

Enabling Data Governance with Limited Resources Using AI & Automation

Webinar

David Kolinek, Product Manager at Ataccama



Agenda

01 Typical Data governance initiative

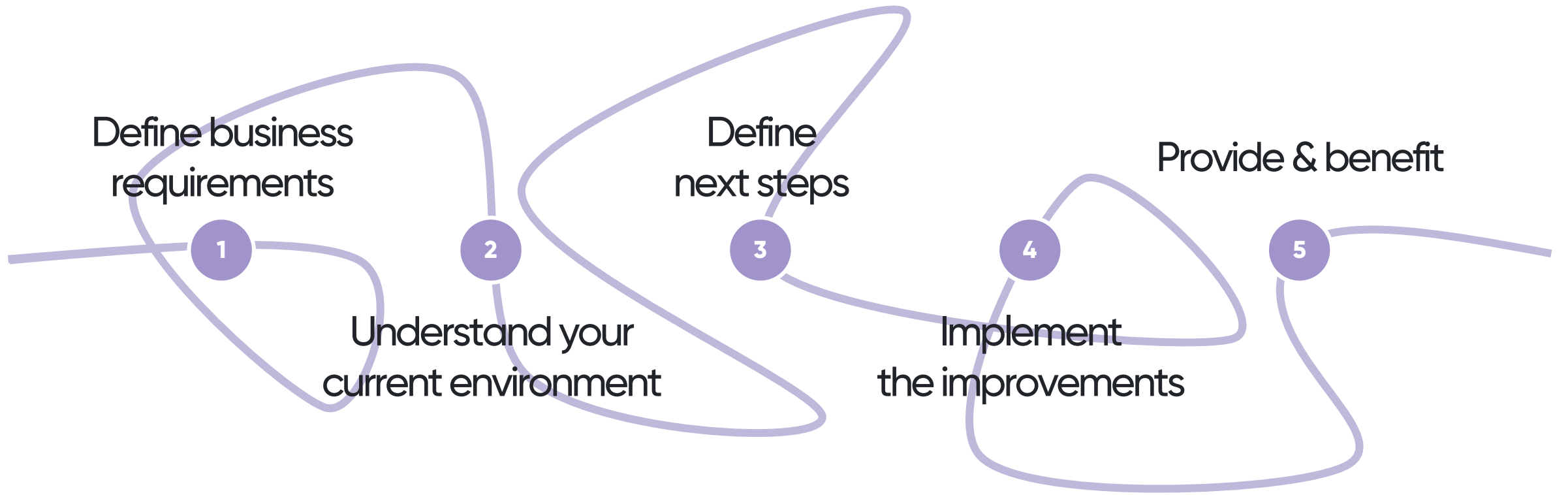
02 Self-driving Data governance initiative

03 Ataccama ONE Walk-through

04 Q&A

Typical

Data governance initiative

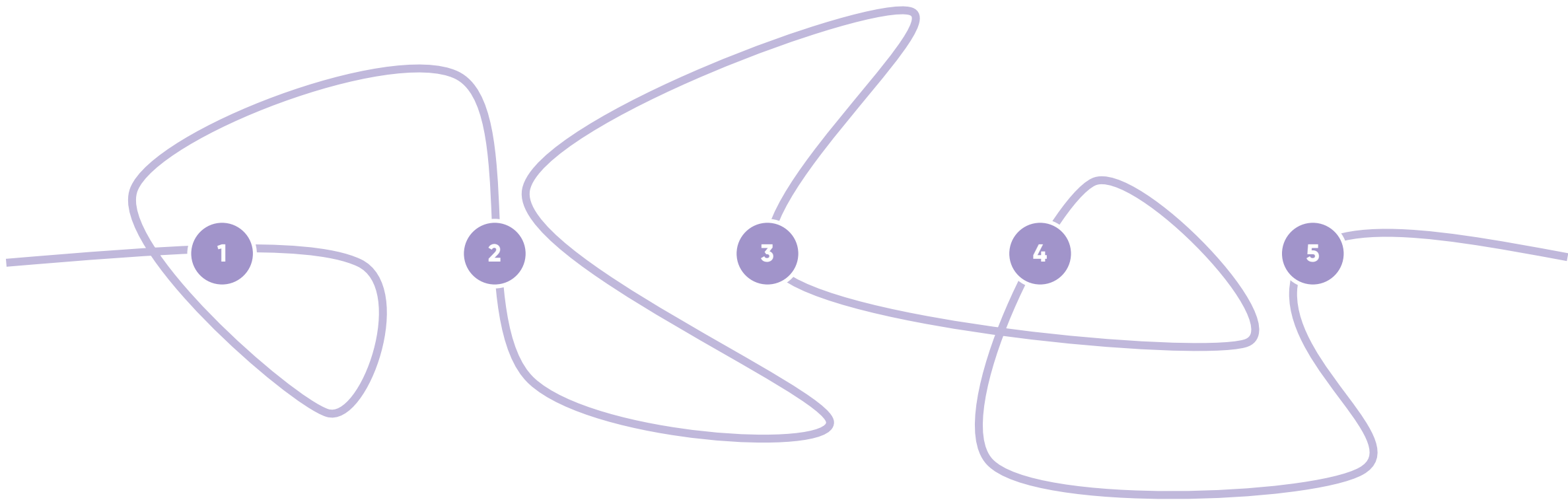


Issues

- › Some owners are not defined
- › Very little metadata is available
- › Different understanding of the same term across multiple teams
- › All previous initiatives failed as no one had energy to keep the metadata up to date
- › And much more...

Self-driving

Data governance initiative





- › Objectives
- › Different policies
- › Business rules
- › ...



Define business requirements

Understand your current environment

Define next steps

Implement the improvements

Provide & benefit

- › Import metadata from all existing sources into one place
- › Automate data classification
- › Use data profiling to capture additional metadata
- › Capture the data lineage and other relationships between data assets
- › Check the data quality for data domains, not only specific assets
- › Complement the rule-based approach with a pattern-based one

Define business requirements

Understand your current environment

Define next steps

Implement the improvements

Provide & benefit

- › Detect gaps in your environment understanding
- › Use metadata for detecting the largest issues
- › Revise platform recommendations on what to do next
- › Prioritize based on business objectives

Define business requirements

Understand your current environment

Define next steps

Implement the improvements

Provide & benefit

- › Leverage metadata for automation
- › Provide a self-service interface to business users

Define business requirements

Understand your current environment

Define next steps

Implement the improvements

Provide & benefit

- › Share the knowledge base you just created with everyone in the organization
- › A data provisioning solution for everyone

Define business requirements

Understand your current environment

Define next steps

Implement the improvements

Provide & benefit

Ataccama ONE Walk-through



Categories

Add new category

POLICY MANAGEMENT

Regulations

Categories

All policies

<input type="checkbox"/>	Name	Last enforced
<input type="checkbox"/>	> Data Protection	20/01/2020
<input type="checkbox"/>	> Data Quality	20/01/2020
<input type="checkbox"/>	> Risk assessment	20/01/2020
<input type="checkbox"/>	> Basel Accords	20/01/2020
<input type="checkbox"/>	> Employee welfare	20/01/2020





Add new regulation

POLICY MANAGEMENT

Regulations

Categories

All policies

Filter regulations or policies

<input type="checkbox"/>	Name	Last enforced
<input type="checkbox"/>	GDPR	20/01/2020
<input type="checkbox"/>	Restrict access to data for external users	20/01/2020
<input type="checkbox"/>	Restrict access to personal data	20/01/2020
<input type="checkbox"/>	Restrict access to SPI data	20/01/2020
<input type="checkbox"/>	Protect employees data	20/01/2020
<input type="checkbox"/>	Protect customers data	20/01/2020
<input type="checkbox"/>	Protect financial data	20/01/2020
<input type="checkbox"/>	Restrict access to data for external users	20/01/2020
<input type="checkbox"/>	Restrict access to PII data	20/01/2020
<input type="checkbox"/>	CCPA	20/01/2020
<input type="checkbox"/>	The Privacy Act 1988	20/01/2020
<input type="checkbox"/>	PIPA	20/01/2020
<input type="checkbox"/>	Bank Secrecy Act (BSA)	20/01/2020
<input type="checkbox"/>	USA PATRIOT Act	20/01/2020





Regulations

Categories

All policies

Description

Personally identifiable information (PII) is data which can be used to identify, locate, or contact an individual and includes information like name, date of birth, place of residence, credit card information, phone number, race, gender, criminal record, age, and medical records. Every organization stores and uses PII, be it information on their employees or customers. Even schools and universities will store the PII of their students, while hospitals will store patient data.

Restrict access

Message 'Assets are hidden because of the GDPR regulation'

WHEN

Asset type is any asset

AND

Term assigned to this asset is any of PII SPI Sensitive

APPLY FOR

Everyone except Risk or Mark Brown

Policy enforcement

Last time enforced	Total
13/01/2020	69

Regulations

GDPR
Regulation on the protection of natural persons with regard to the processing of personal data and on the fr...

13 policies

CCPA
CCPA defines personal information as information that identifies, relates to, describes, is reasonably capable of being associated with, or could ...

2 policies

+ Add to regulation



Create data source

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Type a name, owner, date

<input type="checkbox"/>	Name	Access	Credentials	Terms	Owner	Last import	
<input type="checkbox"/>	CRM			Product Customer GDPR	Lucinda Ramsey	2 weeks ago	
<input type="checkbox"/>	Politic party full			Product Customer	Anna Fali	23. 4. 2018	
<input type="checkbox"/>	Credit cards			Customer Products related data	Sergio Pliego	a week ago	
<input type="checkbox"/>	One MDM portal			Product Customer	Gibby Radki	23. 4. 2018	



Basic information

A business perspective of the data source. It can refer to business purpose, department, location or to a particular user group. It should capture the organization of data sources in your company.

Name your data source *

Describe purpose from the business perspective

Connection

Represents a physical connection to the data source e.g. a database or a filesystem from which you can populate your Data catalog. A data source can consist of more connections.

 Create Another


Basic information

A business perspective of the data source. It can refer to business purpose, department, location or to a particular user group. It should capture the organization of data sources in your company.

Name your data source *

Describe purpose from the business perspective


Connection



MSSQL

jdbc:mssql://ata.production.sheldon.com/erp:1527/<jdbcDatabaseName>

Admin (default)
Team_Marketing



+ Add Connection

Create & Document ▾

Create Another

Basic information

A business perspective of the data source. It can refer to business purpose, department, location or to a particular user group. It should capture the organization of data sources in your company.

Name your data source *

MSSQL backup

Create

Create new data source without any follow-up action, which can be done manually later.

Create & Import

Import all metadata to the Data Catalog without accessing any data.

Create & Discover

Quick, perfect for the data discovery. Runs on the sample of the data.

Create & Document

Import metadata, run quick data discovery and efficiently profile and validate data quality of all relevant assets.

[Empty text input field]

527/<jdbcDatabaseName>

Create & Document ^

Create Another



DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

- Overview
- Catalog Items
- Data Quality
- Data Browser
- Connection Details
- History
- Access

Catalog Items

1 343

↗ 45

Schemas

12

↗ 1

Tables

1 001

↗ 46

Views

342

↘ 1

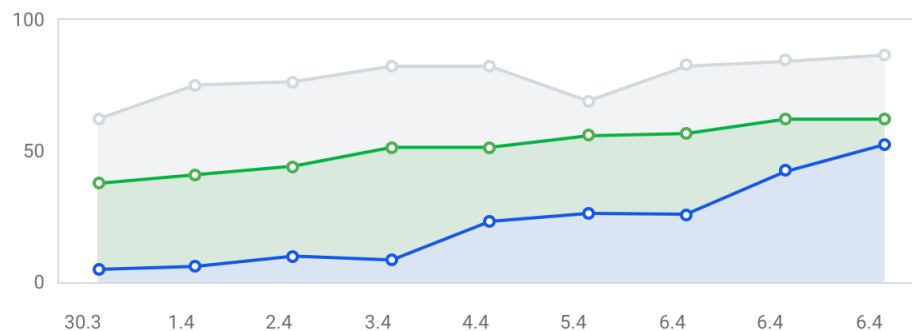
Overall Quality

75%

↘ 3%



Documented Catalog Items Over the Time



General Information

Flight transactions compiles data from a range of different communication channels, including a company's website, telephone, email, live chat and marketing materials.

Owner **DK** David Kolinek

Last import 9/3/2020

Created at 7/3/2020

Detected Glossary Terms

Parent of

Term	Schemas	Tables	Attributes
> PII	6	20	37
> Customer	4	16	18
> Product	2	8	0

Insights 3

Detected Similarities

We have detected similar assets in this Data source. Choose what is the relationship between them.

> flightRecords and flightRecordsProd 94%

> employees_new and employeesOld 84%

> GHY734 and marketing_campaigns_2019 64%



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

<input type="checkbox"/>	Party full CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern	1 273		Virgie Castillo profiled a week ago	<input type="button" value="Profile"/>	
<input type="checkbox"/>	Flight_transactions CRM > Customers > Starbucks Looks like a table with a GDPR content.	29 100		Dustin Francis	<input type="button" value="Profile"/>	
<input type="checkbox"/>	party_full_export CRM > Customers > Starbucks	6 666		Francis Medina	<input type="button" value="Profile"/>	
<input type="checkbox"/>	party_transformed CRM > Customers > Starbucks Looks like a table with a GDPR content.	4 607		Jimmy Farmer hasn't been profiled		
<input type="checkbox"/>	person CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern ● ● +5	23 077		Myrtie Nelson profiled a week ago		
<input type="checkbox"/>	Flight_cancellation CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern	45 222		Aaron Richa... profiled a week ago		



DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

[Overview](#)
[Profile](#)
[Data Quality](#)
[Data Preview](#)
[Data Lineage](#)
[Access](#)
[History](#)

Glossary terms

- PII
- Citizen
- Name
- Date of birth
- Email
- Contact
- SIN
- Gender
- First name

Data Quality Overview

Last evaluated at 9:16 02/14/2020

Evaluate Again

Records

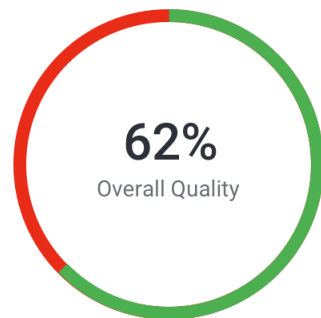
4 886 764

DQ Checks

12

Attributes

5



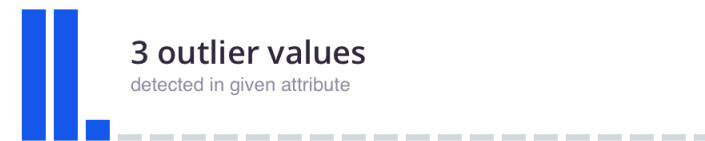
Attributes summary

Name Terms

src_primary_key

Insights

Anomalies detected in attribute birth_date
 NULL, N/A, and 1900-01-01 are outliers.
 Apply automatic cleansing before using this data.






Cleanse data

Unexpected length and record count
 Detected based on previous 14 versions of the profile

Open detail

Relations

- Parent of: Party_full.csv
- Governed by: GDPR policy
- Transformed items: Party_full
- Environments: Partu_full_QA

- person 
- Account_ID 
- birth_date 

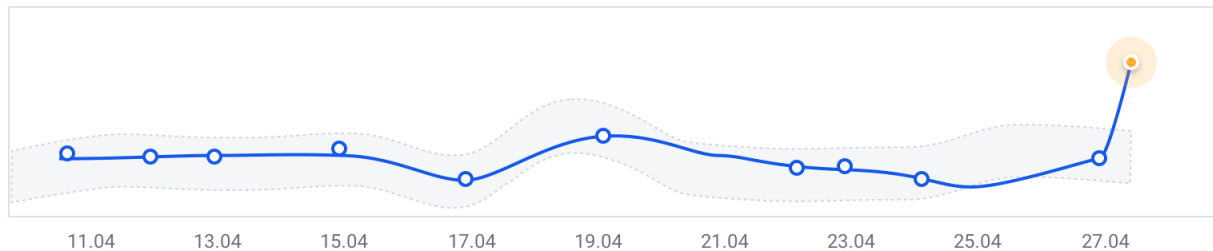
 **person**
Amazon AWS > Person

Anomaly detected by the system

Choose State 

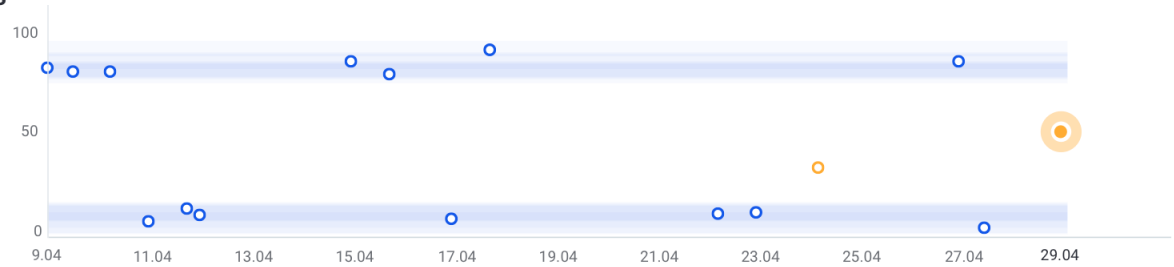
Length

27.04.2020
15
+8



Number of records

27.04.2020





DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

[Overview](#)
[Profile](#)
[Data Quality](#)
[Data Preview](#)
[Data Lineage](#)
[Access](#)
[History](#)

Glossary terms

- PII
- Citizen
- Name
- Date of birth
- Email
- Contact
- SIN
- Gender
- First name

Data Quality Overview

Last evaluated at 9:16 02/14/2020

Evaluate Again

Records

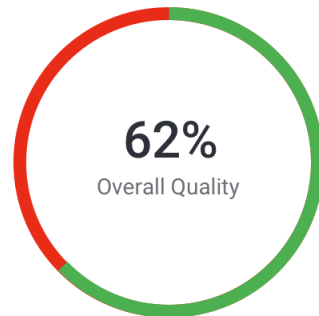
4 886 764

DQ Checks

12

Attributes

5



Attributes summary

Name Terms

src_primary_key

Insights

Anomalies detected in attribute birth_date
 NULL, N/A, and 1900-01-01 are outliers.
 Apply automatic cleansing before using this data.



Cleanse data

Unexpected length and record count
 Detected based on previous 14 versions of the profile

Open detail

Relations

- Parent of: Party_full.csv
- Governed by: GDPR policy
- Transformed items: Party_full
- Environments: Partu_full_QA



Create report

Transform

Export

Use in

DATA CATALOG

Overview Profile Data Quality Data Preview Data Lineage Access History

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Full profile

278636 records 9 attributes Version 112 Last profiled 5 minutes ago

Search

Name	Insights	Glossary Terms	Count	Frequency Analysis	Masks
field_5		[USA] Address		"8500 Leslie street Marham" "8500 Leslei street Toronto L3T 7M8"	"DDDD LLLLLL LLLLLL LLLLLL" Show All +7
field_6		E-mail		"jane.watson@gmail.com" "john.smith@gmail.com"	"LLLL.LLLLLL@LLLLL.LLL" Show All +7
field_8				"2000/01/01" "2008/04/12" "2008/04/08"	"DDDD/DD/DD" "LLLL_LLLL_LLLLLL"
field_2		Year of birth		"NULL" "N/A"	"DDDD-DD-DD" Show All +10
field_1		Gender		"F" "M"	"L" Show All +3
field_0		Full name		"SYSTEM MIGRATION" "L. Proud" "T. Buckingham"	"LLLLL LLLLLL" "L. LLLL" Show All +28
field_3		[USA] Social Security Number		"956155543" "47053392"	"DDDDDDDD" "LLL: DDDDDDDDD" Show All +12
field_7				"59983" "100011" "100010"	"DDDDD" "DDDDDD" Show All +5
field_4				"5526332671941400" "5526332671948710"	"DDDDDDDDDDDDDDDD" Show All +6





Create report

Transform

Export

Use in

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Overview Profile Data Quality Data Preview Data Lineage Access History

Row-level Quality

75%



Aggregated Quality

62%



Records

4 886 764

DQ Checks

12

Attributes

5

Attribute Validity Applied Checks

Search attributes



Attribute	Validity	Records	Valid	Invalid	Aplied Validity Rules	Last Evaluated
T email		578 922	4 923	58 808	Completeness, Email validation	12/12/12
T first_name		578 922	89 635	7 188	Completeness, Name validation	15/12/12
T gender		578 922	14 228	71 347	Completeness, Accuracy	12/12/12
T address		578 922	32 830	97 010	Completeness, Address validation	12/12/12
T SIN		578 922	29139	86 872	Completeness, SIN Validation, Uniqueness	12/12/12





DATA CATALOG

Overview Profile Data Quality **Data Preview** Data Lineage Access History



Catalog Items



Sources

Rules

Rule Relationships



BUSINESS GLOSSARY

Terms

Settings

Full name <small>Full name Pattern</small>	Position <small>Career Enum Reference data</small>	Salary <small>Salary Pattern</small>	Personal address <small>Address Pattern</small>
Frederick Gardner	Software tester	C\$ XXX	Canada XXXXXXXXXXXX
Louise Hudson	UX designer	C\$ XXX	USA XXXXXXXXXXXX
Leah Wolfe	Network engineer	C\$ XXX	Canada XXXXXXXXXXXX
Ray Hicks	Technical support	C\$ XXX	Canada XXXXXXXXXXXX
Eliza Ramos	Technical support	C\$ XXX	Canada XXXXXXXXXXXX
Jacob Bridges	AI researcher	C\$ XXX	Canada XXXXXXXXXXXX
Lelia Collins	Technical consultant	C\$ XXX	Canada XXXXXXXXXXXX
Mary Carroll	Software tester	\$ XXX	USA XXXXXXXXXXXX
Hunter Moran	Technical consultant	C\$ XXX	Canada XXXXXXXXXXXX
William Berry	Business analyst	C\$ XXX	Canada XXXXXXXXXXXX
Jackson Byrd	Network engineer	\$ XXX	USA XXXXXXXXXXXX
Harriett Boyd	Technical support	\$ XXX	USA XXXXXXXXXXXX
Alma Vega	Technical consultant	\$ XXX	USA XXXXXXXXXXXX
Lillie Silva	Business analyst	\$ XXX	USA XXXXXXXXXXXX

Salary is masked based on the Compensation and benefits policy.





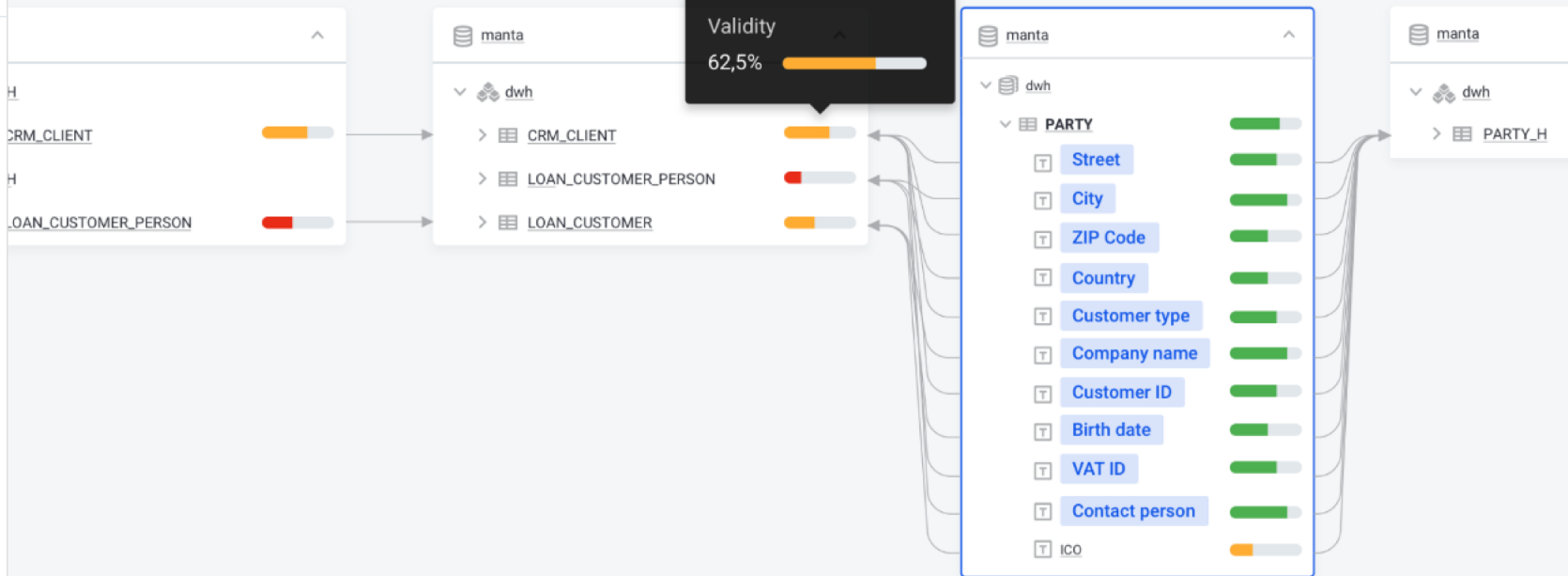
DATA CATALOG

- Catalog Items
- Sources
- Rules
- Rule Relationships

BUSINESS GLOSSARY

- Terms
- Settings

Filter tables, attributes, terms





Dashboard

All Items

Category

Applications

Metrics

CDE

Security

Business terms

Technical terms

Relations

Governed by PII

Synonyms

Aliases

Overview Relationships Occurance Data Quality Statistics Stakeholders Policies History Settings

Business definition

Personal data is defined under the GDPR as "any information which related to an identified or identifiable natural person". Personal data, also known as personal information or personally identifiable information is any information relating to an identifiable person.

Relations

Child of Security Classification

Synonym of SPI

Consist of Full Name Home Address Telephone Date of Birth Gender Web Cookie

Stakeholders

Owner

Peter Williams

Security classification

Internal use only

Data Quality Overview

Last evaluated at 9:16 02/14/2020

Evaluate Again

Records

700M+

Attributes

4 618

Tables

645

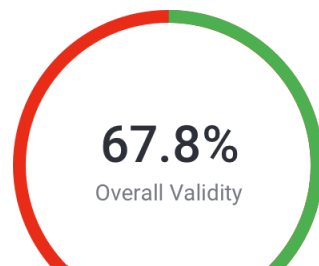
Data Sources

44

DQ Checks

19

Validity



Custom DQ Dimensions

Completeness 89%

Something 63%

Uniqueness 70%

Occurence statistics

Attributes

4 618

Tables

645

Data sources

44

Suggestions

35

Suggested Terms

Confirm All

<input type="checkbox"/>	Name	Suggestions
<input type="checkbox"/>	T attr_247	Address <input type="checkbox"/> <input checked="" type="checkbox"/>
<input type="checkbox"/>	T first_name	First Name <input type="checkbox"/> <input checked="" type="checkbox"/>
<input type="checkbox"/>	T last_name	Last Name <input type="checkbox"/> <input checked="" type="checkbox"/>



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

<input type="checkbox"/>	Party full CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern	1 273		Virgie Castillo profiled a week ago	
<input type="checkbox"/>	Flight_transactions CRM > Customers > Starbucks Looks like a table with a GDPR content.	29 100		Dustin Francis <input type="button" value="Profile"/>	
<input type="checkbox"/>	party_full_export CRM > Customers > Starbucks	6 666		Francis Medina <input type="button" value="Profile"/>	
<input type="checkbox"/>	party_transformed CRM > Customers > Starbucks Looks like a table with a GDPR content.	4 607		Jimmy Farmer hasn't been profiled	
<input type="checkbox"/>	Orders_nov_19 CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern ● ● +5	23 077		Myrtie Nelson profiled a week ago	
<input type="checkbox"/>	Flight_cancellation CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern	45 222		Aaron Richa... profiled a week ago	



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

- Rule
- Data Source
- Profile
- Attribute

[Virgie Castillo](#) profiled a week ago

[Dustin Francis](#)

Looks like a table with a **GDPR** content.

[party_full_export](#) 6 666 [Francis Medina](#)

CRM > Customers > Starbucks

[party_transformed](#) 4 607 [Jimmy Farmer](#) hasn't been profiled

CRM > Customers > Starbucks

Looks like a table with a **GDPR** content.

[Orders_nov_19](#) 23 077 [Myrtie Nelson](#) profiled a week ago

CRM > Customers > Starbucks

GDPR Customer Outlier Enum Pattern ● ● +5

[Flight_cancellation](#) 45 222 [Aaron Richa...](#) profiled a week ago

CRM > Customers > Starbucks

GDPR Customer Outlier Enum Pattern



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

birth

Search

Search **birth** everywhere...

with term **Date of birth**

with term **Birth name**

with term **Birthplace**

Date of birth definition

[Virgie Castillo](#) profiled a week ago

[Dustin Francis](#) [Profile](#)

[party_full_export](#) 6 666 [Francis Medina](#) [Profile](#)
CRM > Customers > Starbucks

[party_transformed](#) 4 607 [Jimmy Farmer](#) hasn't been profiled
CRM > Customers > Starbucks
Looks like a table with a **GDPR** content.

[Orders_nov_19](#) 23 077 [Myrtie Nelson](#) profiled a week ago
CRM > Customers > Starbucks
GDPR **Customer** **Outlier** **Enum** **Pattern** ● ● +5

[Flight_cancellation](#) 45 222 [Aaron Richa...](#) profiled a week ago
CRM > Customers > Starbucks
GDPR **Customer** **Outlier** **Enum** **Pattern**



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

Search input field containing the text 'birth'

Search button

Search **birth** everywhere...

- with term **Date of birth**
- with term **Birth name**
- with term **Birthplace**

Date of birth definition

Virgie Castillo profiled a week ago

Dustin Francis Profile

<input type="checkbox"/>	party_full_export CRM > Customers > Starbucks	6 666	Francis Medina	Profile	
<input type="checkbox"/>	party_transformed CRM > Customers > Starbucks Looks like a table with a GDPR content.	4 607	Jimmy Farmer	hasn't been profiled	
<input type="checkbox"/>	Orders_nov_19 CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern ● ● +5	23 077	Myrtie Nelson	profiled a week ago	
<input type="checkbox"/>	Flight_cancellation CRM > Customers > Starbucks GDPR Customer Outlier Enum Pattern	45 222	Aaron Richa...	profiled a week ago	



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

Term

- Rule
- Data Source
- Profile
- Attribute

Looks like a table with a **GDPR** content.

[Virgie Castillo](#) profiled a week ago

[Dustin Francis](#)

[party_full_export](#) 6 666 [Francis Medina](#)

CRM > Customers > Starbucks

[party_transformed](#) 4 607 [Jimmy Farmer](#) hasn't been profiled

CRM > Customers > Starbucks

Looks like a table with a **GDPR** content.

[Orders_nov_19](#) 23 077 [Myrtie Nelson](#) profiled a week ago

CRM > Customers > Starbucks

GDPR Customer Outlier Enum Pattern ● ● +5

[Flight_cancellation](#) 45 222 [Aaron Richa...](#) profiled a week ago

CRM > Customers > Starbucks

GDPR Customer Outlier Enum Pattern



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

Search bar containing 'Term Date of birth' with a search icon and a clear button.

Search button

- Rule
- Data Source**
- Profile
- Attribute

Virgie Castillo profiled a week ago

Dustin Francis [Profile](#)

Looks like a table with a **GDPR** content.

[party_full_export](#) 6 666 Francis Medina [Profile](#)
CRM > Customers > Starbucks

[party_transformed](#) 4 607 Jimmy Farmer hasn't been profiled
CRM > Customers > Starbucks
Looks like a table with a **GDPR** content.

[Orders_nov_19](#) 23 077 Myrtie Nelson profiled a week ago
CRM > Customers > Starbucks
GDPR Customer Outlier Enum Pattern ● ● +5

[Flight_cancellation](#) 45 222 Aaron Richa... profiled a week ago
CRM > Customers > Starbucks
GDPR Customer Outlier Enum Pattern



DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Search

- ERP - Berlin
- CRM
- ONE MDM Portal
- PostgreSQL DEV Test
- S3 backup bucket

[Virgie Castillo](#)
profiled a week ago

[Dustin Francis](#)
Profile

[party_full_export](#)
6 666
[Francis Medina](#)
Profile

CRM > Customers > Starbucks

[party_transformed](#)
4 607
[Jimmy Farmer](#)
hasn't been profiled

CRM > Customers > Starbucks

Looks like a table with a GDPR content.

[Orders_nov_19](#)
23 077
[Myrtie Nelson](#)
profiled a week ago

CRM > Customers > Starbucks

GDPR Customer Outlier Enum Pattern ● ● +5

[Flight_cancellation](#)
45 222
[Aaron Richa...](#)
profiled a week ago

CRM > Customers > Starbucks

GDPR Customer Outlier Enum Pattern



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

Search

- ERP - Berlin
- CRM
- ONE MDM Portal**
- PostgreSQL DEV Test
- S3 backup bucket

[Virgie Castillo](#) profiled a week ago

[Dustin Francis](#) Profile

[party_full_export](#) 6 666 [Francis Medina](#) Profile
CRM > Customers > Starbucks

[party_transformed](#) 4 607 [Jimmy Farmer](#) hasn't been profiled
CRM > Customers > Starbucks
 Looks like a table with a GDPR content.

[Orders_nov_19](#) 23 077 [Myrtie Nelson](#) profiled a week ago
CRM > Customers > Starbucks
GDPR Customer Outlier Enum Pattern ● ● +5

[Flight_cancellation](#) 45 222 [Aaron Richa...](#) profiled a week ago
CRM > Customers > Starbucks
GDPR Customer Outlier Enum Pattern



DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Search

- Rule
- Data Source
- Profile
- Attribute

[Virgie Castillo](#)
profiled a week ago

[Dustin Francis](#)
Profile

Looks like a table with a **GDPR** content.

<input type="checkbox"/>	party_full_export <small>CRM > Customers > Starbucks</small>	6 666	Francis Medina Profile
<input type="checkbox"/>	party_transformed <small>CRM > Customers > Starbucks</small> Looks like a table with a GDPR content.	4 607	Jimmy Farmer hasn't been profiled
<input type="checkbox"/>	Orders_nov_19 <small>CRM > Customers > Starbucks</small> GDPR Customer Outlier Enum Pattern ● ● +5	23 077	Myrtie Nelson profiled a week ago
<input type="checkbox"/>	Flight_cancellation <small>CRM > Customers > Starbucks</small> GDPR Customer Outlier Enum Pattern	45 222	Aaron Richa... profiled a week ago



KNOWLEDGE CATALOG

DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

Data Catalog

Term: Date of birth
Data source: ONE MDM Portal

Search

person
4,886,764
Virgie Castillo
profiled a week ago

ONE MDM Portal > Master layer

PII
Citizen
Name
Date of birth
Email
Contact
SIN
Gender
First name



DATA CATALOG

Catalog Items

Sources

Rules

Rule Relationships

BUSINESS GLOSSARY

Terms

Settings

[Overview](#)
[Profile](#)
[Data Quality](#)
[Data Preview](#)
[Data Lineage](#)
[Access](#)
[History](#)

Glossary terms

- PII
- Citizen
- Name
- Date of birth
- Email
- Contact
- SIN
- Gender
- First name

Data Quality Overview

Last evaluated at 9:16 02/14/2020

Evaluate Again

Records

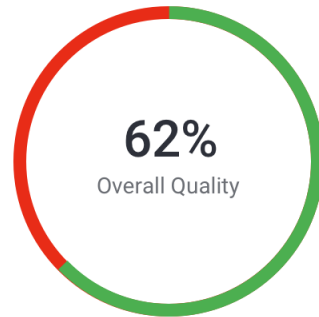
4 886 764

DQ Checks

12

Attributes

5



Attributes summary

Name Terms

src_primary_key

Insights

Anomalies detected in attribute birth_date
 NULL, N/A, and 1900-01-01 are outliers.
 Apply automatic cleansing before using this data.



Cleanse data

Unexpected length and record count
 Detected based on previous 14 versions of the profile

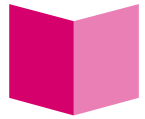
Open detail

Relations

- Parent of: Party_full.csv
- Governed by: GDPR policy
- Transformed items: Party_full
- Environments: Partu_full_QA

Key elements for successful implementation

Foundations for successful Data Governance



Centralized **metadata management** with embedded **best practices**



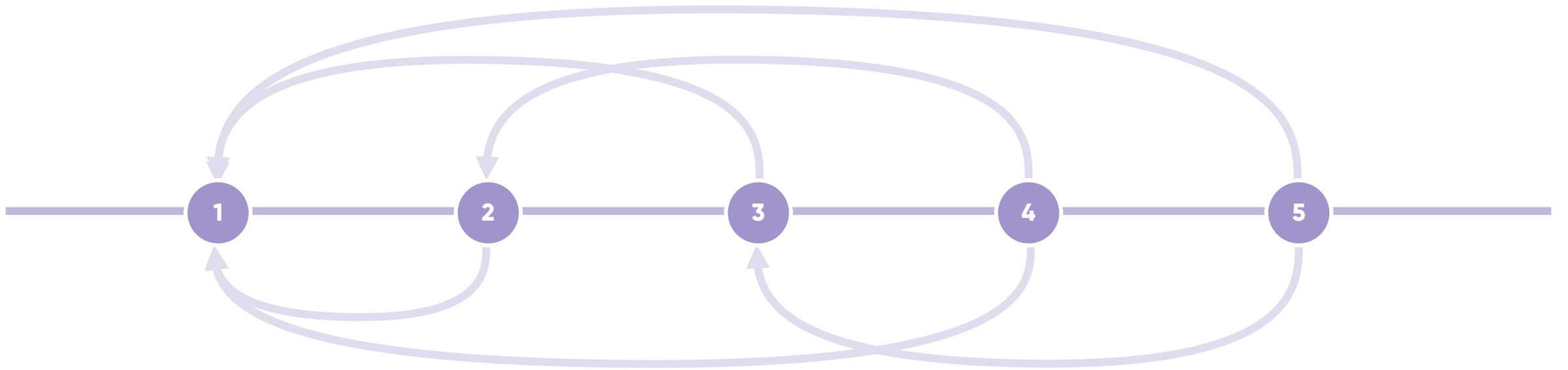
Leveraging AI for metadata capture and automation



Learning from **user interactions**



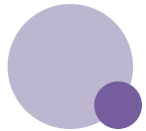
Continuous **improvements** and **synchronization**



Do you want your
data governance initiative
to be successful?

**Don't rely on crowdsourcing.
Rely on automation.**

Schedule a demo



**Data Discovery
& Profiling**



**Metadata Management
& Data Catalog**



**Data Quality
Management**



**Master & Reference
Data Management**



**Big Data Processing
& Data Integration**

info@ataccama.com

Q&A



Digital Conference

Hear from industry analysts,
Ataccama clients & other
market experts

ataccama.com/conference

