OMG - Enterprise Knowledge Graph Forum

DPROD: the bigger picture

Shared meaning for data products

A practical introduction to DPROD, its core model, examples, and adoption path.

Published: 2026-11-04

Contents

About this guide . . . 3
Why shared meaning matters . . . 4
Standards that work together . . . 5
Four building blocks . . . 6
A product in practice . . . 7
When terms need to be explicit . . . 8
Getting started . . . 9
How DPROD evolves . . . 10
Where to go next . . . 11

About this guide

A practical guide for product owners, data leaders, architects, and the teams who publish and consume data.

What this guide covers

- Why shared descriptions matter in a decentralized data ecosystem.
- How products, ports, datasets, and distributions fit together.
- A worked example that connects a product to a shared financial schema.
- The optional role of data contracts and a practical adoption path.

This is an EKGF community overview, not a replacement for the normative specification. DPROD's published OMG baseline and the evolving community work are distinct publications; check the specification versions before selecting a version for implementation.

What is DPROD?

The Data Product Ontology (DPROD) is an Object Management Group (OMG) standard that profiles the W3C Data Catalog Vocabulary (DCAT) to specifically describe Data Products. As organizations increasingly adopt decentralized data architectures like Data Mesh, DPROD provides the standardization needed to ensure interoperability and unlock the full potential of distributed data ecosystems.

Built on established W3C technologies including DCAT, RDF, OWL, SHACL, and PROV, DPROD offers a clear schema for describing data products, ensuring they are discoverable, interoperable, and treated with the same level of accountability as traditional products.

Why shared meaning matters

Inconsistent metadata

Every platform, every team, every vendor describes data products differently. Without a shared schema, consumers spend more time reconciling metadata than using the data.

Limited discoverability

Data products exist in silos. Without standard descriptions, there is no federated search, no cross-domain catalog, and no way for consumers to find what already exists.

Interoperability friction

Data products built on incompatible descriptions cannot be composed. DPROD lets a data product produced on one platform be described, understood, and consumed on another — without rewriting metadata.

Working Group Leadership

DPROD is the collective work of its EKGF workgroup and wider community. Tony Seale is the workgroup chair. Matthias Autrata has contributed especially to Data Contracts; Josh Cornejo and many others have made substantial contributions across DPROD.

The project's documented editors are Matthias Autrata, Marcel Fröhlich, Jacobus Geluk, Pete Rivett, Tony Seale, and Natasa Varytimou.

Its documented authors are Oli Bage, Ben Clinch, Josh Cornejo, Steve Fisher, Andrea Gioia, Charles Ivie, Carl Mattocks, Carlos Tubbax, and Ben Whittam Smith. DPROD also benefits from the workgroup's reviewers, implementers, and other community contributors.

OMG Request for Comments

OMG published the DPROD proposed specification for public comment as part of its standardisation process, focused on improving discoverability, interoperability, and reducing vendor lock-in across data marketplaces. Read the OMG release

Standards that work together

Built on W3C standards

DPROD does not reinvent the wheel. It profiles a stack of established W3C standards so data products inherit decades of work on semantics, validation, and interoperability.

DCAT

Data Catalog Vocabulary

The W3C vocabulary DPROD extends. Provides the core building blocks for describing datasets, distributions, and data services — which DPROD profiles and specialises for data products.

RDF

Resource Description Framework

The graph data model that lets every DPROD concept be identified by a URI and linked to concepts from other vocabularies without translation layers.

OWL

Web Ontology Language

The formal logic layer DPROD uses to define classes and properties with precise semantics — so tools can reason about data products, not just store records about them.

SHACL

Shapes Constraint Language

The validation layer. DPROD ships SHACL shapes that let any tool verify whether a data product description conforms to the standard — no custom validators required.

PROV

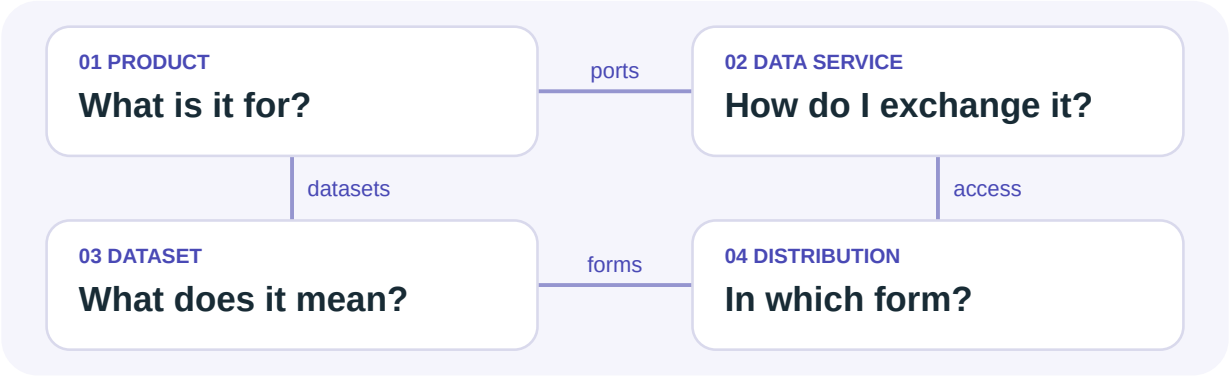
Provenance Vocabulary

The lineage layer. PROV lets DPROD describe how data products were produced, by whom, and from which inputs — essential for trust, audit, and compliance.

Four building blocks

Core Concepts

The building blocks of every DPROD description. Each concept links to its normative definition in the specification.



Data Product

A rational, managed, and governed collection of data with purpose, value, and ownership, meeting consumer needs over a planned lifecycle. Data products have input and output ports, code, and metadata.

Normative definition in the spec

Ports (Data Services)

Digital interfaces that provide access to datasets. Input ports bring data into the product, while output ports share generated data. Ports specify connection details, formats, and link to datasets with shared schemas.

Normative definition in the spec

Distributions

Physical representations of data — CSV, JSON, Parquet, and access methods. A single dataset may have multiple distributions to serve different consumer needs without duplicating the underlying information.

Normative definition in the spec

Datasets

Logical models of the data that can conform to shared standards like FIBO, CDM, or custom ontologies using SHACL or OWL. The schema lives with the data, not in out-of-band documentation.

Normative definition in the spec

A product in practice

UK Bonds in practice

As DCAT Data Services, the DPROD input and output ports can specify connection details, they have distributions that define formats, and link to datasets that conform to shared schemas. In this example, the UK Bonds Data Product includes an output port, which is a RESTful API. This API delivers JSON data conforming to the shared FIBO specification for callable bonds.

Example 1

```
  {
    "@context": "https://ekgf.org/dprod/spec/develop/dprod-simple.jsonld",
    "id": "https://y.com/products/uk-bonds",
    "type": "DataProduct",
    "title": "UK Bonds",
    "description": "UK Bonds is your one-stop-shop for all your bonds!",
    "dataProductOwner": "https://www.linkedin.com/in/tonyseale/",
    "dataProductLifecycleStatus": "https://ekgf.org/dprod/spec/develop/data/lifecycle
-status/Consume",
    "outputPort": {
      "id": "https://y.com/products/uk-bonds/ports/10-year-api",
      "type": "DataService",
      "endpointURL": "https://y.com/uk-10-year-bonds",
      "isAccessServiceOf": {
        "id": "https://y.com/products/uk-bonds/distributions/10-year-json",
        "type": "Distribution",
        "format": "https://www.iana.org/assignments/media-types/application/json",
        "isDistributionOf": {
          "id": "https://y.com/products/uk-bonds/datasets/10-year",
          "type": "Dataset",
          "conformsTo": "https://spec.edmcouncil.org/fibo/ontology/SEC/Debt/Bonds/Cal
lableBond"
        }
      }
    }
  }
```

When terms need to be explicit

A data product is only useful once someone can rely on it. The provider of a product commits to deliver data on a schedule, to keep it conformant to a schema, and to give notice before it changes; the consumer, in turn, agrees to use the data only in permitted ways and to meet obligations of its own, such as reporting usage. In a Data Mesh these commitments are made between many autonomous teams, so they need to be recorded in a form that machines can validate and evaluate, not just in prose.

DPROD Data Contracts is a profile of the W3C Open Digital Rights Language (ODRL) 2.2 odrl-model for exactly this purpose. It uses ODRL's own terms for everything ODRL already defines: a policy, its permissions, prohibitions and duties, the constraints on them, and the parties and assets they refer to.

Optional adoption

The Data Contracts profile is optional. A data product description conforms to DPROD whether or not it carries any policy; an implementation that does process contracts must satisfy the contracts shapes as stated in the Conformance section.

Getting started

1. Reference the JSON-LD context

DPROD ships a stand-alone JSON-LD `@context` so you can describe data products in plain JSON and get full RDF semantics for free. Point your documents at the published context URL:

```
{
  "@context": "https://ekgf.org/dprod/spec/develop/dprod-simple.jsonld",
  "type": "DataProduct",
  "id": "https://example.com/data-products/customer-360",
  "title": "Customer 360",
  "description": "A unified view of the customer across domains.",
  "outputPort": [
    {
      "type": "DataService",
      "id": "https://example.com/data-products/customer-360/api"
    }
  ]
}
```

Any JSON-LD processor will expand these terms into the full DPROD vocabulary automatically — no code generation, no schema compilation.

2. Study the worked examples

The specification includes worked examples covering common patterns: SBA pool rates, equity trades, core data product extensions, data lineage, data quality, data rights, data schemas, and observability ports. Each example is a complete, machine-readable JSON-LD file you can copy and adapt.

Browse examples on GitHub

3. Validate with the SHACL shapes

DPROD ships SHACL shapes that any conforming RDF tool can use to validate your data product descriptions. Use the shapes for the version you are implementing. They check the constraints encoded in those shapes; operational and business requirements may require additional checks.

Browse the shapes in the spec Raw SHACL file

4. Cite the OMG standard

When you reference DPROD in papers, internal documentation, or procurement requirements, use the formal OMG citation:

```
Object Management Group (OMG). Data Product Ontology (DPROD), Version 1.0.
OMG Document Number: dtc/2024-09-01.
https://ekgf.org/dprod/spec/develop/
```

How DPROD evolves

Roadmap

DPROD evolves through formal OMG ballots. Here is where it has been, where it is now, and how the process works.

Timeline

1. Done — 2024

   DPROD 1.0 published for public comment

   OMG released the DPROD proposed specification as part of its standardisation process. The 1.0 version is frozen and serves as the official reference point.

2. Done — 2024

   OMG Request for Comments completed

   Community feedback gathered through the OMG RFC cycle informed a backlog of improvements now being addressed in subsequent ballots.

3. In progress — Ballot 4

   Modelling and clarity improvements

   Fixes to class identifiers in the generated spec, protocol modelling, lifecycle status values, SHACL shape hygiene, and worked examples for data rights. Approved individual issues are bundled for a single vote.

4. Planned — Later ballots

   Deeper alignment and new examples

   More worked examples, tighter alignment with neighbouring OMG and W3C vocabularies, and tooling improvements. Priorities are set by the working group and driven by community issues.

Where to go next

Choose the depth you need

For a first conversation

- DPROD project website: purpose, concepts, adoption, and roadmap.
- Core concepts: the main building blocks in plain language.

For an implementation

- Specification versions: choose the publication appropriate for your work.
- Worked examples: concrete product descriptions and optional contract patterns.
- Source repository: ontology, shapes, contexts, issues, and contribution history.

For the underlying standards

- DCAT: catalogs, datasets, distributions, and data services.
- JSON-LD: JSON documents with linked-data meaning.
- SHACL: validation of RDF descriptions.
- ODRL: machine-readable permissions, prohibitions, and duties.