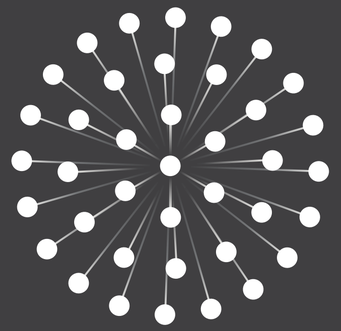


AIDA
USER GROUP

Artificial Intelligence,
Data, and Analytics
User Group

Open Data Contract Standard

v1.0.3

Jean-Georges Perrin • Co-founder & President • jgp@jgp.ai

Jean-Georges "jgp" Perrin

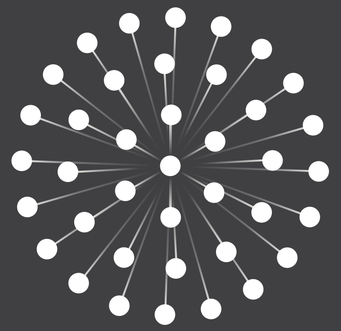
AIDA · Artificial Intelligence Data Analytics User Group

ARCH SAFe 5

SP SAFe 5

IBM Enterprise Design Thinking Practitioner

IBM Champion
Lifetime Achievement
IBM



AIDA
USER GROUP

Artificial Intelligence, Data, and Analytics User Group

As **Artificial Intelligence is transforming society** and businesses, AIDA User Group has the ambition to **share knowledge**, **connect people**, and **provide guidance** through this next era of computing while embracing peace and the highest levels of **ethics**.

Join AIDAUG for free at http://aidaug.org/join

What is a data contract?

- Creates a link between data producers and data consumers.
- Creates a link between a logical representation of the data and its physical implementation.
- Describes meta meta data: rules, quality, and behavior.

What are the problems we are trying to solve?

- Normalizing and keeping documentation relevant.
- Describing service-level expectations.
- Easing data & tools integration.
- Enabling data product thinking.

AIDA USER GROUP

Artificial Intelligence, Data, and Analytics User Group

Benefits of a standard

- Boosting innovation.
- Simplifying integration.

https://aidaug.org/odcs

AIDA
USER GROUP

Artificial Intelligence,
Data, and Analytics
User Group

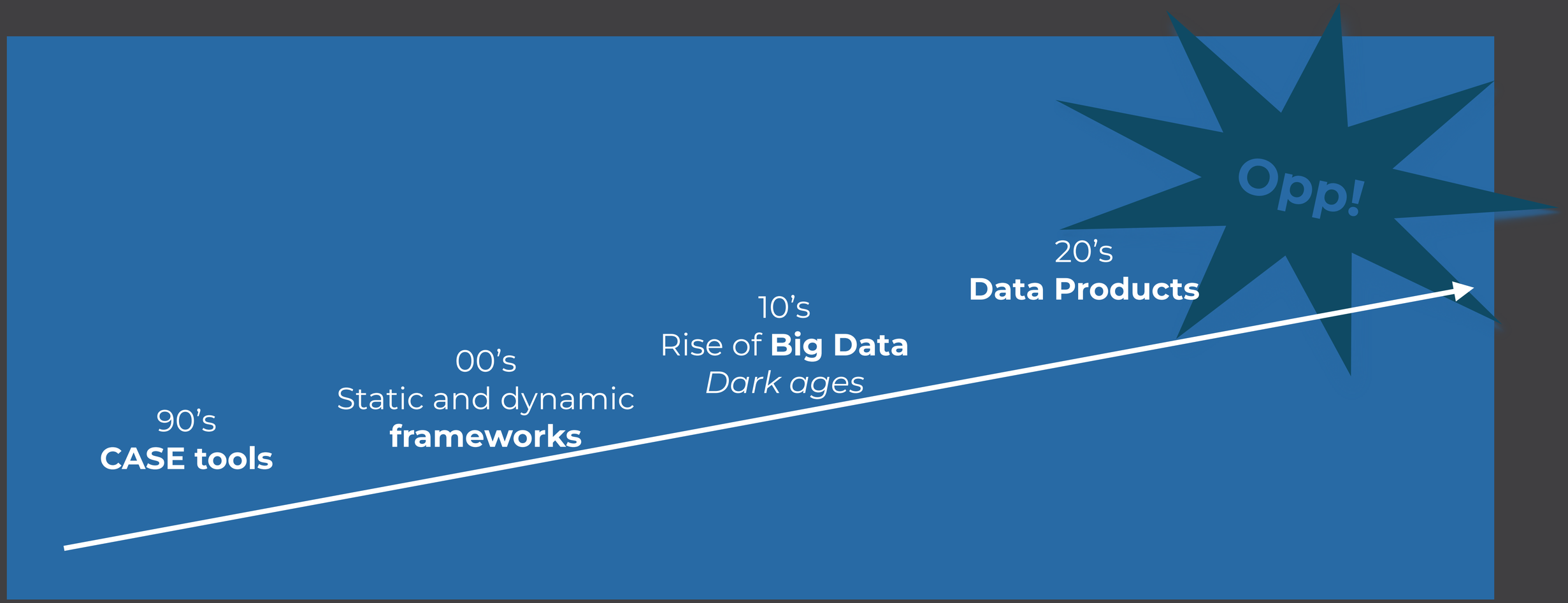
Who uses the Open Data Contract Standard?

Various levels of maturity in:

- Fortune 500: peer-to-peer payment leader, major cable company.
- Startups: several data-oriented startups in Europe and in NA (Among them: Raito, Innoq).
- Mid-market companies: in the US & Europe (Among them: ProfitOptics, Innoboost).
- An active user group behind the standard (AIDA User Group).
- A community (Data Mesh Learning).

Active recruitment of users & vendors underway!

History of data-related development

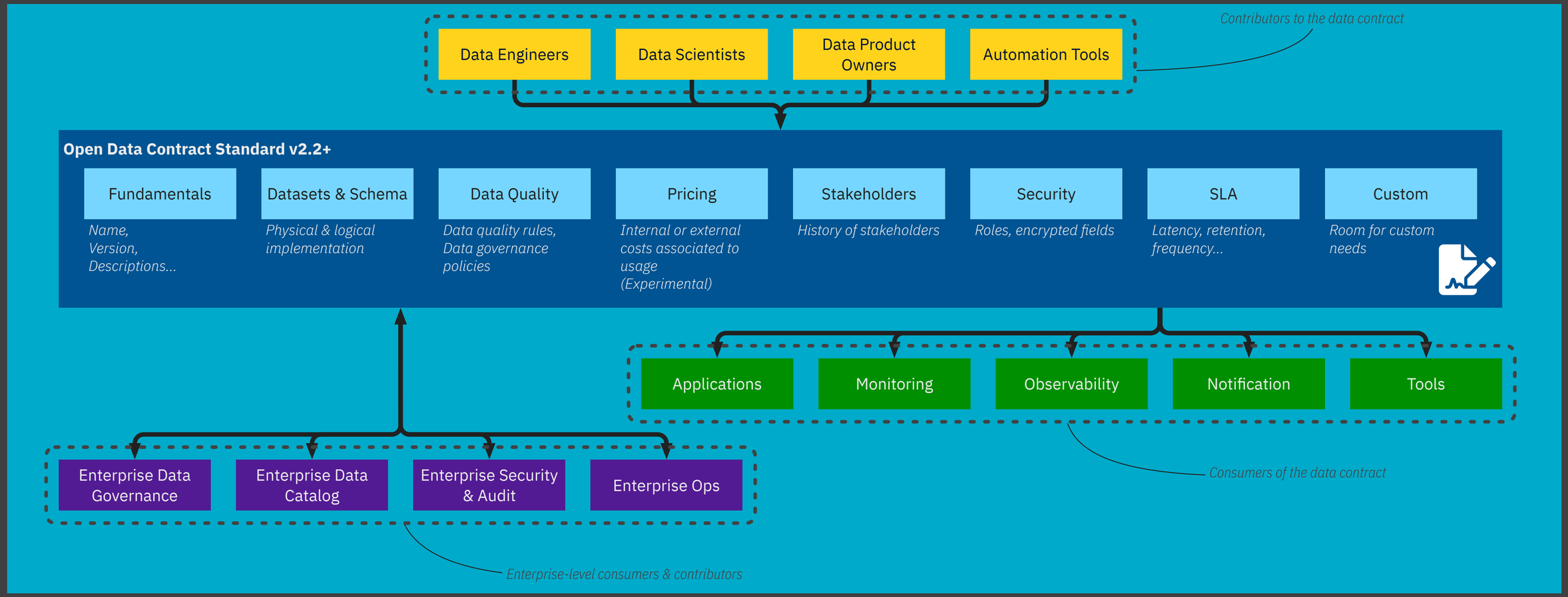


A popular standard

	PayPal (Until v2.1)	ODCS (From v2.2)	Cumulative
Stars	360	27	387
Forks	38	44	44
Contributors	5	11	11

https://aidaug.org/odcs

Architecture & Content



Implementation

- YAML file, both human -and computer-readable.
- Language agnostic.

```yaml
- table: tbl
  description: Provides core payment metrics
  dataGranularity: Aggregation on columns txn_ref_dt, pmt_txn_id
  columns:
    - column: txn_ref_dt
      businessName: Transaction reference date
      logicalType: date
      physicalType: date
      description: Reference date for the transaction. Use this date in reports and
        aggregation rather than txn_mystical_dt, as it is slightly too mystical.
      sampleValues:
        - 2022-10-03
        - 2025-01-28
```

Table name

Mapping from logical to physical world

AIDA USER GROUP

```yaml
- table: tbl
  columns:
    - column: rcvr_cntry_code
      businessName: Receiver country code
      logicalType: string
      physicalType: varchar(2)
      authoritativeDefinitions:
        - url: https://collibra.com/asset/748f-71a5-4ab1-bda4-8c25
          type: Business definition
        - url: https://github.com/myorg/myrepo
          type: Reference implementation
```

Multiple links to enterprise definitions

AIDA
USER GROUP

Evolution of this project

Here

	Phase 0	Phase 1	Phase 2	Phase 3	Phase 4
Standard	Open source by PayPal and developed by AIDA	Publish ODCS v2.x	Animate a working group Publish ODCS v3.x (lic. change)	Publish **Data Product** Standard (ODPS) v1.x Evolution of ODCS	Publish **Data Mesh** Standard (ODMS) v1.x Evolution of ODCS & ODPS
Tools	No	ODCS parser	ODCS validator	Updated & new parsers & validators	Updated & new parsers & validators

AIDA
USER GROUP

Artificial Intelligence,
Data, and Analytics
User Group

Communication plan

- 9/11 IBM TechXchange
- 9/20 Big Data London (Chat with exhibitors)
- 9/27 IBM TechTalk (Online APAC)
- 9/28 IBM TechTalk (Online Americas)
- 10/15 All Things Open (Raleigh, NC, USA) - LF announcement
- Q4: Articles
- 2024: Several conferences, articles, at least one **book** (Implementing Data Mesh, O'Reilly)

Roadmap post LF incubation

- Create a set of Open Standards, around:
  - Data Contracts.
  - Data Products.
  - Data Mesh.
- Host open-source tooling.
- Provide guidance around licensing & operations.

Request to the Linux Foundation

Requesting the support of the TAC to host the project in LF AI & Data as a Sandbox project.

Appendix - Evolution of DC

- Support for unstructured documents
- Support for IoT data
- More details on the relation between data
- …