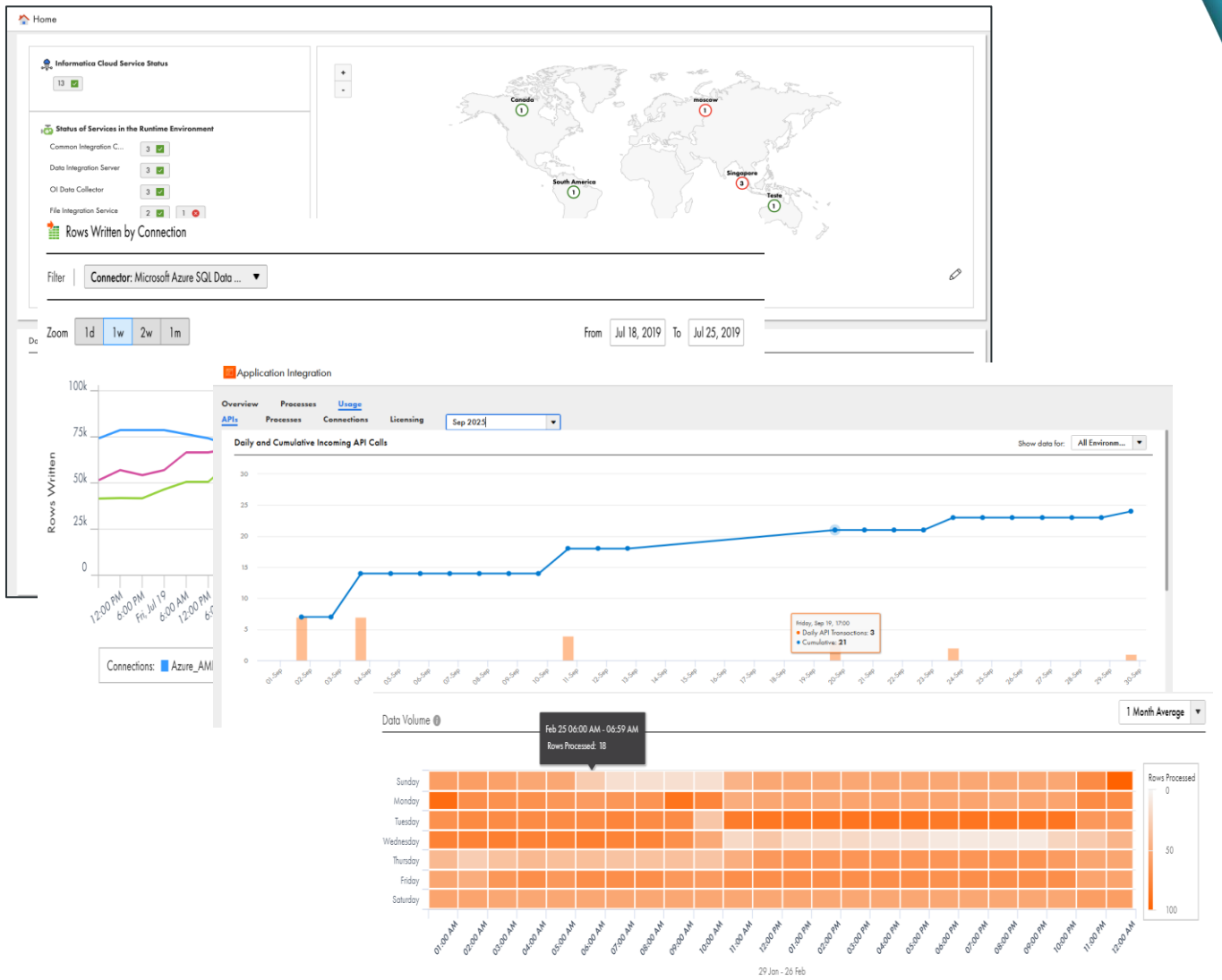
- Validate data pipeline in sandbox and pre-prod environments
- Automate testing and validation in sandbox, pre-prod orgs before deploying changes in production
- Detect data quality issues and anomalies
- Get alerted and Drill down and troubleshoot jobs that contribute to resource spikes
- Visualize task runs and detect anomalies
- User and user group analytics



Operate

Continuous Operations

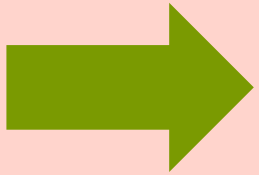
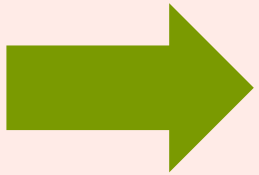
- Intelligent operational analytics
- CLAIRE powered self heal, Auto Tune, auto scale, smart shutdown
- Alerting at infrastructure, data, and service levels
- View trends across sources and connections
- Gain real-time visibility and alerts into API, data volume, job, resource, project level monitoring
- Heat map to identify critical times and peak hours for resource planning
- AI powered monitoring – built in optimization, cost governance



Driving Business Agility through DataOps

Key Takeaways

Key Takeaways

Key Pillars of DataOps	How Informatica Supports DataOps
Data Integration 	Cloud Data Integration and Replication Cloud Application and Integration ✓ Automated batch & real-time ingestion ✓ Scalable connectivity & pipelines ✓ Pipeline orchestration & automation ✓ Data transformation ✓ Monitoring & error handling ✓ Secure data transfer
Data Quality 	Cloud Data Integration and Replication Cloud Data Quality and Profiling Cloud Data Governance and Catalog ✓ Data profiling, validation, and cleansing ✓ Standardization and enrichment ✓ Automated quality checks ✓ Root cause analysis ✓ Quality metrics and reporting ✓ Seamless pipeline integration

Key Takeaways

Key Pillars of DataOps	How Informatica Supports DataOps
Data Observability	Cloud Data Governance and Catalog Cloud Data Quality ✓ Real-time monitoring and end-to-end visibility ✓ Alerting and anomaly detection ✓ Root cause analysis ✓ Performance and freshness metrics ✓ Dashboarding and reporting
Data Governance	Cloud Data Governance and Catalog Cloud Data Access Management ✓ Policy management and stewardship ✓ Data cataloging and metadata management ✓ Access controls and audit trails ✓ Regulatory compliance ✓ Data lineage tracking

Key Takeaways

Key Pillars of DataOps	How Informatica Supports DataOps
Data Delivery/Democratization	Cloud Marketplace Cloud Data Governance and Catalog ✓ Self-service access and data catalogs ✓ Role-based access control ✓ Data APIs and transformation for usability ✓ Data sharing and collaboration ✓ Automated delivery with usage monitoring

Thank You

Where data & AI come to

™

Informatica