Data Observability, Monitoring and Data Quality (DOMAQ®)

Background

Exploding data sources and growing data insight consumption needs are making enterprise data pipelines complex. Business users, data analysts, data engineers and data scientists want to use more data from a wider variety of sources, often on a real-time basis, to improve decision-making. Data teams struggle to build and architect systems that can meet such changing needs and at the same time ensure the best quality of data. Most enterprise data teams still spend the bulk of their time firefighting daily operational issue, which means that the quality of the data often takes a back seat. The problem gets worse as data volumes grow and data pipelines get increasingly complex.

Fresh Gravity's DOMAQ Solution

Fresh Gravity's DOMAQ tool enables business users, data analysts, data engineers and data architects to detect, predict, prevent, and resolve issues, sometimes in an automated fashion, that would otherwise break production analytics and AI. It takes the load off from the enterprise data team by ensuring data is constantly monitored, data anomalies are automatically detected, and future data issues are proactively predicted without any manual intervention.

This comprehensive data observability, monitoring, and data quality tool is built on top of Databricks platform to ensure optimum scalability and use AI/ML algorithms extensively for accuracy and efficiency.

Key Features of Fresh Gravity's DOMAQ Tool

Comprehensive Data Monitoring & Data Observability

- Provides automated AI-driven Data Quality Reports & out-of-the-box AI support for all automated and Custom Analysis such as Custom Metric Analysis, Custom Validation Rules Analysis, and Comparison Analysis

- Determines whether the data is good or bad, explains why it is bad and provides recommendations on how to fix it

- Comes with 100+ out-of-the-box data quality metrics that cover column, length, frequency, pattern, and duplicity analysis , along with options for data drill down for further review. It also provides AI-driven actionable insights that help in root cause analysis

- Ability to define custom metrics tailored to unique business requirements

- Provides set alerts based on custom validation rules that makes it easy to test, validate and notify data against the organization's business rules. It also shares reasons for the problems and helps data steward make faster decisions

- Capable of bringing in multiple data sources and data quality metrics together to get a side-by-side comparison

- Ability to create customizable reports that are generated on a pre-defined schedule with options to share and download

- A Low-code/No-code data observability platform. It uses machine learning to infer and learn what your data looks like, proactively profiles & identify data issues, assess its impact, and alerts through integrations to common operations systems. It also allows root cause analysis and time series analysis

Seamless Data Management

- Ability to seamlessly connect to Snowflake, Redshift, BigQuery, & other popular OLAP & OLTP sources along with object stores like ADLS and S3

- Comes with auto generated pipelines (both one-time & incremental) to pull data from various sources to avoid impacting source systems during data analysis

- Helps users keep their data quality dashboards relevant by executing them on the latest data

- Comes with a unique feature that allows data quality checks on data sets obtained by combining discrete data sources

- Easy and intuitive UI to dive deeper into the DQ reports along with a sample data download option

- Flexible and customizable deployment options that support both Client or Fresh Gravity owned & managed deployments

- Provides granular access control mechanism to suit enterprise data security requirements
- AI/ML based metadata classification that allows mapping the incoming data dynamically and intelligently to a canonical model
- Helps simulate data cleansing operations without impacting source systems
- Enables data SMEs to make more informed decisions before they decide to harmonize the cleansed data back to source systems

Self-service UI That Encourages Collaboration

- Helps streamline day-to-day data operations for users through out-of-the-box task-based collaborative workflows
- Supports logically segregated work area for individual users with options to share with the team for better collaboration
- Allows assigning data review tasks to dedicated teams for immediate attention to review/fix data anomalies
- Ensures optimum reusability by using features that allow sharing custom queries and rules across the team

Seamless Deployment & Best-in-Class Security

- DOMAQ is a multi-tenant (DEV/TEST/PROD) SaaS offering where secure tenants can store, organize, & manage data
- Keeps customer data discrete across different data partners or customers
- Allows customer data to be accessed only via the Platform by proper authentication through the User Interface or through the DOMAQ API
- For Administrative tasks, DOMAQ staff's access to the Tenantis controlled through role-based access control (RBAC), enforcing the principle of Least Priviledge
- Data in REST is encrypted to AES-256-bit encryptionor hosting vendor encryption capabilities
- Follows RESTful architecture and utilizes secure electronic means to exchange data with customers using HTTPS
- Encrypted data is transmitted over the public Internet using TLS 1.2.
- DOMAQ REST API supports OAuth 2.0/OpenIDConnect (OIDC) Single Sign On (SSO) that allows clients to verify the identity of the end user based on the authentication performed by an authorization server.
- Comes with a fully managed infrastructure that allows enterprises to focus only on the data.
- Supports no impact, automated deployment & upgrade.
- Provides options for row and column level data encryption/masking
- Supports easy integration with enterprise SSO
- Maintains an activity log of all system activity including login, create, read, update, and delete activity
- Ensures business continuity and disaster recovery through distribution of Platform database nodes across a minimum of three AWS and GCP zoned as well as multiple servers within a zone
- Runs on auto scalable spark based Databricks engines and can be seamlessly scaled up and down based on data volume
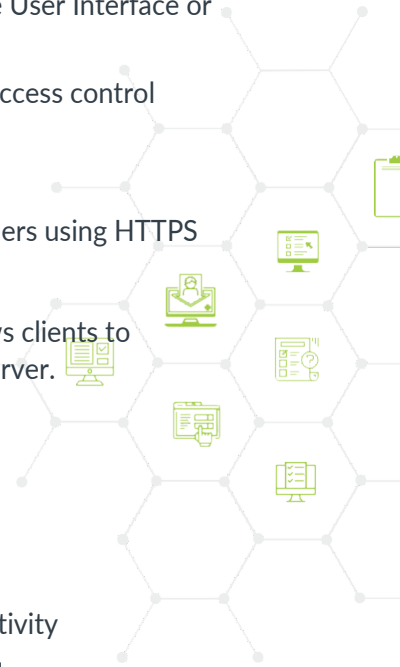
About Fresh Gravity

Founded in 2015, Fresh Gravity is an innovation-driven Data & Analytics consulting firm that helps businesses make data-driven decisions. We believe in the transformative power of data using technologies to help clients. To know more about us and our offerings, contact us at info@freshgravity.com