

Data quality intelligence

Accessible and Enhanced Data Quality



DATA QUALITY INTELLIGENCE

Data Quality Intelligence represents an **innovative** practice that integrates cutting-edge Artificial Intelligence techniques with data quality principles.

Using Advanced Data Quality solutions, we can **detect anomalies**, perform **data remediation**, boost our data through **data enrichment** and even perform **data quality controls** using natural language by leveraging state of the art AI solutions.



ANOMALY DETECTION



ANOMALY DETECTION



DATA REMEDIATION

DATA ENRICHMENT



DATA AS A SOURCE OF KNOWLEDGE

The ability to improve and enrich data through analysis of the information it contains, as well as the capacity to forecast additional information based on the existing dataset.



FIELD EXTRACTION FROM TEXT

Generative AI allows to **populate fields** in a dataset using unstructured textual descriptions.



KNOWLEDGE EXTRACTION FROM TEXT

With **generative AI**, knowledge can be extracted directly from textual fields.

ANOMALY DETECTION



DETERMINISTIC CONTROLS

By using deterministic checks, data integrity can be improved by systematically detecting **null values**, eliminating **duplicates**, identifying **outliers**, and rigorously validating the data sets provided.



ARTIFICIAL INTELLIGENCE CONTROLS

Using generative AI, anomalies within a dataset can be automatically detected. The AI autonomously analyzes the records, identifying discordant data.



NATURAL LANGUAGE CONTROLS

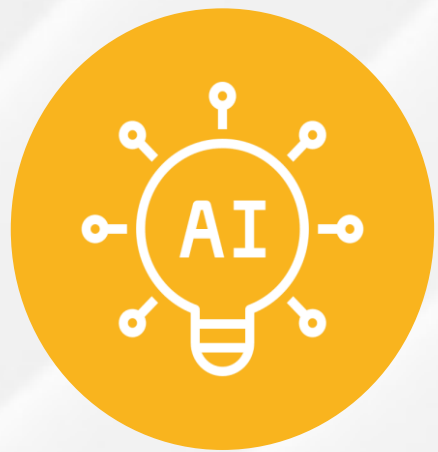
Through **Generative AI**, **Data Quality checks** specified in **Natural Language** are submitted to the data quality intelligence solution, which conducts the verification and returns the anomalies.

DATA REMEDIATION



SIMILARITIES BETWEEN RECORDS

Whenever an anomaly is detected in a record, through ML the correct record with the greatest similarity can be found and the value of that record can be used to correct the anomaly.



GENERATIVE AI CORRECTION

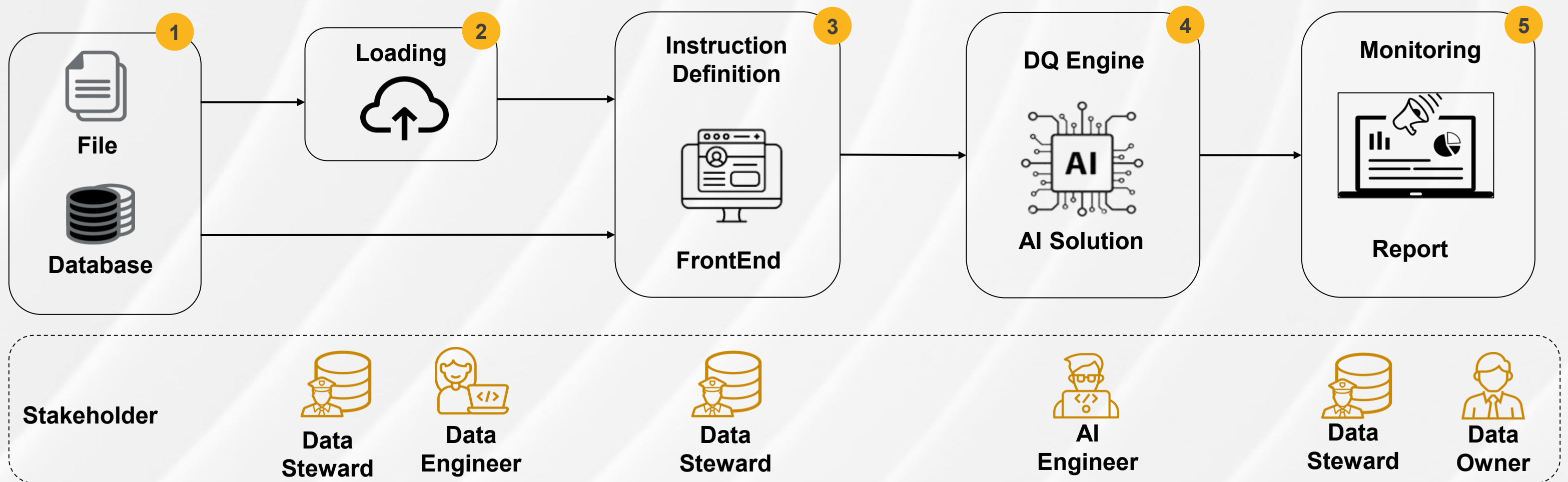
Using generative AI, one can automatically **correct anomalous data** by standardizing formatting or completing missing fields based on available information.



MACHINE LEARNING PREDICTION

Using advanced Machine Learning models, it is possible to **predict** the value of a record detected as anomalous, apply regression models, classification techniques, and even time series analysis.

OUR SOLUTION



The solution relies on a tailored front-end **User Interface**, empowering users to craft scenarios for data enrichment, specify desired data quality checks, and establish effective data remediation policies. The solution is based on the following steps:

- **Data Loading:** data is directly ingested from a Database (1) or manually loaded in the solution (2)
- **Instructions Definition:** Users decide which of the earlier tasks to carry out, putting them in control of the whole process (3)
- **DQ Engine:** the task is carried out by the Solution (4)
- **Monitoring:** a report is generated showcasing the key outcomes achieved by the process (5)

TOOLS



**AZURE
OPENAI**



**AZURE MACHINE
LEARNING**

BUSINESS GOAL

BUSINESS STRATEGIES MAXIMIZATION

High-quality data ensures that decisions generated by business systems are accurate, offering more precise information to guide strategies, decreasing errors.

STAKEHOLDERS AND CUSTOMERS TRUST

Reliable data is essential for building trust. Having reliable data foster confidence in the organization's services, leading to stronger relationships and loyalty.

OPERATIONAL EFFICIENCY

High-quality data improves efficiency by reducing time spent correcting errors or managing inconsistent data. to focus on core activities and enhance overall performance.

COMPLIANCE

Data integrity and accuracy is paramount for regulatory compliance. By ensuring that the data used for decision-making is correct, businesses can meet regulatory requirements.



Contact us!



Claudio Motisi

c.motisi@reply.it



Giulio Cammarata

g.cammarata@reply.it

WWW.CLUSTERREPLY.IT