



innoQ Deutschland GmbH

Krischerstraße 100

40789 Monheim am Rhein

+49 2173 3366-0

info@innoq.com

www.innoq.com

Data Mesh Manager

Product Description and Scope

Product	Data Mesh Manager
Author	Jochen Christ
Version	1.2
Status	Final
Classification	Restricted

1 Introduction

The scope of this document is to give a product description for Data Mesh Manager (datamesh-manager.com) and defining the scope of the product, additional modules and professional services.

Content

1	Introduction	2
2	Data Mesh Manager	4
2.1	Core Features	5
2.2	Core Concepts	6
2.2.1	Data Contracts	6
2.2.2	Data Products	8
2.3	Enterprise Data Marketplace	10
2.4	Data Quality Enforcement	16
2.5	Data Governance (AI)	18
2.6	Authentication and Integrations	20
3	Customizations	21
4	Professional Services	22
4.1	Support	22
4.2	Additional Customization	23
4.3	Consulting	24

2 Data Mesh Manager

Data Mesh Manager is a web application to manage data contracts, data products, and data governance.

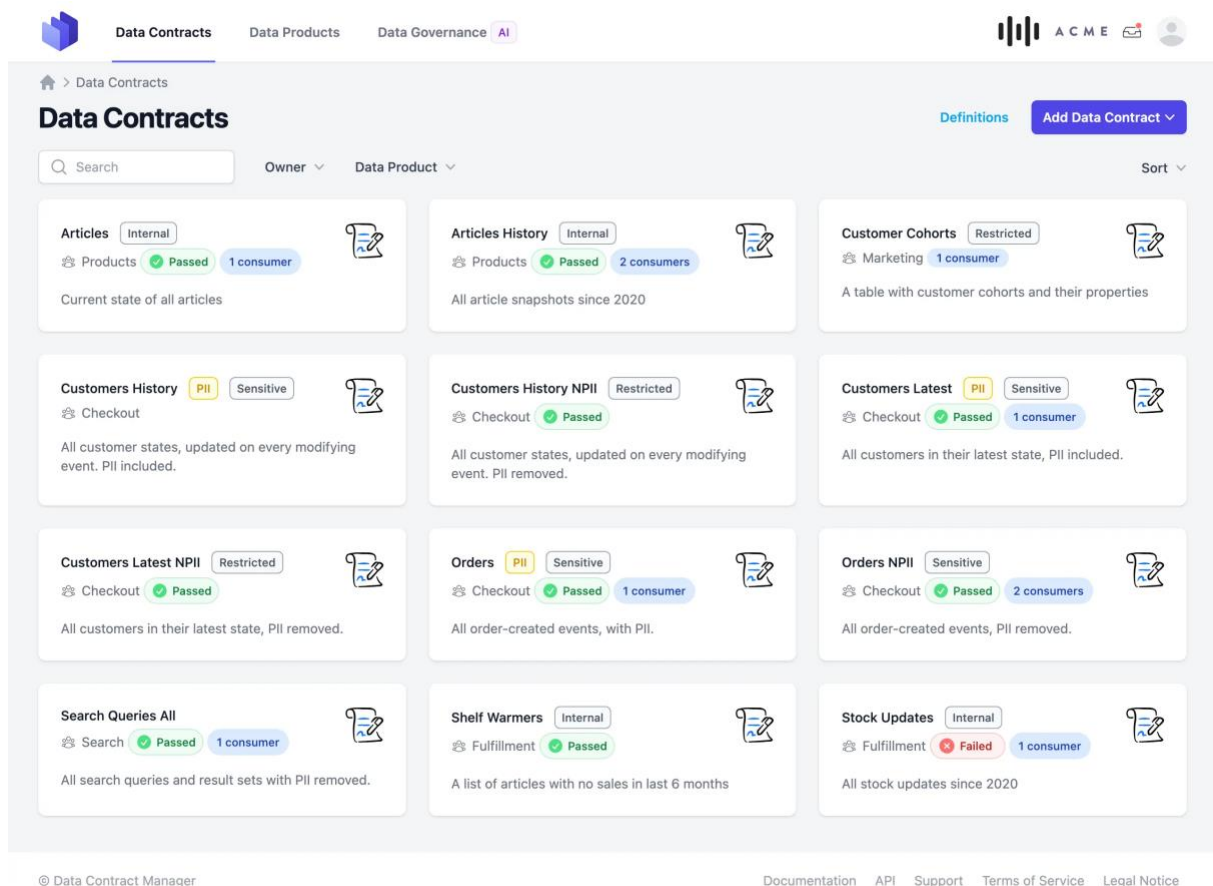


Figure 1: Data Contracts Screenshot

2.1 Core Features

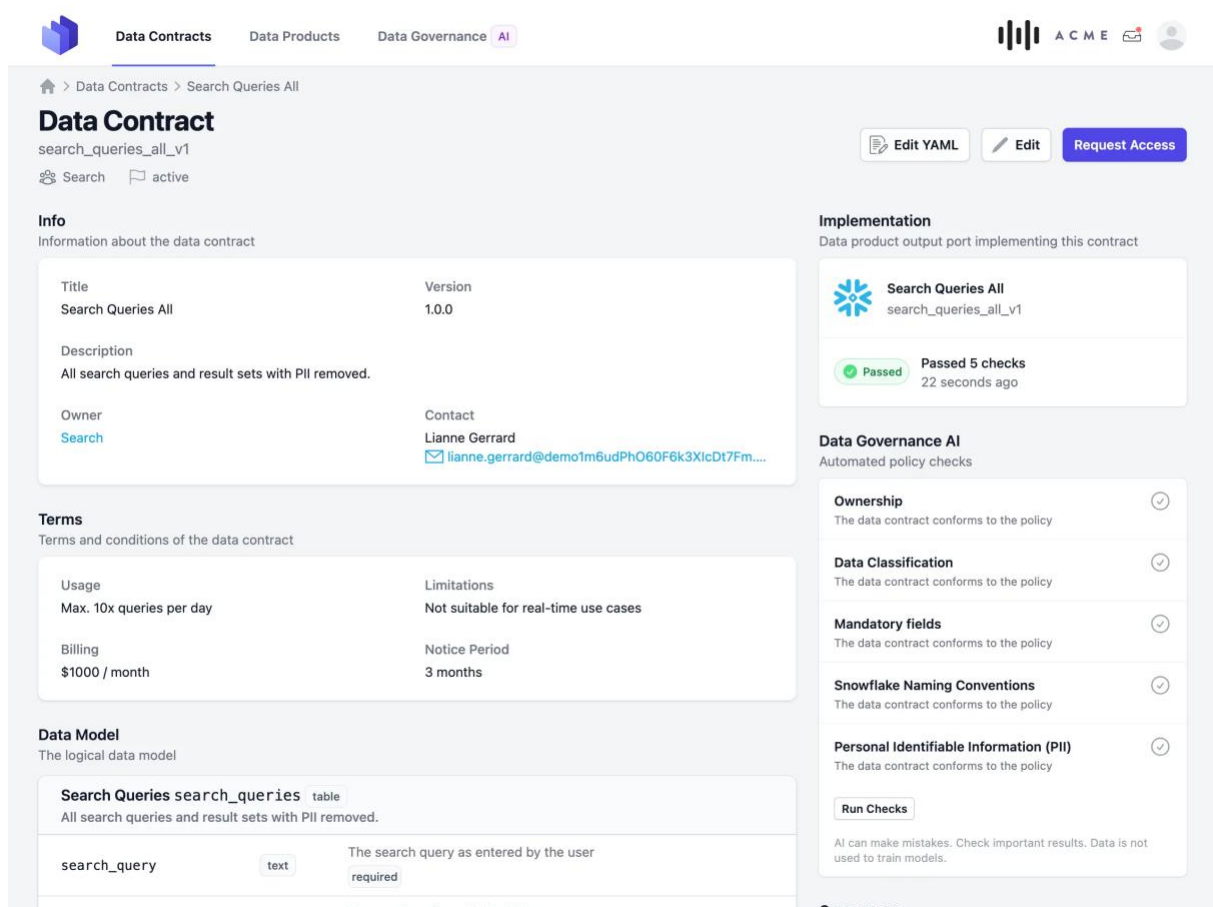
The application uses data contracts and data product metadata to provide these core features:

1. Enterprise Data Marketplace
2. Data Quality Enforcement
3. Data Governance Automation

2.2 Core Concepts

2.2.1 Data Contracts

A data contract is a document that defines the structure, format, semantics, quality, and terms of use for exchanging data between a data provider and their consumers. A data contract specifies a data set.



The screenshot shows a web interface for managing data contracts. The top navigation bar includes 'Data Contracts', 'Data Products', and 'Data Governance'. The main content area is titled 'Data Contract' for 'search_queries_all_v1'. It includes an 'Info' section with details like Title, Version (1.0.0), Description, and Owner. A 'Terms' section lists usage limits and billing. A 'Data Model' section shows a table schema for 'search_queries'. On the right, an 'Implementation' section shows a 'Search Queries All' data product with a 'Passed' status. Below that, a 'Data Governance AI' section lists policy checks for Ownership, Data Classification, Mandatory fields, Snowflake Naming Conventions, and Personal Identifiable Information (PII), all of which are passed. A 'Run Checks' button is also present.

Figure 2: Data Contract Screenshot

Features:

- **Data Contract Discovery**
An overview over available data contracts, including search and filter
- **Data Contract HTML view**
Make data contracts accessible for non-technical users

- **New Data Contract Wizard**

Guide users through data contract creation process

- **Data Contract YAML editor**

To edit data contracts directly, with code-completion and schema check.

- **Data Quality**

Integrate and view test results from Data Contract CLI directly in the web application

2.2.2 Data Products

A data product is a logical grouping of data assets on the data platform around a business concept with a defined ownership and lifecycle.

A data product implements one or multiple data contracts.

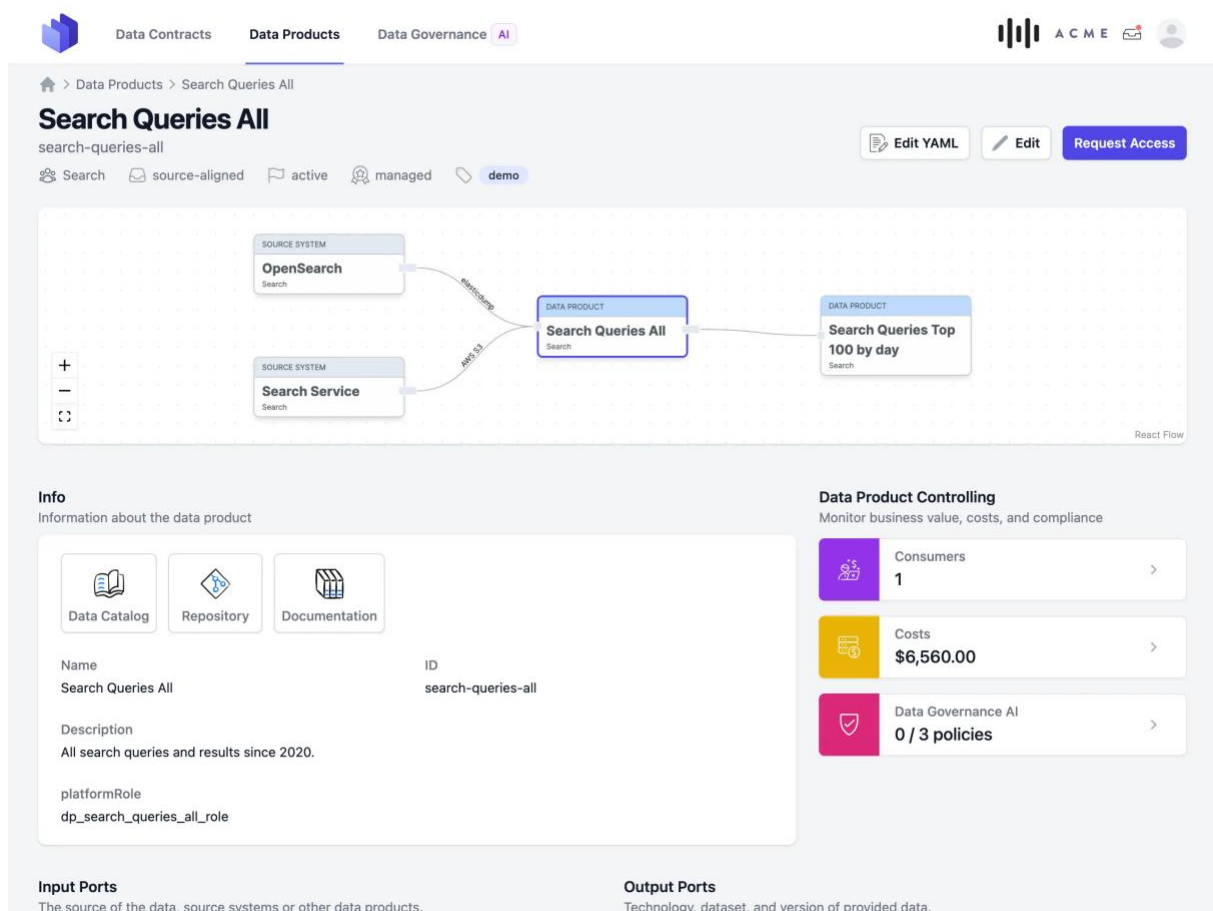


Figure 3: Data Product Screenshot

Features:

- **Data Product Management**
Define data product metadata
- **Output Ports**
Define output ports of data products, including the server details

- **Data Product Cost Management**

Track costs of compute, storage, maintainance and other resources per data product

2.3 Enterprise Data Marketplace

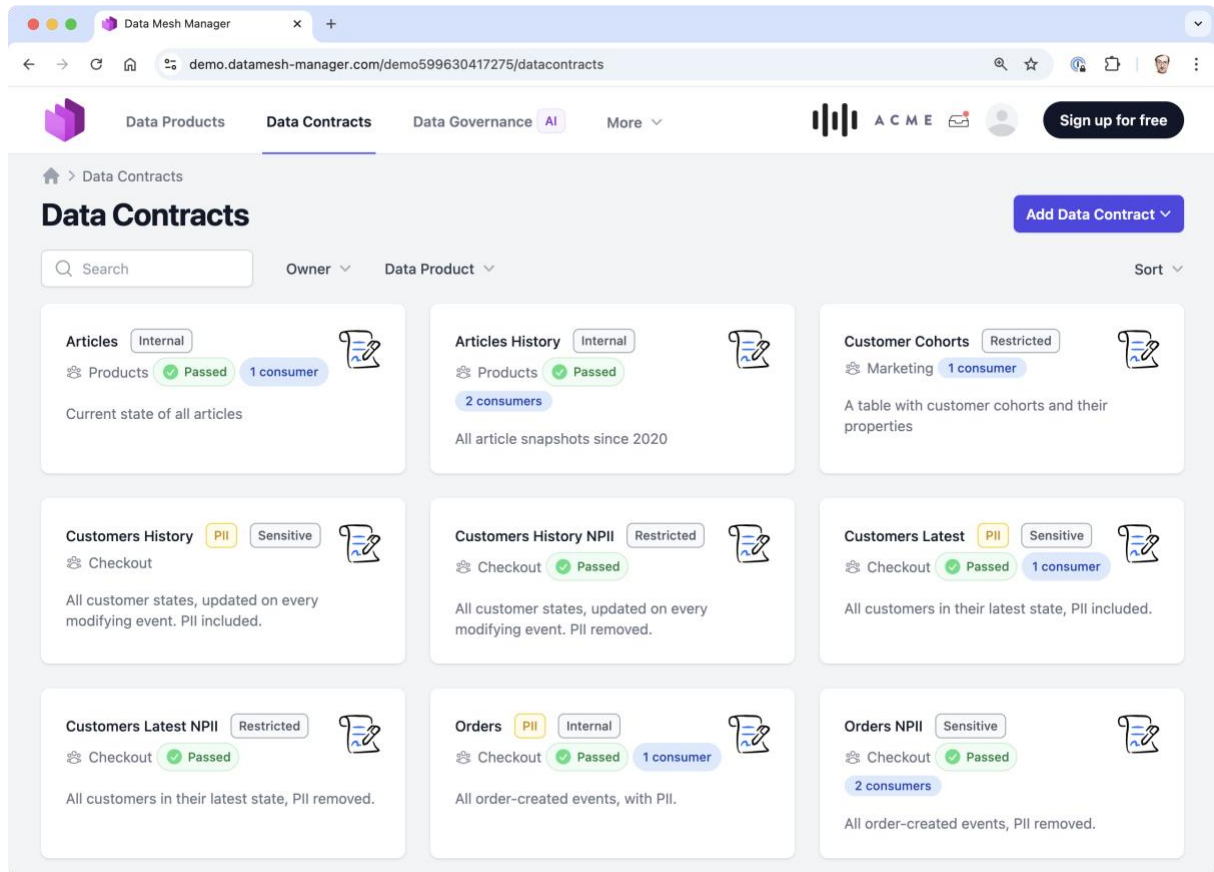
Data contracts contain all the information (including data model, data quality, guaranteed service level, terms and conditions, example data) that a data consumer needs to know to evaluate if a data set is suitable for their use case. This makes them a perfect foundation to build a data marketplace by combining them on a .

Data Mesh Manager contains the features to build a data marketplace:

1. Discovery and Catalog
2. Details Page, rendered as HTML
3. Request and Approval Workflow
4. Data Governance Checks
5. Access and Permission Automation

Example workflow:

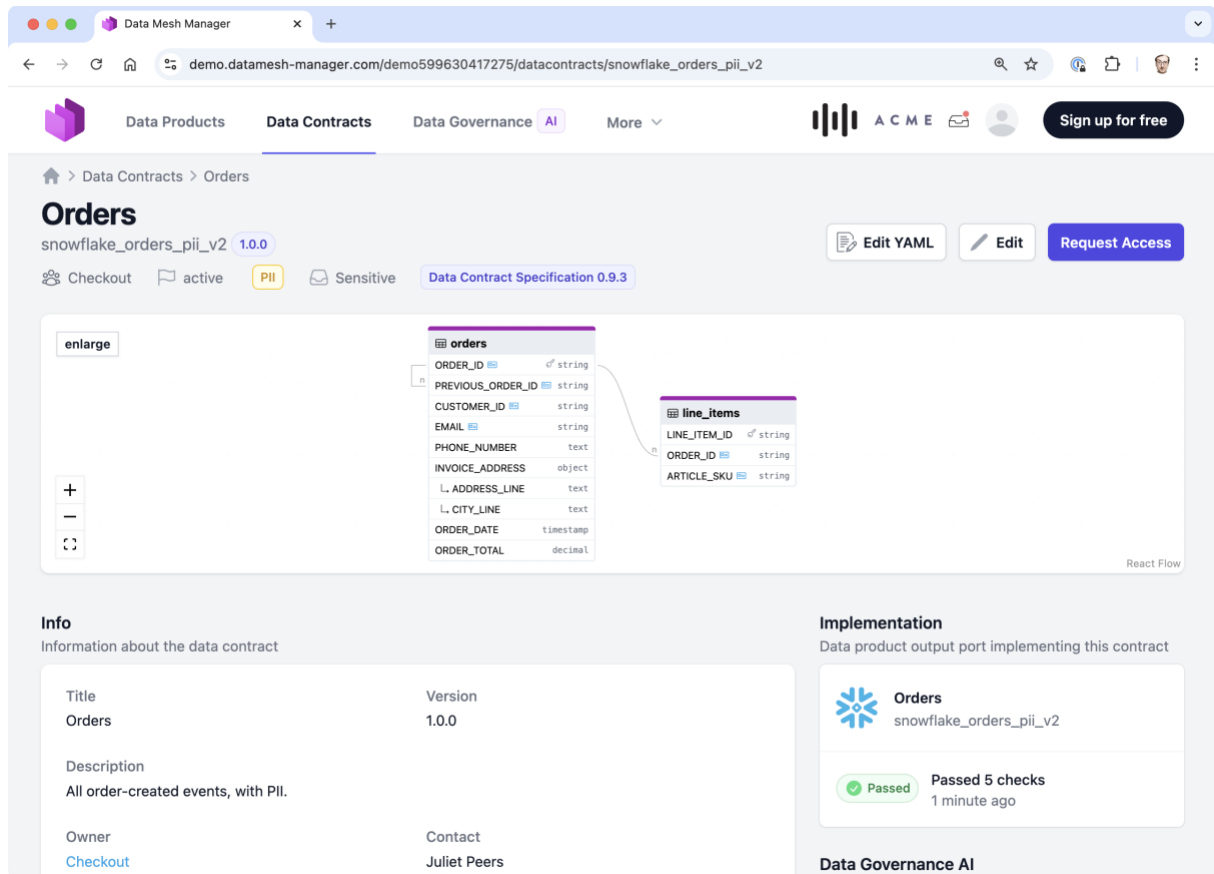
1. Discover available data, with search and filter capabilities:



The screenshot shows the Data Mesh Manager interface. The browser address bar displays `demo.datamesh-manager.com/demo599630417275/datacontracts`. The navigation bar includes tabs for Data Products, Data Contracts (selected), Data Governance, and AI. A search bar and filters for Owner and Data Product are present. A grid of Data Contracts is displayed, each with a title, status, and description.

Contract Name	Category	Status	Consumers	Description
Articles	Internal	Passed	1 consumer	Current state of all articles
Articles History	Internal	Passed	2 consumers	All article snapshots since 2020
Customer Cohorts	Restricted	Passed	1 consumer	A table with customer cohorts and their properties
Customers History	PII, Sensitive	Passed	-	All customer states, updated on every modifying event. PII included.
Customers History NPPI	Restricted	Passed	-	All customer states, updated on every modifying event. PII removed.
Customers Latest	PII, Sensitive	Passed	1 consumer	All customers in their latest state, PII included.
Customers Latest NPPI	Restricted	Passed	-	All customers in their latest state, PII removed.
Orders	PII, Internal	Passed	1 consumer	All order-created events, with PII.
Orders NPPI	Sensitive	Passed	2 consumers	All order-created events, PII removed.

2. Evaluate the data contract details page



The screenshot shows the 'Orders' data contract details page in the Data Mesh Manager. The page is titled 'Orders' and shows the contract 'snowflake_orders_pii_v2' with version '1.0.0'. It includes a 'Data Contract Specification 0.9.3' and a 'React Flow' diagram showing the relationship between 'orders' and 'line_items' tables. The 'orders' table has fields: ORDER_ID (string), PREVIOUS_ORDER_ID (string), CUSTOMER_ID (string), EMAIL (string), PHONE_NUMBER (text), INVOICE_ADDRESS (object), L_ADDRESS_LINE (text), L_CITY_LINE (text), ORDER_DATE (timestamp), and ORDER_TOTAL (decimal). The 'line_items' table has fields: LINE_ITEM_ID (string), ORDER_ID (string), and ARTICLE_SKU (string). The page also includes an 'Info' section with details about the contract, an 'Implementation' section showing the contract is passed, and a 'Data Governance AI' section.

Orders
snowflake_orders_pii_v2 1.0.0

Checkout active PII Sensitive Data Contract Specification 0.9.3

orders

Field	Type
ORDER_ID	string
PREVIOUS_ORDER_ID	string
CUSTOMER_ID	string
EMAIL	string
PHONE_NUMBER	text
INVOICE_ADDRESS	object
L_ADDRESS_LINE	text
L_CITY_LINE	text
ORDER_DATE	timestamp
ORDER_TOTAL	decimal

line_items

Field	Type
LINE_ITEM_ID	string
ORDER_ID	string
ARTICLE_SKU	string

Info
Information about the data contract

Field	Value
Title	Orders
Version	1.0.0
Description	All order-created events, with PII.
Owner	Juliet Peers

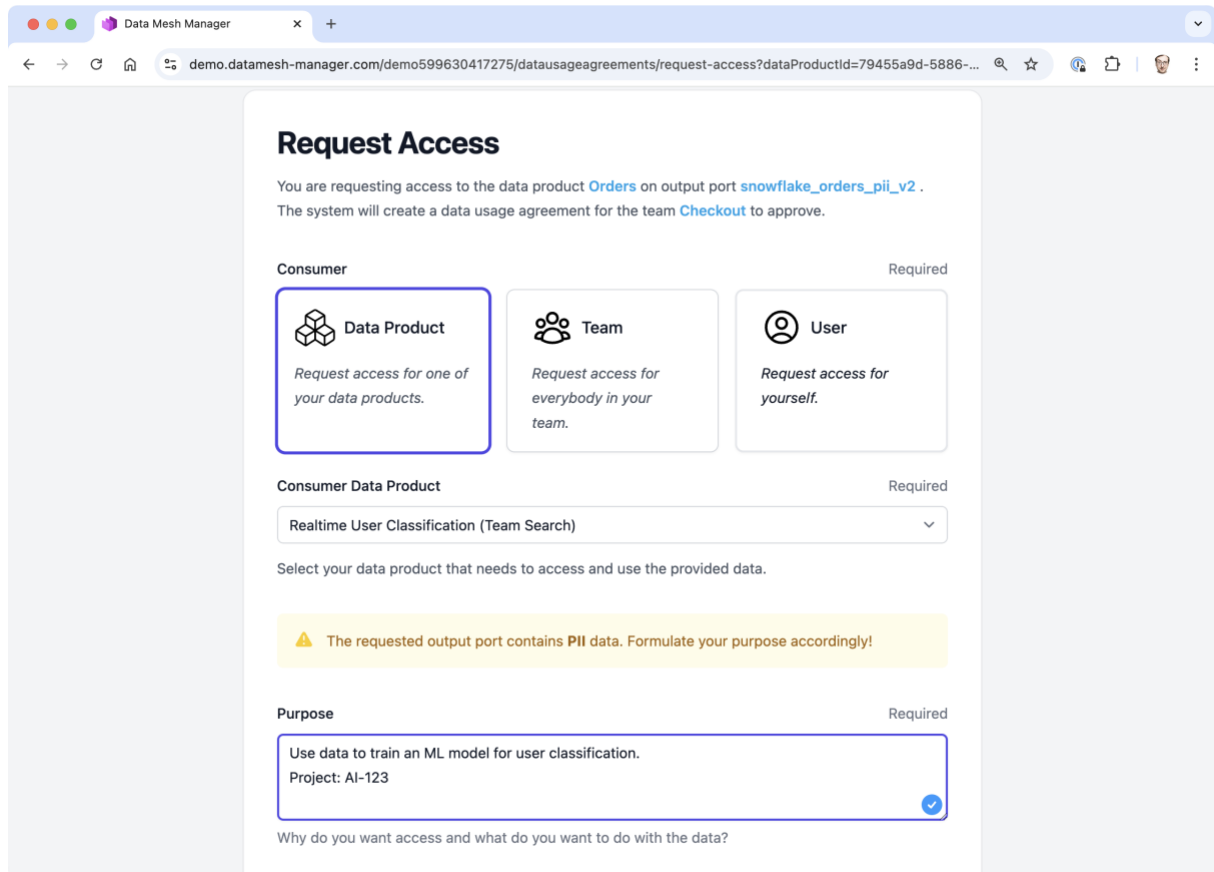
Implementation
Data product output port implementing this contract

Orders
snowflake_orders_pii_v2

Passed Passed 5 checks
1 minute ago

Data Governance AI


3. A data consumer requests access for themselves as an individual or team, or for a consuming data product.





Request Access

You are requesting access to the data product **Orders** on output port **snowflake_orders_pii_v2** .
The system will create a data usage agreement for the team **Checkout** to approve.

Consumer Required


Data Product
Request access for one of your data products.


Team
Request access for everybody in your team.


User
Request access for yourself.

Consumer Data Product Required

Realtime User Classification (Team Search) ▼

Select your data product that needs to access and use the provided data.

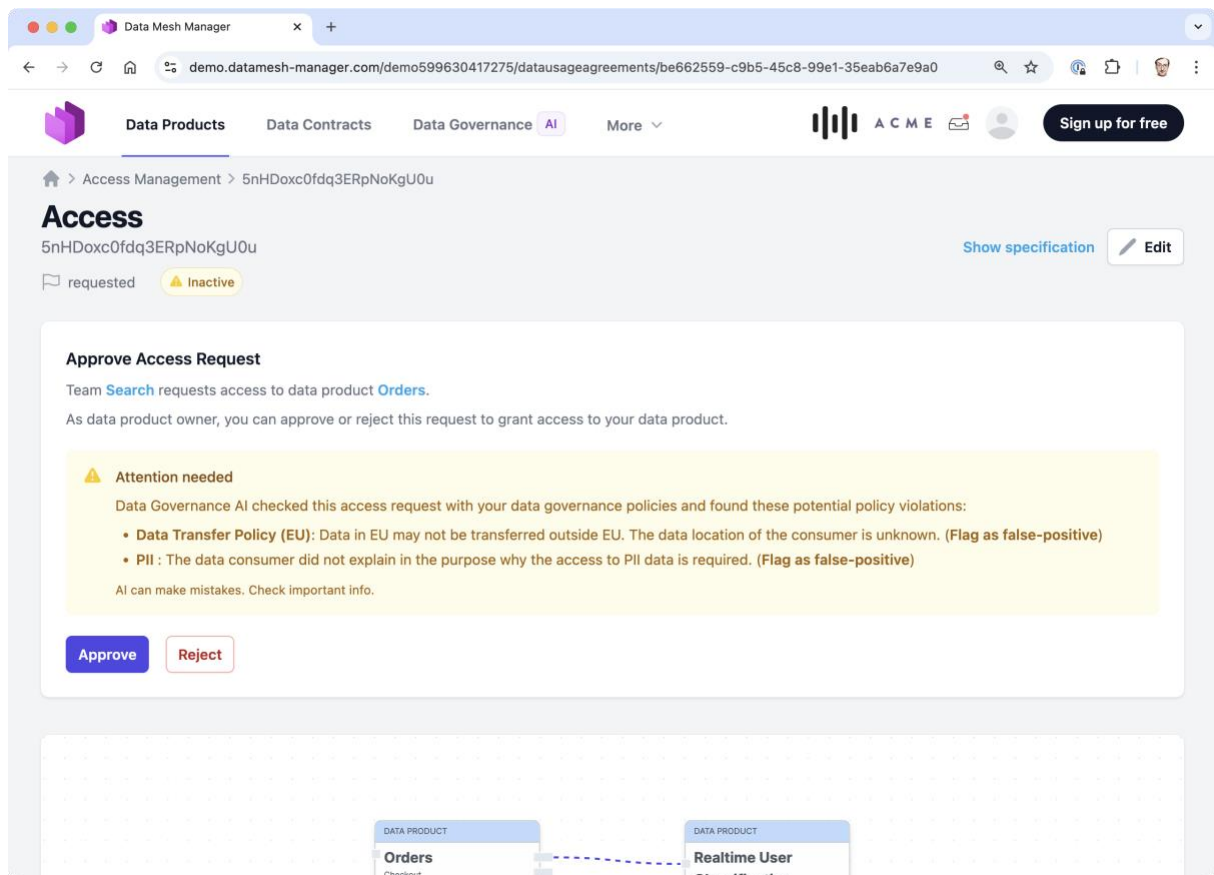
⚠ The requested output port contains PII data. Formulate your purpose accordingly!

Purpose Required

Use data to train an ML model for user classification.
Project: AI-123 ✓

Why do you want access and what do you want to do with the data?

4. Data Product owner gets a notification to approve the access request. The application checks all metadata and data governance policies if there are any data governance or compliance violations. Final decision has the data product owner.

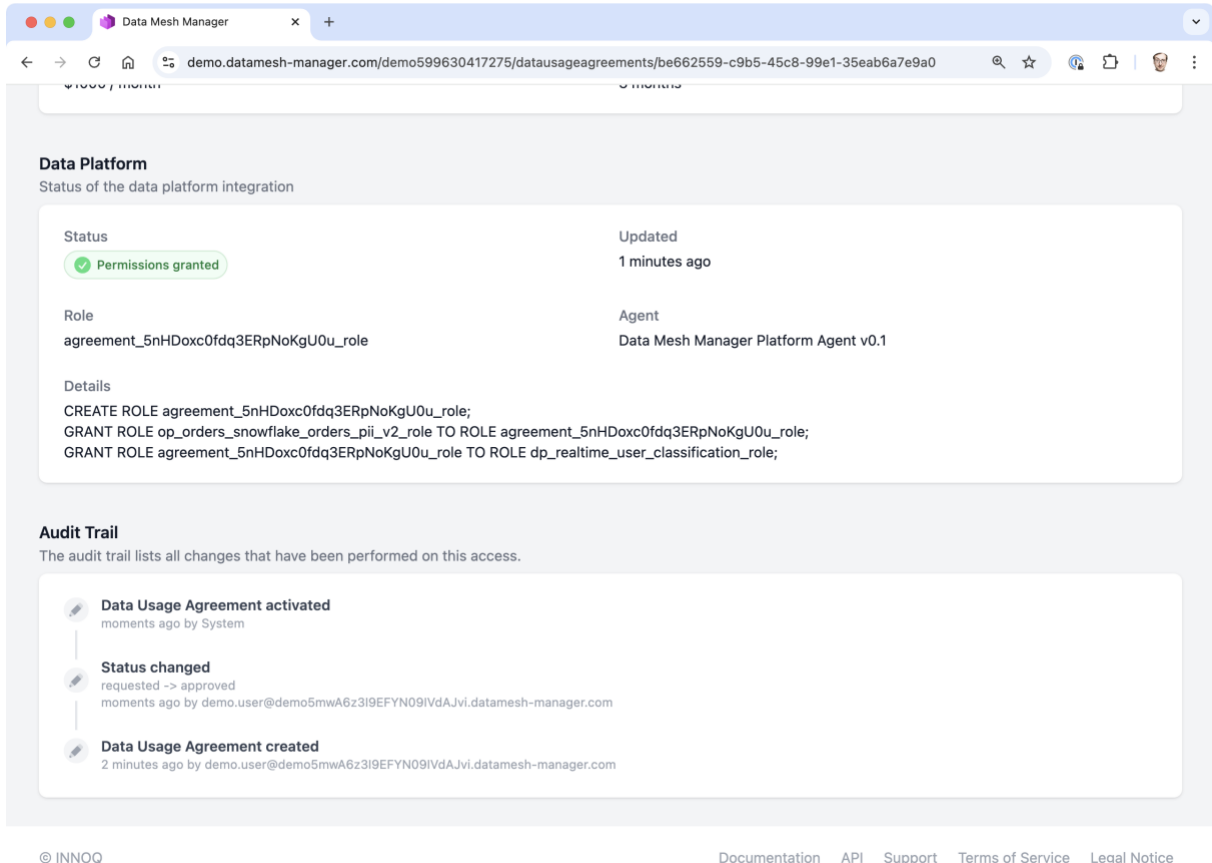


The screenshot shows the 'Data Mesh Manager' web application. The main content area displays an 'Access' request for the data product 'Orders'. The request is in a 'requested' state and is currently 'Inactive'. A warning message indicates that the Data Governance AI has identified potential policy violations:

- Data Transfer Policy (EU):** Data in EU may not be transferred outside EU. The data location of the consumer is unknown. (Flag as false-positive)
- PII:** The data consumer did not explain in the purpose why the access to PII data is required. (Flag as false-positive)

Below the warning, there are 'Approve' and 'Reject' buttons. At the bottom, a diagram illustrates the data flow between two data products: 'Orders' (Checkout) and 'Realtime User Classification'.

5. After approval, the physical access to the data set (e.g. security roles, IAM permissions) can be auto-created with the API integration for a full self-service market place.



The screenshot shows a web browser window with the URL `demo.datamesh-manager.com/demo599630417275/datausageagreements/be662559-c9b5-45c8-99e1-35eab6a7e9a0`. The page title is "Data Mesh Manager".

Data Platform
Status of the data platform integration

Status Permissions granted	Updated 1 minutes ago
Role agreement_5nHDoxc0fdq3ERpNoKgU0u_role	Agent Data Mesh Manager Platform Agent v0.1

Details
 CREATE ROLE agreement_5nHDoxc0fdq3ERpNoKgU0u_role;
 GRANT ROLE op_orders_snowflake_orders_pii_v2_role TO ROLE agreement_5nHDoxc0fdq3ERpNoKgU0u_role;
 GRANT ROLE agreement_5nHDoxc0fdq3ERpNoKgU0u_role TO ROLE dp_realtime_user_classification_role;

Audit Trail
The audit trail lists all changes that have been performed on this access.

- Data Usage Agreement activated**
moments ago by System
- Status changed**
requested -> approved
moments ago by demo.user@demo5mwA6z3i9EFYN09IVdAJvi.datamesh-manager.com
- Data Usage Agreement created**
2 minutes ago by demo.user@demo5mwA6z3i9EFYN09IVdAJvi.datamesh-manager.com

© INNOQ Documentation API Support Terms of Service Legal Notice

2.4 Data Quality Enforcement

Data Mesh Manager has native integration with the open-source Data Contract CLI tool to test if data products are in line with the associated data contract specification.

The tests can be integrated into the data product's CI/CD pipeline as a quality gate for stable and high-quality data products.

The results are published to Data Mesh Manager.

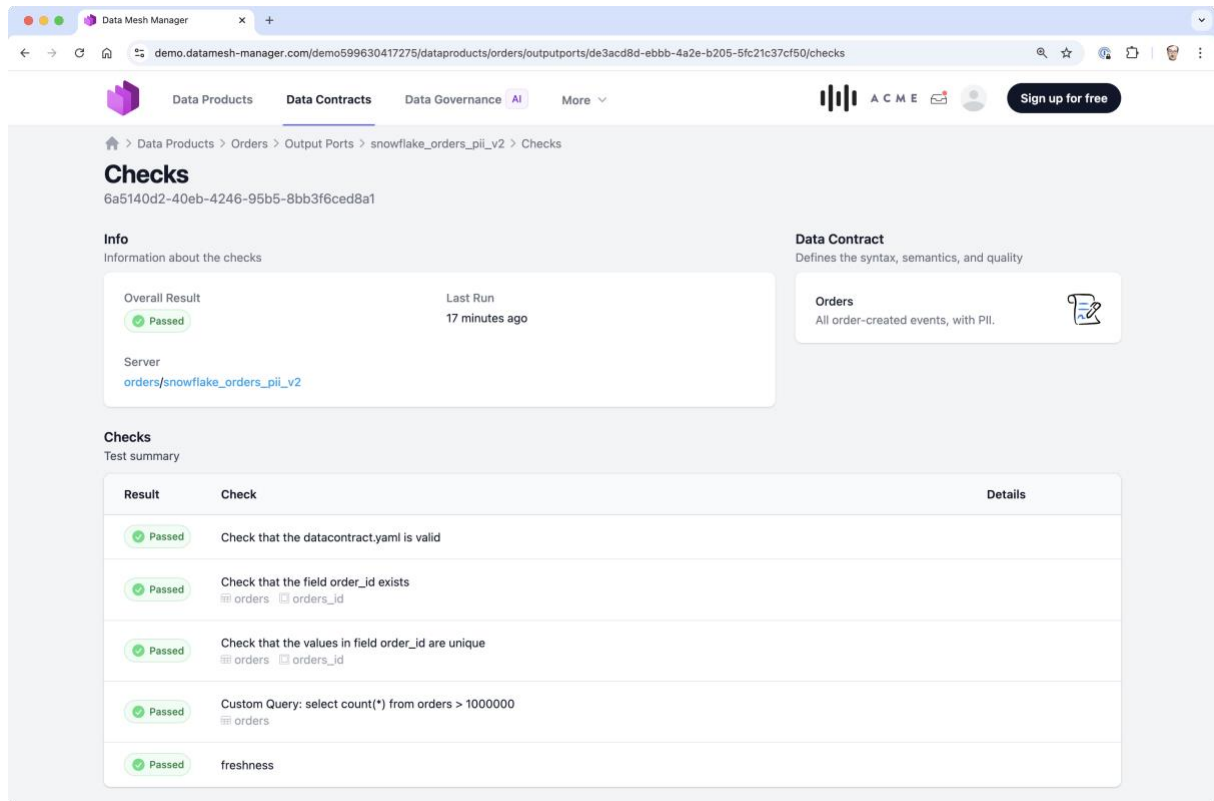
Data Contract CLI test execution:

```
jochen@Jochens-MacBook-Pro-2 ~ % datacontract test https://datacontract.com/examples/orders-latest/datacontract.yaml
Testing https://datacontract.com/examples/orders-latest/datacontract.yaml
```

Result	Check	Field	Details
passed	Check that JSON has valid schema	orders	All JSON entries are valid.
passed	Check that JSON has valid schema	line_items	All JSON entries are valid.
passed	Check that field order_id is present	orders	
passed	Check that field order_timestamp is present	orders	
passed	Check that field order_total is present	orders	
passed	Check that field customer_id is present	orders	
passed	Check that field customer_email_address is present	orders	
passed	Check that field processed_timestamp is present	orders	
passed	row_count >= 5	orders	
passed	Check that required field order_id has no null values	orders.order_id	
passed	Check that unique field order_id has no duplicate values	orders.order_id	
passed	duplicate_count(order_id) = 0	orders.order_id	
passed	Check that required field order_timestamp has no null values	orders.order_timestamp	
passed	Check that required field order_total has no null values	orders.order_total	
passed	Check that field customer_id has a min length of 10	orders.customer_id	
passed	Check that field customer_id has a max length of 20	orders.customer_id	
passed	Check that required field customer_email_address has no null values	orders.customer_email_address	
passed	Check that required field processed_timestamp has no null values	orders.processed_timestamp	
passed	Check that field lines_item_id is present	line_items	
passed	Check that field order_id is present	line_items	
passed	Check that field sku is present	line_items	
passed	values in (order_id) must exist in orders (order_id)	line_items.order_id	
passed	row_count >= 5	line_items	
passed	Check that required field lines_item_id has no null values	line_items.lines_item_id	
passed	Check that unique field lines_item_id has no duplicate values	line_items.lines_item_id	
passed	Check that field sku matches regex pattern <code>^[A-Za-z0-9]{8,14}\$</code>	line_items.sku	

```
data contract is valid. Run 26 checks. Took 6.339405 seconds.
```


Results published to Data Mesh Manager:

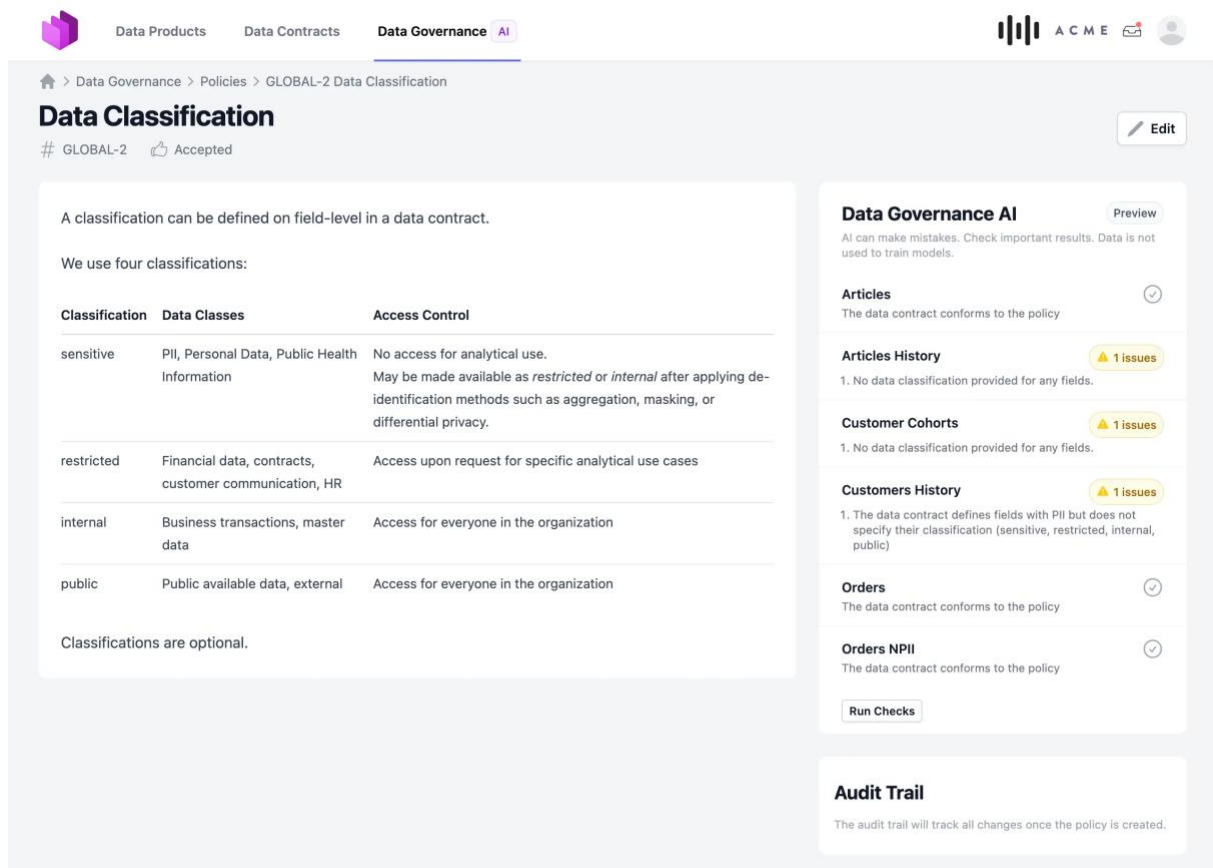


The screenshot shows the Data Mesh Manager web interface. The browser address bar displays the URL: `demo.datamesh-manager.com/demo599630417275/dataproducts/orders/outputports/de3acd8d-ebbb-4a2e-b205-5fc21c37cf50/checks`. The navigation bar includes links for Data Products, Data Contracts (active), Data Governance, and AI. The main content area is titled "Checks" and shows the overall result as "Passed". It also displays the last run time as "17 minutes ago" and the server path as `orders/snowflake_orders_pii_v2`. A "Data Contract" section defines the syntax, semantics, and quality for "Orders". Below this, a table lists the checks performed, all of which passed.

Result	Check	Details
Passed	Check that the datacontract.yaml is valid	
Passed	Check that the field order_id exists	orders orders_id
Passed	Check that the values in field order_id are unique	orders orders_id
Passed	Custom Query: select count(*) from orders > 1000000	orders
Passed	freshness	

2.5 Data Governance (AI)

With the Data Governance module in Data Mesh Manager a governance group can define data governance policies and the system automatically checks the conformance of data contracts and data products using state-of-the-art AI technologies. The use of AI is optional and can be fully disabled. AI models can also be hosted in private cloud tenants (e.g. Azure OpenAI), or fully on-premise (e.g. Ollama).



Data Classification

GLOBAL-2 Accepted

A classification can be defined on field-level in a data contract.

We use four classifications:

Classification	Data Classes	Access Control
sensitive	PII, Personal Data, Public Health Information	No access for analytical use. May be made available as <i>restricted</i> or <i>internal</i> after applying de-identification methods such as aggregation, masking, or differential privacy.
restricted	Financial data, contracts, customer communication, HR	Access upon request for specific analytical use cases
internal	Business transactions, master data	Access for everyone in the organization
public	Public available data, external	Access for everyone in the organization

Classifications are optional.

Data Governance AI Preview

AI can make mistakes. Check important results. Data is not used to train models.

Articles ✓
The data contract conforms to the policy

Articles History 1 issues
1. No data classification provided for any fields.

Customer Cohorts 1 issues
1. No data classification provided for any fields.

Customers History 1 issues
1. The data contract defines fields with PII but does not specify their classification (sensitive, restricted, internal, public)

Orders ✓
The data contract conforms to the policy

Orders NPII ✓
The data contract conforms to the policy

[Run Checks](#)

Audit Trail

The audit trail will track all changes once the policy is created.

Figure 4: Data Governance AI screenshot

Features:

- **Computational Governance in Natural Language**
Ownership, naming conventions, mandatory fields, data classification, access control, and more. Document policies in natural language.
- **AI-driven Policy Checks**
Data Governance AI will automatically check if policies are correctly applied in data products and data contracts.
- **RAG Architecture**
Data Governance AI uses a Retrieval-Augmented Generation process to retrieve additional context information from Data Mesh Manager when needed.
- **Recommendations**
The system generated recommendations what the team can do to adhere to the policies.
- **Feedback**
Users can flag false-positives to ignore for reinforced learning

2.6 Authentication and Integrations

Data Mesh Manager Enterprise integrates with standard authentication systems and includes a role-based access control system.

- **Single Sign-On (SSO)**

Single-Sign On for users, supporting Azure Entra ID (Azure Active Directory), Okta, and others using OpenID Connect (OIDC).

- **Role-based Access Control**

Users can only create and update data contracts and data products for their own team.

For further integrations, Data Mesh Manager comes with a **REST-API** to create and retrieve all resources that are managed in the application.

An **Event-API** can be subscribed to get notified and trigger actions when resources were created or updated.

3 Customizations

Every company is different. Business models are unique. Every software architecture has special needs. We're dedicated to make Data Mesh Manager and subsequent modules perfectly fit for customer's needs and its integration into existing data landscape.

Customizations and feature requests will be charged on a Time & Material basis according to customer's requirements.

4 Professional Services

Get expert guidance and services from the team that knows Data Contracts best.

4.1 Support

	SLA
Help Desk	Unlimited by Email
Troubleshooting	Yes
Response Time P1 - Critical	1 Business Day
Response Time P2 - Moderate	3 Business Days
Response Time P3 - Minor	3 Business Days

Priority Levels

P1 - Critical - Immediate global issue that involves security, or complete failure of the application.

Example: Data Mesh Manager no longer starts.

P2 - Moderate - Significant issue that is impairing a single user or less than 25% of users from getting their work done.

Example: An error in the data contract form prevents to save new data contracts.

Workaround is possible by using the YAML editor

P3 - Minor - A minor or cosmetic issue, that while slightly annoying does not impair a user being able to get their work done.

Example: A data contract's description is too long and overlaps with other UI elements.

Support Hours

Monday through Friday 9am-5pm CET.

Other SLAs can be agreed.

4.2 Additional Customization

The application can further be customized for customer's needs.

Examples:

- Additional fields in data contract editor and view
- Additional fields in data product editor and view
- Custom data contract template
- Automated permission provisioning in data platform
- Automated cost reporting from data platform components
- End-to-end set-up for integrating other software solutions
- Send notifications to MS Teams channels

For every customization request, we will create a quote of the expected effort and time.

4.3 Consulting

Since its founding, INNOQ has specialized in technological and methodological consulting for the software architectures of business-critical IT systems. We know only too well the decisive role played by architecture for each individual system and for the overall system landscape. Our range of services is accordingly broad: from independent analysis of your existing environment, to selection of suitable technologies for solid new and continued development, to complete conception and design of successful systems. The environment and the requirements of our customers are always individual, but the focus is always on your competitive edge. To achieve this, we employ a broad range of concepts, methods, and software architecture templates, which we subject to continuous review.

INNOQ authored datacontract.com, datamesh-architecture.com and ml-ops.org, and is thought leader in modern data management.

We offer consulting services around data contracts, e.g.:

- Facilitation of use-case driven data contract workshops with data product owners and data consumers
- User trainings
- Prioritization and implementation of additional Data Contract CLI features
- Data platform integration
- Data Products implementation

And more services are as offered on <https://data-ai.innoq.com/>