

DATA CONTRACTS

Adapting to evolving AI-driven workflows,
ensuring data quality and integrity



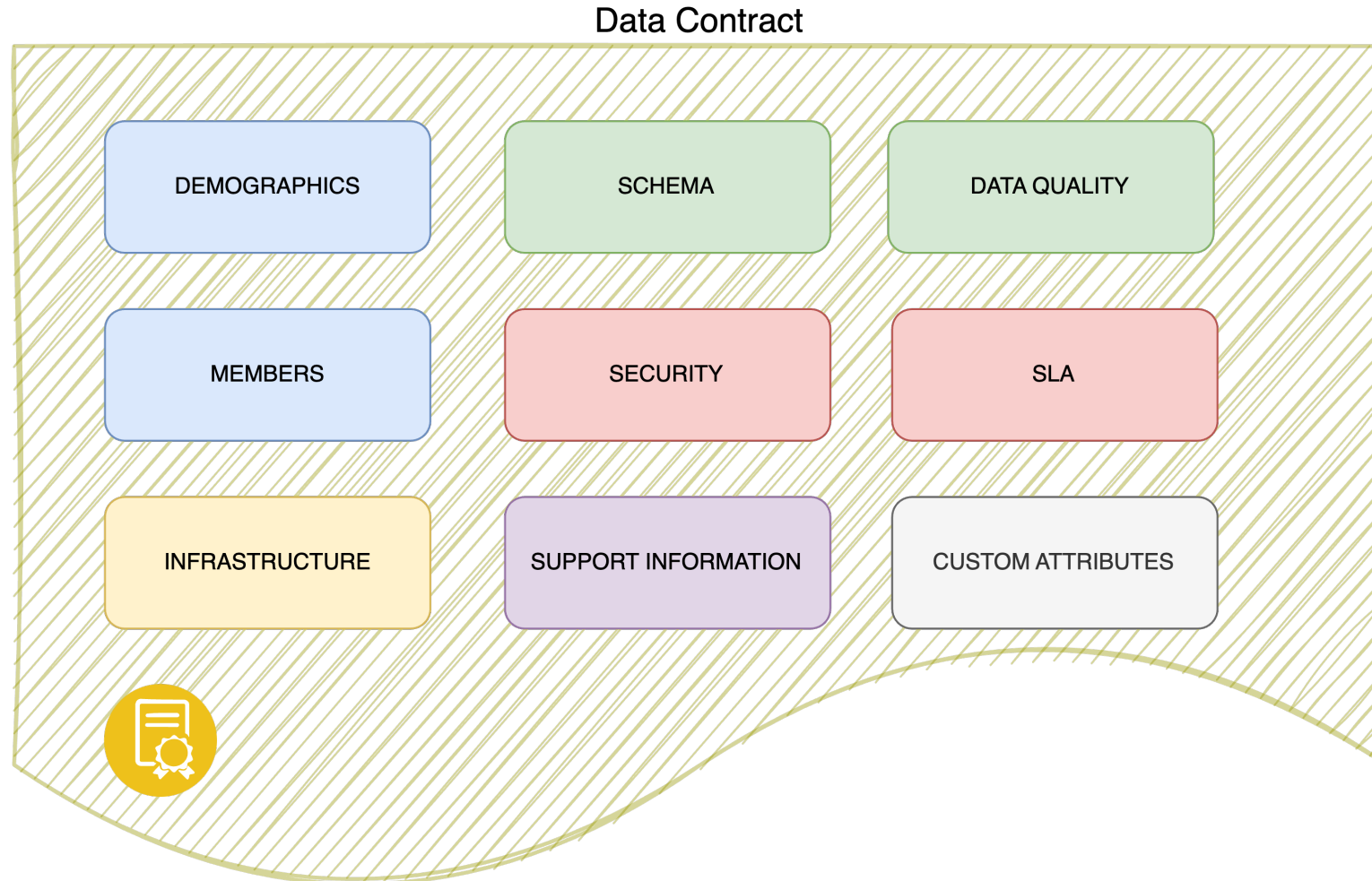
WHY DATA CONTRACT?

- Today, businesses increasingly rely on data so ensuring its consistency, reliability, and accessibility across departments is a growing challenge
- The main problem data industry faces is the lack of ownership and enforceability of producers of data assets entering data platforms (e.g. the analytical database)
- Data tools like data catalog, data observability, or data transformation tool focus on empowering data teams but don't enable collaboration amongst data teams and upstream data producers, who are often software engineers
- The solution of this problem is to introduce data contracts and the tools for enforcing them.

WHAT IS DATA CONTRACT

- As any other business contract it can be interpreted as formal agreement, in this case between data producers—those who create or modify data—and data consumers—those who use the data
- Data contracts enable software engineers who produce the data to be aware of their consumers data dependencies by embedding enforcement within their already established CI/CD workflows.
- When data contract is violated, engineers have the ability to:
 - prevent a breaking change
 - provide the business impact
 - tag the team best equipped to deal with the violation of data contract.

DATA CONTRACT STRUCTURE

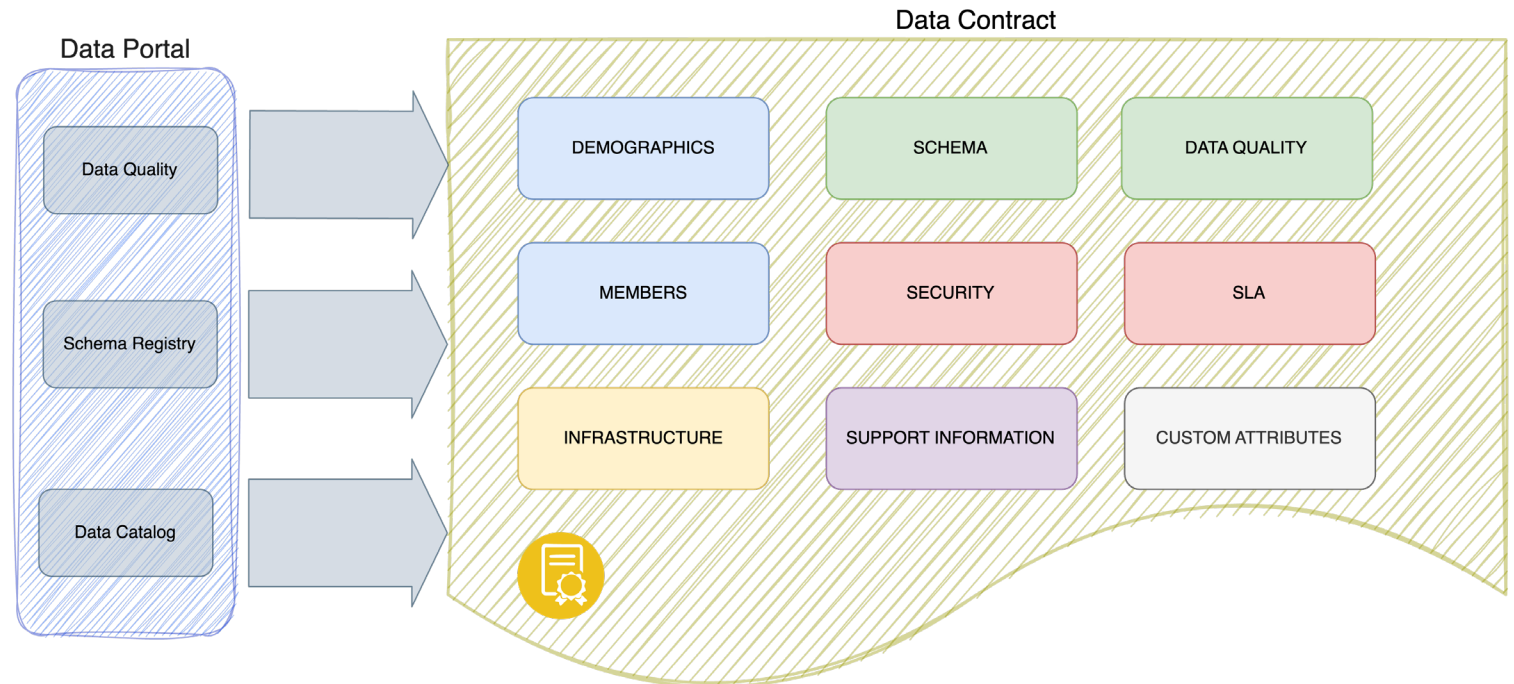


DATA CONTRACT STRUCTURE - DETAILS

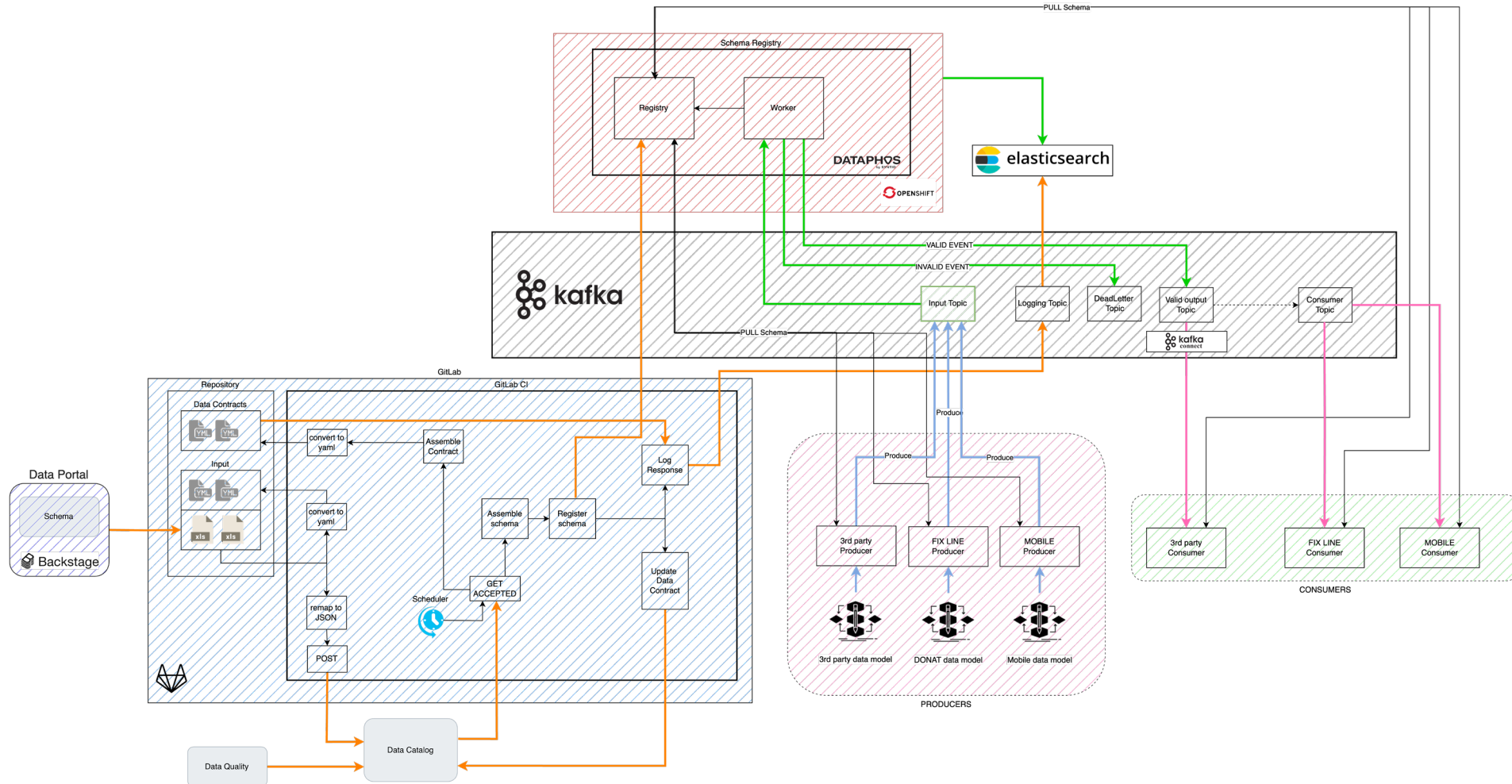
- Demographics or Fundamentals
 - name, version, ID, description, status, domain, data product ...
- Schema
 - Business and physical representation
 - Schema header information
 - Schema properties
- Data Quality
 - data quality rules & parameters
- Members
 - team members or stakeholders
- Support information
 - Communication channel and details
- Security
 - roles needed to access the dataset
- SLA
- Infrastructure
 - specific environment and a specific technology for physical location of data
- Custom

WHO CREATES/DEFINES DATA CONTRACT AND HOW

- Data Engineers
- Data Scientists
- Data Producers
- Data Consumers
- Data Product Owners
- Tools for process automation



IMPLEMENTATION OF DATA CONTRACT – THE REAL DEAL



Thank you!

