



05 Aug, 2025

# AI Governance using Informatica CDGC

- Arya Jemo, Principal Solution Architect, IPS
- Barbara Bukowska, Team Lead, IPS

Where data & AI come to **LIFE**

# Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available on our [Success Portal](#) - where you can download the **slide deck** for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

# Feature Rich Success Portal



Bootstrap trial and  
POC Customers



Enriched Customer  
Onboarding  
experience



Product Learning  
Paths and Weekly  
Expert Sessions



Informatica  
Concierge



Tailored training and  
content  
recommendations

# More Information



## Success Portal

<https://success.informatica.com>



## Communities & Support

<https://network.informatica.com>



## Documentation

<https://docs.informatica.com>



## University

<https://www.informatica.com/in/services-and-training/informatica-university.html>

# Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

# Table of contents

- AI Market Adoption
- AI Regulations
- AI Governance
- Data Governance
- Demo of Cloud Data Governance and Catalog
- Q&A



# AI Adoption Is Growing Rapidly

## Organizations using GenAI in multiple business units

Over the past 18 months

25%  
18-mo  
incre

## Spending on GenAI

Has increased to \$644bn

76%  
2025

## Most advanced GenAI initiatives

Meeting or exceeding expectations

74%  
with



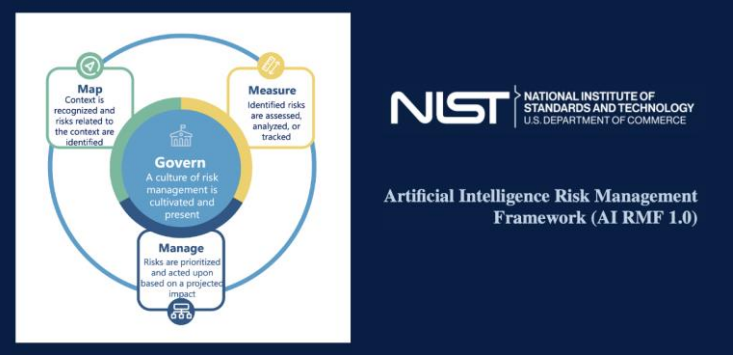
Generative AI is going to reinvent virtually every customer experience we know and enable altogether new ones about which we've only fantasized.

**Andy Jassy, CEO**  
AWS, '24 Letter to Shareholders

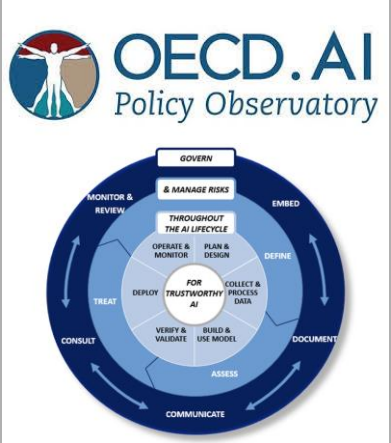
1. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>
2. <https://www.gartner.com/en/newsroom/press-releases/2025-03-31-gartner-forecasts-worldwide-genai-spending-to-reach-644-billion-in-2025>
3. <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-state-of-gen-ai-q4.pdf>

# AI Governance

Framework of policies, processes and systems that ensure the safe and effective development and deployment of AI technologies



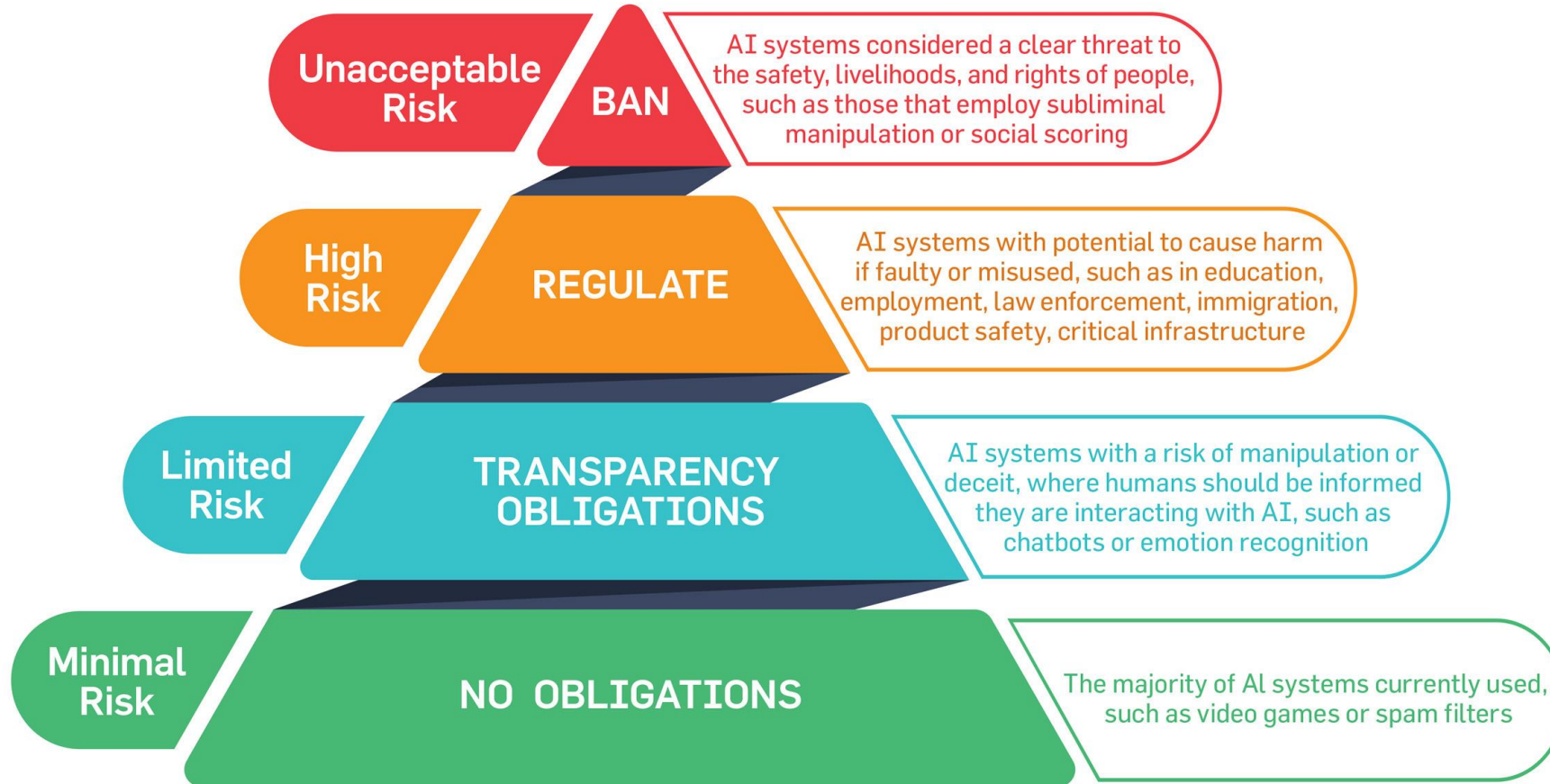
**ISO**  
Information technology — Artificial intelligence — Management system  
Published (Edition 1, 2023)



# EU AI Act



The first-ever legal framework on AI, the Act classifies AI systems into four levels of risk based on their purposes and use cases, with corresponding levels of regulatory scrutiny and developer's obligation.



<https://cacm.acm.org/news/governments-setting-limits-on-ai/>

# EU Artificial Intelligence Act



Let's deep dive into one risk category



## Requirements for providers of high-risk AI systems (Art. 8–17)

- Establish a **quality management system** to ensure compliance.
- Design their high risk AI system for **record-keeping** to enable it to automatically record events relevant for identifying national level risks and substantial modifications throughout the system's lifecycle.
- Establish a **risk management system** throughout the high risk AI system's lifecycle
- Conduct **data governance**, ensuring that training, validation and testing datasets are relevant, sufficiently representative and, to the best extent possible, free of errors and complete according to the intended purpose.
- Draw up **technical documentation** to demonstrate compliance and provide authorities with the information to assess that compliance.
- Provide **instructions for use** to downstream deployers to enable the latter's compliance.
- Design their high risk AI system to allow deployers to implement **human oversight**.
- Design their high risk AI system to achieve appropriate levels of **accuracy, robustness, and cybersecurity**.

# AI Governance Made Simple

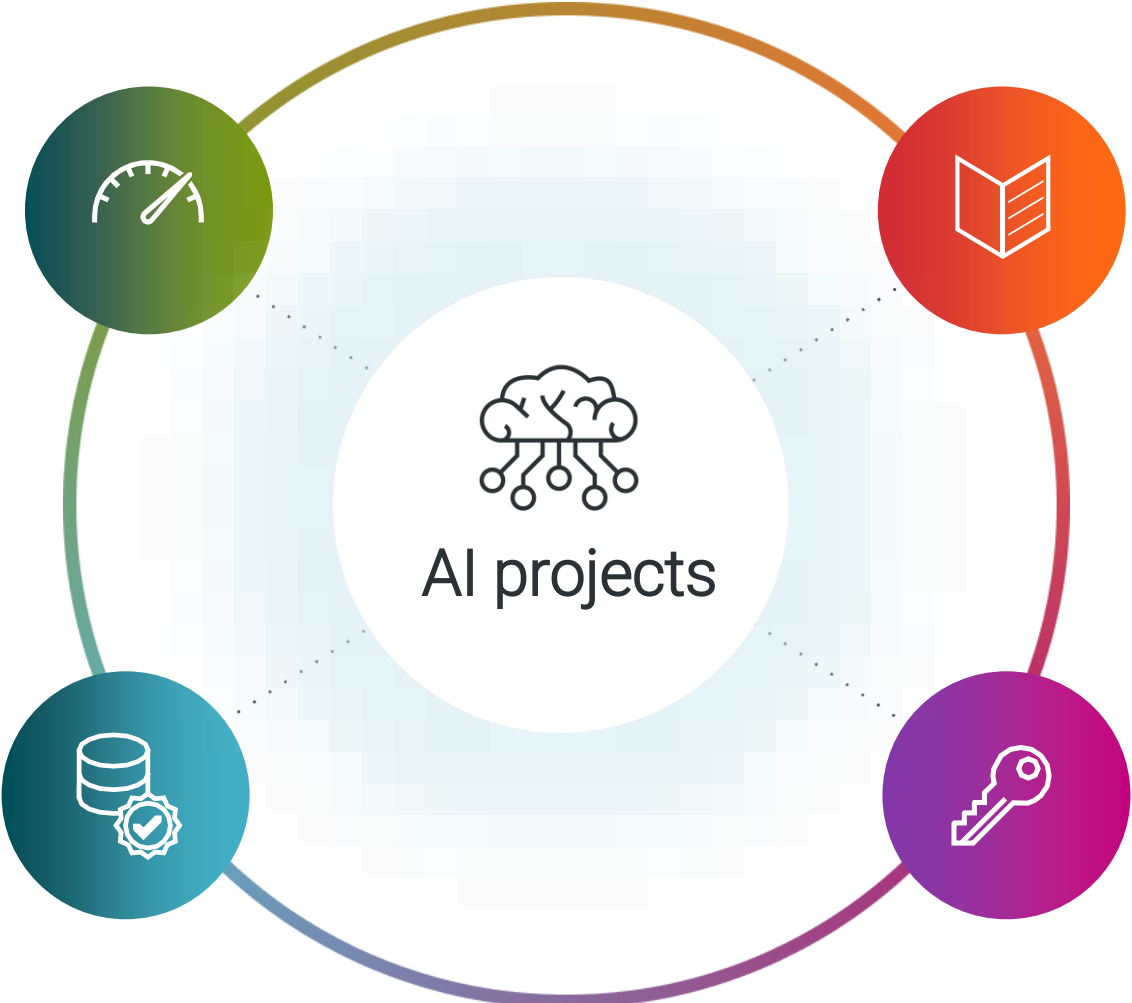
Build confidence in your AI initiatives with trust to accelerate ROI

Observe

Inventory

Deliver

Control

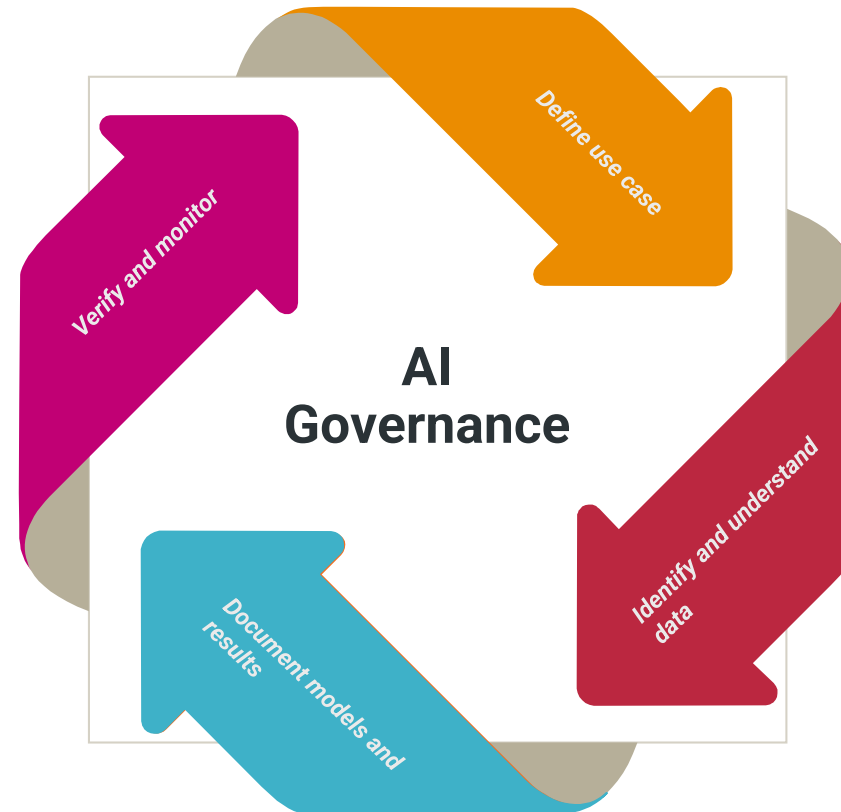


# AI Governance

AI Governance involves a wide range of aspects, from ensuring data privacy and fairness to fostering transparency and stakeholder engagement

## *Define use case*

- Catalog Business terms to enable clear communication
- Clarify problem statement and business objectives
- Identify stakeholders and expected impact
- Define success metrics and ethical considerations



## *Verify and monitor*

- Perform model testing (accuracy, fairness, robustness)
- Conduct ethical and bias audits
- Get stakeholder review and approvals

## *Identify and understand data*

- Collect and catalog relevant data sources
- Assess data quality, privacy, and bias concerns
- Ensure data compliance with regulations (e.g., GDPR)

## *Document models and results*

- Document model architecture, data lineage, and assumptions
- Define validation and testing protocols

# Example: Executing the AI strategy

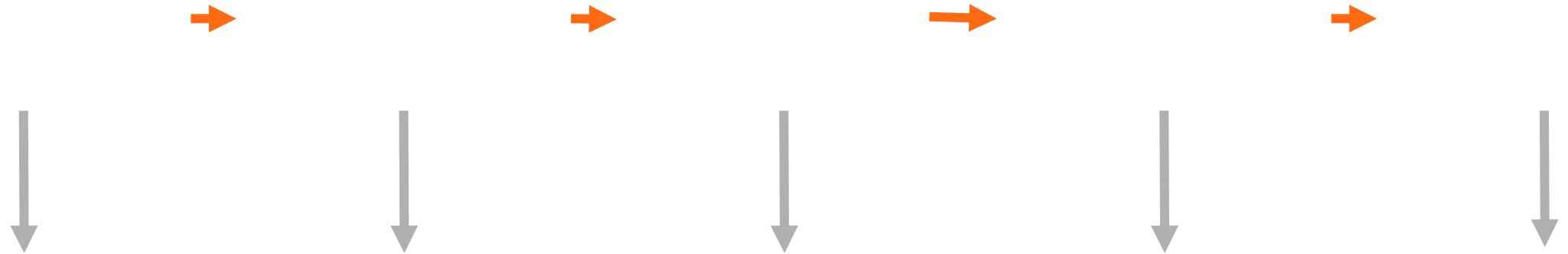


**DALE**

Chief Data and Analytics Officer

Ensure that Data and AI Assets can be utilized safely to accelerate AI initiatives

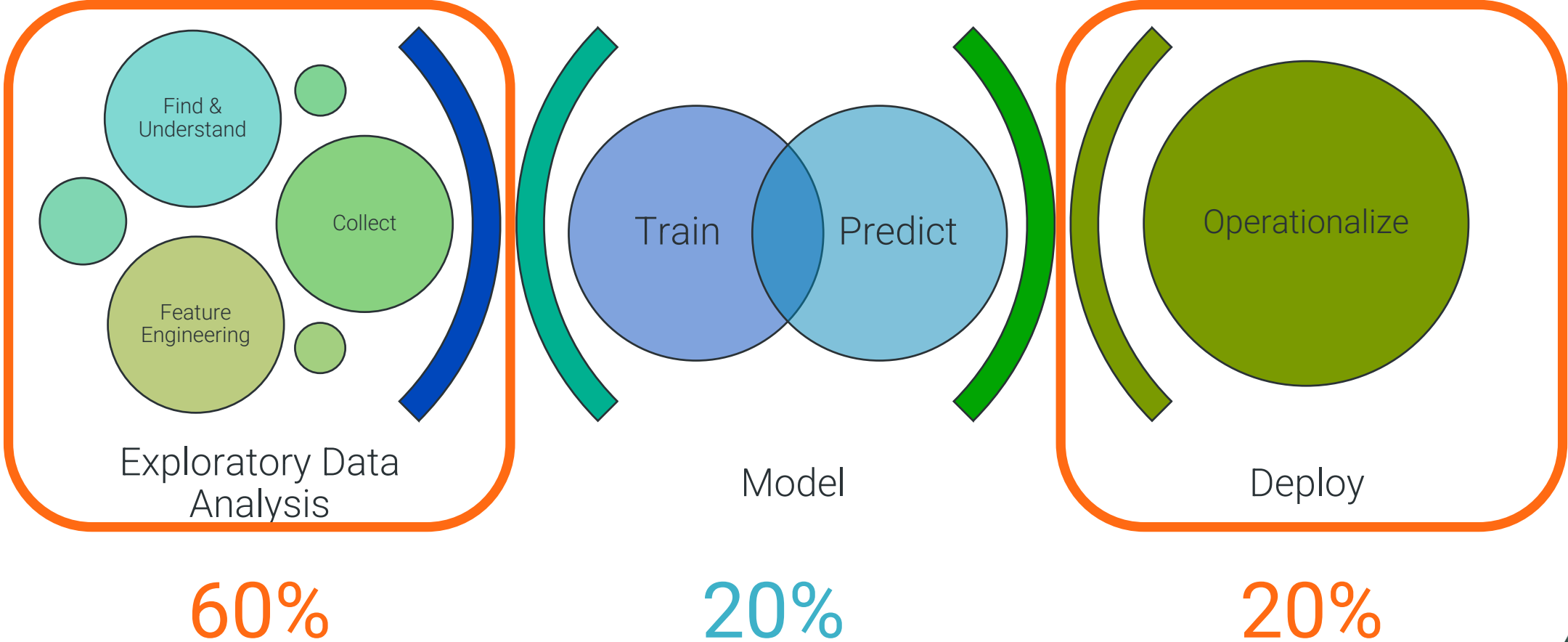
**OFFENSIVE**



**DEFENSIVE**

# The AI Workflow

Data is the most important part of your AI initiatives



An aerial, high-angle photograph of a busy pedestrian crossing. The scene is filled with a dense crowd of people walking in various directions. Long, dark shadows are cast across the pavement, indicating a low sun position. The text is centered over the image in a bold, white, sans-serif font.

**AI is the brain that makes sense of  
data, the vital blood of your  
company**

# AI intensifies the call for comprehensive data governance

## Data Scale and Velocity

AI models require large volumes of diverse datasets for training. Matching the demands of AI with manual efforts and legacy tools is impractical

## Traceability & Accountability

For AI model transparency, companies need clear visibility across the data landscape and understanding of where each data element can be sourced from

## Emerging Legislative Regulations

Emerging AI regulations necessitate the need for robust data discovery solutions to demonstrate compliance

## Responsible AI & AI Ethics

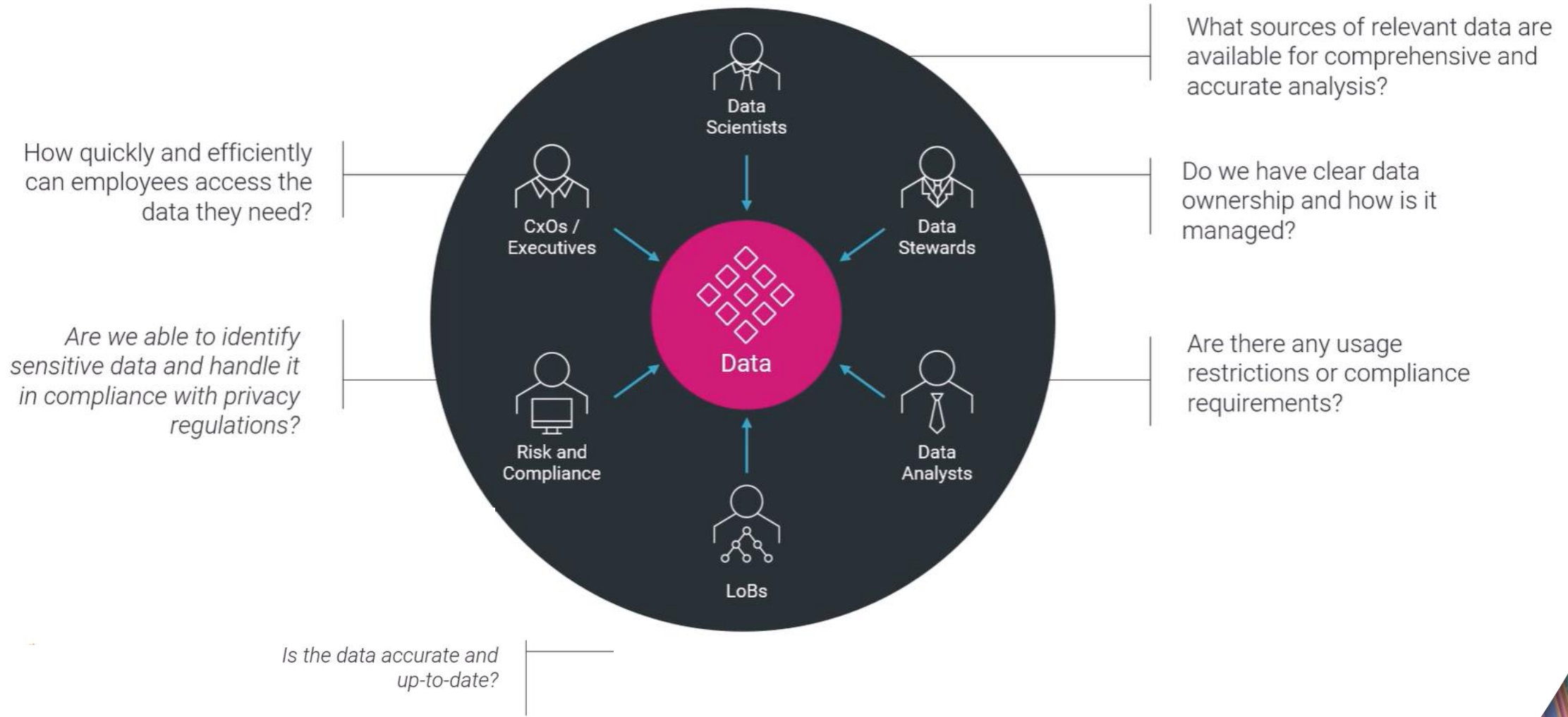
Principles for responsible use of data that is fair and free of biases demand deep discoverability into data sources and origins

## Privacy & Security

Identification and classification of sensitive information accurately from large volumes of data is crucial for implementing data protection measures

# Journey to AI Governance starts with Data

## Finding Trustworthy Data is the First Step and It Impacts Everyone



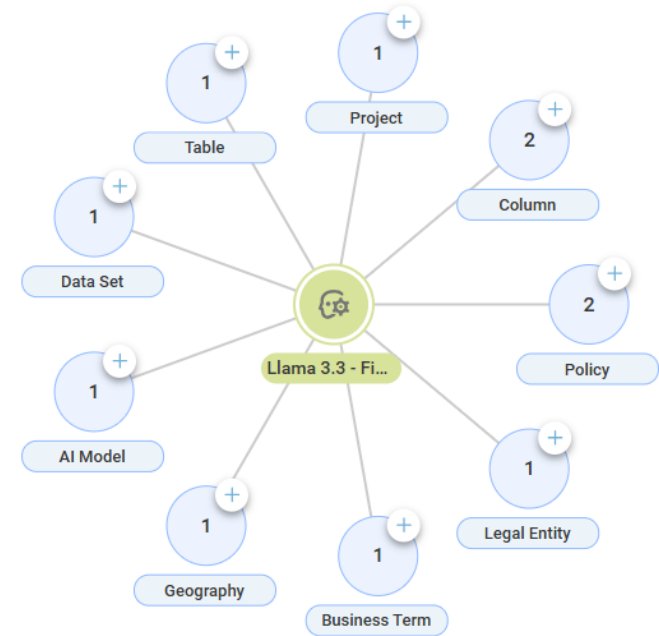


# Demo

Where data & AI come to **LIFE**

# Demo of Cloud Data Governance and Catalog

- AI Governance Assets
- Dashboard for Artificial Intelligence Compliance Officer
- New AI features of CDGC with July 2025 release



Data Governance and Catalog | Adaptive | Start typing search terms for suggestions.

### AI Compliance - AICO Dashboard

Dashboard for Artificial Intelligence Compliance Officer that should help them to ensure usage of AI/ML is compliant with company ethics & regulations.

#### EU AI Act

The first-ever legal framework on AI, the Act classifies AI systems into four levels of risk based on their purposes and use cases, with corresponding levels of regulatory scrutiny and developer's obligation.

<https://cacm.acm.org/news/governments-setting-limits-on-ai/>

#### Process how to comply with EU AI Act

[This process flow contains steps by steps instructions you need to follow to ensure your AI Model / System will comply with the last EU AI Act](#)

#### PII Assets Related To AI Models

Business Term	34
Subdomain	26
Process	17
Data Set	10
Policy	6
AI Model	2
Metric	2
Project	2
Domain	1

Group By: Asset Type

#### DQ Scores - AI Models Input DataSets

Group By: Threshold Result

#### AI Regulations To Comply

Name	Description
US - AI Bill of Rights	Reference : Blueprint for an A
OpenAI Usage Regulations	should help to understand us
Authorized AI/ML Business Usage Purpose	We are listing ALL usage purp
EU Artificial Intelligence Act	The first-ever legal frameworl

#### "Non-High-Risk" Models using PII Data

Name	Description
Personalized Product Offers (PPD)	ML Model that aims to provic
Customer Churn Prediction	Customer retention is one of

#### In Review AI Assessments

Name	Description
Model Assessment - Customer Churn Pred	this project contains assesse

# New features of CDGC

Where data & AI come to 

# New features of AI Governance in CDGC of July 2025 Release

AI Governance is Gaining Traction - Driven by regulation / natural need to govern data assets

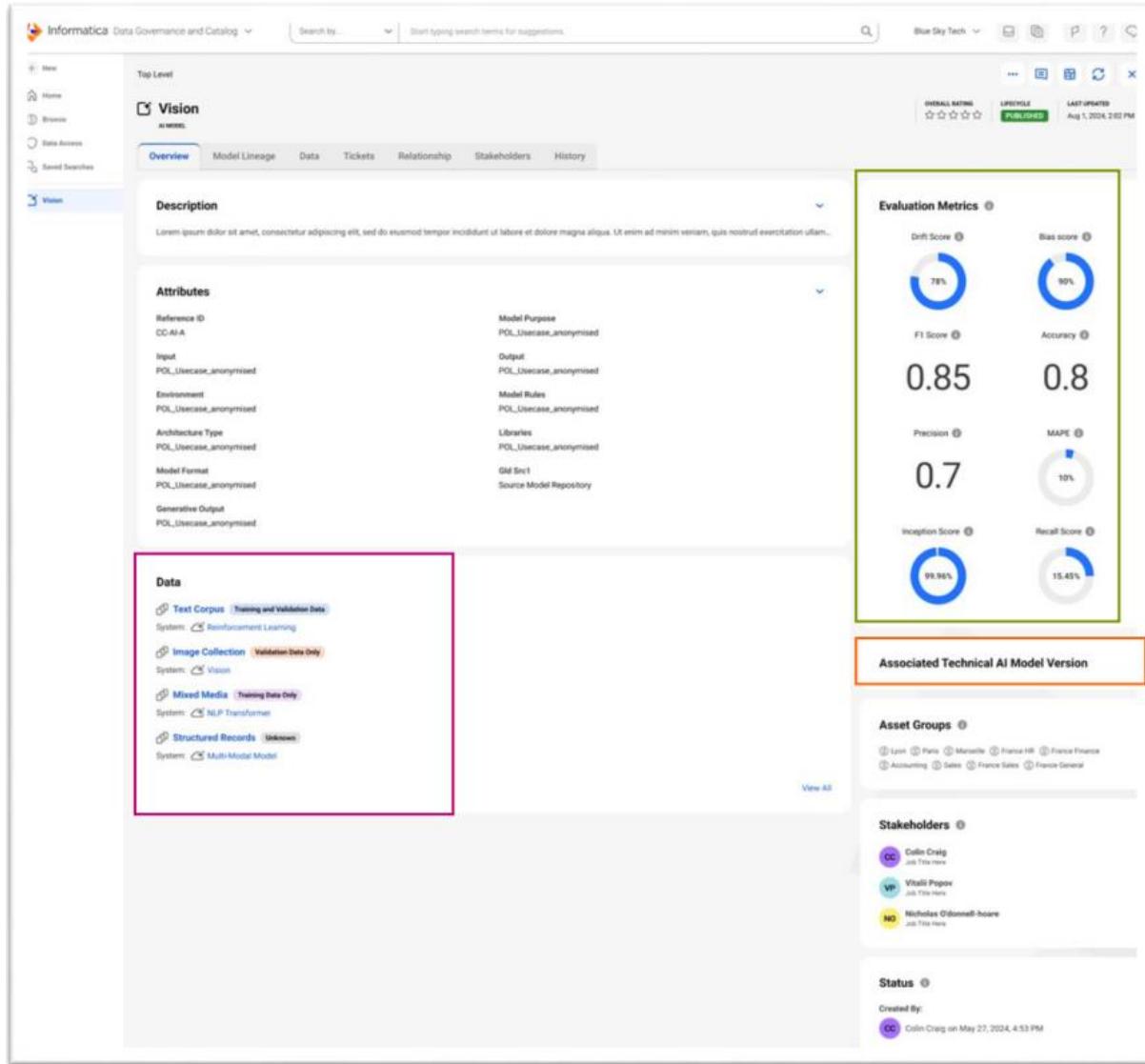
## Preparation and Approval – AI Model Improvements

- Create an inventory of your AI Models
- Understand their provenance / lineage
- Understand data sets used to train / evaluate
- User Configurable evaluation metrics

## Operational Deployment – AI System (New Business Asset)

- Understand how, where and why AI Capabilities are delivered to your organization
- Linked to the models that they use (and often reuse in multiple systems)

# More effective AI Model asset



## Configurable “Evaluation” Metrics (in MCC)

- Not just Drift and Bias
- Talk about what’s relevant for your model
- New visualizations

## Data Set Relationships

- New relationships to Data Sets allow us to show training /validation relationship
- (Deprecated AI Model x Data Element: pre-upgrade action required)

## Link to (new) scanned AI Models

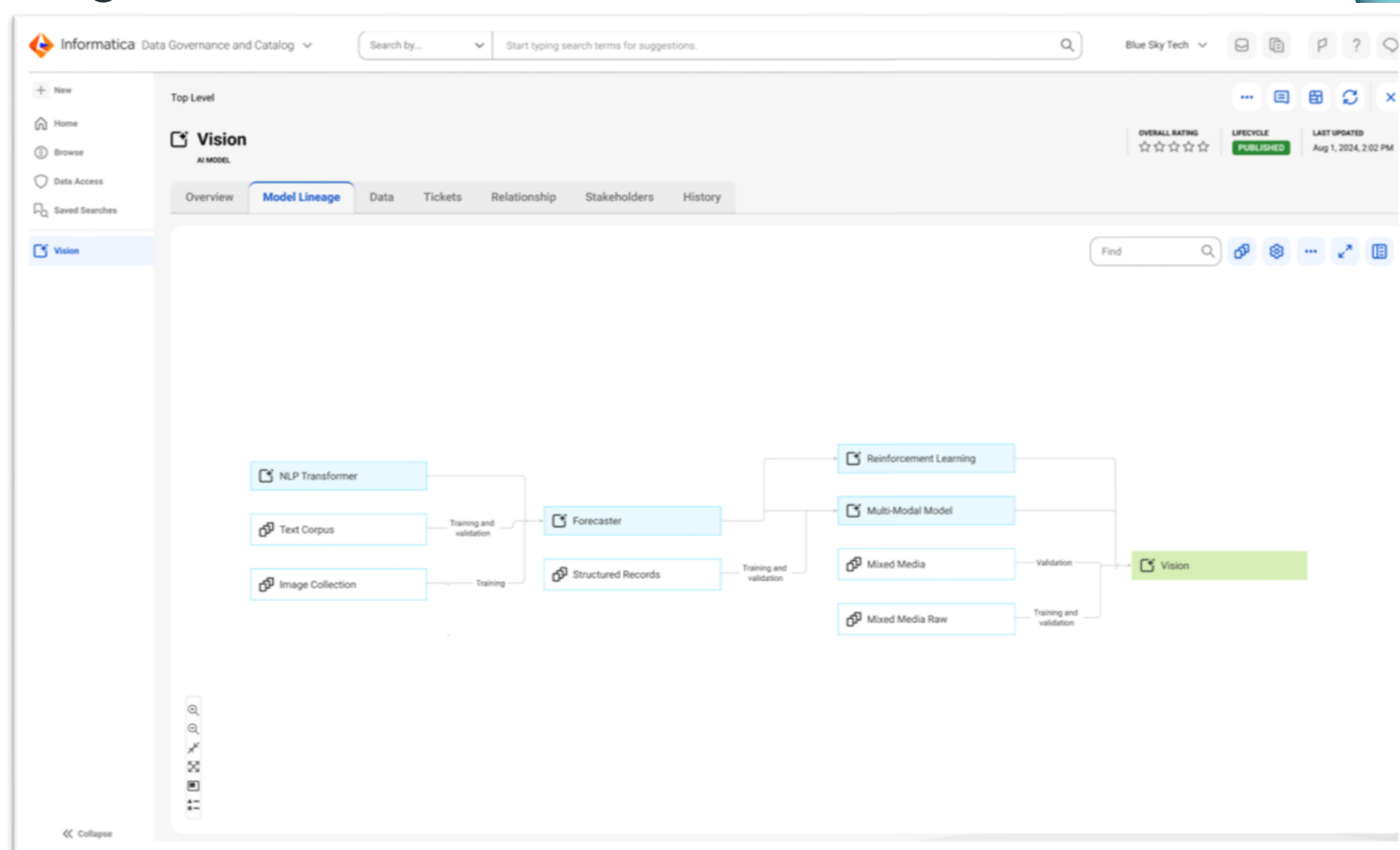
- Promotion from scanned assets (Databricks notebooks only in July) to talk about models of most interest to the business

# AI Model Lineage

Model lineage – from Foundation model to trained, approved model fit for the Use Case

Uses

- Changed behaviour for AI Model x AI Model relationship
- Includes data sets, using new AI Model x Data Set relationships



# New Asset - AI System

## AI System describes deployed AI Capabilities

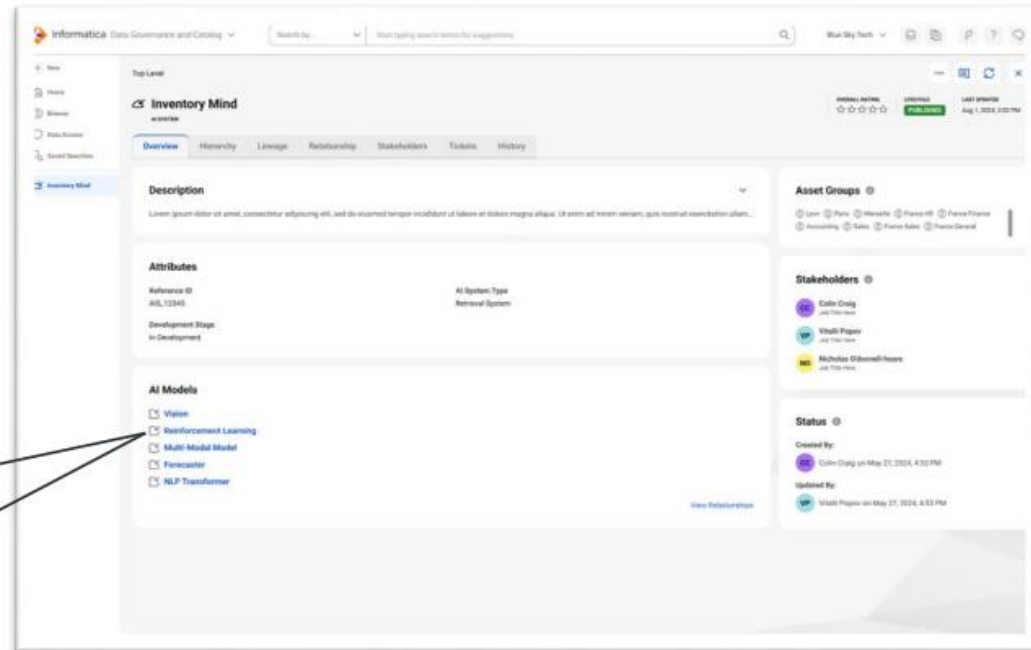
### AI System

- Conversational (Chat Bots)
- RAG
- Agentic

Shows the models used in that system

**AI Models**

- Vision
- Reinforcement Learning
- Multi-Modal Model
- Forecaster
- NLP Transformer

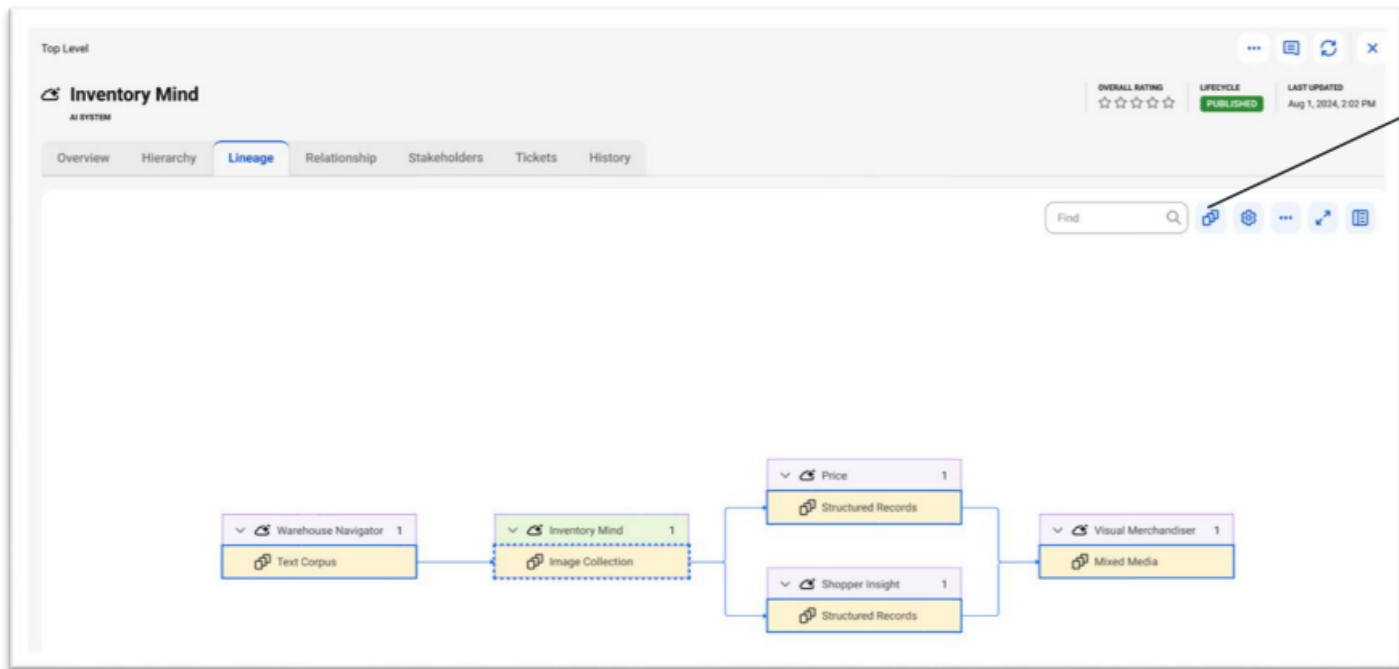


# New Asset - AI System

## Data Pipeline option – part of traditional lineage

### AI System

- Can be part of lineage where it is part of a pipeline – uses Data Sets
- AI Models/ AI Projects as overlay



The 'Overlays' configuration panel is shown, with an arrow pointing from the 'Inventory Mind' asset in the lineage diagram to this panel. The panel is titled 'Overlays' and contains several sections of checkboxes:

- Glossary**
  - Business Term
  - Domain
  - Subdomain
  - Metric
- Data Quality**
  - Aggregated
  - Correctness
  - Validity
  - Completeness
  - Consistency
  - Uniqueness
  - Integrity
  - Currency
- Others**
  - Data Marketplace
- Business Impact**
  - Policy
  - Process
  - AI Model
  - Project

At the bottom right of the panel are 'Apply' and 'Cancel' buttons.

# Where data & AI come to



# Demo 1

## [How to Govern AI Models with Cloud Data Governance and Catalog](#)

### AI Compliance - AICO Dashboard

Default dashboard for Artificial Intelligence Compliance Officer that should help them to ensure usage of AI/ML is compliant with company ethics & regulations.

Shared by: JA

#### Purpose

AI/ML is broadly used inside our company, building our own model or leveraging public/private ones. Our team is responsible to ensure we are compliant with coming/existing regulations & company ethics defined [here](#)



#### "Non-High-Risk" Models using PII Data

Name	Description
Customer Churn Prediction	Customer retention is one of
Personalized Product Offers (PPO)	ML Model that aims to provid
Fraud detection	AI-powered fraud detection s

1 - 3 of 3 Items

#### My Tasks

Ticket title	Ticket status	Asset
<a href="#">Change required to Internal Data Policy</a>	OPEN	Intern
<a href="#">Manage asset Policy Personally Identifiabl</a>	RESOLVED	Person

1 - 2 of 2 Items

#### AI Regulations To Comply

Name	Description
<a href="#">EU Artificial Intelligence Act</a>	
<a href="#">OpenAI Usage Regulations</a>	should help
<a href="#">Authorized AI/ML Business Usage Purpose</a>	We are listin
<a href="#">US - AI Bill of Rights</a>	Reference : BU

1 - 4 of 4 Items

#### Process to comply with EU AI Act

[This process flow contains steps by steps instructions you need to follow to ensure your AI Model / System will comply with the last EU AI Act](#)



#### In Review AI Assessments

Name	Description
<a href="#">Model Assessment - Customer Churn Pred</a>	this project contains assess

1 - 1 of 1 Items

#### AI Models without stakeholder

# Marie, Head of AI @ La Baguette Inc.

- Sebastian's model is now in production and has significantly increased the average basket.
- But Sebastian has left the company. As any developer, he was not rigorous regarding documentation.
- Marie want to be sure she will be able to explain how AI is used in the company:
  - What AI is running?
  - For which usages?
  - What data we use? Is there sensitive ones?



- New...
- Home
- Browse
- Data Access
- Saved Searches

### Browse

- Glossary
- Data Catalog
- System
- Process
- Policy
- Project
- AI Model**
- Business Area
- Legal Entity
- Geography
- Regulation

**AI Model (31)** Find ↕ ⓘ ⚙️

Name ↑	Description
<a href="#">AI_0023</a>	-
<a href="#">ARIMA</a>	ARIMA is an acronym for "autoregressive integrated moving avera...
<a href="#">ChatGPT</a>	Under the EU AI Act, ChatGPT and similar generative AI models are...
<a href="#">Codestral</a>	A language model specifically designed for coding tasks.
<a href="#">Customer Churn Prediction</a>	Customer retention is one of the primary growth pillars for product...
<a href="#">DALL-E</a>	DALL-E is a AI system that can create realistic images and art from...
<a href="#">Deep Neural Networks</a>	Deep neural networks, or deep learning networks, have several hid...
<a href="#">Embeddings</a>	Embeddings are a numerical representation of text that can be use...
<a href="#">Fraud detection</a>	AI-powered fraud detection systems are using machine learning al...
<a href="#">GPT-3.5</a>	GPT-3.5 models can understand and generate natural language or ...
<a href="#">gpt-4</a>	More capable than any GPT-3.5 model, able to do more complex ta...

1 - 25 of 31 Items Page 1 of 2 Items Per Page: 25

**AI\_0023**  
AI MODEL

[Overview](#) [Related Assets \(0\)](#)

**Description:**  
-

**PUBLISHED**

**Hierarchy:** ?  
Top Level

**Custom Attributes:** ?  
Architecture Type: Model Format:

[Show More](#)

**Last Updated:**  
Nov 21, 2024, 10:46 AM

**Reference ID:** ?  
AI-36

**Stakeholders:** ?  
-

**Related Assets:** ?

<< Collapse

- New...
- Home
- Browse
- Data Access
- Saved Searches
- Customer Churn
- Personally Ident...
- CUSTOMER
- CRM\_CUSTOMER
- Customer Churn ...

Top Level

## Customer Churn Prediction

- Overview
- Relationships**
- Data
- Stakeholders
- Tickets
- History

OPEN TICKETS: 1 (1 task overdue) !

OVERALL RATING: ★★★★★ (1) Add Rating

LIFECYCLE: PUBLISHED Switch to Draft

LAST UPDATED: Nov 21, 2024, 10:50 AM

### Description

Customer retention is one of the primary growth pillars for products with a subscription-based business model. Competition is tough in the SaaS market where customers are free to choose from plenty of providers even within one product category. Several bad experiences – or even one – and a customer may quit. And if droves of unsatisfied customers churn at a clip, both material losses and damage to reputation would be enormous. Model has been initially built using Order Data dataset

Related Dataset: <https://cdgc.dm-us.informaticacloud.com/asset/bb121141-6d2a-4289-8b7d-b14902b8449c?tabId=relationships>

### Attributes

**AI Model Purpose:**  
Classification. The goal of classification is to determine to which class or category a data point (customer & Order Data in our case) belongs to.

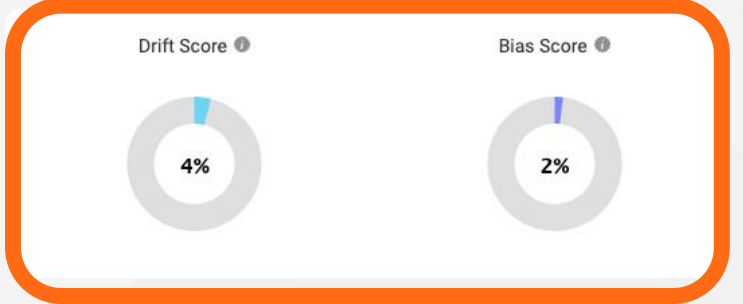
**Environment:**  
Training

**Model Rules:**

$$\Pi = Na [ \underbrace{\beta\gamma(b - c_{\text{contact}} - c_{\text{incentive}})}_{\text{Retained churners}} + \underbrace{\beta(1 - \gamma)(-c_{\text{contact}})}_{\text{Non-retained churners}} + \underbrace{(1 - \beta)(-c_{\text{contact}} - c_{\text{incentive}})}_{\text{False would-be churners}} ] - A$$

True would-be churners

[Link more details](#)



- > Contains
- > Is Part Of

### Responsibilities

- Governance Administrators  
Governance Administrator
- Data Scientists  
Data Scientist

[View All](#)

### Status

- + New...
- Home
- Browse
- Data Access
- Saved Searches
- Customer Churn ...

Top Level

### Customer Churn Prediction

AI MODEL

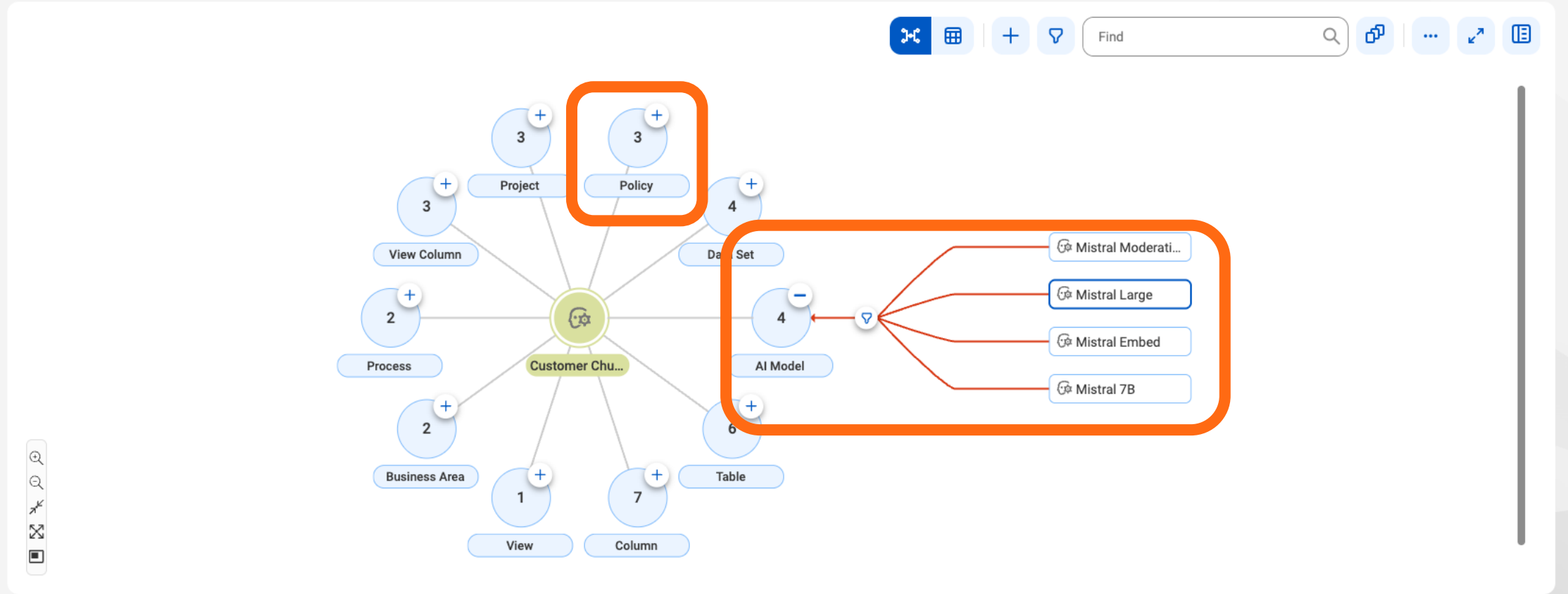
- Overview
- Relationships**
- Data
- Stakeholders
- Tickets
- History

OPEN TICKETS  
1 (1 task overdue) !

OVERALL RATING  
★★★★★(1)

LIFECYCLE  
**DRAFT**  
Switch to Published

LAST UPDATED  
Nov 26, 2024, 12:07 PM



<< Collapse

- + New...
- Home
- Browse
- Data Access
- Saved Searches
- Customer Churn
- Personally Ident...
- CUSTOMER
- CRM\_CUSTOMER
- Customer Churn ...

Top Level

### Customer Churn Prediction

AI MODEL

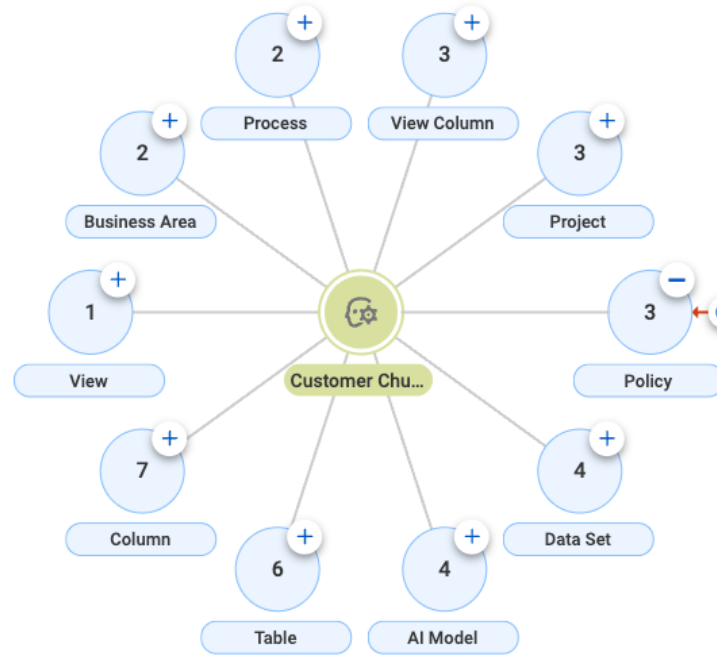
- Overview
- Relationships**
- Data
- Stakeholders
- Tickets
- History

OPEN TICKETS  
1 (1task overdue) !

OVERALL RATING  
★★★★★(1)  
Add Rating

LIFECYCLE  
**PUBLISHED**  
Switch to Draft

LAST UPDATED  
Nov 21, 2024, 10:50 AM



**Personally Identifiable Information (PII) Data Terms of Use POLICY**

There is personal data contained within the data collection so it should be treated inline with privacy regulations unless [Show more](#)

- Terms of Use /
- Personally Identifiable Information (PII) Data Terms of Use

- New...
- Home
- Browse
- Data Access
- Saved Searches
- Customer Churn
- Personally Ident...
- CUSTOMER
- CRM\_CUSTOMER
- Customer Churn ...

Terms of Use /

### Personally Identifiable Information (PII) Data Terms of Use

POLICY

OVERALL RATING: ★★★★★ (1) [Add Rating](#)

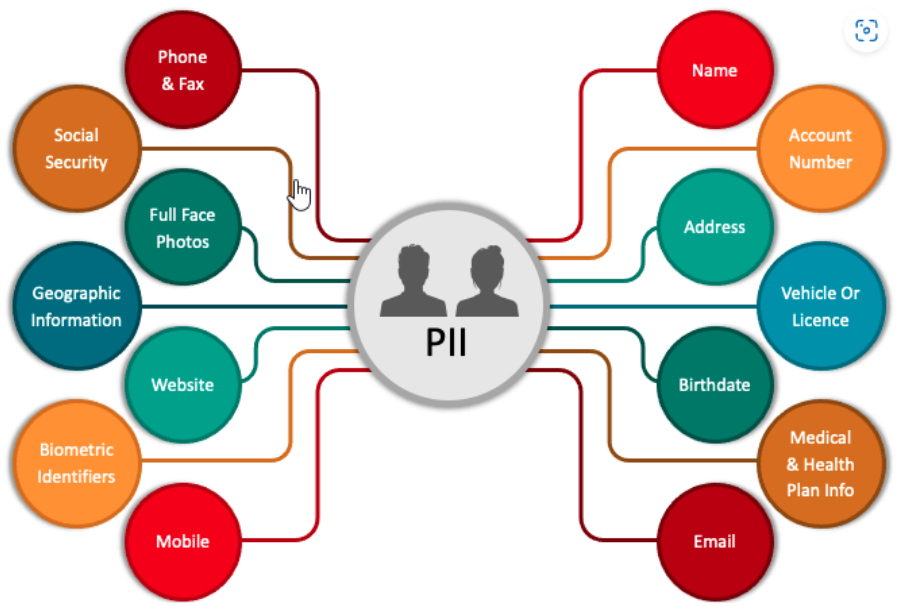
LIFECYCLE: **PUBLISHED**

LAST UPDATED: May 18, 2022, 5:55 PM

- Overview
- Hierarchy
- Relationships**
- Data Quality
- Stakeholders
- Tickets
- History

#### Description

There is personal data contained within the data collection so it should be treated inline with privacy regulations unless specific additional treatments are evidenced.



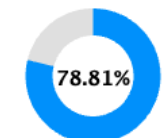
Show less

#### Stakeholders

- RD Rebecca Drake**  
Data Steward
- Governance Administrators**  
Governance Administrator
- Regulatory Office**  
Regulatory Officer

[View All](#)

#### Data Quality



↑ +0%

Mar 10, 2025, 9:24 AM

#### Rules Against Target



Good (1964)  
Acceptable (150)  
Not Acceptable (503)

[View Details](#)

- New...
- Home
- Browse
- Data Access
- Customer Churn
- Personally Ident...
- CUSTOMER
- CRM\_CUSTOMER
- Customer Churn ...

Top Level

# Personally Identifiable Information (PII) Data Terms of Use

## Privacy of Personal Data Policy

POLICY

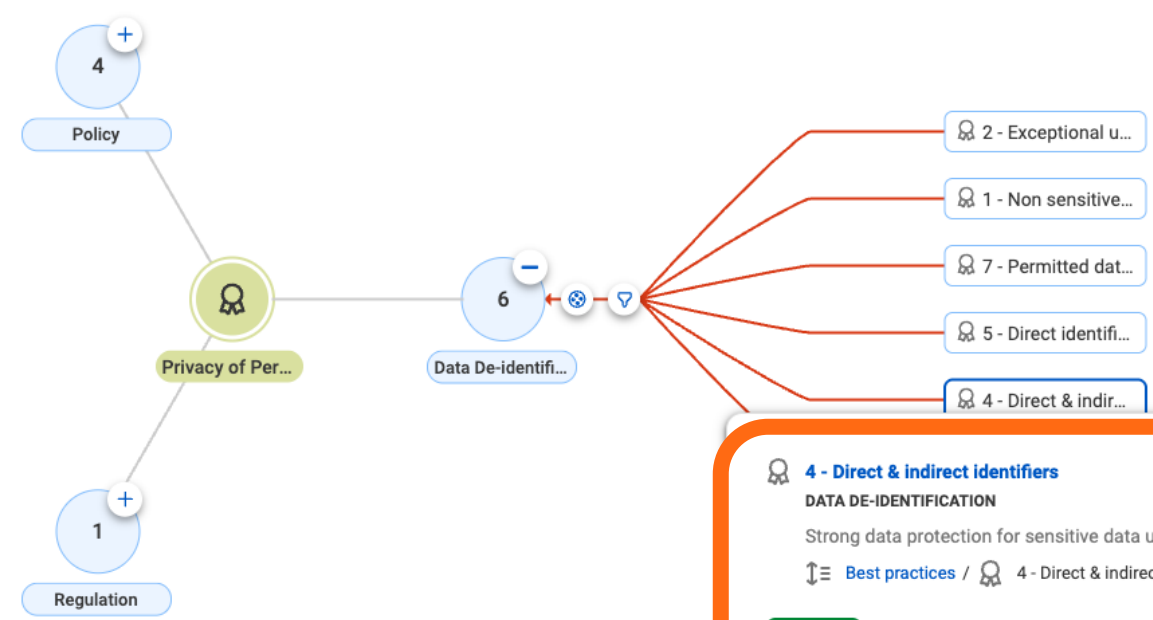
OVERALL RATING  
☆☆☆☆  
Add

LIFECYCLE  
**PUBLISHED**

LAST UPDATED  
Feb 15, 2024, 8:50 AM

- Overview
- Hierarchy
- Relationships**
- Data Quality
- Stakeholders
- Tickets
- History

Find



**4 - Direct & indirect identifiers**

**DATA DE-IDENTIFICATION**

Strong data protection for sensitive data usages context

Best practices / 4 - Direct & indirect identifiers

**Published**

<< Collapse

- + New...
- Home
- Browse
- Data Access
- Saved Searches
- Privacy of Person...
- 4 - Direct & indire...

Best practices /

### 4 - Direct & indirect identifiers

DATA DE-IDENTIFICATION

LIFECYCLE	STATUS	RULES	LAST UPDATED
PUBLISHED	ENABLED	2	Mar 11, 2025, 10:02 AM

- Overview
- Conditions
- Rules
- Relationships
- Stakeholders
- Tickets
- History

#### Rules (2)

##### PII direct identifiers

##### Conditions

Or

Usage Context is any of DAM - Project with third party DAM - Project on new products & s... DAM - Management analysis

##### Field Level De-identification

Then

Class Tech Email Address protect with EMAIL

Class Location Coordinates protect with STARS6

<< Collapse

- New...
- Home
- Browse
- Data Access
- Saved Searches
- Privacy of Person...
- 4 - Direct & indire...

Best practices /

### 4 - Direct & indirect identifiers

DATA DE-IDENTIFICATION

LIFECYCLE: PUBLISHED STATUS: ENABLED RULES: 2 LAST UPDATED: Mar 11, 2025, 10:02 AM

- Overview
- Conditions
- Rules
- Relationships
- Stakeholders
- Tickets
- History

#### Rules (2)

Then

- Class Tech Email Address protect with EMAIL
- Class Location Coordinates protect with STARS6
- Class Location Street Address protect with STARS6
- Class Person Full Name protect with STARS6
- Class Person Identifier protect with STARS6
- Class Person Last Name protect with NAME
- Class Person Phone Number protect with KEEP4FROMLEFT
- Class Tech Internal ID str protect with ID

<< Collapse

- New...
- Home
- Browse
- Data Access
- Saved Searches

## Data Access Management

- Data Access Control
- Data Filter
- Data De-identification

- Data De-identification (7)
- Data Protection
- Precedence Tier

+ Find [ ] [ ] [ ] [ ] [ ]

Name ↑	Precedence Tier	Lifecycle	Status	Updated On	Descr
1 - Non sensitive information	Best practices	PUBLISHED	✓ ENABLED	Mar 10, 2025, 12:...	Define
2 - Exceptional use authorization	Best practices	PUBLISHED	✓ ENABLED	Mar 5, 2025, 1:24...	All cla
3 - Special category data	Best practices	PUBLISHED	✓ ENABLED	Mar 11, 2025, 10:...	All cla
4 - Direct & indirect identifiers	Best practices	PUBLISHED	✓ ENABLED	Mar 11, 2025, 10:...	Strong
5 - Direct identifiers only	Best practices	PUBLISHED	✓ ENABLED	Mar 5, 2025, 1:44...	Mediu
6 - Data transfer - Approved countr...	Best practices	DRAFT	⊖ DISABLED	Mar 5, 2025, 1:17...	TBD w
7 - Permitted data usage	Best practices	PUBLISHED	✓ ENABLED	Mar 11, 2025, 10:...	Low d

### 1 - Non sensitive infor... DATA DE-IDENTIFICATION

- Overview
- Related Assets (2)

**Description:**  
Define all classified data element that can be visible for all

PUBLISHED ✓ ENABLED

**Last Updated:**  
Mar 10, 2025, 12:24 PM

**Effective Date:**  
Mar 3, 2025, 4:00 PM

**Tier:**  
Best practices

**Reference ID:**  
ADI-1

**Stakeholders:**

GD Grace Day  
Data Access Owner

**Related Assets:**

<< Collapse

- New...
- Home
- Browse
- Data Access
- Saved Searches
- DWH\_CDAM
- Predicted\_Custo...
- bfsi\_demo
- Person\_v3

DWH\_CDAM / unity-na-dw / bfsi\_demo /

(Add Business Name)

TABLE Person\_v3

OVERALL RATING  
☆☆☆☆☆  
Add

LIFECYCLE  
PUBLISHED

DATA MARKETPLACE  
PUBLISHED

LAST UPDATED  
Mar 11, 2025, 9:22 AM

- Overview
- Contains
- Lineage
- Relationships
- Data Quality
- Data Observability
- Stakeholders
- Attributes
- Tickets
- History

Data Elements (22)

Curate Find [Filter] [Sort] [Settings]

#	Name	Type	Glossaries	Classifications	CLAIRE™ Re... Null   Distinct
1	businessId	Column		Tech Internal ID str	
2	sourceSystem	Column		Business Entity Code	
3	status	Column		Business Type Status	
4	createdBy	Column		Person Full Name	
5	updatedBy	Column		Person Full Name	
6	mergedBusinessId	Column		Other Unknown	
7	firstName	Column	First Name (98.7%)	Person First Name	
8	lastName	Column		Person Last Name	
9	fullName	Column	Full Name (98.5%)	Person Full Name	
10	prefixName	Column		Person Title	
11	maritalStatus	Column	Marital Status (98.5%)	Person Marital Status	
12	birthDate	Column		Person Birth Date	
13	gender	Column	Gender (100%)	Person Gender	

businessId COLUMN

- Overview
- Related Assets (2)

Description:

PUBLISHED HIGH

Hierarchy: DWH\_CDAM / ... / Person\_v3

Related Data Collections: Customer segmentation & credit rating Preview

Data Quality:

Data Observability Events:

Profiling Stats:

Stakeholders:

Data Element Classifications:

<< Collapse