



**Master Data Management &
Data Governance Conference
Europe**

**Master Data Management &
Data Governance Conference Europe**

9 - 12 May 2023, London

***Please Score and Comment on this Session and
Speak Via the Conference App***



Data Governance for Transactional Data: The Data Police, or a Data Service?

Leonor Van Der Beek
Head of Data Governance, Tata Steel Netherlands

Pedro Pinho
Managing Partner, Apgar

What we do.

Gartner

APGAR is identified by Gartner as **1 of the 18 major MDM External Services Providers** in the 2021 MDM Market Guide for the second year in a row.

APGAR is listed in Gartner's Toolkit for "Vendor Identification and Selection Guide for **AI and Data and Analytics (D&A) Service Providers**"

Data Advisory.



- Data Literacy
- **Data Governance**
- Data Architecture
- Data for Green

Data Enforcement.



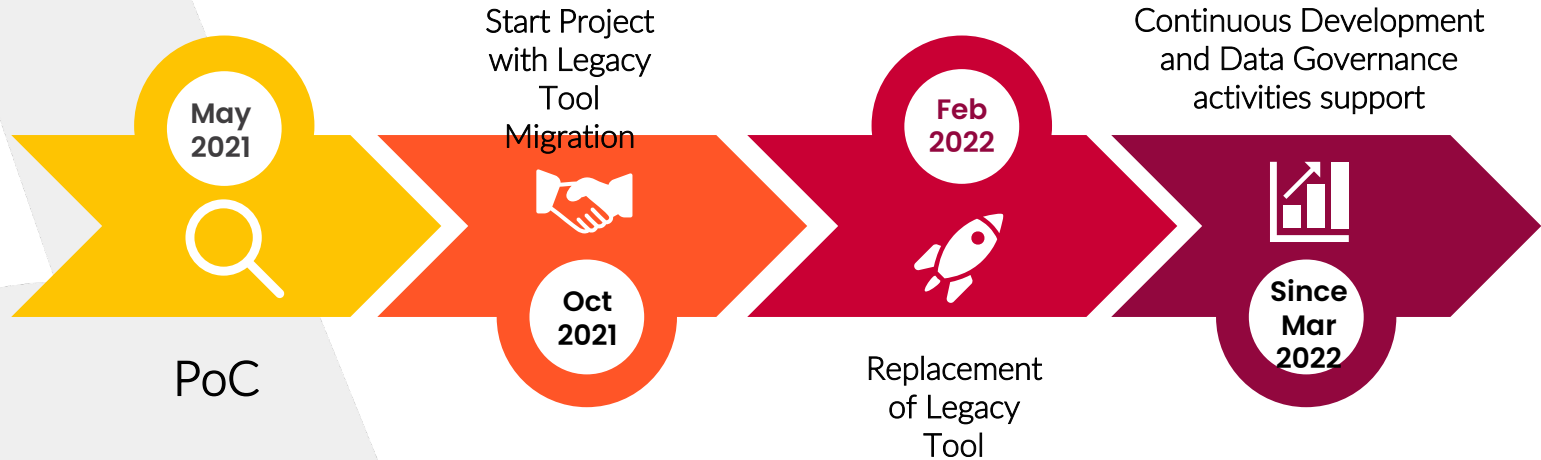
- **Metadata Management**
- Master Data Management
- Data Integration
- Data Preparation
- Advanced Analytics
- Data Ops
- Production Support

Data Acceleration.



- MDM for Finance
- MDM for Business Partner
- Reference Data Management
- **Data Catalog**
- Data Privacy Management
- AI Driven Operations

Our journey.



TATA STEEL



Data governance for transactional data: the data police, or a data service?

Leonoor van der Beek,

Head of Data Governance
Tata Steel Netherlands

DATA
DRIVEN **STEEL**

Agenda

- Data @ Tata Steel Netherlands
- How to position Data Governance
- Effect of positioning on
 - Data Governance roles & responsibilities
 - Policies and processes
 - Tooling
- Wrap up



~7 mln t

Annual production of
steel products



9k

Employees



€3,5 bln

Annual net turnover



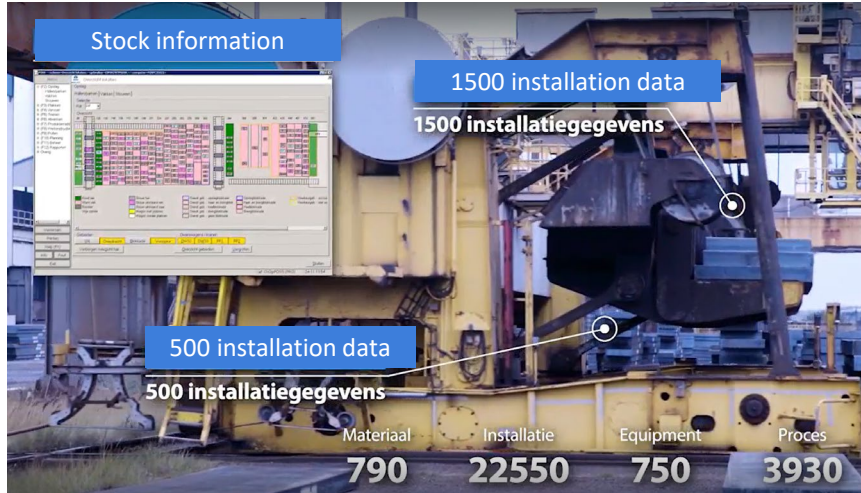
750 ha

Business area

Welcome to Tata Steel in IJmuiden

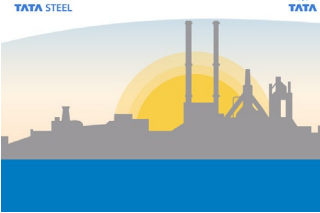


Data at the foundation of operation & transformation



Data at the foundation of operation & transformation

Sustainable profit transformation



2016

Vision for Advanced Analytics

FOR DISCUSSION

worldsteel
ASSOCIATION

WHAT'S NEW

1 April 2019

Press release

Tata Steel IJmuiden awarded leading advanced analytics steel plant in the world

Today, Tata Steel IJmuiden received the WSA award for the leading advanced analytics enabled steel plant in the world. Hub director Hans vd Berg comments: "as leadership team we embraced the opportunities of advanced analytics and digital early on. Over the past years we build a local Tech community, update our IT architecture and drove an holistic program deploying AA on topics like process and quality control, predictive maintenance, procurement of raw materials and planning optimization"

2017

WEF recognition

fd. Mijn laatste Krant Meer ▾

nieuws nieuws

Heiko Jessayan 2 min

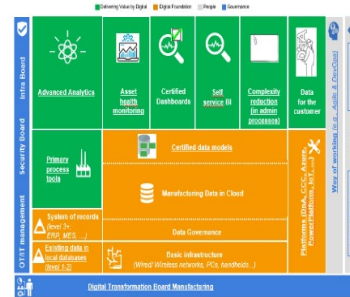
INDUSTRIE

World Economic Forum roemt Tata Steel IJmuiden om industriële data-analyse

Tata Steel in IJmuiden is door het World Economic Forum (WEF) verkozen tot voorbeeldbedrijf voor de toepassing van zogeheten advanced analytics, een complexe vorm van data-analyse. Tata gebruikt de advanced analytics om de eigen industriële processen te verbeteren.

2019

Digital Roadmap



2021

Data Driven Steel



2022

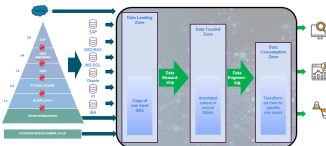
Data Governance developed along with digital

Local Data Governance



2016

Data lake becomes 1st central DG initiative



2018

Central DG in Data Virtualisation



2020

2000th unique table governed

<p>2036 Tables</p> <p>Total number of tables registered</p>	<p>70 Data Owners</p> <p>Users identified as Data owners.</p>
<p>133 Data Stewards</p> <p>Users identified as Data Stewards.</p>	<p>149 Projects</p> <p>Use cases registered</p>

2022

Start of central TSN-wide DG



2023

Data Governance: the data police or a data service?

Data Police



- Top Down
- Compliance and/or risk oriented
- DG tasks & responsibilities embedded in functions
- Policy centered

Data Service



- Use case driven
- Innovation & insight oriented
- DG tasks & responsibilities embedded in roles
- Value centered

Irene

BI-specialist

“Easy access to reporting data is key to me”



Johan

sr. manager

“I don't have time to discuss the data – I need to be able to trust them blindly.”



Steven

finance business partner

• “The data just needs to be there and be complete”



Gregory

process technologist

• “Every day new and urgent questions”



Joyce

information manager

• “Innovation should never compromise stability.”



Misha

AA-specialist

• “I'd like to be as self-sufficient as possible”



- “Easy access to reporting data is key to me”



Irene - BI-specialist

Irene, age 32, works at the Hot Strip Mill. The unit's management team counts on her for reliable, timely reports and dashboards to assess performance of the plant and to zoom in on potential issues. Irene knows the local data systems very well. Both Irene and her “customers” have their own favourite BI tool – not necessarily the same. Some of her dashboards are used during “ochtendwijdning” or displayed on screens in the control room. She needs to know how her data is provided. While many reports are HSM specific, they often require data from up/downstream plants or external sources. ExCo would like her to align her reports with those of other plants, but there are so many HSM specific aspects, that Irene doubts that's realistic.

TOOLS

Aqar DG,
PowerBI, Spotfire
SAP, RDBMSs
Excel, TDV

SKILLS

Visualisation
SQL
Data modeling

DATA CONSUMPTION

TDV; MESBI
DB2
Data lake: process data
SAP
Local databases

INFORMATION GENERATION

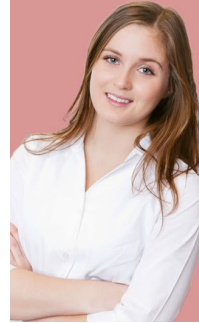
PowerBI reports
Spotfire reports
Custom tables

CORE NEEDS

Find and understand data
Combine data from different sources
High quality data
Easy access to data from BI tool
Performance, availability, timeliness of data
Insight whether data from other plants is calculated the same way

Who is this persona?

- “Easy access to reporting data is key to me”



Irene - BI-specialist

WHAT IS THERE FOR ME

- Searching & inspecting data & data products
 - ADC is the one stop-shop to find TSN data & data products
 - ADC offers explanations of what data(products) are and where to find them
 - ADC provides insight in the lineage of the (master) data and its relations with other data
 - ADC provides insight in characteristics & quality of the data
- Data & data access
 - There is one, clear, process for requesting data access, no matter where the data resides
 - Once I am authorized to access certain data, I can do this no matter what tools I use
 - From my reporting tool, I can access all the data sources I need – or request them to be made available: This includes central data sources (e.g. data lake) and local data sources (e.g. data warehouses)
 - There is a clear and consistent process for requesting data to be made available, no matter what the source of the data is, or where I would like the data to be made available
- Creating custom dashboards
 - I can prepare datasets for optimal use in certain applications. Data can be merged/filtered/transformed to fit the purpose.
- Creating reports/dashboards
 - There are clear guidelines for how to create, maintain, and document data products

- There is a clear process for how to certify a dashboard or report when relevant
- It is easy for others to find and request access to my reports, and to see whether they are certified
- Training
 - Data Governance introduction
 - Citizen Developer Security Awareness
 - Data analyst training
 - TDV training
 - Information and Report design
 - Spotfire training / PowerBI training

HOW I BEHAVE

- As a report or dashboard (product) owner, I take responsibility for its quality and documentation, I make sure I have all the required knowledge, and I manage changes/maintain it well
- I certify my data products where needed
- I share data, reports, and dashboards only in compliance with their classification
- I provide descriptions and documentation for the custom tables or views that I create, and for new source data requested by me

What do we offer to & expect from this persona?

What does your service offer?

The screenshot displays the TSMLE Data Governance web application. The left sidebar contains navigation options: Home, Data Catalog, Tables, Projects, Data Access, and Admin. The main content area is titled "Bucket Ore Preparation Detailed Information" and includes a sub-header "Confidential-Restricted data from the ore preparation process". A "Edit Bucket" button is visible at the top. Below this, there are tabs for "Tables", "Data Owners", "Members", and "AD Groups Logs", with "Members" currently selected. A disclaimer states: "For maintenance and bug fixing purposes, administrators of the underlying infrastructure may also have access to confidential data." A search bar is present above a table listing members and their access types.

ID	Name	Access Type
A153429	Henry van den	Consumer
A166892	Christel	Consumer Data Engineer
A167437	Menno	Consumer Data Engineer
A169888	Beek, Leonor van der	Data Lake Admin
A177956	Nelson	Data Lake Admin
A179986	Emre	Data Lake Admin
A181891	Karin van der	Data Lake Admin
A189951	Vincent	Data Lake Admin
A205501	.ars	Consumer
A208694	Mohamed	Data Lake Admin
A290627	.Alexander	Consumer
A295923	.Selish	Exploration
A342306	.Awad	Data Lake Admin
A348163	.Borja	Data Lake Admin
A408799	Erik van den	Consumer
A412635	Niels	Consumer
A427591	.Martien	Consumer Data Engineer

- Data security & insight in data access/use

What does your service offer?

Request access to Restricted Enterprise Data



Select one of the buckets listed below:

Bucket Name	Description
<input type="radio"/> Blast Furnaces	Confidential-Restricted data from the Blast Furnaces
<input type="radio"/> Central Maintenance Sourcing	Confidential-Restricted data from the Central Maintenance Sourcing group
<input type="radio"/> Certificate Monitor	Confidential-Restricted output data of the certificate monitor
<input type="radio"/> Coke Quality	Confidential-Restricted output tables for the R&D project concerning Coke Quality
<input type="radio"/> Commercial	Confidential-Restricted, customer/account related data.
<input type="radio"/> Commercial DnA	Restricted data from CDnA which is made available through MDC and/or TDV
<input type="radio"/> DnA	Bucket for Restricted DnA internal (admin) data, such as log data.

- Data security & insight in data access/use
- One place to request access to any data set

What does your service offer?



My Inbox ⁰



Data Catalog



Start searching on the Tata Steel Data Catalog

The data catalog is an enterprise information resource providing visibility into the many data assets and data elements which are made available for (limited) sharing on our IT platforms. New data is being added every week, so check back often to see if the dataset you are looking for is already available.

Enter a keyword or click a category below. Use * for wildcard search.



Object

Data objects are e.g. tables or collections of sensor tags. You can search for technical table names or words in the title and description.

Business Term

Business terms are definitions or business descriptions of fields. They list all related fields, often in various tables in multiple zones/applications.

Field

Fields are e.g. table columns or sensor tags. You can search on the technical name of the column/tag, or on words describing the data.

Data Entity

Data Entities are business metadata associated to objects. They list objects in e.g. Landing and Trusted Zone or Databricks and TDV.

- Data security & insight in data access/use
- One place to request access to any data set
- Find any data available – and what it means
- Insight in the data itself
- Overview of data quality
- ...

Non-pervasive policies & processes (where possible)

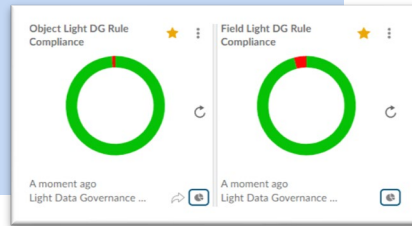
Data Governance Light

- Requirements
 - Data Owner
 - Data sensitivity classification
- When
 - Always

vs

Full Data Governance

- Requirements
 - Data Governance Light
 - Data Steward
 - All mandatory metadata table
 - All mandatory metadata column
- When
 - All data shared for use in data products in PRD



Support it with the tools your users need (look-up)

Business View

Categories: Object, Application, Business Term, Field, Resource, Container, Resource Cluster, Data Entity

Field Filters: Object, Container, Application, Resource (MDC data lake - Public, Co), Source System, Data Zone (Trusted Zone), Data Seward, Data Owner, Data Sensitivity

Search: Field (3081 results)

Field Cards:

- Date end stringing cycle**: Description: Date time end of stringing cycle. Different settings can be used during one cycle. Last Update Date: 2023-02-24 09:29:29. Technical Name: dt_start_spoed. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Date data insert**: Description: Date data registration in table Time stamps in DRB are always in local time, so CET in winter or CEST in summer. When... Last Update Date: 2023-02-24 09:29:29. Technical Name: dt_insert. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Date start treatment**: Description: Start of treatment according to administration. This means that at least the operator has accepted the request of the system. Last Update Date: 2023-02-24 09:29:29. Technical Name: dt_start_treat. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Date start stringing cycle**: Description: Date time start of stringing cycle. Different settings can be used during one cycle. Last Update Date: 2023-02-24 09:29:29. Technical Name: dt_start_spoed. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Gas type**: Description: Type of gas (argon or nitrogen). Last Update Date: 2023-02-24 09:29:29. Technical Name: gas_type. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Gas quantity**: Description: Amount of gas added per stringing cycle. Last Update Date: 2023-02-24 09:29:29. Technical Name: gas_quantity. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Plug performance 1**: Description: Stringing performance plug 1. Indicator given by operator at end of treatment. Three most used ratings: good (good), matp. Last Update Date: 2023-02-24 09:29:29. Technical Name: mt_working_plug1. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Plug performance 2**: Description: Stringing performance plug 2. Indicator given by operator at end of treatment. Three most used ratings: good (good), matp. Last Update Date: 2023-02-24 09:29:29. Technical Name: mt_working_plug2. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Installation ID ladle treatment**: Description: Type of installation (indicated by two letters) and specific installation and location indicated by two numbers. Last Update Date: 2023-02-24 09:29:29. Technical Name: installation_id. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Heat number large**: Description: Large sequence number allocated to the steel. Last Update Date: 2023-02-24 09:29:29. Technical Name: heat_number_large. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Method**: Description: Details per stringing cycle, such as used porous plug or method. Last Update Date: 2023-02-24 09:29:29. Technical Name: method. Object Technical Name: dt_rs_rs_pbdb_spoed.
- Type of stringing**: Description: Two types of stringing: regular stringing by porous plugs or stringing. Last Update Date: 2023-02-24 09:29:29. Technical Name: type_of_stringing. Object Technical Name: dt_rs_rs_pbdb_spoed.

Display On Business View: True, False, All

Classic View

Classic View

Table Columns: Object Name, Object Technical Name, Technical Name, Name, Description, Main

Table Rows:

- Object Name: 99011_2019, Object Technical Name: aant_uit_brdf_os, Technical Name: aant_uit_brdf_os, Name: Number of strips over width of, Description: Number of strips over width of, Main: Num
- Object Name: 99011_2019, Object Technical Name: gkrmetechard_os, Technical Name: gkrmetechard_os, Name: Side trimming Indicator, Description: Indication of whether a coil has, Main: Side
- Object Name: 99011_2019, Object Technical Name: installat_os_yrbout, Technical Name: installat_os_yrbout, Name: OS installation number, Description: Numeric part of installation no. OS is, Main: OS
- Object Name: 99011_2019, Object Technical Name: installat_os, Technical Name: installat_os, Name: OS Installation type, Description: Alphabetic part of the installat, OS is, Main: OS
- Object Name: 99011_2019, Object Technical Name: rot_ident, Technical Name: rot_ident, Name: Coil identifier, Description: Coil identification number, Main: Coil
- Object Name: 99011_2019, Object Technical Name: ts_bew, Technical Name: ts_bew, Name: Timestamp of reporting a proc, Description: Timestamp of receiving messag, Time, Main: Time
- Object Name: 99011_2019, Object Technical Name: ts_out_os, Technical Name: ts_out_os, Name: Timestamp physical output Rec, Description: Timestamp physical output Rec, Time, Main: Time
- Object Name: 99011_2019, Object Technical Name: verkinging_os, Technical Name: verkinging_os, Name: Elongation percentage coil by s, Description: Elongation in percent of the co, Elong, Main: Elong
- Object Name: 99011_2019, Object Technical Name: versieun_rnl, Technical Name: versieun_rnl, Name: Version number of the coil in e, Description: Version number of coil; e.g. no, Veru, Main: Veru
- Object Name: 99011_2019, Object Technical Name: v1076, Technical Name: v1076, Name: Additional properties on conRt, Description: List of extra properties that are, Add, Main: Add
- Object Name: 99011_2019, Object Technical Name: ps_kl_nummer, Technical Name: ps_kl_nummer, Name: Unique grey identifier Identif, Description: Unique grey identifier sequent, Uniq, Main: Uniq
- Object Name: 99011_2019, Object Technical Name: ps_kl_versieun, Technical Name: ps_kl_versieun, Name: Unique grey identifier for a ver, Description: Unique grey identifier sequent, Uniq, Main: Uniq
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Count of defect class 21, Description: Total count of defects of class: 2, Count, Main: Count
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Count of defect class 22, Description: Total count of defects of class: 2, Count, Main: Count
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Count of defect class 24, Description: Total count of defects of class: 2, Count, Main: Count
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Count of defect class 946, Description: Total count of defects of class: 1, Count, Main: Count
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Count of defect class 947, Description: Total count of defects of class: 1, Count, Main: Count
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Hills & Mueller coil identificat, Description: Hills & Mueller coil identificat, Hills, Main: Hills
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Coil Length Parameter, Description: Length of the coil material, len, Coil, Main: Coil
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Incoming coil identification, Description: Identification number of the coil, Inco, Main: Inco
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Coil side, Description: Coil side Bottom = 1, Top = 0, Coil, Main: Coil
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: 1st most dominant repeating di, Description: Most dominant repeating dist, 1st d, Main: 1st d
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: 2nd most dominant repeating di, Description: Second dominant repeating dist, 2nd d, Main: 2nd d
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: 3rd most dominant repeating di, Description: 3th dominant repeating distan, 3rd d, Main: 3rd d
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: 4th most dominant repeating di, Description: 4th dominant repeating distan, 4th d, Main: 4th d
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: The number of the coil and th, Description: The number of the coil and th, Param, Main: Param
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Coil Thickness Parameter, Description: Thickness of the coil material, l, Coil, Main: Coil
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Start process date and time, Description: Date and time of begin of proc, Start, Main: Start
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Coil Width Parameter, Description: Coil width as measured by Part, Coil, Main: Coil
- Object Name: 99011_2019, Object Technical Name: m3_beatstratches_gc, Technical Name: m3_beatstratches_gc, Name: Hot metal ladle target weight, Description: Target weight of hot metal in d, Hot, Main: Hot

Search and Filter Options: Search, Add a criterion, MDC LZ, MDC TZ, All criteria match, Resource, Data Zone, Trusted Zone, Add a criterion, TSV PRD Publication, month attributes

- Different types of users have different requirements & UX wishes
- One size fits all does not exist



Support it with the tools your users need (edit)

Manual web form updates

Search > Field > change_from_bd_reason > Update Field

Field: change_from_bd_reason

Actions +

CRUD Data Terms Data Set Related Document Source Fields Target Fields Tracking

1 error

Identification

Technical Name	change_from_bd_reason	Data Zone	Trusted Zone
Name	change_from_bd_reason	Container	tu_manufacturing_mhc_res_fr
Description		Resource	PRD_DATABRICKS_MDC_RES
Client Technical Name	v30_complant_view	Installation	
Client Name	v30_complant_view	Typology	[not defined]
Client Path	252_manufacturing_mhc_res_fr_fr_complant	Data Processing Activity	[not defined]
Translation	No occurrence defined.	Dashboard URL	
Class Type	[not defined]		
Main Business Term	[not defined]		
Main Data Entity	[not defined]		
Input Type	[not defined]		
Application	PRD_DATABRICKS_MDC_RES		
Source System	Oracle		

Governance

Data Owner	1. nicole.verhoef@tatateel.com - Verhoef - Nicole
Data Steward	1. nanda.van-der-ven@tatateel.com - Van-der-ven - Nanda

Save Revert

Excel based batch imports

ID	Name	Description	...
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			

Self-service batch annotation

https://datacatalog.eu.tatasteel.com/ebx-ui/ui/custom/ebx-perspective-Home/action/27

TATA STEEL

Show Message > Display message

Cancel Submit for validation

No errors were found while processing your draft annotations. Please submit for validation.

- Different types of users have different requirements & UX wishes
- One size fits all does not exist



Recap: Data Governance - the data police or a data service?

- There is more than 1 way to position DG within an organization.
- The choice you make impacts all aspects of DG
- Which positioning to choose depends on the context (organization, data, maturity)

Data Police



- Top Down
- Compliance and/or risk oriented
- DG tasks & responsibilities embedded in functions
- Policy centered

Data Service



- Use case driven
- Innovation & insight oriented
- DG tasks & responsibilities embedded in roles
- Value centered

Thanks.



Leonoor van der Beek
Head of Tata Steel Netherlands

<https://www.linkedin.com/in/leonoorvanderbeek/>



Pedro Pinho
Managing Partner Apgar

<https://www.linkedin.com/in/pedromariapinho/>



**Master Data Management &
Data Governance Conference
Europe**

**Master Data Management &
Data Governance Conference Europe**

9 - 12 May 2023, London

***Please Score and Comment on this Session and
Speak Via the Conference App***