DATA GOVERNANCE AND MASTER DATA MANAGEMENT CONFERENCE EUROPE

11 - 14 March 2024 |London, UK

Please score and comment on this session and speaker in the event mobile app

IRMUK

Order Up! Delivering a Global Data Governance Framework

Our Journey in Data Governance

March 2024

Who are we

David Fernandez

Global Data Governance Manager

JUST EAT Takeaway.com

Deloitte. Santander Bank of America

Let's connect here

Kirill Skaletskiy

Global Senior Data Quality Professional

JUST EAT Takeaway.com

adidas X5

Let's connect here

Data Governance at JET

1 A bit about JET

2 Our Data Governance Journey

3 Data Governance use cases

We are a global leader in online food delivery...

20
Countries

891m
Orders[1]

€26.4bn
GTV[1]

699k
Partners[1]

94m
Active Consumers[2]

Note 1: Full year 2023
Note 2: Full year 2022

...with a track record of integration and growth

2000

Thuisbezorgd.nl

Jitse Groen launches Thuisbezorgd.nl in The Netherlands

2007
First international expansion with entry into Germany and Belgium

2011
Name change to Takeaway.com

2012
First external investment of €13m from Prime Ventures

2014
Acquisitions of Lieferando.de and Pyszne.pl

2016
IPO at Euronext Amsterdam; first Scoober city launched

2018
Acquisition of 10bis in Israel, reflecting entry into B2B market

2019
Acquisition of Lieferheld, pizza.de and Foodora from Delivery Hero

2020

2021
Acquisition of Grubhub in the United States

Acquisition of Bistro.sk in Slovakia

JUST EAT Takeaway.com

2001

Launches in Denmark

Just Eat

2006
Just Eat launches in UK and moves HQ

2009
First external investment of £10m from Index Ventures

2014
IPO at London Stock Exchange

2015
Acquisition of Menulog in Australia

2016
Acquisition of SkipTheDishes in Canada

2018
Acquisition of Hungry House in the UK

2022 - We realized we have a case for Data Governance

Data without governance is like a bag of marbles without a bag – you'll end up with a colorful mess rolling all over the place, causing chaos and probably tripping a few people along the way

A new journey....

2000

Thuisbezorgd.nl

Jitse Groen launches Thuisbezorgd.nl in The Netherlands

2001

Launches in Denmark

Just Eat

2006

Just Eat launches in UK and moves HQ

2007

First international expansion with entry into Germany and Belgium

2009

First external investment of £10m from Index Ventures

2011

Name change to Takeaway.com

2012

First external investment of €13m from Prime Ventures

2014

Acquisitions of Lieferando.de and Pyszne.pl

2014

IPO at London Stock Exchange

2015

Acquisition of Menulog in Australia

2016

IPO at Euronext Amsterdam; first Scoober city launched

2016

Acquisition of SkipTheDishes in Canada

2018

Acquisition of 10bis in Israel, reflecting entry into B2B market

2018

Acquisition of Hungry House in the UK

2019

Acquisition of Lieferheld, pizza.de and Foodora from Delivery Hero

JUST EAT Takeaway.com

2020

2021

Acquisition of Grubhub in the United States

Acquisition of Bistro.sk in Slovakia

2022

Kirill and David join Data and Analytics at JET

The **Data Governance journey begins…**

Chapter 1: In a desert of ungoverned data

We started by Managing our Risks

Risk Identification

Identifying **potential risks** associated with a particular activity, process, or business area

Risk Mitigation

Implementing measures to **reduce the likelihood** and/or impact of each risk

Control Monitoring

Ongoing monitoring of risk and control environment to ensure risk mitigation measures are working as intended and identify changes in the risk profile

Objective

Systematic + comprehensive approach to managing risks, all significant risks are **identified**, **evaluated**, and **managed** effectively - automated where possible

Risk management process is integrated into the **day-to-day activities** of the organization

Align with external requirements

Controls

Frameworks:

- General IT Controls
- Order to Cash
- Internal Data and Analytics

Auditing

- Bi-Annually
- Quarterly
- Monthly

Control Lifecycle

Standard Operating Procedure

- Risk **identified**, and **accountability** assigned
- Control **designed** and **documented**
- Control **implemented** and **tested**

Ongoing monitoring

- **Automate** monitoring of control
- **Assess** provided evidence

Control Review

- Periodic **review** of applicability of control
- **Update** SOP if required

Deprecate control

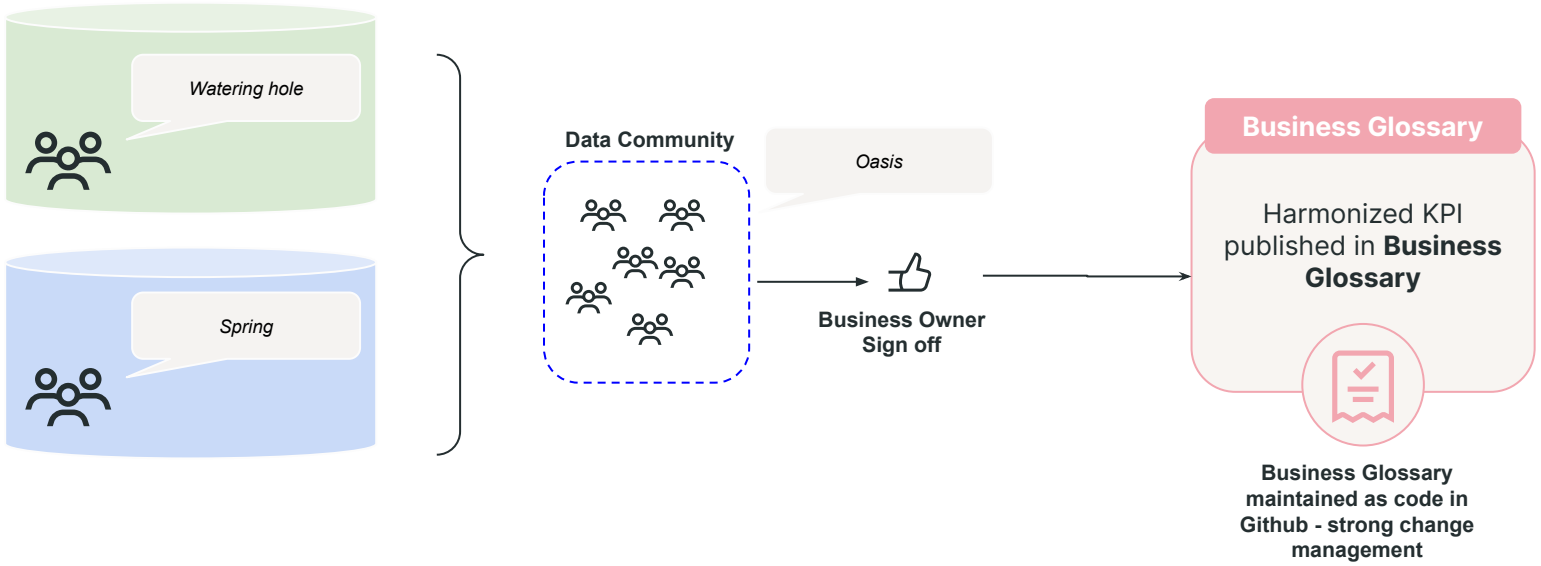
- **Remove** control from the Risk and Control Matrix

End of Chapter 1: We survived! Finding an oasis

Chapter 2: Speaking the same language

Harmonizing our Business Glossary in a shared language

Standardize and unify business glossary terminology: improve communication, alignment, accuracy of measurement, accountability, and simplify reporting



End of Chapter 2: We have a common understanding

Chapter 3: What does our environment look like

JET before Data Catalog: A maze of unknown unknowns

JET with Data Catalog: Charting the Uncharted

- Exercise in **transparency**
- Bazaar of **data knowledge**, everything is free!
- One-stop shop to **understand JET data ecosystem**

Introducing Datahub: Navigating our data landscape

Datahub

Centralized data management platform enabling effective data discovery, sharing, and collaboration, helping drive more value from data

Data Discovery

Intuitive **search** and **exploration** of metadata to identify and understand relevant data assets

Data Dictionary

Inventory of **physical data assets** and their attributes, including Google BigQuery, Tableau, etc...

Data Lineage

Upstream and **downstream** record of data flow throughout its lifecycle, enabling impact analysis

Business Glossary

Repository for **harmonized business elements** (KPIs and metrics) across all business domains.

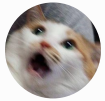
Data Quality

Integrates with **Great Expectations** and **dbt-tests** and visualizes data validation results

Datahub Use Case

“Hey Kirill, I want to know how many Active Corporates we had last month, can you please help me here? ASAP”

Manager

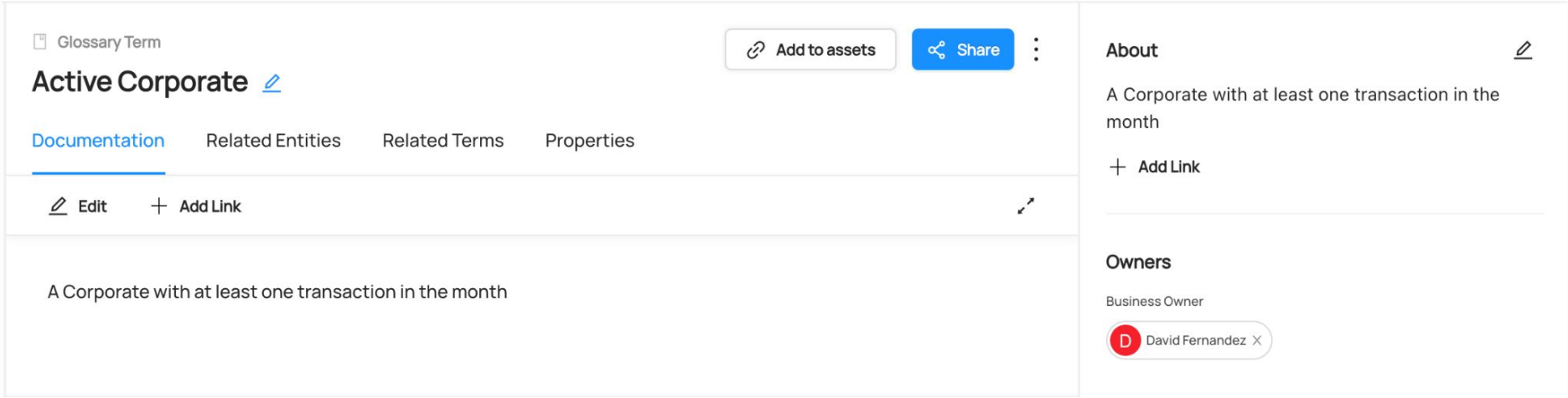


“Sure”

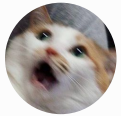
Kirill

“Kirill just joined JET and lacks knowledge about Active Corporates and how to perform calculations related to it, but heard about Datahub”

Business Glossary



Datahub Use Case



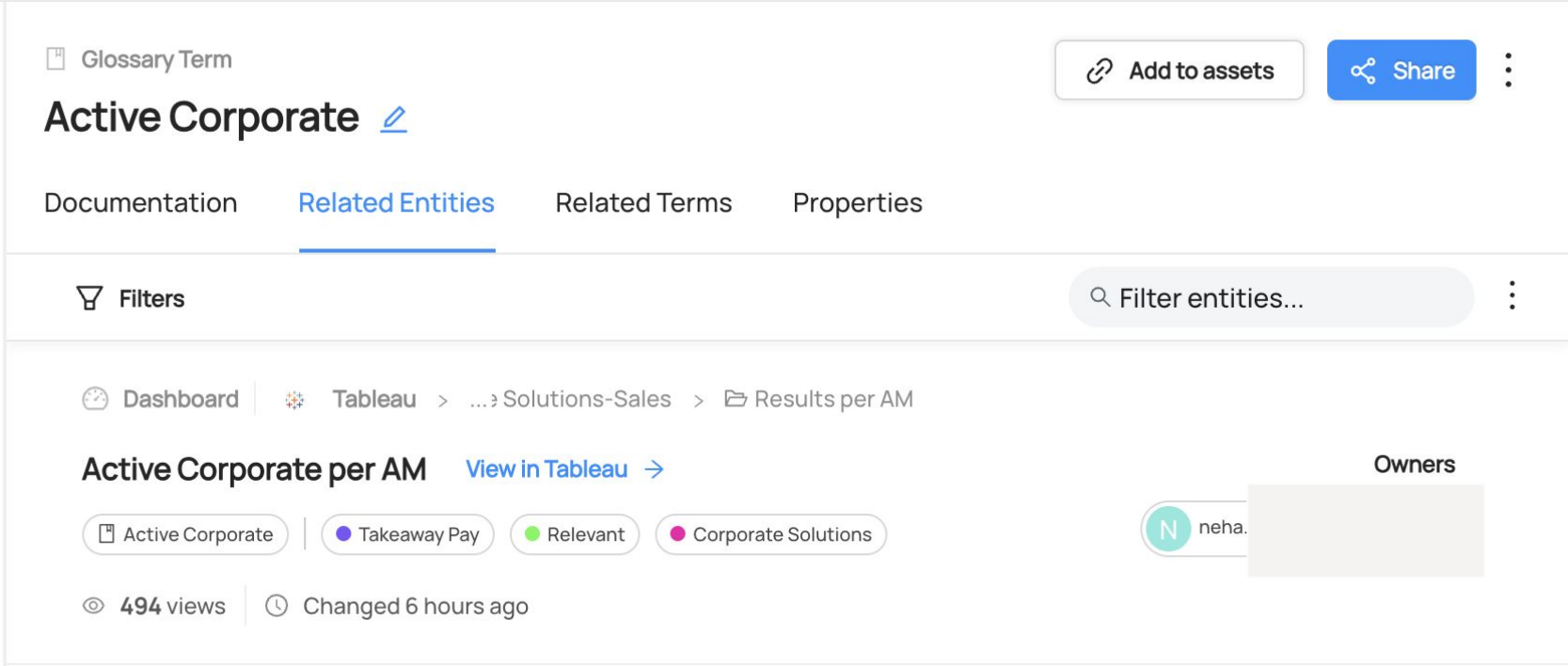
Hi David, I noticed you're the Owner of Active Corporate KPI. Could you assist me in locating Tableau dashboards featuring this specific KPI?

Kirill

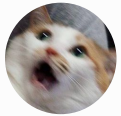
“Sure, check the Related Entities tab in Datahub”

David

Business Glossary



Datahub Use Case



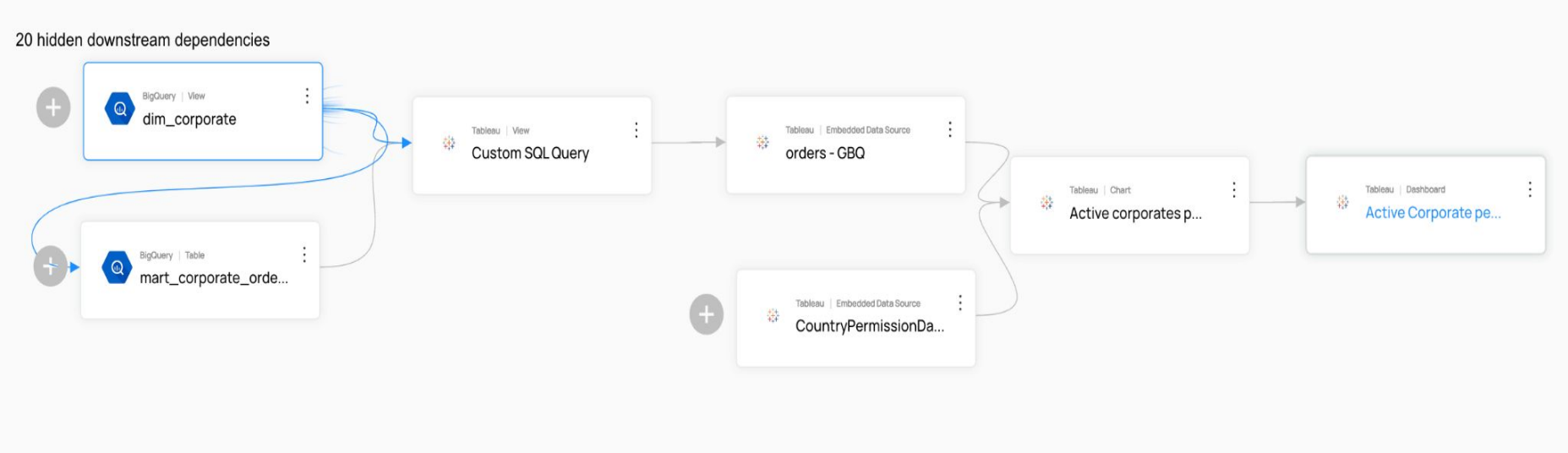
Hi, I noticed you're the Owner of Active Corporate per AM dashboard in Tableau. Could you assist me in locating source tables for this dashboard?

Kirill

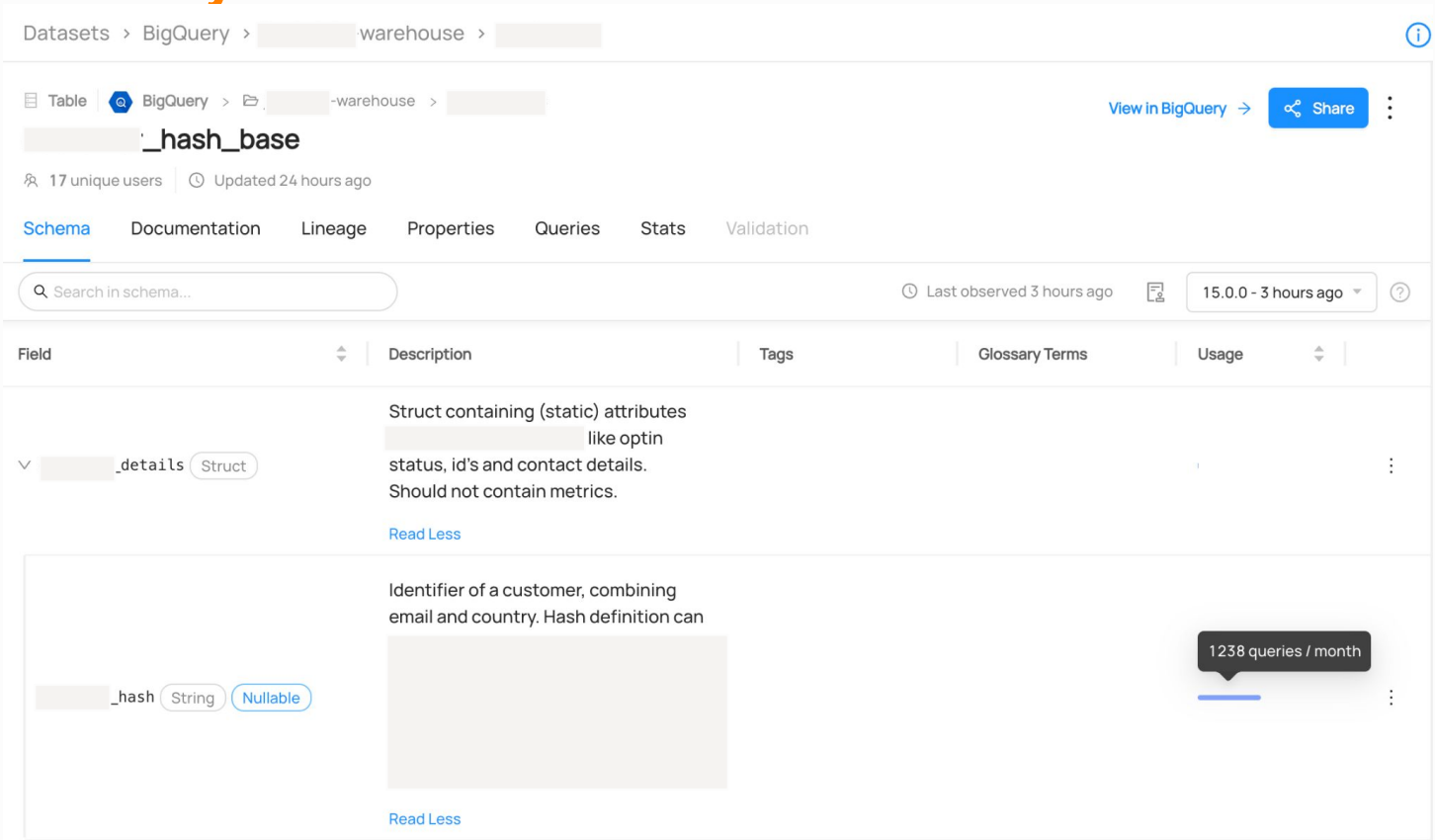
“Sure, check the Lineage tab in Datahub”

Neha

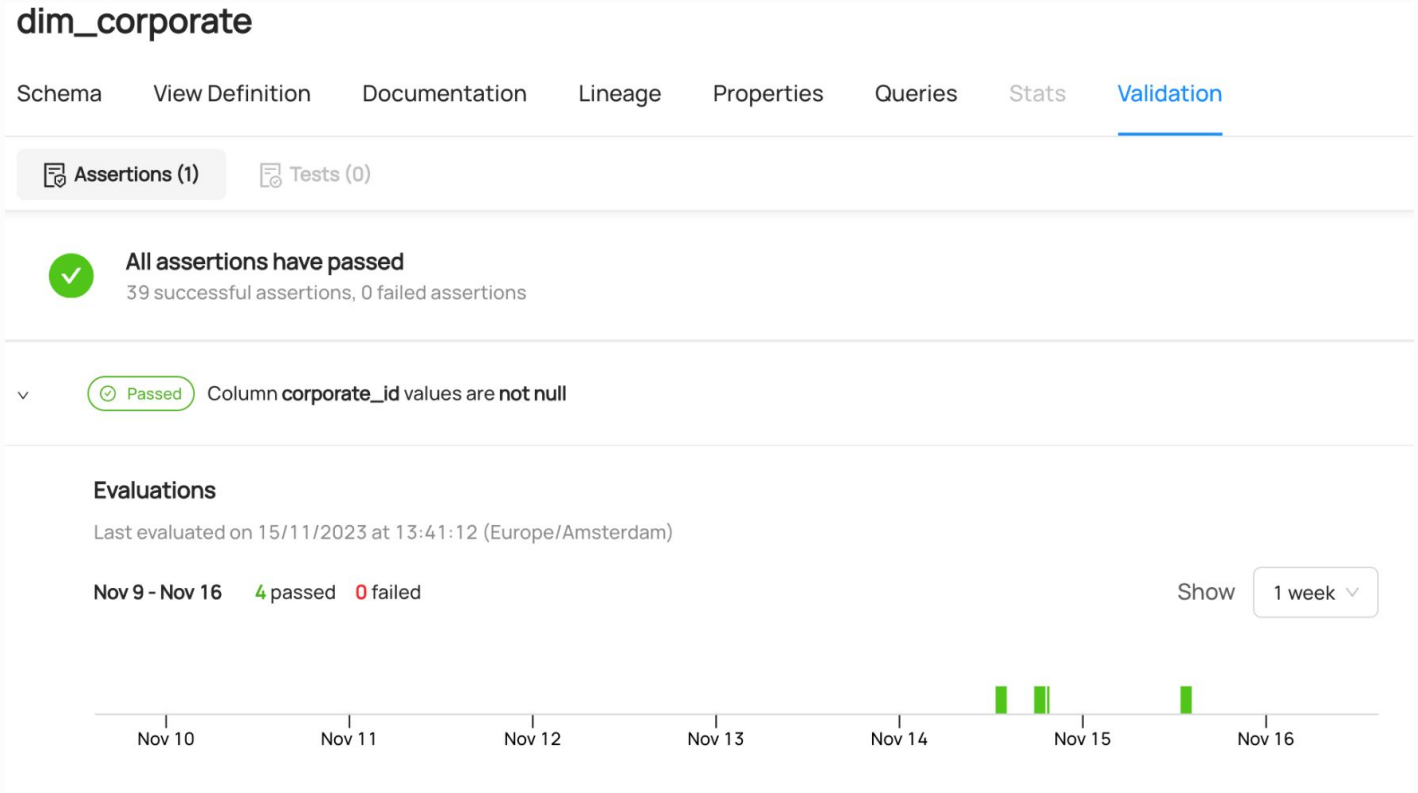
Data Lineage



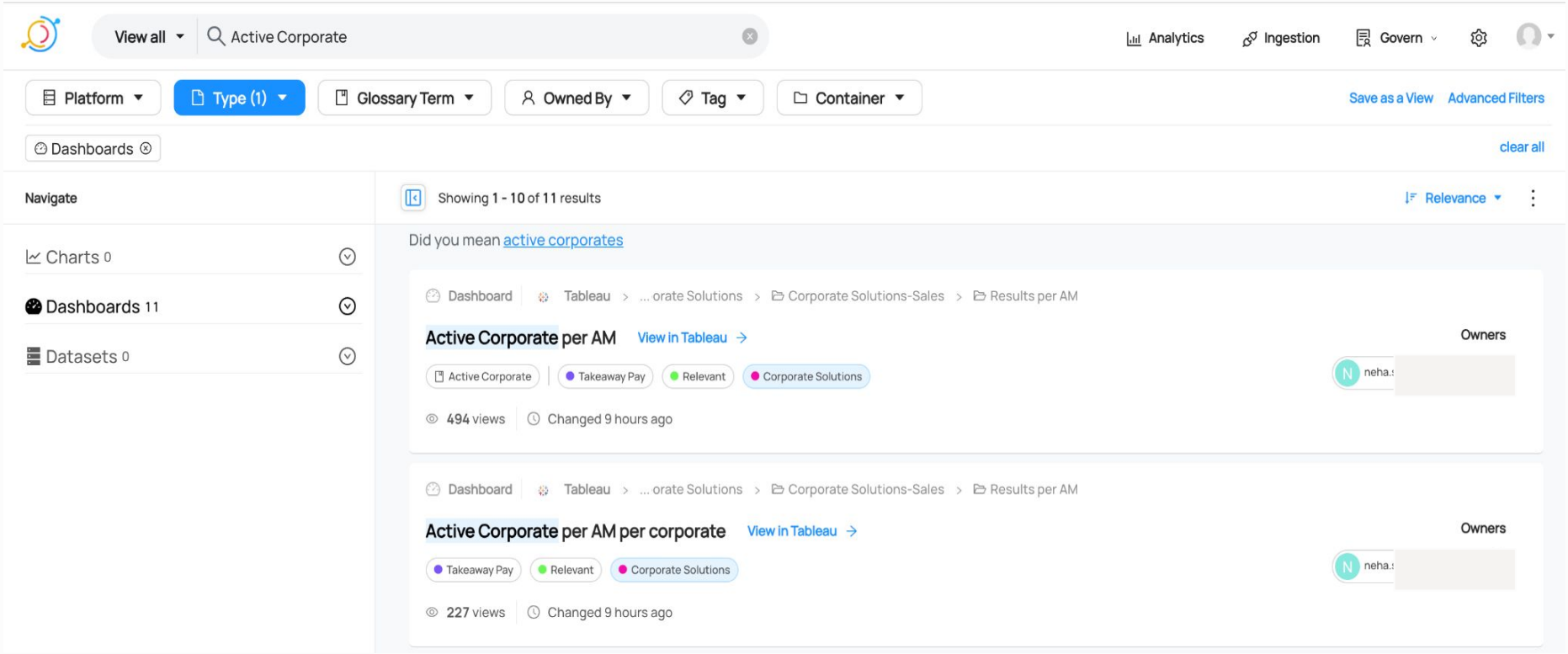
Data Dictionary



Data Quality



Search


View all
Active Corporate
Analytics
Ingestion
Govern
Platform
Type (1)
Glossary Term
Owned By
Tag
Container
Save as a View
Advanced Filters
Dashboards
clear all
Navigate
Showing 1 - 10 of 11 results
Relevance
Did you mean active corporates
Charts 0
Dashboards 11
Datasets 0
Dashboard
Tableau
... orate Solutions
Corporate Solutions-Sales
Results per AM
Active Corporate per AM
View in Tableau
Owners
Active Corporate
Takeaway Pay
Relevant
Corporate Solutions
494 views
Changed 9 hours ago
Active Corporate per AM per corporate
227 views

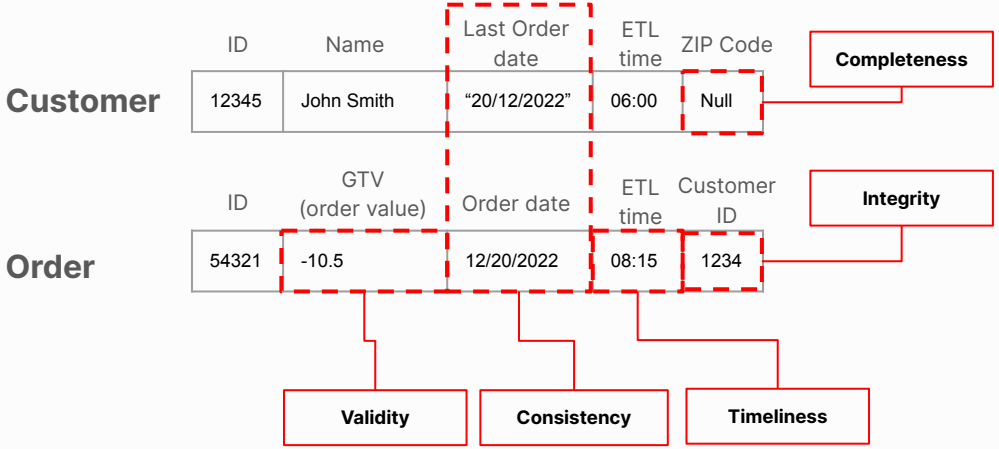
End of Chapter 3: We mastered our surroundings

Chapter 4: Identifying the good apples

Data Quality is a fundamental pillar of Data Governance: it represents the level of trust JET can have in the metrics

Data quality is the **measure** of how well suited **data meets** its intended **requirements**

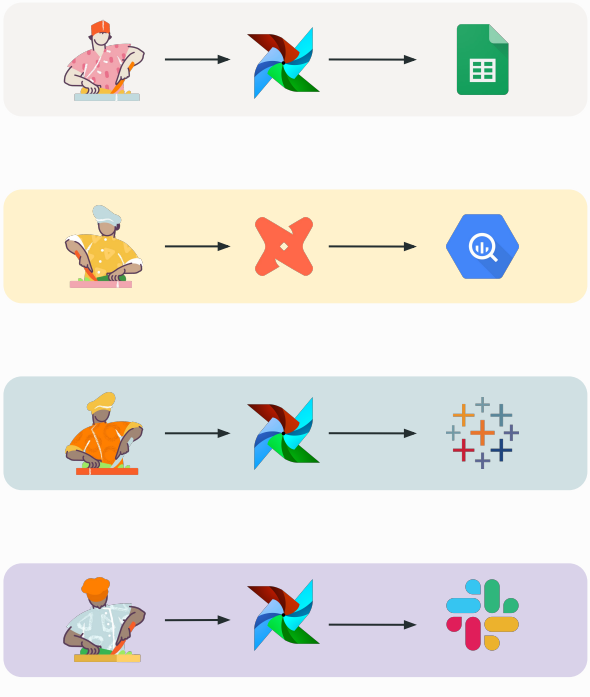
Measures of data quality are based on data quality **dimensions**



Completeness	Validity	Consistency
The degree to which data **contains required attributes** and sufficient **number of records**	The degree to which data conform to a set of **business rules**	The degree to which **data aligns** or is in uniformity with **another table/system**

Integrity	Timeliness
The degree to which data conform to **data relationship** rules and **duplicates**	The degree to which we have the data **at the right time**

Our approach to Data Quality (DQ) in the past has been very siloed



- Teams operating **independently** using **multiple validation tools**
- No Data Quality **standards** or **best practices**
- No **single source of truth** for all validations
- **Chats** are filled with Data Quality questions

The Road is Data Quality Program to ensure data is fit for consumption and meets needs of data consumers

Ambition

To **consolidate**, **standardize**, **monitor** and **improve** all data quality activities in JET under the **Data Quality Framework**

Goals

Get overview on existing DQ validations in a single place

Standards and specifications for quality controls as part of the data lifecycle

Continuous processes to measure, monitor, and report on data quality levels

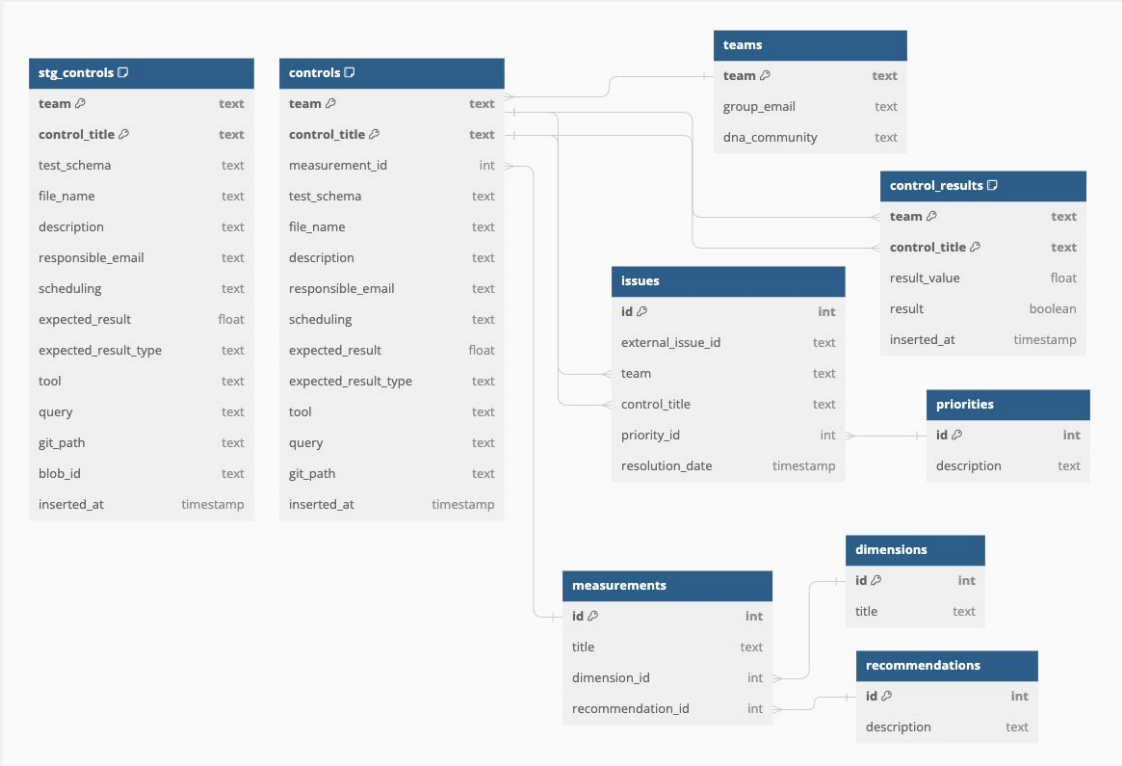
48 Measurements to understand Data Quality

Measurement ID	Data Quality dimension	Measurement Type	Description	DQ Control Example	Granularity
6	Completeness	Field completeness - non-nullable fields	Ensure all non-nullable fields are populated	Identify all orders with a null order date	Field
32	Consistency	Consistent amount field data by aggregated date	Compare amount aggregated by date to historical total and percentage	Identify a 5% deviation from last month's GTV to this month's GTV	Field
40	Integrity	Parent / child referential integrity	Confirm referential integrity between parent / child tables	Identify all orders with a city id that is not in city table	Table
42	Timeliness	Timely delivery of data for processing	Compare actual time of data delivery to scheduled data delivery	Identify all tables where the delivered time is greater than the SLA	Table

DQF Agnostic-schema. Bridging everyone!

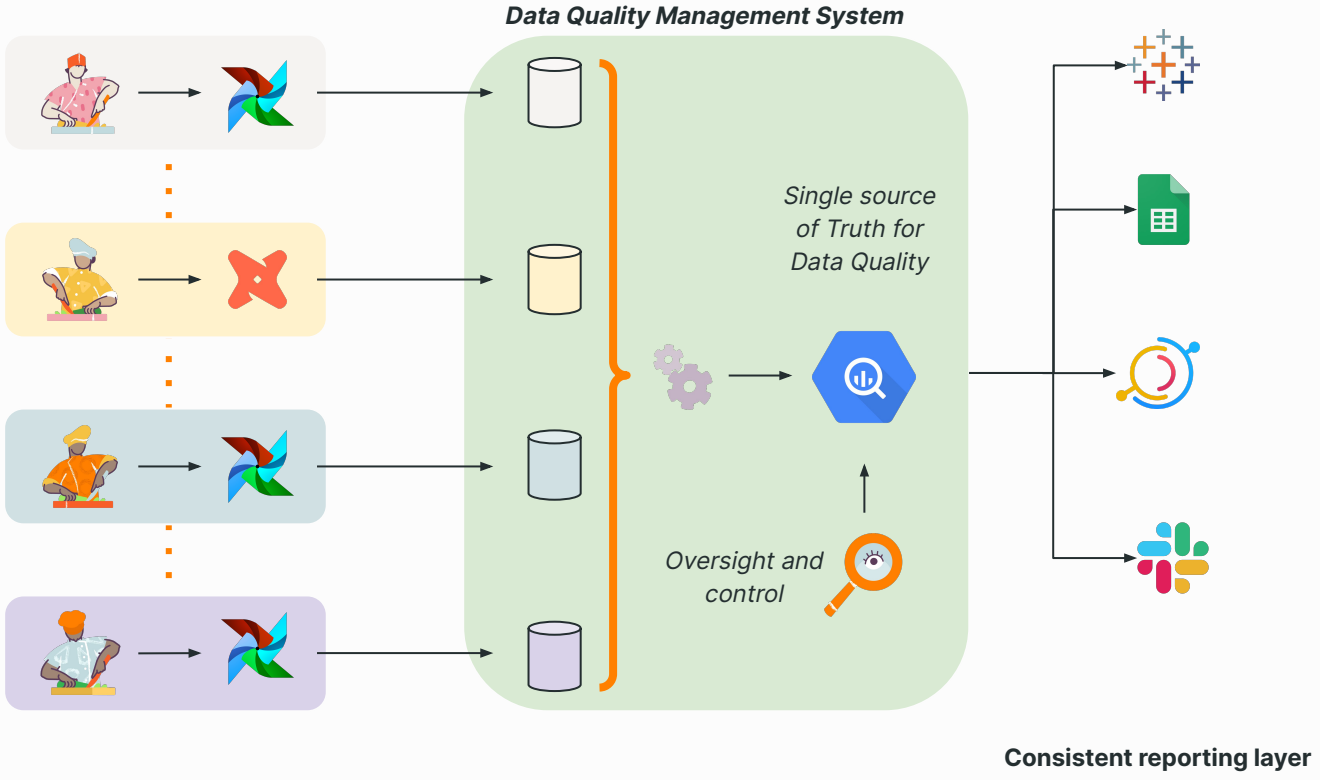
1 Controls
2 Control results
3 Teams
4 Measurements
5 Dimensions
6 Recommendations
7 Jira issues

Pending!



This framework is the catalyst for our next-gen data quality journey

- Defined **DQ measurements** linked to 5 dimensions for tracking and identifying gaps
- Set **standards** for Data Quality, including **metadata**
- Provided **training material**



End of Chapter 4: Dinner is served!

Chapter 5: Who gets what

The Data Access Framework (DAF) requirements are determined by the internal policies and JET's tech landscape

Inputs

- **Data Classification Policy**
- **Data Owner Guidelines**
- **Infrastructure-as-code**

Data Classification

Framework for classifying data based on its sensitivity, value and critically to JET

Data Owner guidelines

Outlining the accountabilities of Data Ownership, including **granting access rights to data**

GCP Infrastructure

JET is using Infrastructure-as-code approach, managing access on dataset level in Github

Unpacking the Data Classification Policy

Public

Access is provided to all individuals working at JET

Internal

Access is provided to all individuals working at JET

Restricted

Access is provided on a role base automatically

Confidential

Access is provided on a role base with a valid business reason

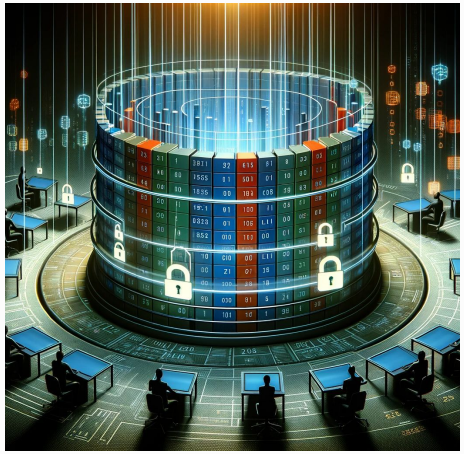
Rolling out the DAF in 3 Phases will ensure its success

Phase 01

Dataset/Bucket-level access

Phase 02

Column-level access



Phase 03

Row-level access

End of Chapter 5: Food for everyone

Chapter 6: All the world's an oasis

Data Owners Accountabilities

Data Access
Managing and reviewing access rights

Data Quality
Ensuring data is of sufficient high quality to meet the expectations of their consumers

Metadata
Maintenance of up-to-date metadata

Cost
Understanding the impact on cost of their data

Data Retention
Ensuring data retention periods are met

Data Labeling
Using tags to label data appropriately

Data Owners Accountabilities (1)

01

Data Access

Managing access permissions to their data, which involves **classifying** their data based on the Data Classification Policy and **assessing** users' business needs to ensure access is provided on a 'need-to-know' basis.

The **Data Access Framework** will provide Data Owners the capabilities to achieve this.

02

Data Quality

Ensuring that data is of high quality to the degree that it **meets the expectations** and needs of data consumers. That is, if the data is **fit for the purposes** to which the consumers want to apply it.

The **Data Quality Framework** will provide Data Owners the tooling and best practices to realize this.

03

Metadata

Updating and accurately maintaining the metadata for their data, that involves ensuring the **descriptive information** about the data is current and precise, thereby enhancing its discoverability and context.

Datahub will provide Data Owners with the platform to oversee and share their metadata.

Data Owners Accountabilities (2)

04

Cost

Comprehending the **financial impact** related to the storage and processing of their data, including optimizing these costs by implementing effective data management strategies.

The **Data Platforms** teams provide Data Owners the capabilities to meet these requirements.

05

Data Retention

Upholding **compliance** with the Data Retention Policy, which includes adhering to prescribed data retention periods by ensuring the timely disposal of data once the retention period concludes.

The **Data Platforms** teams provide Data Owners the capabilities to meet these requirements.

06

Data Labeling

Labeling their data in compliance with specific policy requirements, which entails **accurate classification** and **tagging** of data to reflect its nature, sensitivity, and usage, thus enabling effective management and appropriate security controls.

The **Data Platforms** teams provide Data Owners the capabilities to meet these requirements.

End of Chapter 6: We've built ourselves a thriving village!

Chapter 7: *Where we're going we don't need roads*

AI Specific Risks

Accuracy and quality

Inaccurate predictions

Unreliable outcomes

Hallucinations

Ethics and fairness

Biases and discrimination

Denial of equal opportunity

Unethical and unlawful behaviour / misuse

Explainability and visibility

Could result insufficient transparency and lack of accountability

Unpredictable, unexplainable consequences or behaviours

Data Protection

Compliance challenges due to lack of visibility and the nature of the technology

Difficulties to ascertain user expectations and the impact on the individual

Intellectual property

Content ownership

Training data

Challenges with public data

Open source components

AI Vendor Risks

Insufficient knowledge and audit possibilities

Possible data / AI breaches

Unlawful data processing/ sharing

Third parties claims

Regulatory effects

Security and safety

Direct interaction with end-users

Data management and access control challenges

Vulnerabilities and exposure to hacking and attacks

Impact on humans, society and the environment

Principles for fair, responsible and safe use with AI

Main Principles

- Accountability
- Transparency & visibility
- Explainability & interpretability
- Fairness & ethical standards
- Data quality
- Data Protection
- Safety and Information Security

Method

- Secured operation environment
- Review and approve AI Tools & Use cases
- Testing & verification

  Including biases, non-discrimination and errors.
- Privacy and Security by Design

Responsible and Safe Use of AI

How AI Risk Assessment (AIRA) works at JET?

Business use case

Business has a use case, and completes AIRA questions

Joint Review

AIRA is jointly reviewed by relevant stakeholders, for example InfoSec, DPO Office and Data Governance teams.

Risks & Mitigation Actions

A first draft of the risks and mitigation actions is shared with the business.

Sign off

The risks and mitigation actions need to be signed off based on the risk level.

Support

InfoSec, DPO Office and Data Governance teams support the business with the implementation of the identified mitigation actions.

End of Chapter 7: Bringing value to all

Truth is often stranger than fiction

Presented in a very linear fashion for narrative

Truth is many initiatives working in parallel

Data Governance as a breathing mechanism, not a choking one

Unlocking the potential of data to drive growth, innovation, and competitive advantage in a structured and secure manner

Example Use Case - Data Analyst

Situation: I'm a *Data Analyst* and want to create a new dashboard

Guidelines:

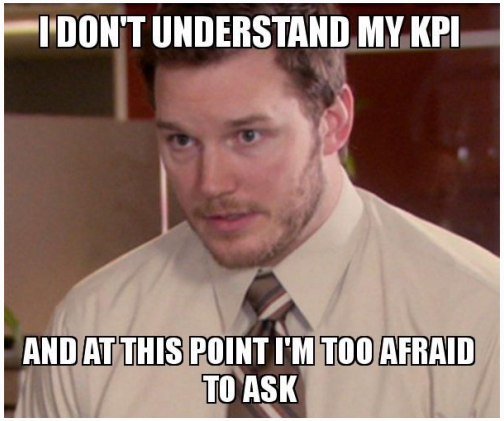
1. Verify from **Datahub** that a similar dashboard does not already exist
2. Use the **Data Access Framework** to correctly tag the dashboard and identify the stakeholders who should have access
3. Once published, ensure that the metadata is correct and uploaded to **Datahub**

Example Use Case - Business User

Situation: I'm a *Business User* and want to understand what a KPI means, and where to consume it from

Guidelines:

1. Search for the KPI in **Datahub**'s Business Glossary, and find the physical data asset of that KPI
2. Use the **Data Quality Framework** to understand the level of quality of that KPI
3. Use **Datahub**'s lineage functionality to identify datasets/reports that use that KPI
4. Go through steps in **Data Access Framework** to access the data or reports



Example Use Case - Data Engineer

Situation: I'm a *Data Engineer* and want to build a dataset with a new KPI

Guidelines:

1. Understand from **Datahub** the definition of the new KPI, that should have been published in the Business Glossary
2. Create a new dataset as per the requirements from the **Data Access Framework**
3. Implement data quality validations on the dataset that are aligned with the expectations of the data consumers and that feed into the **Data Quality Framework**
4. Review the lineage for the KPI in **Datahub**, ensure it reflects the definition from the Business Glossary

Thank you

Thank you!!

David Fernandez

Global Data Governance Manager

JUST EAT Takeaway.com

Let's connect here

Kirill Skaletskiy

Global Senior Data Quality Professional

JUST EAT Takeaway.com

Let's connect here

DATA GOVERNANCE AND MASTER DATA MANAGEMENT CONFERENCE EUROPE

11 - 14 March 2024 |London, UK

Please score and comment on this session and speaker in the event mobile app

IRMUK