Open Standards for Data Products

DR. SIMON HARRER
CO-FOUNDER AND CEO @ ENTROPY DATA

Entropy Data

SummerSOC

/in/simonharrer

WHO IS SPEAKING

Dr. Simon Harrer

> I'm a software engineer into data.
> I build tools that help data,
> humans, and agents work together.

- Co-founder & CEO - Entropy Data
- Author - datamesh-architecture.com
- Maintainer - Data Contract CLI
- Maintainer - data-landscape.com
- TSC Member - Bitol (ODCS & ODPS) @ LF
- Contributor - Open Semantic Interchange (OSI)

Hochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

University of Stuttgart

Data Contracts: Structuring Promises and Expectations in Data Exchange
Conference IEEE International Conference on Web Services (ICWS) 2026 · with Laura Schuiki, Arif Wider

KING'S College LONDON

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

TILBURG UNIVERSITY

KING'S College LONDON

Hochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

Data Product MCP: Chat with your Enterprise Data
M Tonnarelli, F Scaramuzza, S Harrer, LW Dietz
arXiv preprint arXiv:2601.08687

Automating Data Governance with Generative AI
LW Dietz, A Wider, S Harrer
Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 8 (1), 760-771

Entropy Data

Entropy Data is a data product marketplace based on data contracts and semantics to make high-quality data available for agents and humans.

6 product engineers and a lot of tokens. ISO 27001 certified.
No VC. Customers in US, AU, CH, UK, and EU. And profitable. :-)

Let’s peek into the future

Let’s learn standards

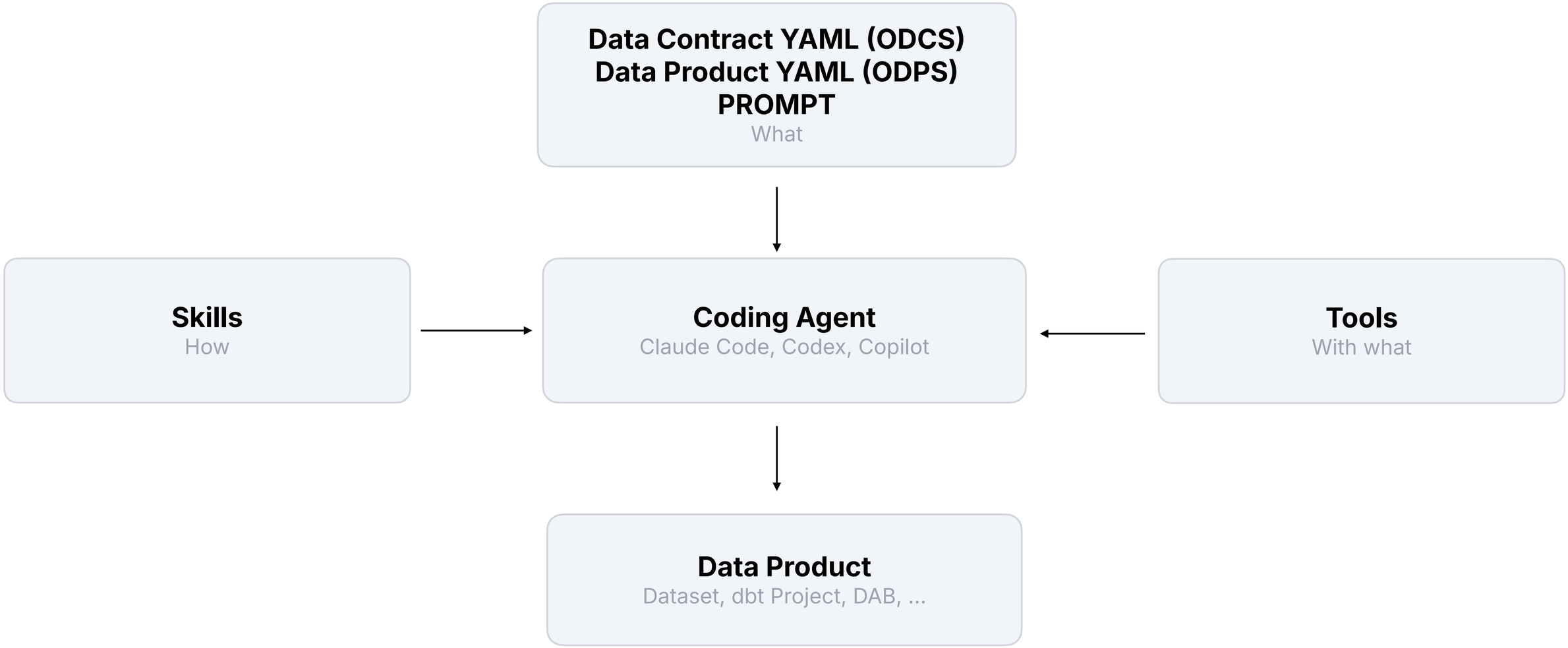
Back to the future

What if we already were using Open Standards for Data Products?

Let me give you a glimpse into the future…

BUILDING DATA PRODUCTS UPON OPEN STANDARDS

Let Coding Agents code



Open Standards for Data Products

Open Standards for Data Products

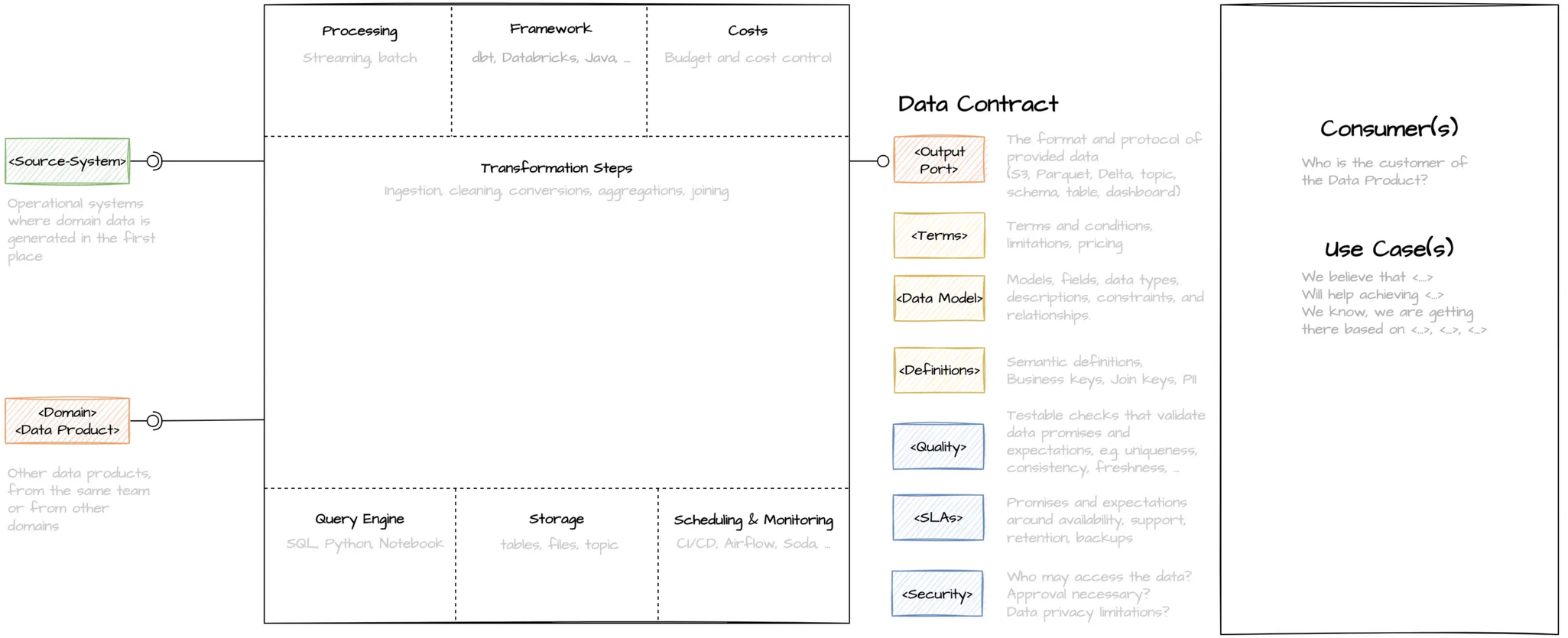
What Is Data Mesh?

	Domain Ownership	Data as a Product	Self-serve Data Platform	Federated Governance
Strategic Domain-driven Design	Domain Bounded Context	Product Thinking	Domain-agnostic	Context Mapping
Socio-technical Perspective	Domain Teams	Data Product by Domain Team	Data Platform Team	Governance Group
Technology	Operational & Analytical Data	Interoperability Interfaces	Self-serve Data Platform	Policy Automation

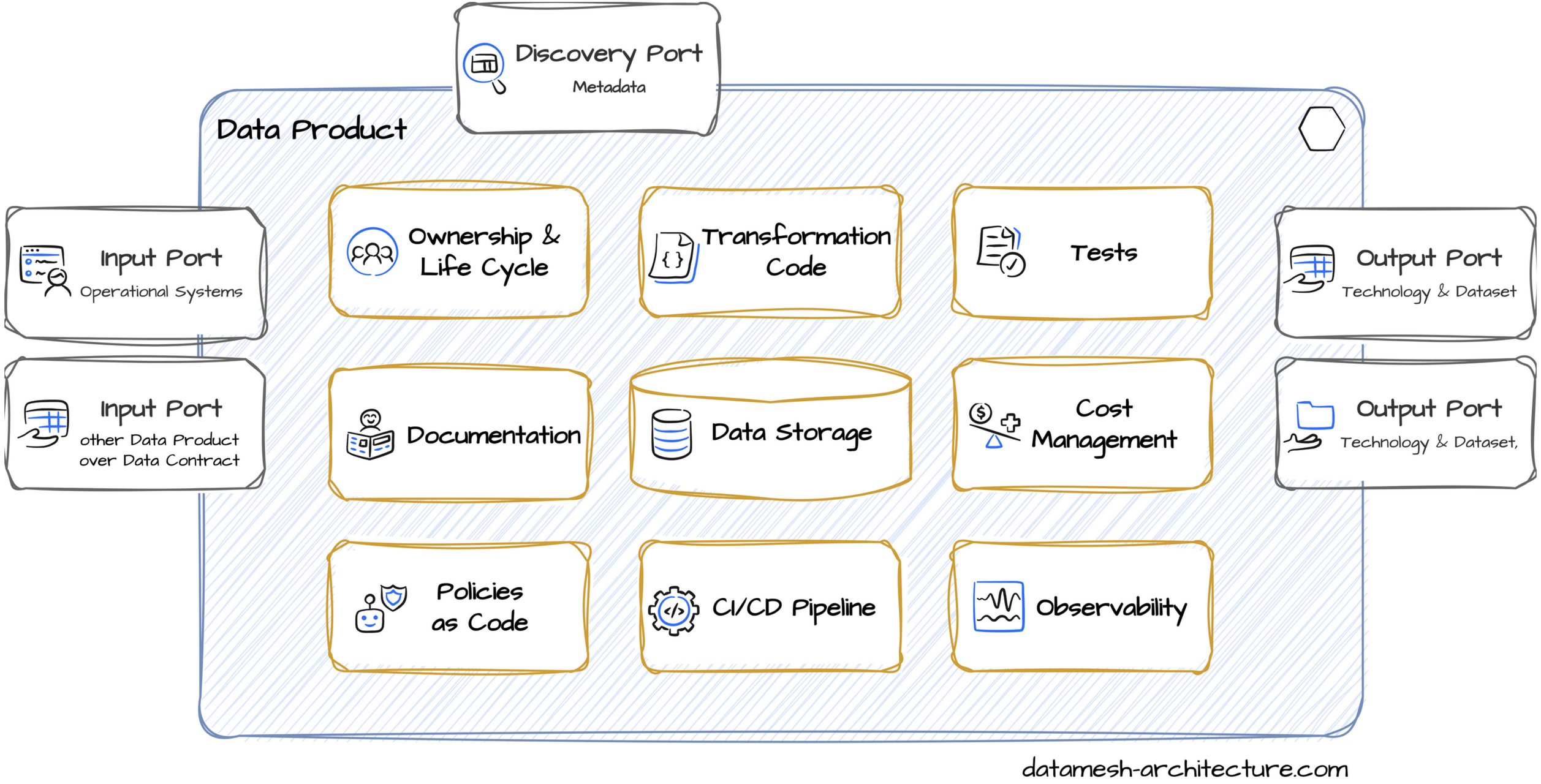
datamesh-architecture.com

Domain	Data Product	Date

Data Product Architecture



Classification	Ubiquitous Language	Follow-up
The archetype of the provided data (source-aligned, aggregate, consumer-aligned)	Context-specific domain terminology (relevant for Data Product)	e.g. create a Data Contract YAML (datacontract.com) Date:


Discovery Port
Metadata
Data Product
Input Port
Operational Systems
Input Port
other Data Product
over Data Contract
Ownership &
Life Cycle
Transformation
Code
Tests
Output Port
Technology & Dataset
Documentation
Data Storage
Cost
Management
Output Port
Technology & Dataset,
Policies
as Code
CI/CD Pipeline
Observability
datamesh-architecture.com

Open Standards for Data Products

Open Standards
for Data Products

What is a Standard?
And why do we need them?

PUBLIC SERVICE ANNOUNCEMENT:

OUR DIFFERENT WAYS OF WRITING DATES AS NUMBERS CAN LEAD TO ONLINE CONFUSION. THAT'S WHY IN 1988 ISO SET A GLOBAL STANDARD NUMERIC DATE FORMAT.

THIS IS ***THE*** CORRECT WAY TO WRITE NUMERIC DATES:

2013-02-27

THE FOLLOWING FORMATS ARE THEREFORE DISCOURAGED:

02/27/2013 02/27/13 27/02/2013 27/02/13
20130227 2013.02.27 27.02.13 27-02-13
27.2.13 2013.II.27. 27½-13 2013.158904109
MMXIII-II-XXVII MMXIII $\frac{\text{LVII}}{\text{CCCLXV}}$ 1330300800
((3+3)×(111+1)-1)×3/3-1/3^3 ~~2013~~
10/11011/1101 02/27/20/13 $\overset{2}{0}\overset{3}{1}\overset{1}{2}\overset{4}{3}\overset{}{7}$ $\overset{}{5}\ \overset{}{67}\ \overset{}{8}$

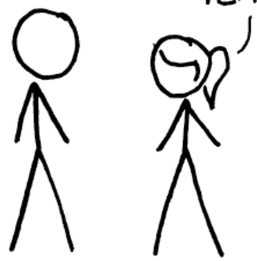
ONE ... TWO ... *THREE!* ✗ DEPRECATED
ONE ... TWO ... THREE ... *GO!* ✗ DEPRECATED
} TOO EASY TO MIX UP
THREE ... TWO ... ONE ... *GO!* ✓ ISO STANDARD

IF I WERE IN CHARGE OF ISO, THE FIRST THING I'D DO WOULD BE TO STANDARDIZE THE WAY PEOPLE COUNT OUT LOUD BEFORE DOING SOMETHING IN SYNC.

HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION: THERE ARE 14 COMPETING STANDARDS.

14?! RIDICULOUS! WE NEED TO DEVELOP ONE UNIVERSAL STANDARD THAT COVERS EVERYONE'S USE CASES.



SOON:

SITUATION: THERE ARE 15 COMPETING STANDARDS.

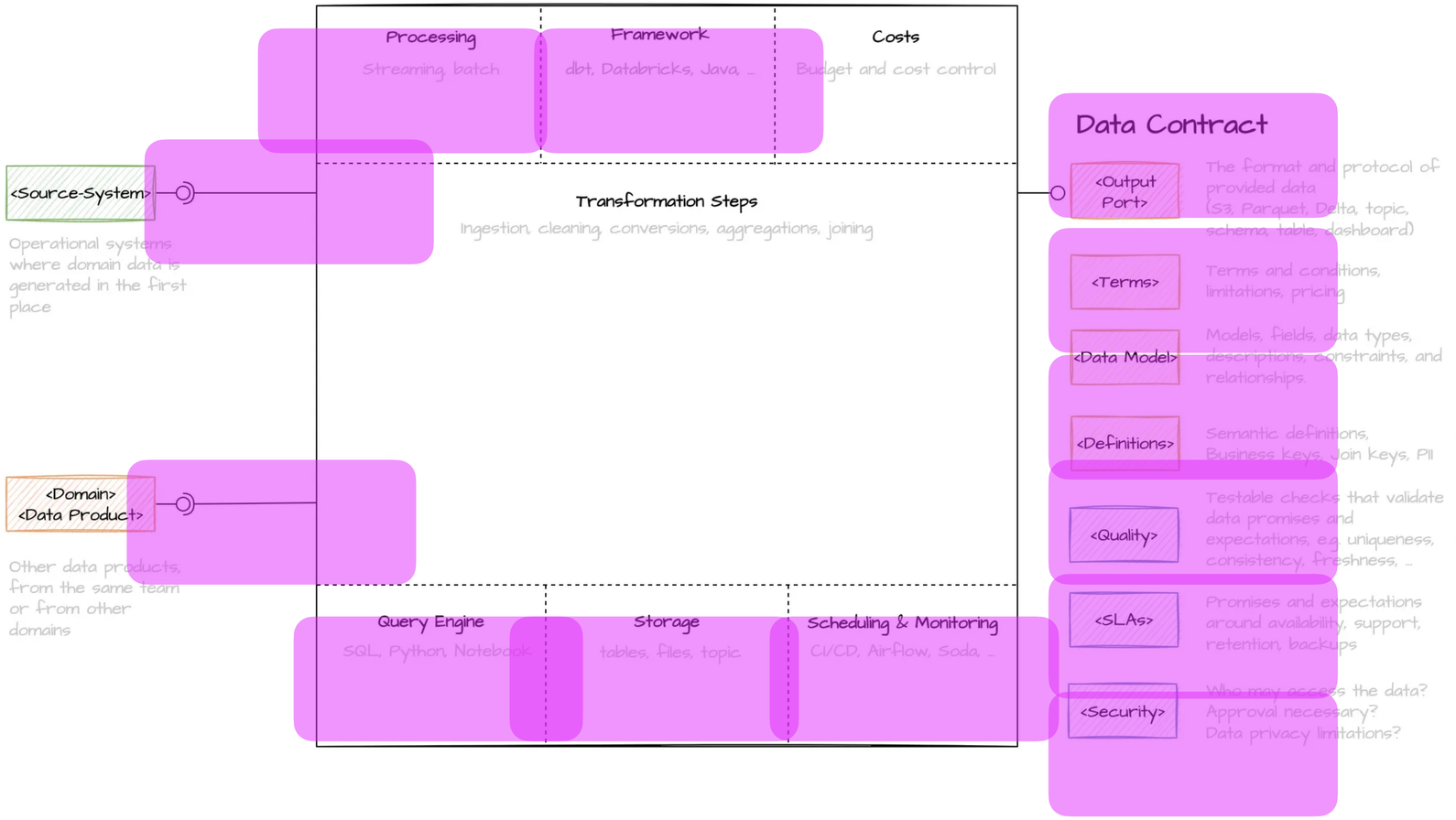
Source: all three from xkcd.com

De Facto vs. De Jure	Usage
Initiative vs. Vendor	Owner
Many vs. One	Control
Open vs. Paywall	$$$

Open Standards for Data Products

Domain	Data Product	Date

Data Product Architecture



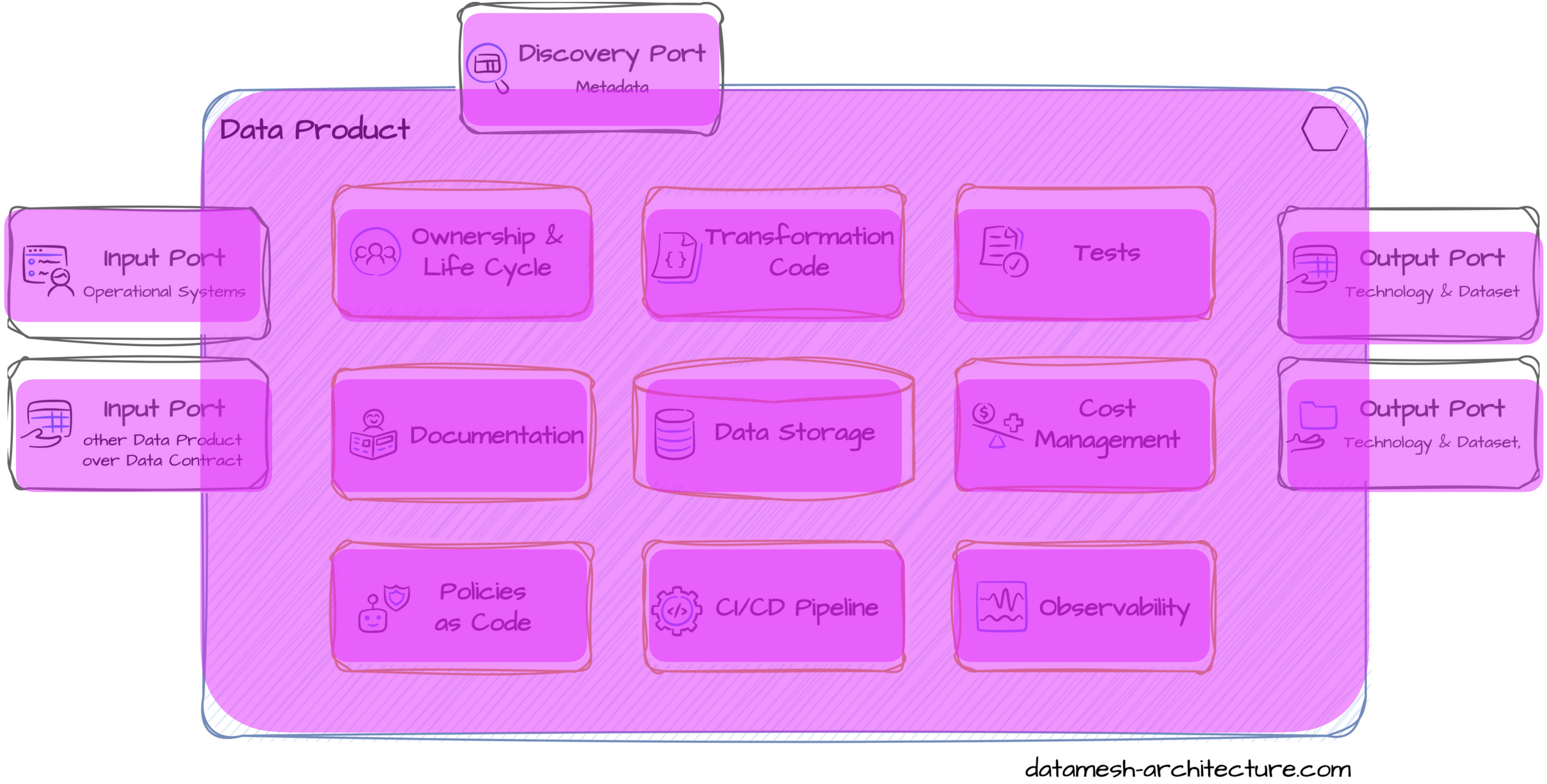
Consumer(s)

Who is the customer of the Data Product?

Use Case(s)

We believe that <...>
Will help achieving <...>
We know, we are getting there based on <...>, <...>, <...>

Classification	Ubiquitous Language	Follow-up
The archetype of the provided data (source-aligned, aggregate, consumer-aligned)	Context-specific domain terminology (relevant for Data Product)	e.g. create a Data Contract YAML (datacontract.com) Date:



datamesh-architecture.com

Open Standards

Data Landscape

An opinionated, interactive map of the open standards.

The open standards that power a modern data architecture, organised by what they describe. Inspired by the CNCF Landscape and the ThoughtWorks Tech Radar, and Simon's talk on open standards. Click any standard to learn more, or download as PDF.

> *"Ohhh this is amazing!! You don't know how much that's helping immediately at a project I'm currently on!"*
>
> Tiankai Feng

ADOPT 37 | SITUATIONAL 21 | ASSESS 15 | CAUTION 7 | Single-vendor specs are muted | Search standards... | Landscape | Radar | Table

DEFINITION — HOW DATA IS DESCRIBED

CONTRACTS 6

- ADOPT — ODCS — BITOL @ LF
- ADOPT — OpenAPI — LF
- ADOPT — AsyncAPI — LF
- ADOPT — GraphQL — LF
- SITUATIONAL — gRPC — CNCF
- SITUATIONAL — OData — OASIS

DATA PRODUCTS 4

- ADOPT — ODPS — BITOL @ LF
- ASSESS — DPDS — ODM
- ASSESS — ODPS — LF
- ASSESS — DPROD — OMG

SCHEMA 7

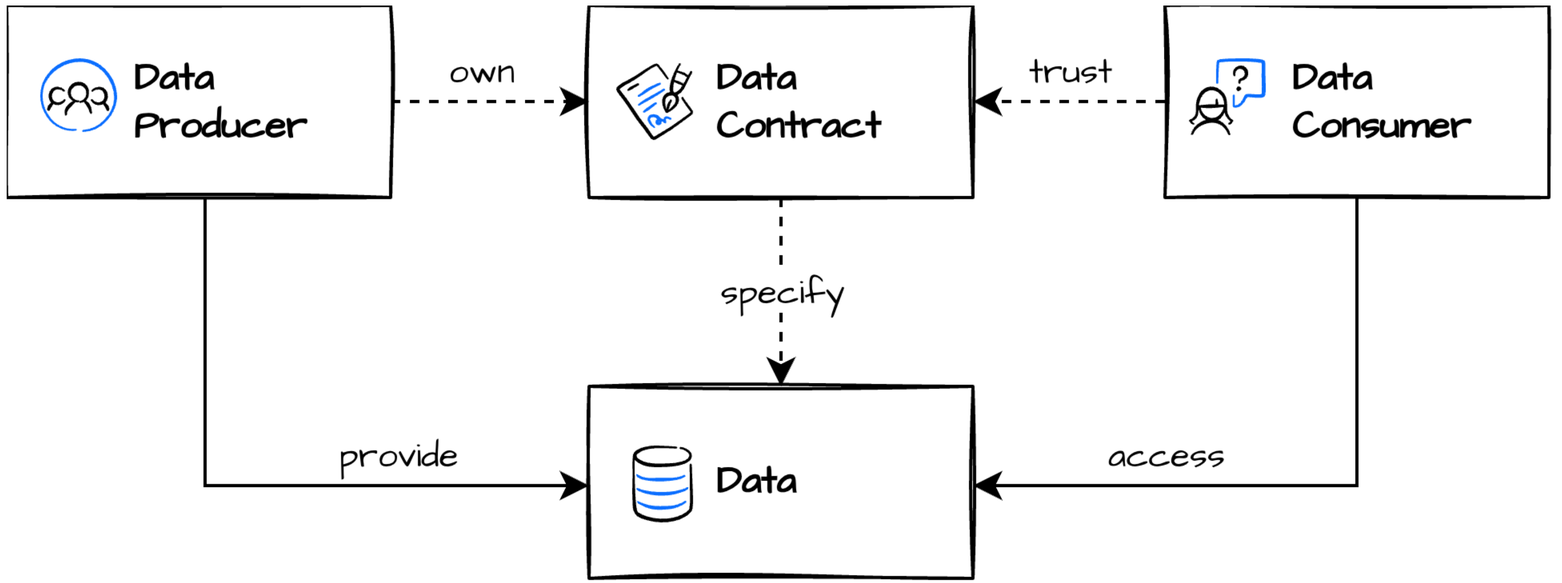
- ADOPT — XML Schema — W3C
- ADOPT — JSON Schema — IETF
- ADOPT — SQL DDL — ISO/IEC
- ADOPT — AVRO Schema — ASF
- ADOPT — Protobuf — GOOGLE
- ASSESS — LinkML — LINKML
- ASSESS — Table Schema — FRICTIONLESS

SEMANTICS 8

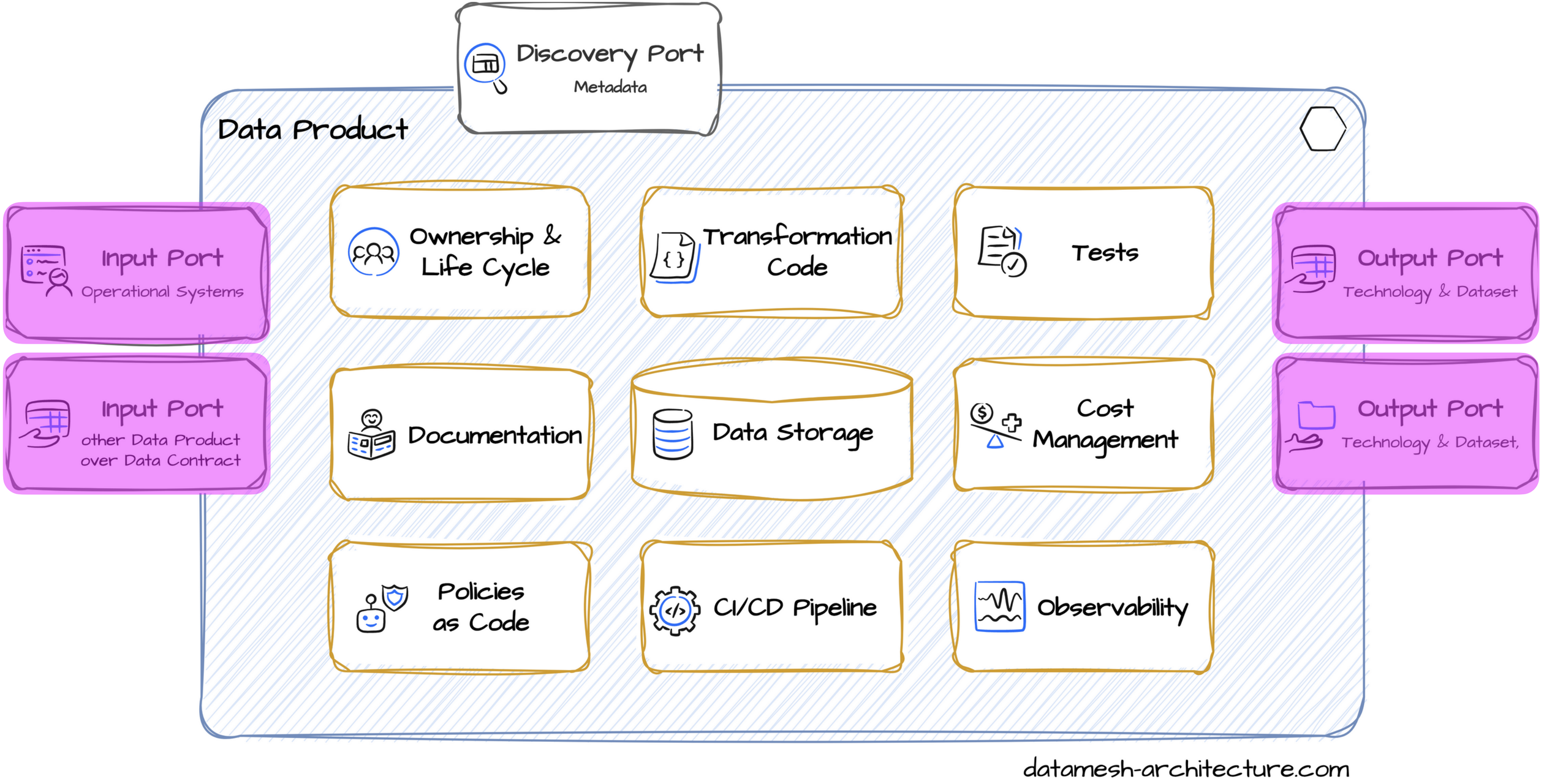
- SITUATIONAL — RDF/OWL — W3C
- SITUATIONAL — DCAT — W3C
- SITUATIONAL — SKOS — W3C
- SITUATIONAL — SHACL — W3C
- SITUATIONAL — JSON-LD — W3C
- ASSESS — OSI — OSI INITIATIVE
- ASSESS — schema.org — W3C
- ASSESS — ShEx — W3C

Standards for Data Contracts

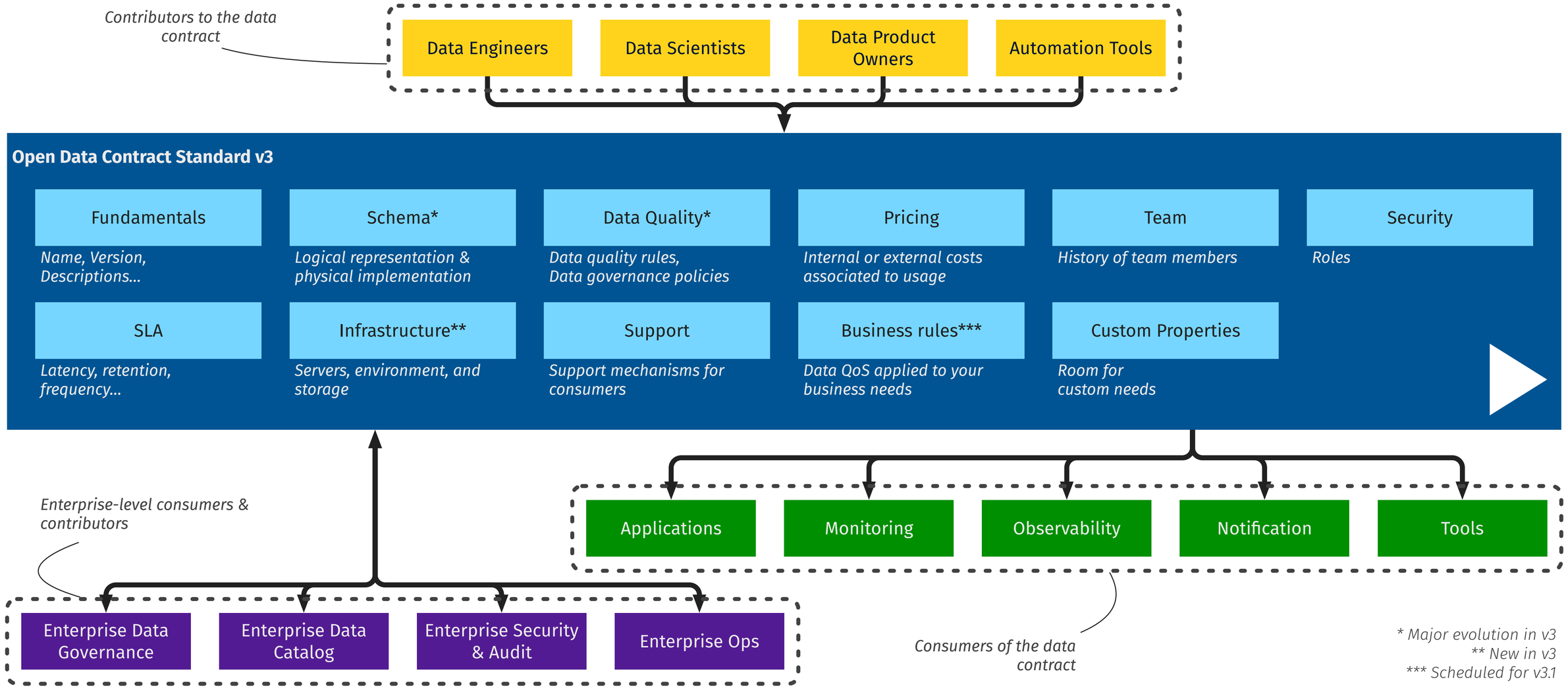
Think “API Spec”



A data contract is a document that defines the ownership, structure, semantics, quality, and terms of use for exchanging data between a data producer and their consumers. Think of an API, but for data.


Discovery Port
Metadata
Data Product
Input Port
Operational Systems
Input Port
other Data Product
over Data Contract
Ownership &
Life Cycle
Transformation
Code
Tests
Documentation
Data Storage
Cost
Management
Policies
as Code
CI/CD Pipeline
Observability
Output Port
Technology & Dataset
Output Port
Technology & Dataset,
datamesh-architecture.com

Once upon a time, every company had their own format



https://github.com/bitol-io/open-data-contract-standard

Bitol | LF AI & DATA

Fundamentals

Every data contract starts with its fundamentals: the version of the standard, unique identifier, name, data contract version number, and status. This metadata identifies the contract and enables versioning.

Schema

The schema defines the structure and semantics of your data. Here we define the table structure with its columns. Each property includes technical details, business semantics, and governance attributes.

Data Quality

Data quality rules ensure that data meets expectations. Define checks like valid value constraints, row count thresholds, and custom SQL validation logic. These can be tested automatically with the Data Contract CLI.

Team

Document who owns and maintains the data contract. Include team name, description, and members with their roles. Add support channels, how data consumers can reach out to the owners.

orders-v1.odcs.yaml

```yaml
apiVersion: v3.1.0
kind: DataContract
id: orders
name: Orders
version: 1.0.0
status: active
```

Fundamentals

Every data contract starts with its fundamentals: the version of the standard, unique identifier, name, data contract version number, and status. This metadata identifies the contract and enables versioning.

Schema

The schema defines the structure and semantics of your data. Here we define the table structure with its columns. Each property includes technical details, business semantics, and governance attributes.

Data Quality

Data quality rules ensure that data meets expectations. Define checks like valid value constraints, row count thresholds, and custom SQL validation logic. These can be tested automatically with the Data Contract CLI.

Team

Document who owns and maintains the data contract. Include team name, description, and members with their roles. Add support channels, how data consumers can reach out to the owners.

orders-v1.odcs.yaml

```yaml
apiVersion: v3.1.0
kind: DataContract
id: orders
name: Orders
version: 1.0.0
status: active
schema:
  - name: orders
    physicalType: TABLE
    description: All successful and cancelled web-shop order
    properties:
      - name: order_id
        logicalType: string
        primaryKey: true
      - name: customer_id
        businessName: "Customer Identifier"
        logicalType: string
        physicalType: VARCHAR(10)
        required: true
        examples:
          - "C-10000001"
        classification: internal
        tags:
          - pii:true
      - name: order_total
        logicalType: integer
        physicalType: INTEGER
```

standard, unique identifier, name, data contract version number, and status. This metadata identifies the contract and enables versioning.

Schema

The schema defines the structure and semantics of your data. Here we define the table structure with its columns. Each property includes technical details, business semantics, and governance attributes.

Data Quality

Data quality rules ensure that data meets expectations. Define checks like valid value constraints, row count thresholds, and custom SQL validation logic. These can be tested automatically with the Data Contract CLI.

Team

Document who owns and maintains the data contract. Include team name, description, and members with their roles. Add support channels, how data consumers can reach out to the owners.

Terms of Use

Define the purpose, usage guidelines, and limitations for your data. This helps consumers understand what they can and cannot do with the data, establishing clear governance boundaries.

SLAs

```yaml
name: Orders
version: 1.0.0
status: active
schema:
  - name: orders
    physicalType: TABLE
    description: All successful and cancelled web-shop order
    properties:
      - name: order_id
        logicalType: string
        primaryKey: true
      - name: order_status
        logicalType: string
        quality:
          - type: library
            description: Only valid order statuses
            metric: invalidValues
            arguments:
              validValues:
                - pending
                - shipped
                - cancelled
            mustBe: 0
    # table-level quality checks
    quality:
      - type: sql
        description: Expect 100k+ rows
        query: select count(*) from orders
        mustBeGreaterThan: 100000
```

standard, unique identifier, name, data contract version number, and status. This metadata identifies the contract and enables versioning.

Schema

The schema defines the structure and semantics of your data. Here we define the table structure with its columns. Each property includes technical details, business semantics, and governance attributes.

Data Quality

Data quality rules ensure that data meets expectations. Define checks like valid value constraints, row count thresholds, and custom SQL validation logic. These can be tested automatically with the Data Contract CLI.

Team

Document who owns and maintains the data contract. Include team name, description, and members with their roles. Add support channels, how data consumers can reach out to the owners.

Terms of Use

Define the purpose, usage guidelines, and limitations for your data. This helps consumers understand what they can and cannot do with the data, establishing clear governance boundaries.

SLAs

orders-v1.odcs.yaml

```yaml
apiVersion: v3.1.0
kind: DataContract
id: orders
name: Orders
version: 1.0.0
status: active
schema:
  - name: orders
    physicalType: TABLE
    description: All successful and cancelled web-shop order
    properties:
      - name: order_id
        logicalType: string
        primaryKey: true
team:
  name: sales
  description: Owned by the "Sales" team
  members:
    - username: john@example.com
      name: John Doe
      role: Owner
support:
  - channel: "#sales-support"
    tool: "slack"
    url: "https://example.slack.com/archives/C123456789"
    description: "Support and collaboration"
```

Data Quality

Data quality rules ensure that data meets expectations. Define checks like valid value constraints, row count thresholds, and custom SQL validation logic. These can be tested automatically with the Data Contract CLI.

Team

Document who owns and maintains the data contract. Include team name, description, and members with their roles. Add support channels, how data consumers can reach out to the owners.

Terms of Use

Define the purpose, usage guidelines, and limitations for your data. This helps consumers understand what they can and cannot do with the data, establishing clear governance boundaries.

SLAs

Service Level Agreements define non-functional guarantees. Data consumers can match these with their use-case requirements.

Servers

Finally, specify where the data lives. The server configuration includes connection details for different environments.

orders-v1.odcs.yaml

```yaml
apiVersion: v3.1.0
kind: DataContract
id: orders
name: Orders
version: 1.0.0
status: active
description:
  purpose: "Provides order and line item data for analytics
  usage: "Used by analytics team for sales analysis and busi
  limitations: "Contains only the last 2 years of data"
  customProperties:
    - property: "sensitivity"
      value: "secret"
      description: "Data contains personally identifiable in
  authoritativeDefinitions:
    - url: "https://entropy-data.com/policies/gdpr-complianc
      type: "businessDefinition"
      description: "GDPR compliance policy for handling cust
schema:
  - name: orders
    physicalType: TABLE
    properties:
      - name: order_id
        logicalType: string
        primaryKey: true
team:
  name: sales
```

Team

Document who owns and maintains the data contract. Include team name, description, and members with their roles. Add support channels, how data consumers can reach out to the owners.

Terms of Use

Define the purpose, usage guidelines, and limitations for your data. This helps consumers understand what they can and cannot do with the data, establishing clear governance boundaries.

SLAs

Service Level Agreements define non-functional guarantees. Data consumers can match these with their use-case requirements.

Servers

Finally, specify where the data lives. The server configuration includes connection details for different environments.

Full Example

Here is the full data contract bringing all the elements together: fundamentals, a schema with two tables, data quality rules, team, terms of use, servers, and custom properties.

orders-v1.odcs.yaml

```yaml
apiVersion: v3.1.0
kind: DataContract
id: orders
name: Orders
version: 1.0.0
status: active
slaProperties:
  - property: availability
    value: 99.9%
    description: Data platform uptime guarantee
  - property: retention
    value: 1
    unit: year
    description: Data will be deleted after 1 year
  - property: freshness
    value: 24
    unit: hours
    element: orders.order_date
    description: Within 24 hours of order placement
  - property: support
    value: business hours
    description: Support only during business hours
```

Team

Document who owns and maintains the data contract. Include team name, description, and members with their roles. Add support channels, how data consumers can reach out to the owners.

Terms of Use

Define the purpose, usage guidelines, and limitations for your data. This helps consumers understand what they can and cannot do with the data, establishing clear governance boundaries.

SLAs

Service Level Agreements define non-functional guarantees. Data consumers can match these with their use-case requirements.

Servers

Finally, specify where the data lives. The server configuration includes connection details for different environments.

Full Example

Here is the full data contract bringing all the elements together: fundamentals, a schema with two tables, data quality rules, team, terms of use, servers, and custom properties.

Reference: Open Data Contract Standard

orders-v1.odcs.yaml

```yaml
apiVersion: v3.1.0
kind: DataContract
id: orders
name: Orders
version: 1.0.0
status: active
schema:
  - name: orders
    physicalType: TABLE
    properties:
      - name: order_id
        logicalType: string
        primaryKey: true
servers:
  - server: postgres
    type: postgres
    environment: prod
    host: aws-1-eu-central-2.pooler.supabase.com
    port: 6543
    database: postgres
    schema: dp_orders_v1
```

Okay, we've created that YAML. Now what?

Code Generation

- Java
- Python in Pydantic
- dbt Models and Sources
- SQL DDL and Queries

Test

- Compare contract with real data
- Breaking data detection in PR
- Breaking metadata detection
- Continuous Monitoring

Metadata Distribution

- Metastores: Hive, …
- Data Catalogs: Colibra, …
- Data Marketplace: Entropy Data, …
- Software Catalogs: LeanIX, …

Automate all the things!

Infrastructure Provisioning

- Output Port (S3 Bucket, …)
- Input Port (dbt sources.yml)
- Transformations (anonymisation)
- Access Control (IAM permissions)

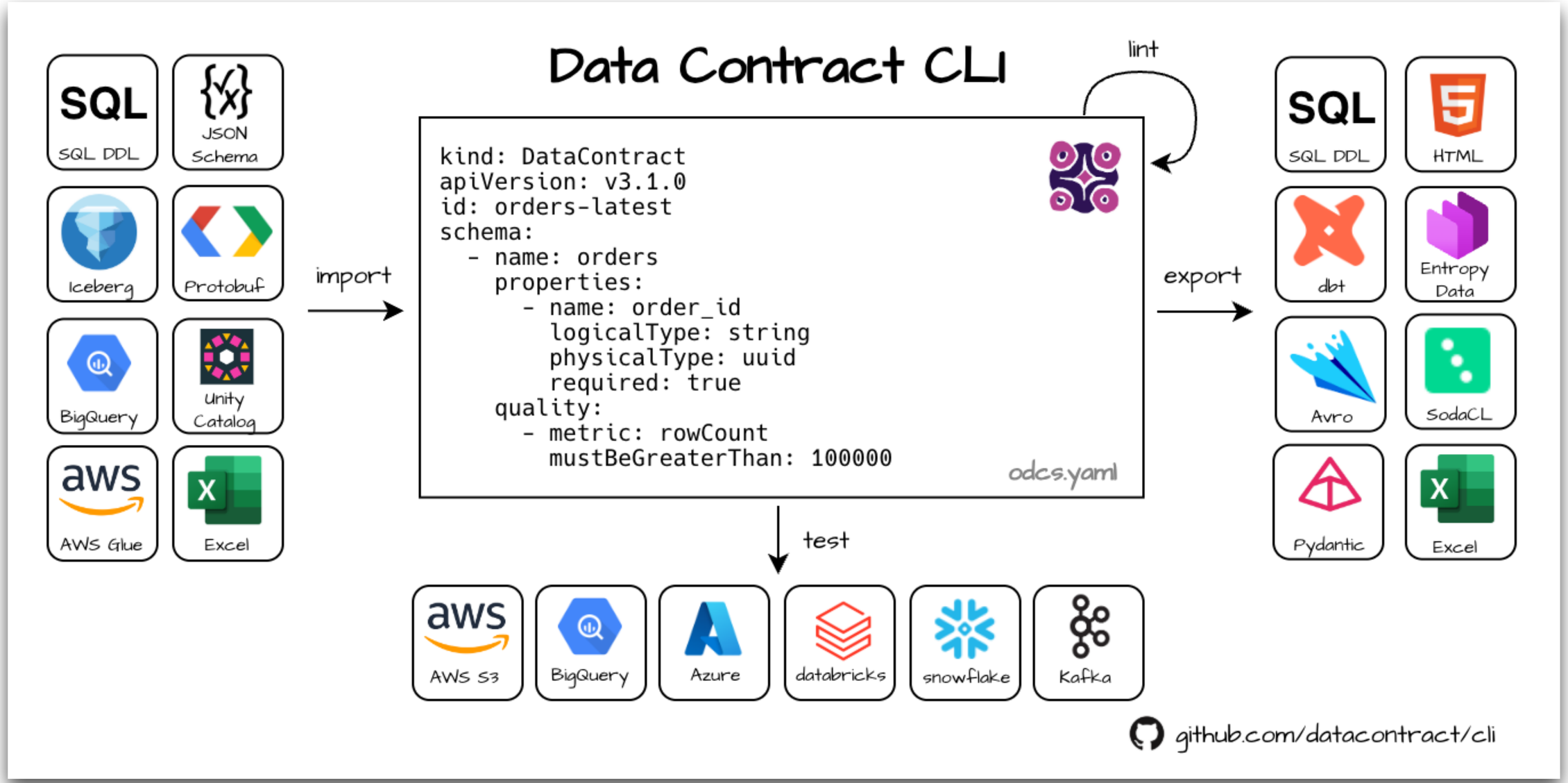
Collaboration

- Contract-First (in workshop)
- Data-First (import from …)
- Semantics

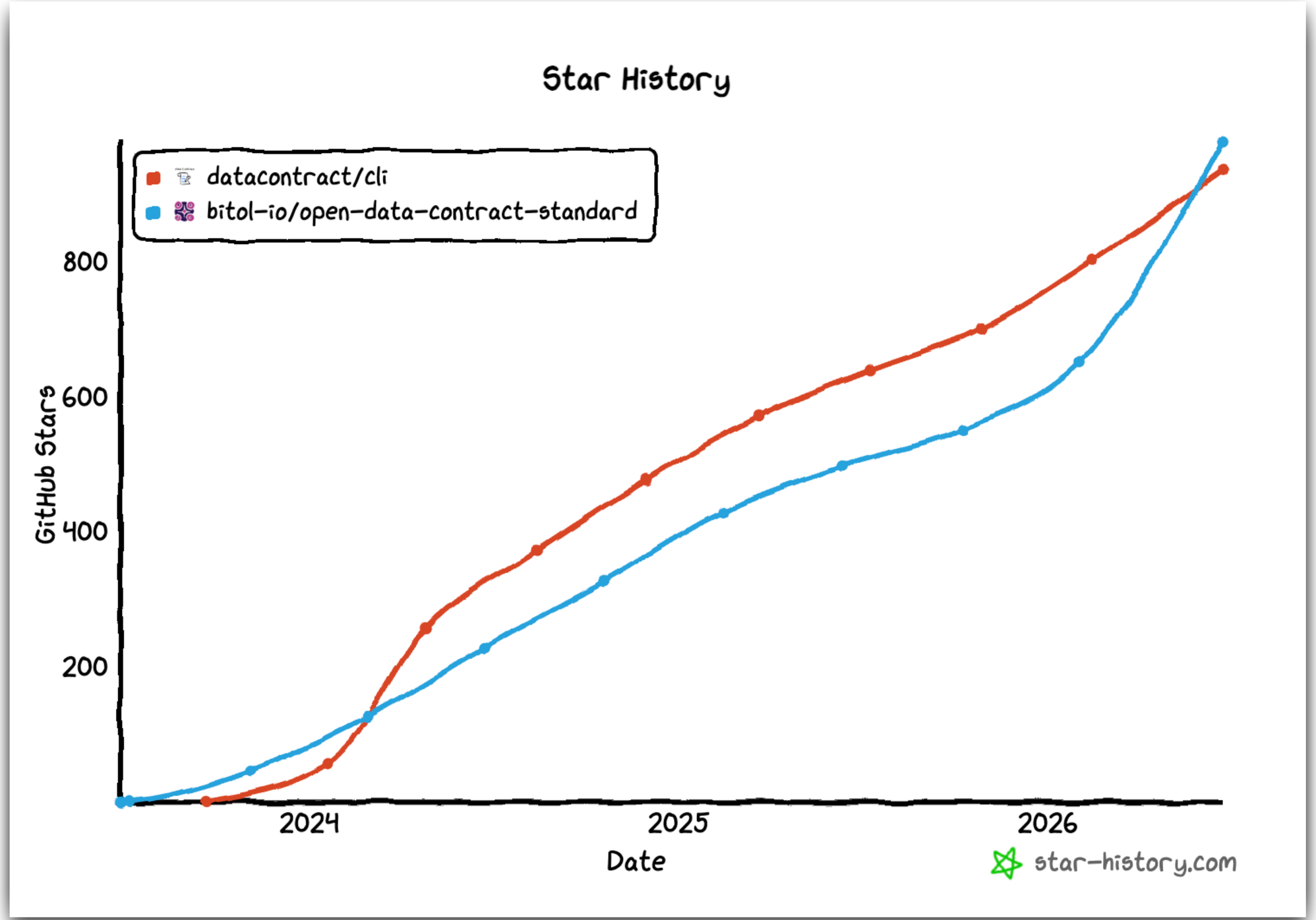
Governance

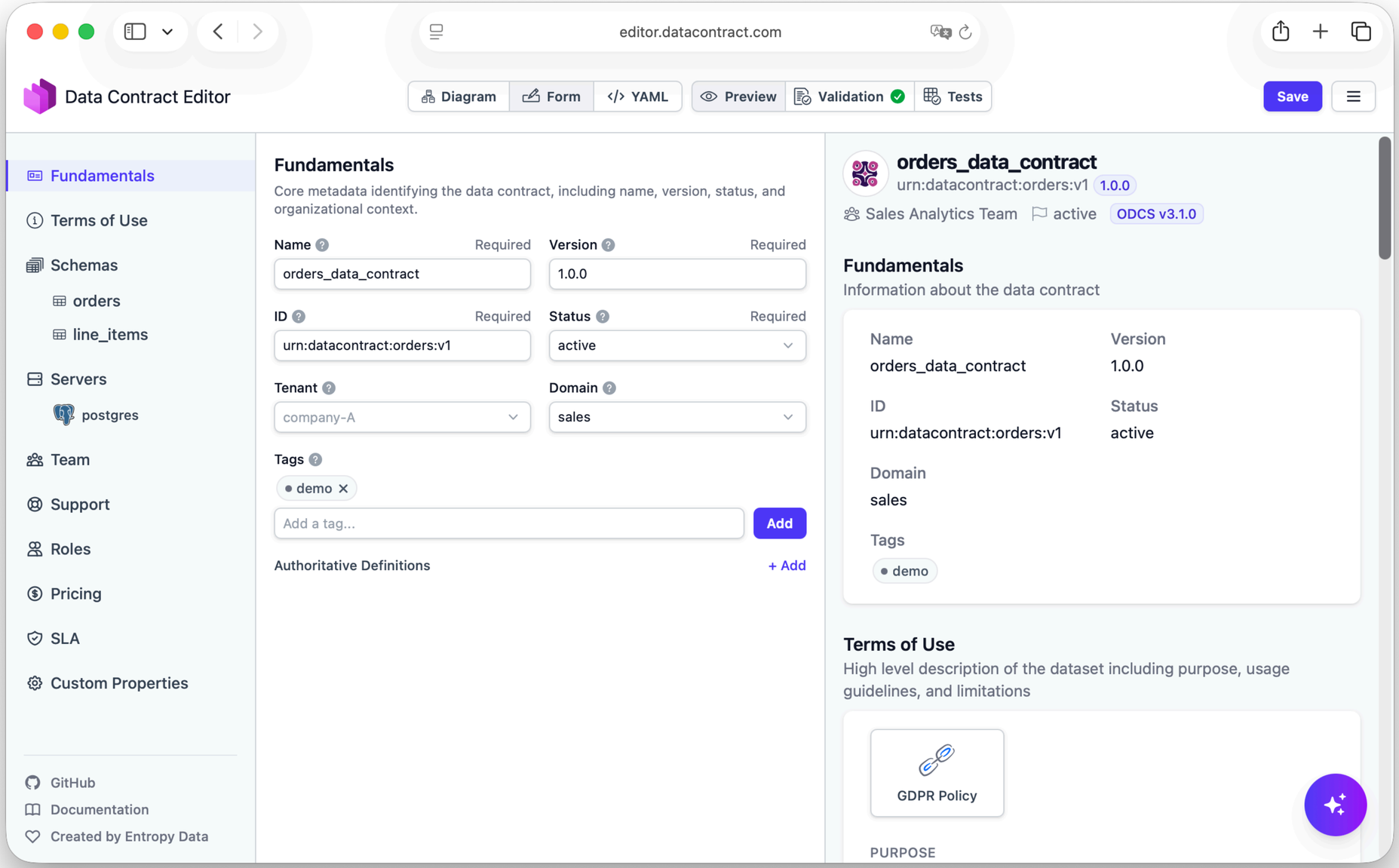
- Policies (naming conventions, …)
- Schema Evolution (Notice period)
- Usage Agreements
- Approval Workflows

Open Source Tooling

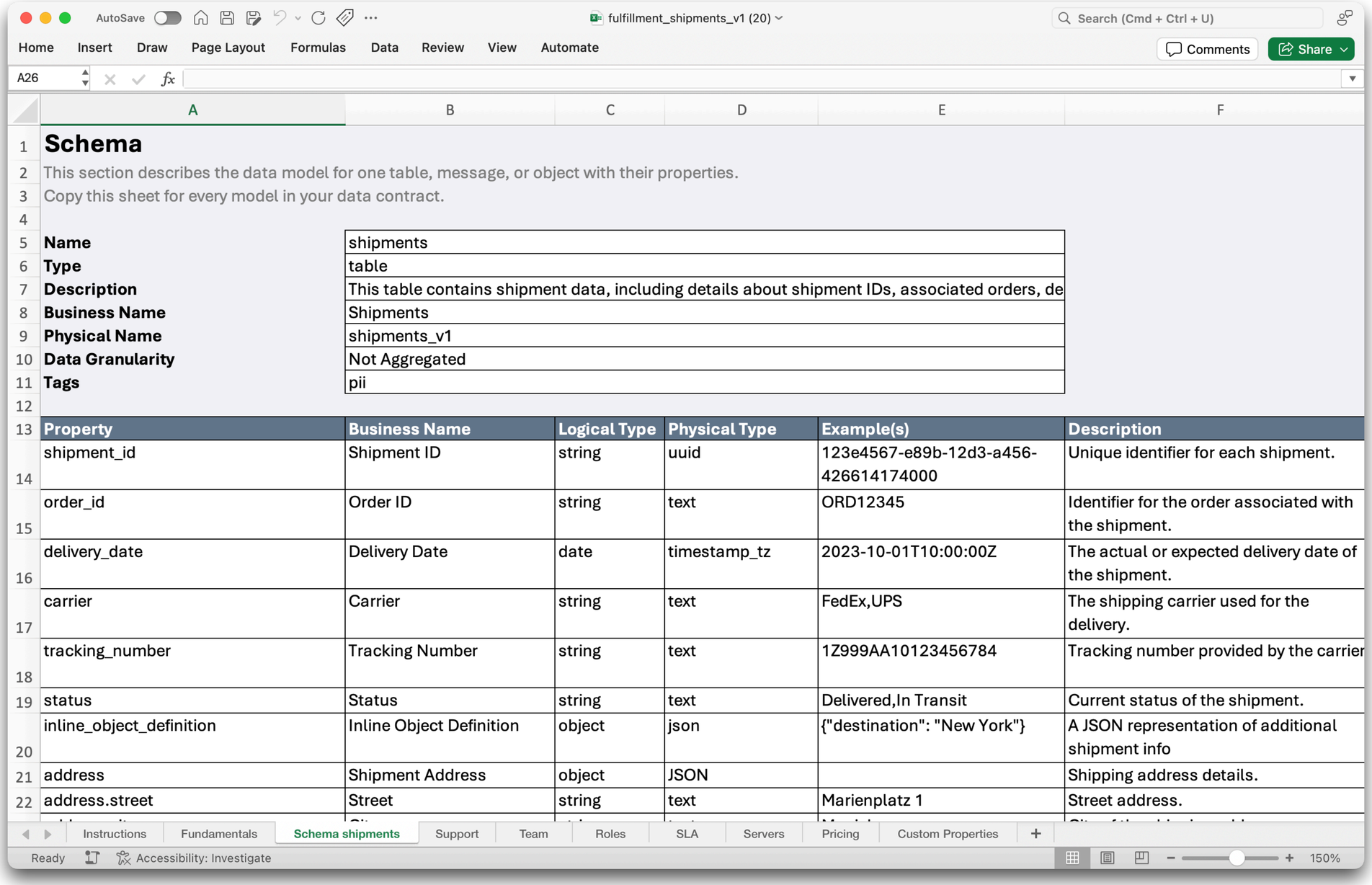


github.com/datacontract/datacontract-cli


Star History
datacontract/cli
bitol-io/open-data-contract-standard
800
600
400
200
GitHub Stars
2024
2025
2026
Date
star-history.com



github.com/datacontract/datacontract-editor



Schema

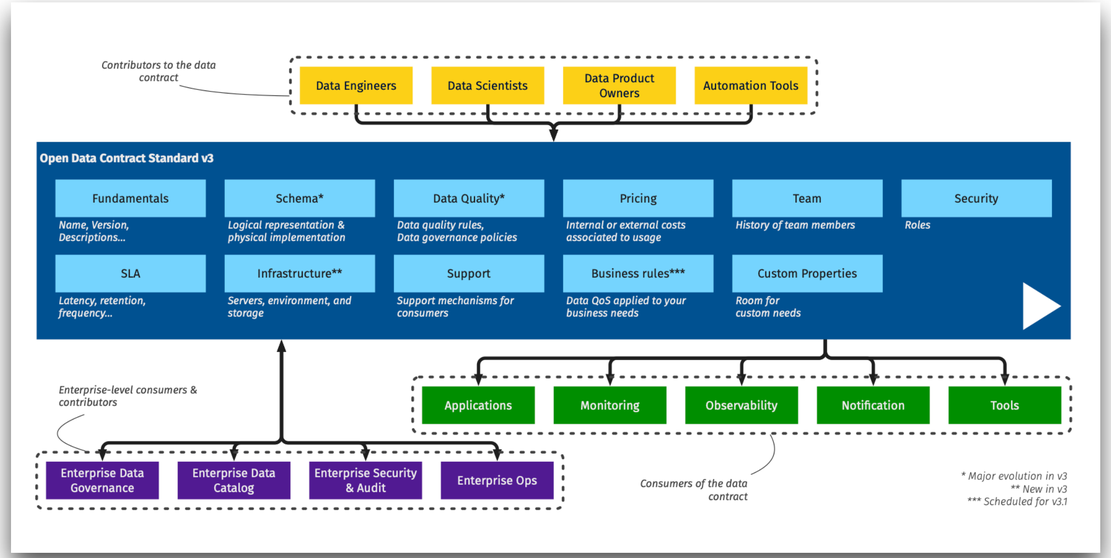
This section describes the data model for one table, message, or object with their properties.
Copy this sheet for every model in your data contract.

Name	shipments
Type	table
Description	This table contains shipment data, including details about shipment IDs, associated orders, de
Business Name	Shipments
Physical Name	shipments_v1
Data Granularity	Not Aggregated
Tags	pii

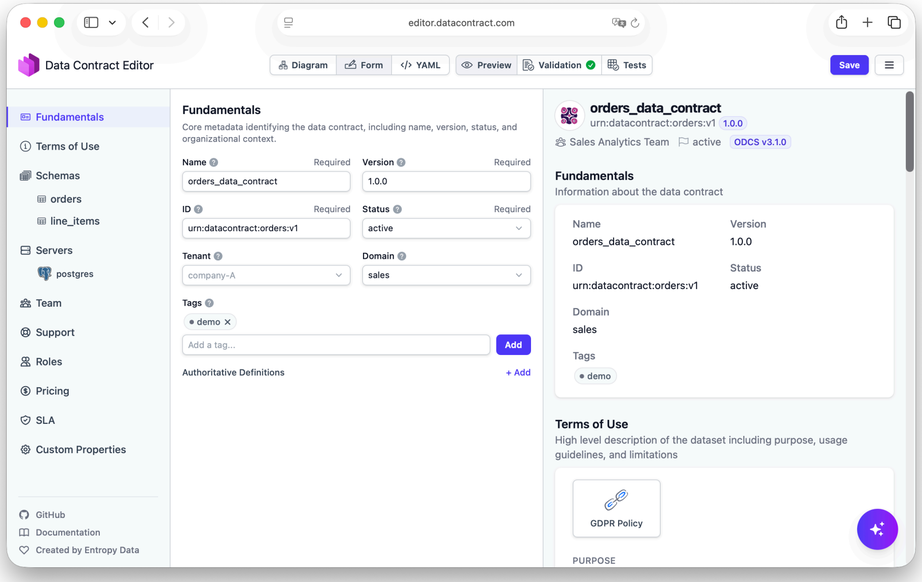
Property	Business Name	Logical Type	Physical Type	Example(s)	Description
shipment_id	Shipment ID	string	uuid	123e4567-e89b-12d3-a456-426614174000	Unique identifier for each shipment.
order_id	Order ID	string	text	ORD12345	Identifier for the order associated with the shipment.
delivery_date	Delivery Date	date	timestamp_tz	2023-10-01T10:00:00Z	The actual or expected delivery date of the shipment.
carrier	Carrier	string	text	FedEx,UPS	The shipping carrier used for the delivery.
tracking_number	Tracking Number	string	text	1Z999AA10123456784	Tracking number provided by the carrier
status	Status	string	text	Delivered,In Transit	Current status of the shipment.
inline_object_definition	Inline Object Definition	object	json	{"destination": "New York"}	A JSON representation of additional shipment info
address	Shipment Address	object	JSON		Shipping address details.
address.street	Street	string	text	Marienplatz 1	Street address.

github.com/datacontract/open-data-contract-standard-excel-template

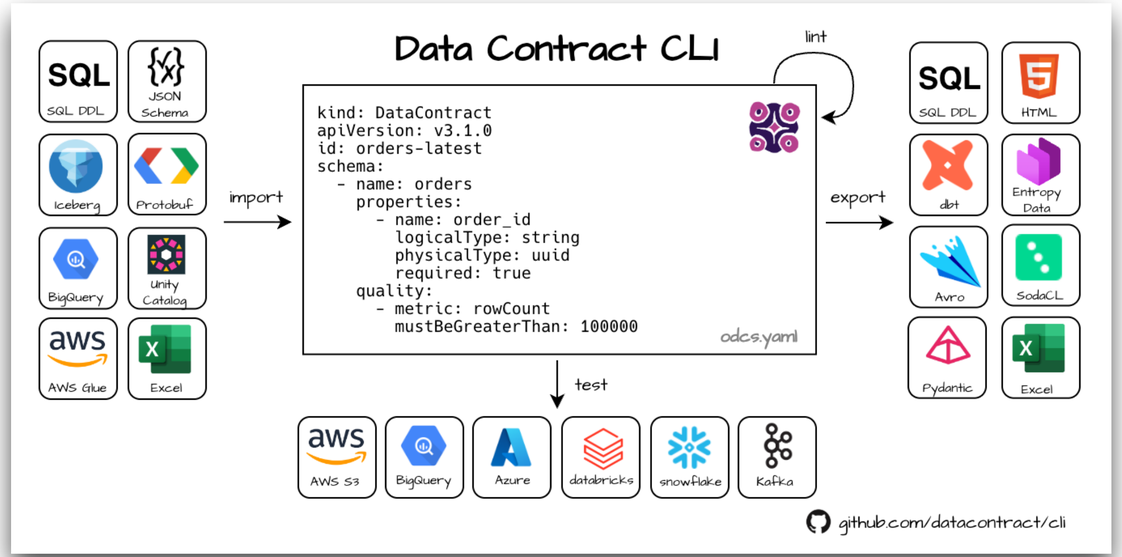
Open Data Contract Standard



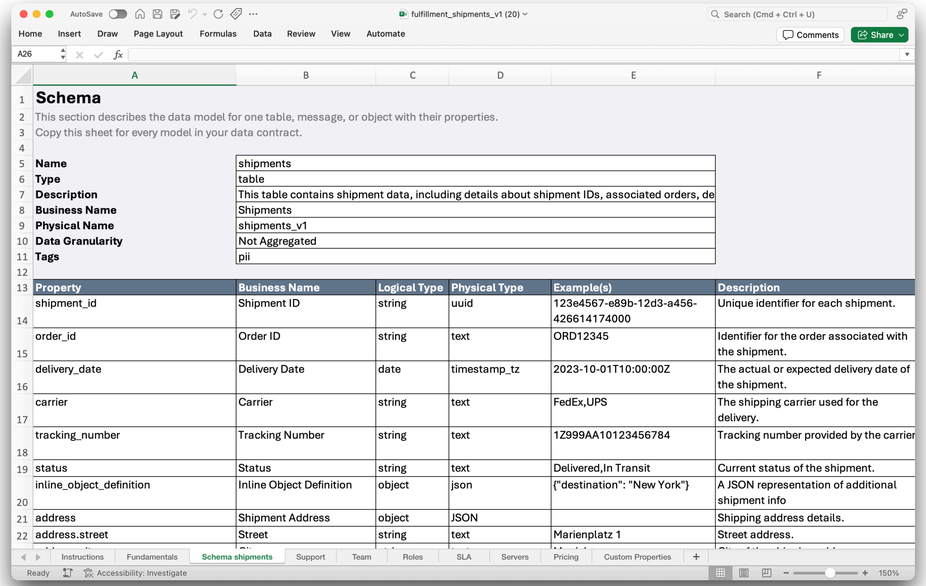
Data Contract Editor

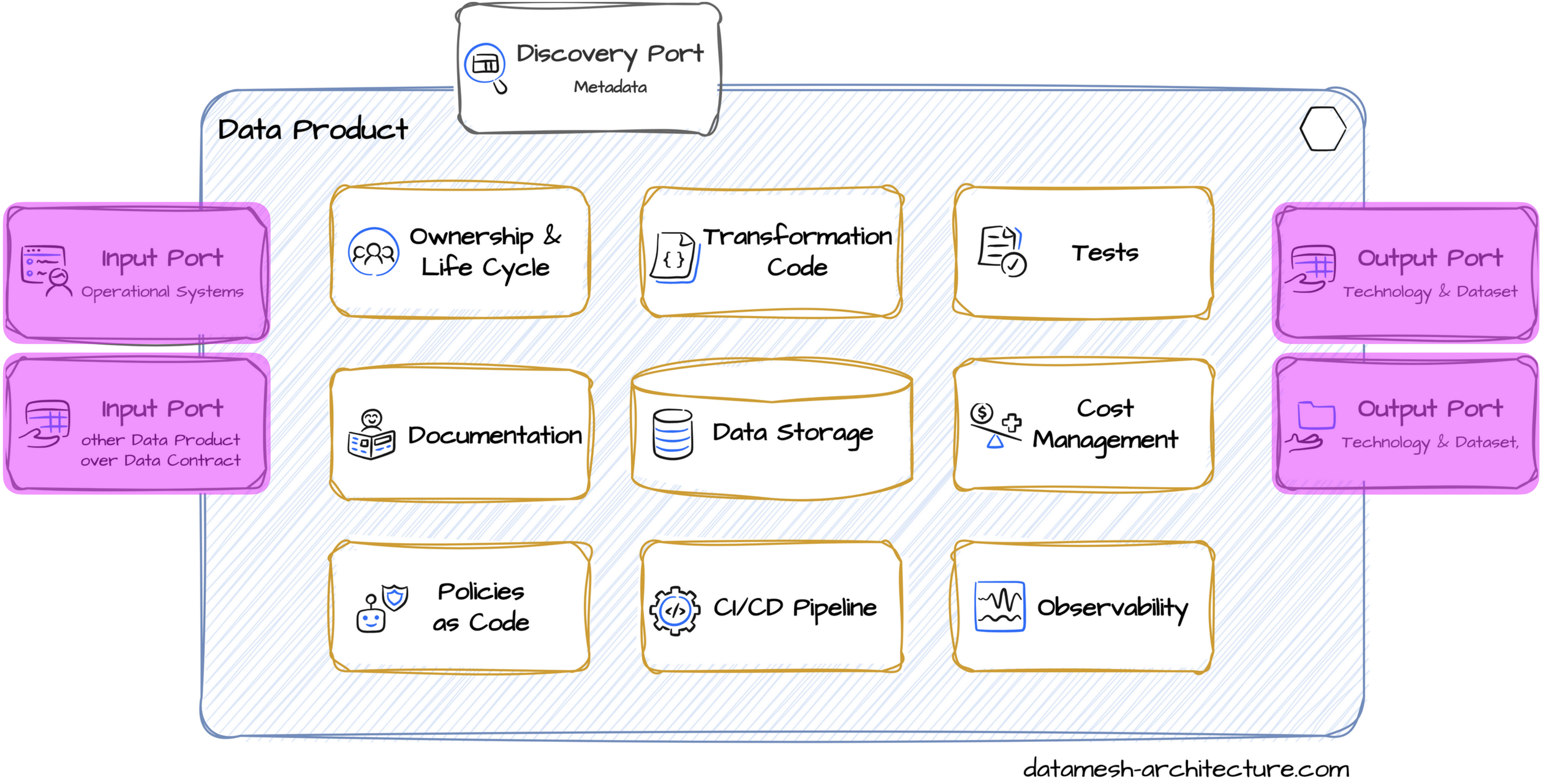


Data Contract CLI



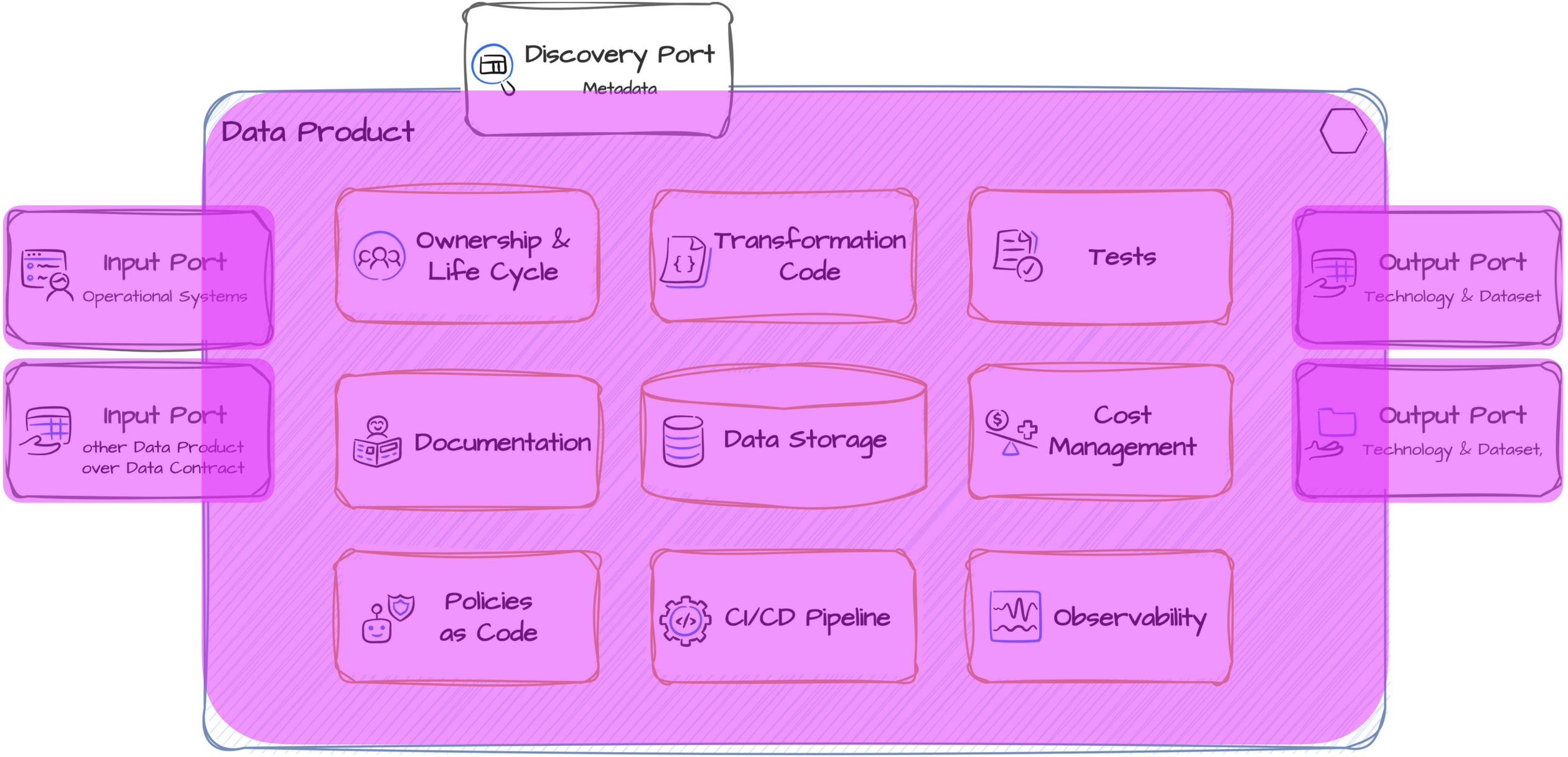
Data Contract Excel Template




Discovery Port
Metadata
Data Product
Input Port
Operational Systems
Input Port
other Data Product
over Data Contract
Ownership &
Life Cycle
Transformation
Code
Tests
Documentation
Data Storage
Cost
Management
Policies
as Code
CI/CD Pipeline
Observability
Output Port
Technology & Dataset
Output Port
Technology & Dataset,
datamesh-architecture.com

Standards for Data Products

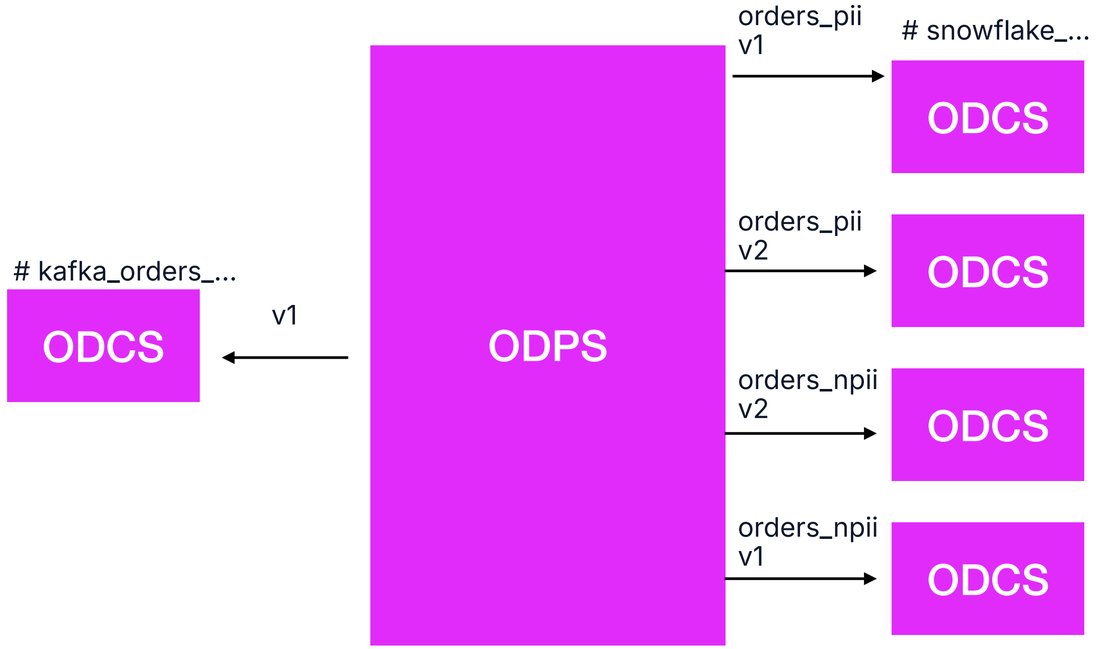
Think “Service”



datamesh-architecture.com

THE BUILDING BLOCKS IN THE DATA WORLD

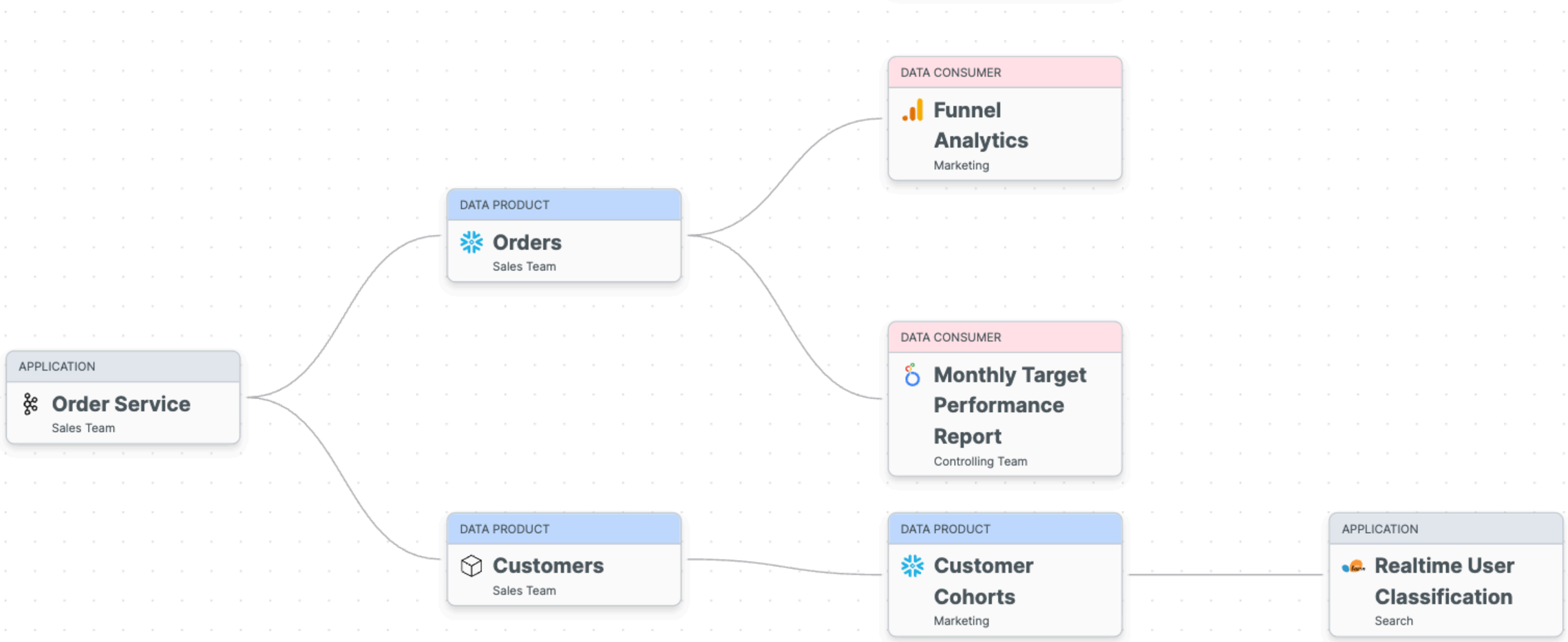
Open Data Product Standard




```yaml
apiVersion: v1.0.0
kind: DataProduct
name: Orders
id: orders
status: active
description:
  purpose: Successful customer orders in the webshop. All orders since 2020-01-01.
tags: [demo, featured]
inputPorts:
  - contractId: kafka_orders_created_datacontract
    version: 1
outputPorts:
  - name: orders_pii
    description: "All order-created events, with PII."
    version: 1
    contractId: snowflake_orders_pii_datacontract
  - name: orders_pii
    description: "All order-created events, with PII."
    version: 2
    contractId: snowflake_orders_pii_datacontract
  - name: orders_npii
    description: "All order-created events, without PII data."
    version: 1
    contractId: snowflake_orders_npii_datacontract
```

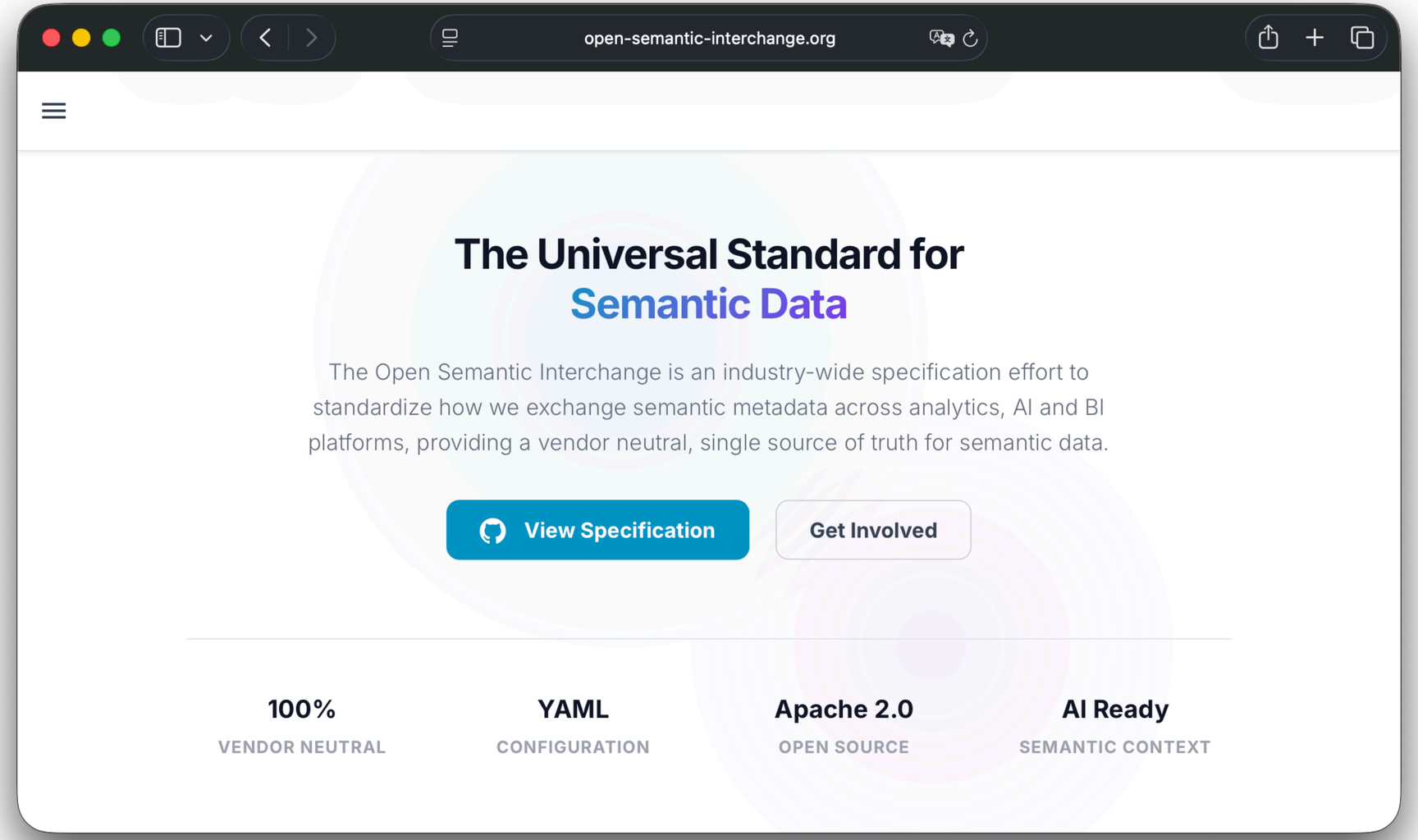
THE BUILDING BLOCKS IN THE DATA WORLD

Open Data Product Standard



Standards for Semantics

Open Semantic Interchange (OSI)



From OSI to Apache Ossie 🦘

Khushboo Bhatia 8:42 PM

Hey Everyone,

We are very excited to announce that Open Semantic Interchange(OSI) has been accepted into the Apache Incubator under a new name: Apache Ossie (Incubating). We chose Ossie to avoid confusion with other open-source projects sharing the OSI acronym. While our name, mascot and governance home have changed, our core mission remains exactly the same: providing a robust specification for defining business context including semantic layer and ontology.

Apache Ossie (Incubating) is now more than a semantic metadata standard, it is the complete business context layer that AI agents need to reason over data correctly. Moving to the Apache Software Foundation ensures Ossie continues as a vendor-neutral standard for expressing this rich, agent-ready context across BI platforms, query engines, and AI agents alike.

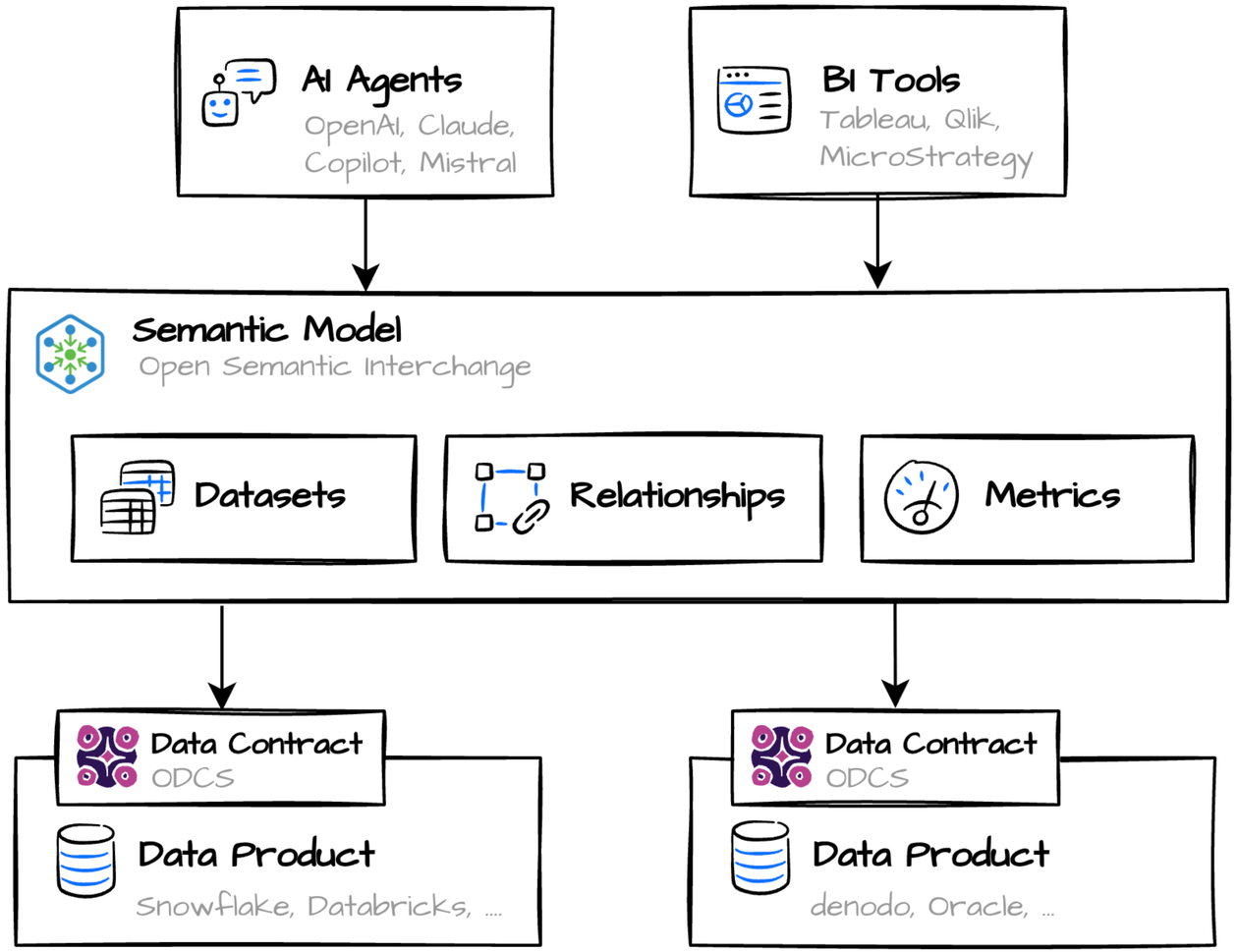
We are currently transitioning our project infrastructure to the ASF and invite everyone to join our consensus-driven community to help shape the future of semantic interoperability. We will post updates right here as soon as the new resources and repositories are ready. All current open PRs will automatically migrated. Thank you for the contributions.

https://github.com/open-semantic-interchange/OSI/discussions/164

👏 4 🎉 2

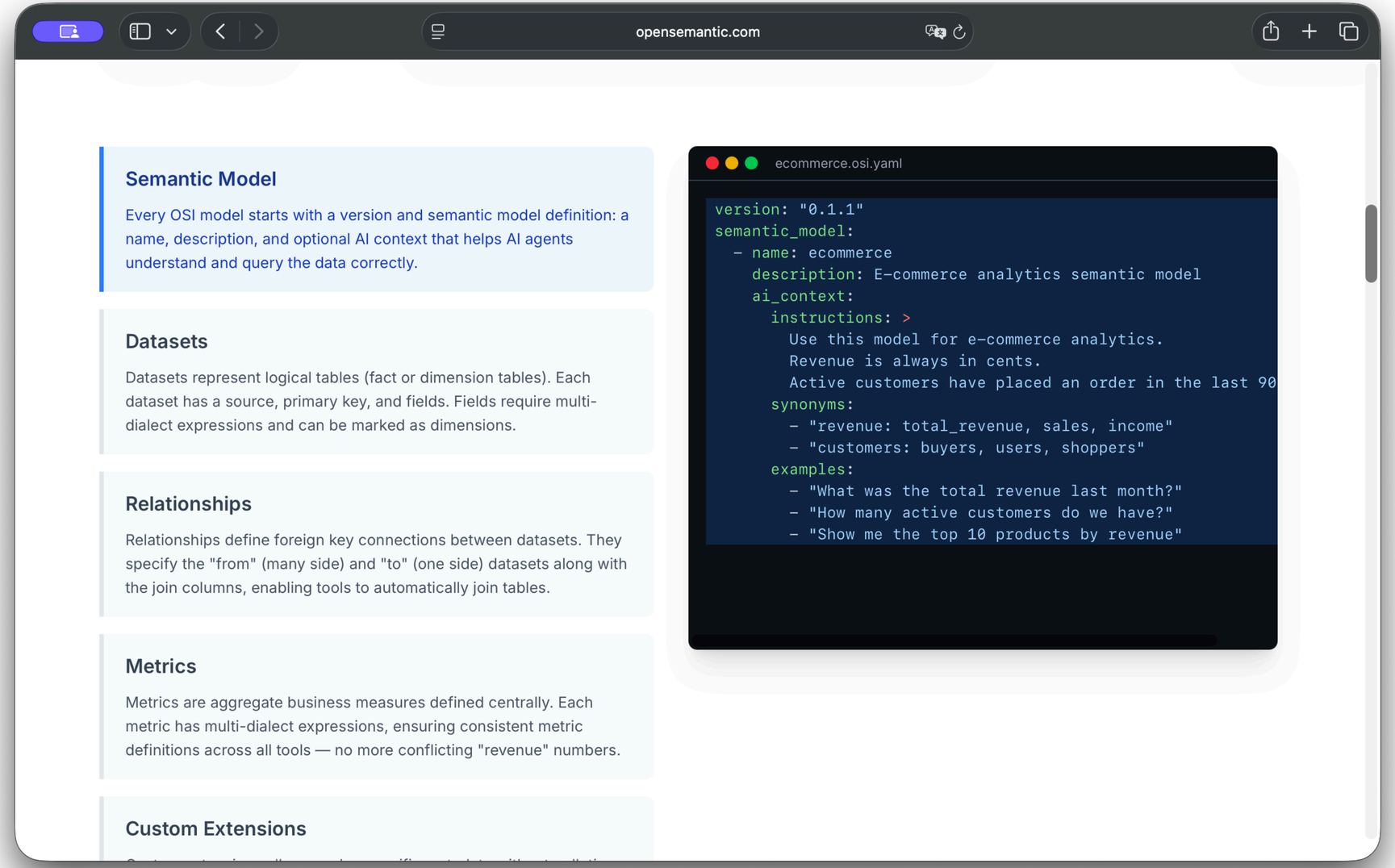
Open Semantic Interchange (OSI)

V0.2.0.DEV (NOT YET STABLE)

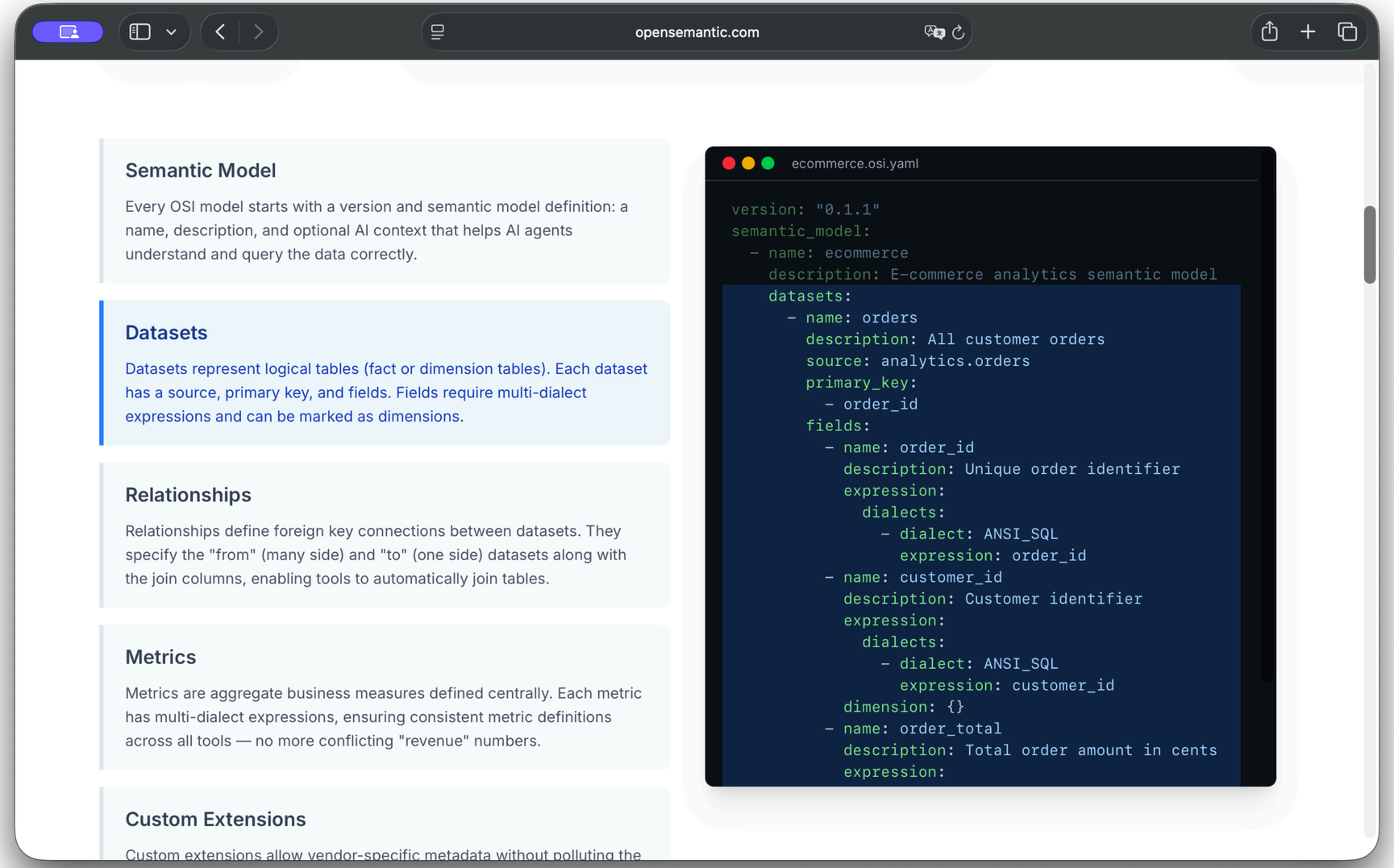


https://opensemantic.com/

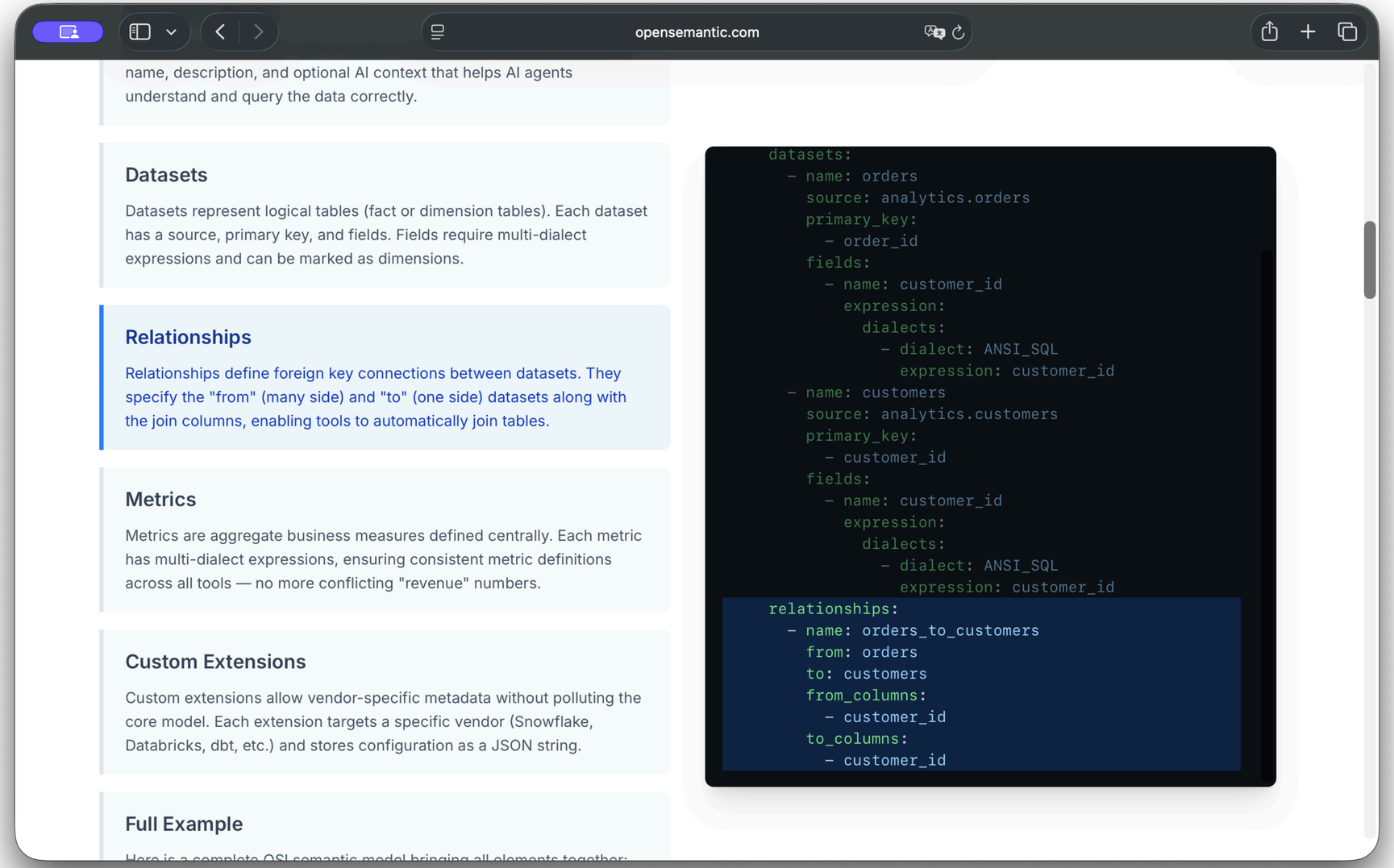
Open Semantic Interchange (OSI)



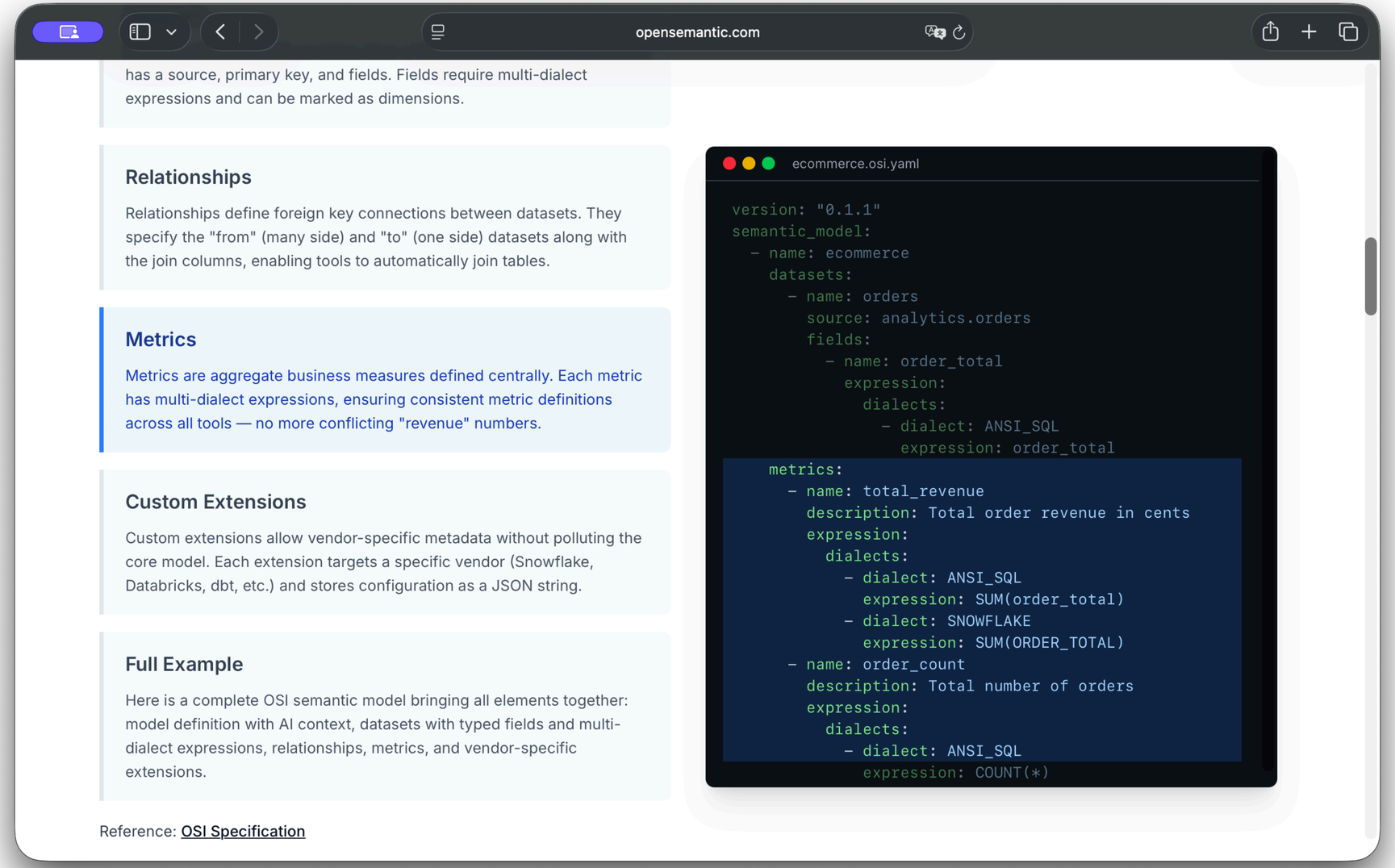
Open Semantic Interchange (OSI)



Open Semantic Interchange (OSI)



Open Semantic Interchange (OSI)



Open Semantic Interchange (OSI)

V0.2.0.DEV (NOT YET STABLE)

	Data Contract (ODCS)	Semantic Model (OSI)
Schema / Dataset	✅	✅
Relationships	✅	✅
Metrics	(upcoming)	✅
Dynamic Fields	(upcoming)	✅
Custom Properties / Vendor Extensions	✅	✅
Terms of Use	✅	
Quality & SLAs	✅	
AI Context	(upcoming)	✅

https://opensemantic.com/

Open Semantic Interchange (OSI)

V0.2.0.DEV (NOT YET STABLE)

Working groups on

- Advanced Metrics & Expression Language
- Composability
- Catalog Integration
- **Ontology representation → already 0.2.0.DEV**
- Model converters and developer tools

And a lot of power behind it

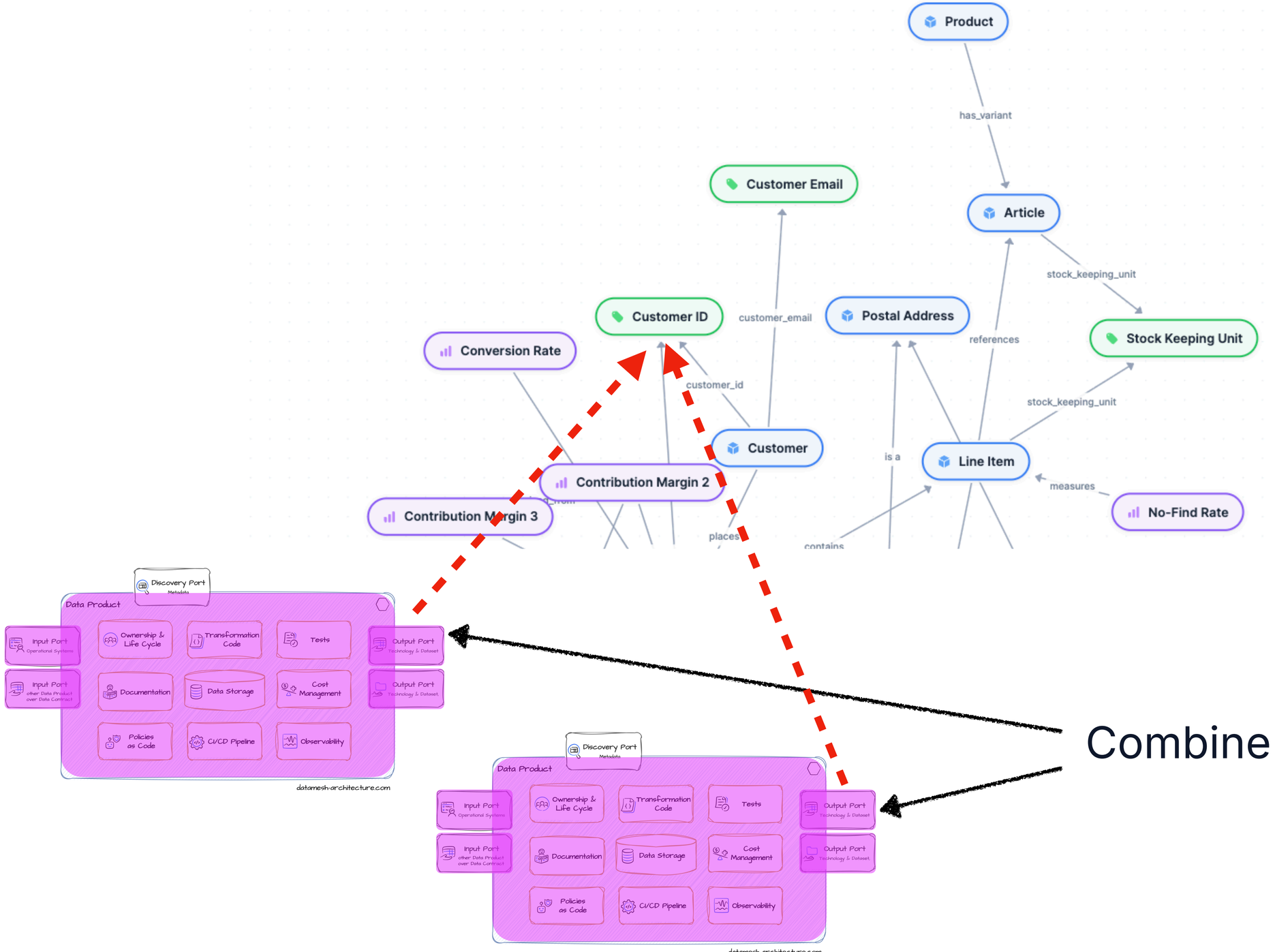
Open Semantic Interchange (OSI)

V0.2.0.DEV (NOT YET STABLE)

Ontology YAML

```yaml
- concept:
    id: "customer"
    name: "Customer"
    type: "EntityType"
    description: "A natural person who places orders in the online shop. Aligned with Go
    iri: "http://www.entropy-data.com/ns/main/Customer"
    group: "customers"
    custom_properties:
      owl:equivalentClass: "http://schema.org/Person"
      name@de: "Kunde"
      name@fr: "Client"
      description@de: "Eine natürliche Person, die Bestellungen im Online-Shop aufgibt.
      description@fr: "Une personne physique qui passe des commandes dans la boutique en
  relationships:
  - id: "customer.account_status"
    name: "account_status"
    roles:
    - concept: "account_status"
    verbalizes:
    - "{Customer} has {account_status}"
    type: "hasProperty"
```

Data Contracts point to those concepts via IRI


Product
has_variant
Customer Email
Article
stock_keeping_unit
Customer ID
customer_email
Postal Address
references
Stock Keeping Unit
Conversion Rate
customer_id
stock_keeping_unit
Customer
is a
Line Item
Contribution Margin 2
measures
Contribution Margin 3
No-Find Rate
places
contains
Discovery Port
Metadata
Data Product
Input Port
Operational Systems
Ownership & Life Cycle
Transformation Code
Tests
Output Port
Technology & Dataset
Input Port
other Data Product over Data Contract
Documentation
Data Storage
Cost Management
Output Port
Technology & Dataset
Policies as Code
CI/CD Pipeline
Observability
datamesh-architecture.com
Discovery Port
Metadata
Data Product
Input Port
Operational Systems
Ownership & Life Cycle
Transformation Code
Tests
Output Port
Technology & Dataset
Input Port
other Data Product over Data Contract
Documentation
Data Storage
Cost Management
Output Port
Technology & Dataset
Policies as Code
CI/CD Pipeline
Observability
Combine

Open Standards

Data Landscape

An opinionated, interactive map of the open standards.

The open standards that power a modern data architecture, organised by what they describe. Inspired by the CNCF Landscape and the ThoughtWorks Tech Radar, and Simon's talk on open standards. Click any standard to learn more, or download as PDF.

> *"Ohhh this is amazing!! You don't know how much that's helping immediately at a project I'm currently on!"*
>
> Tiankai Feng

ADOPT 37 | SITUATIONAL 21 | ASSESS 15 | CAUTION 7 | Single-vendor specs are muted | Search standards... | Landscape | Radar | Table

DEFINITION — HOW DATA IS DESCRIBED

CONTRACTS 6

- ADOPT — ODCS — BITOL @ LF
- ADOPT — OpenAPI — LF
- ADOPT — AsyncAPI — LF
- ADOPT — GraphQL — LF
- SITUATIONAL — gRPC — CNCF
- SITUATIONAL — OData — OASIS

DATA PRODUCTS 4

- ADOPT — ODPS — BITOL @ LF
- ASSESS — DPDS — ODM
- ASSESS — ODPS — LF
- ASSESS — DPROD — OMG

SCHEMA 7

- ADOPT — XML Schema — W3C
- ADOPT — JSON Schema — IETF
- ADOPT — SQL DDL — ISO/IEC
- ADOPT — AVRO Schema — ASF
- ADOPT — Protobuf — GOOGLE
- ASSESS — LinkML — LINKML
- ASSESS — Table Schema — FRICTIONLESS

SEMANTICS 8

- SITUATIONAL — RDF/OWL — W3C
- SITUATIONAL — DCAT — W3C
- SITUATIONAL — SKOS — W3C
- SITUATIONAL — SHACL — W3C
- SITUATIONAL — JSON-LD — W3C
- ASSESS — OSI — OSI INITIATIVE
- ASSESS — schema.org — W3C
- ASSESS — ShEx — W3C

What if we already were using Open Standards for Data Products?

Let us check where the future is now


⌥⌘1
-zsh
simonharrer@Simons-MacBook-Pro dp_demo_new_snowflake_dp %

Other Standards for Data Products

Open Standards

Data Landscape

An opinionated, interactive map of the open standards.

The open standards that power a modern data architecture, organised by what they describe. Inspired by the CNCF Landscape and the ThoughtWorks Tech Radar, and Simon's talk on open standards. Click any standard to learn more, or download as PDF.

> *"Ohhh this is amazing!! You don't know how much that's helping immediately at a project I'm currently on!"*
>
> Tiankai Feng

ADOPT 37 | SITUATIONAL 21 | ASSESS 15 | CAUTION 7 | Single-vendor specs are muted | Search standards... | Landscape | Radar | Table

DEFINITION — HOW DATA IS DESCRIBED

CONTRACTS 6

- ADOPT — ODCS — BITOL @ LF
- ADOPT — OpenAPI — LF
- ADOPT — AsyncAPI — LF
- ADOPT — GraphQL — LF
- SITUATIONAL — gRPC — CNCF
- SITUATIONAL — OData — OASIS

DATA PRODUCTS 4

- ADOPT — ODPS — BITOL @ LF
- ASSESS — DPDS — ODM
- ASSESS — ODPS — LF
- ASSESS — DPROD — OMG

SCHEMA 7

- ADOPT — XML Schema — W3C
- ADOPT — JSON Schema — IETF
- ADOPT — SQL DDL — ISO/IEC
- ADOPT — AVRO Schema — ASF
- ADOPT — Protobuf — GOOGLE
- ASSESS — LinkML — LINKML
- ASSESS — Table Schema — FRICTIONLESS

SEMANTICS 8

- SITUATIONAL — RDF/OWL — W3C
- SITUATIONAL — DCAT — W3C
- SITUATIONAL — SKOS — W3C
- SITUATIONAL — SHACL — W3C
- SITUATIONAL — JSON-LD — W3C
- ASSESS — OSI — OSI INITIATIVE
- ASSESS — schema.org — W3C
- ASSESS — ShEx — W3C

But if you only remember one thing…

Agents make building cheap.
Describing data is the new bottleneck.

Open standards do the describing.
⇒ Open tooling. No vendor lock-in.

ENTROPY DATA

Thank you! Questions?

Interested in doing research together?

Happy to talk. :-)

Contact: simon.harrer@entropy-data.com

Links:

data-landscape.com

datacontract.com

entropy-data.com