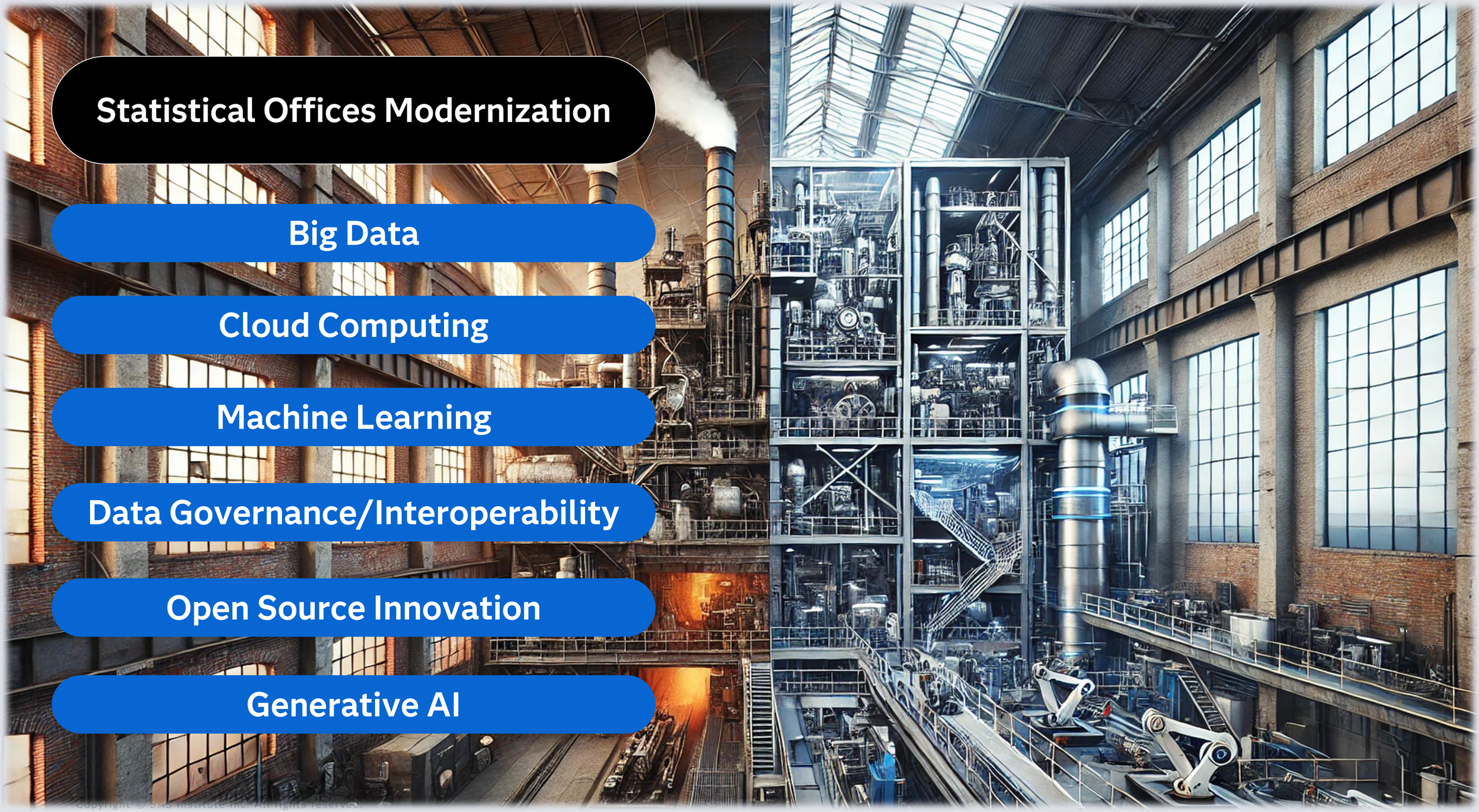
SAS presentation for Statistical Offices

Data and AI platform as a catalyst for statistical innovation

SAS Corporate

SAS Viya as a catalyst for innovation

1. Statistical Offices Modernization
2. Overview SAS Viya innovations
3. Key take-aways



Focus is on overcoming key challenges

Computational issues with big data.

Data sets continue to grow, and problems get more complex every day. You need a way to prepare and analyze diverse data, improve performance and get fast answers to any analytical question.

Pressure to provide scalable solutions within constrained budgets and rising cloud costs.

IT organizations with unlimited budgets for technologies just don't exist. Does an affordable, agile enterprise architecture exist that can keep pace with user demands?

Lack of analytical governance.

Open source technologies are dispersed, need customized support and often aren't integrated. This lack of control over data and analytics heightens risks associated with maintenance, hidden costs, governance and compliance.

Ever-changing everything.

With constantly evolving business requirements and an onslaught of emerging technologies, how do you determine the best long-term solution that can adjust to your changing needs?

Difficulty managing portfolios of diverse technologies.

Multiple tools are used for managing data and applications across organizations. As a result, it's hard to trace data lineage and know which version of a model is in use, whether it reflects current deployments and if it's being refreshed as needed.

Suboptimal technologies impede innovation

Slow tools and technology silos make it impossible to consistently discover insights and move from innovation to operations in a timely manner.

SAS Viya: An Open Platform for trustworthy Statistics, Data and AI

A Cloud-Native, Cloud-Agnostic Multi-language Data and AI platform

SAS Viya as a catalyst for innovation

1. Statistical Offices Modernization
2. Overview SAS Viya innovations
3. Key take-aways


Python
Python
R
SAS
1. OPEN-SOURCE/ML ACCELERATION

Open-Source Acceleration

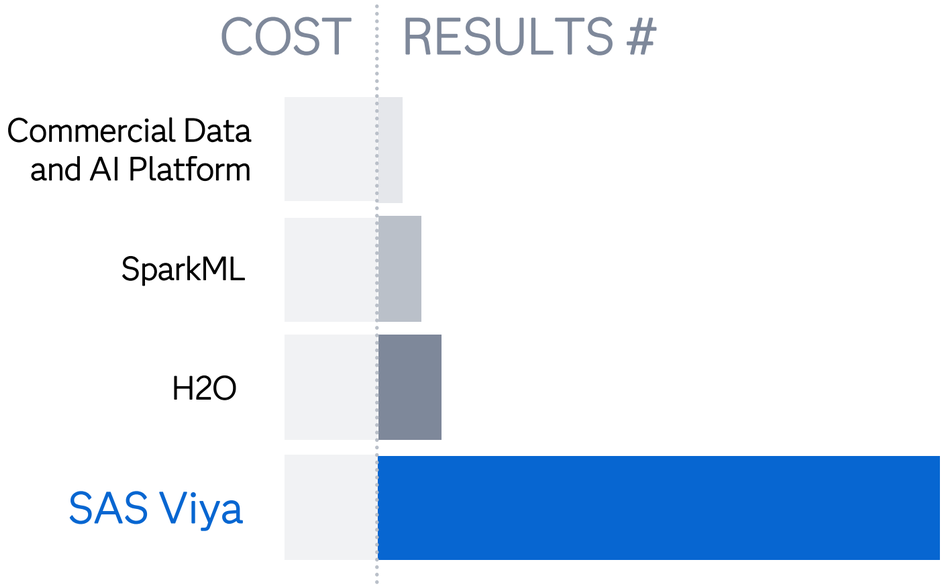
From SAS proprietary programming language to Python and R API's

```
from sklearn.ensemble import RandomForestClassifier

gb = GradientBoostingClassifier(n_estimators=100,
                                max_depth=5,
                                min_samples_leaf=1,
                                max_features=None,
                                learning_rate = 0.1,
                                subsample = 1.0,
                                random_state=0,
                                calc_feature_importances=True)
gb.fit(X_train, y_train)
    return go(f, seed, [])
```

```
from sasviya.ml.tree import GradientBoostingClassifier

gb = GradientBoostingClassifier(n_estimators=100,
                                max_depth=5,
                                min_samples_leaf=1,
                                max_features=None,
                                learning_rate = 0.1,
                                subsample = 1.0,
                                random_state=0,
                                calc_feature_importances=True)
gb.fit(X_train, y_train)
  return go(f, seed, [])
}
```

Distributed, Cost-Efficient Algorithms

Do More With the Same Infrastructure/Cloud Budget

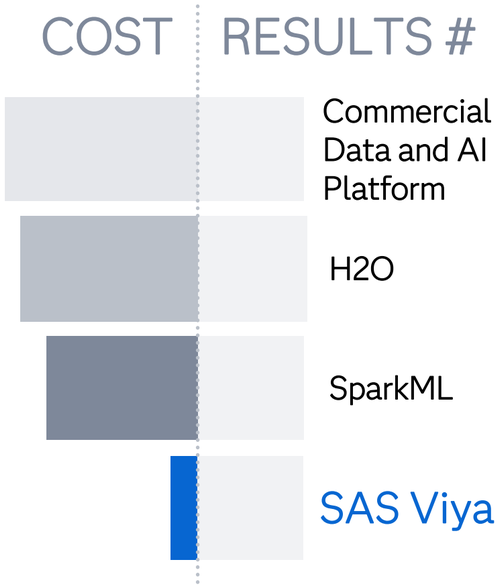


SAS Viya delivers results

30x faster on avg at the same cost

The Futurum Group

Do the Same Work at a Lower Infrastructure/Cloud Cost



SAS Viya delivers

+86% lower operating costs

SAS Viya Workbench Architecture

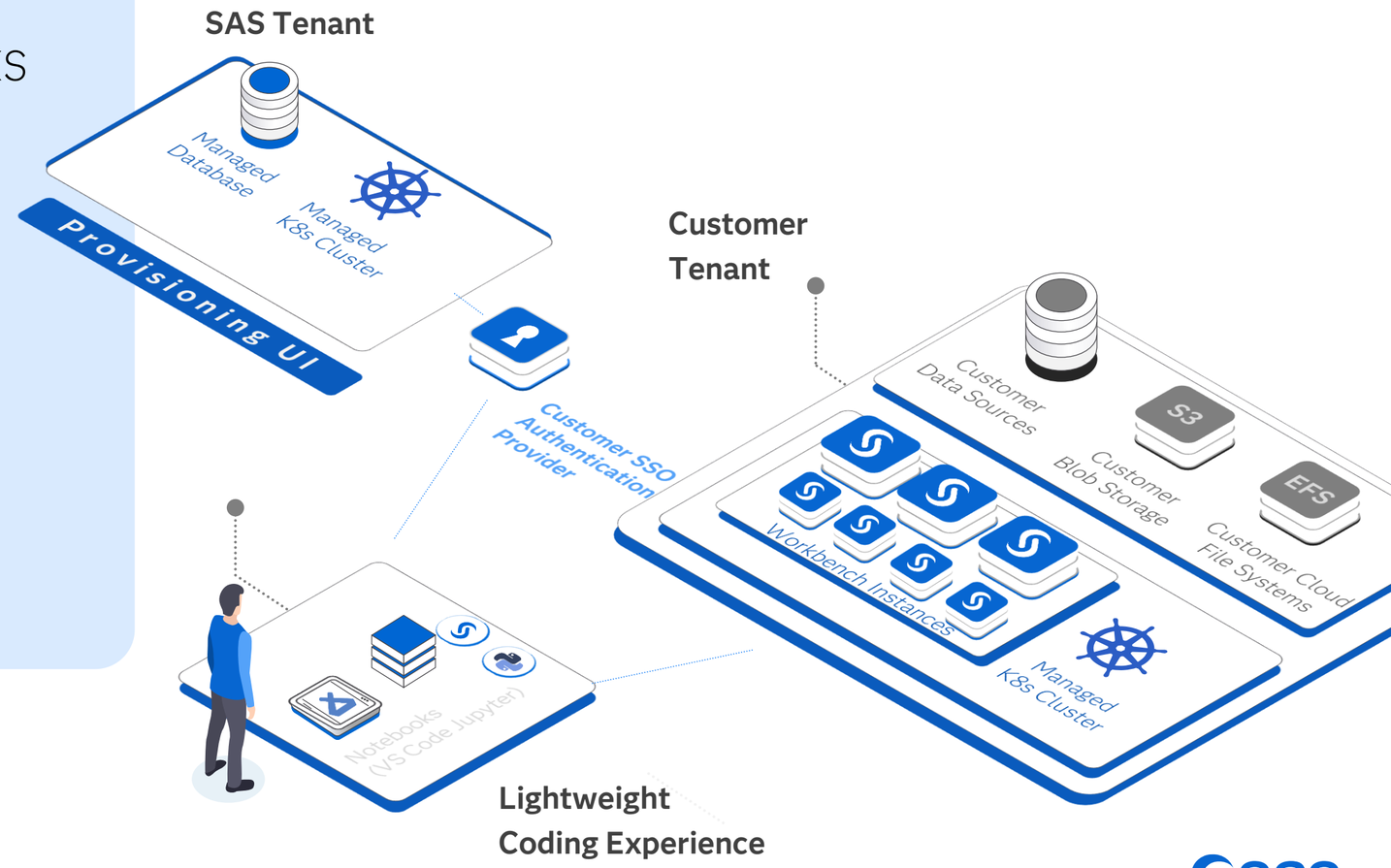
A light-weight canvas for developers or coders

- Self-service on-demand coding environment
- 100% elastic compute cloud, built on AWS EKS
- *(MS Azure is on the roadmap)*
- Managed deployment, update and backup
- VS Code and Jupyter Notebook interfaces

  (Enterprise Guide is on the roadmap)
- SAS and Python compute runtimes

  (R is on the roadmap)
- Coverage on most SAS9/Viya procs

  (SAS ETS/IML/….. on the roadmap)
- Availability of popular Python packages



SAS® Viya® Workbench

2. TRUSTWORTHY DATA AND AI

Trustworthy AI Landscape

ACTIVATION

Oversight
AI Governance, Strategy, and Enforcement

Operations
SOPs w/ Supporting Infrastructure

Compliance
Performance and Risk Management

Culture
Ethically Systemic Norms & Practices

TECHNOLOGY

Data Management
- Data Quality
- Variable Metadata
- Data Preparation
- Data Asset Catalog

Explanation
- Natural Language Explanation
- Explainable ML
- Counterfactual Explanation
- Surrogate Model Interpretation
- Causal Inference

Detection
- Bias Detection
- Fairness Assessment

Privacy and Security
- Privacy Preservation
- Model Security
- Autonomy Preservation
- Consent & Control

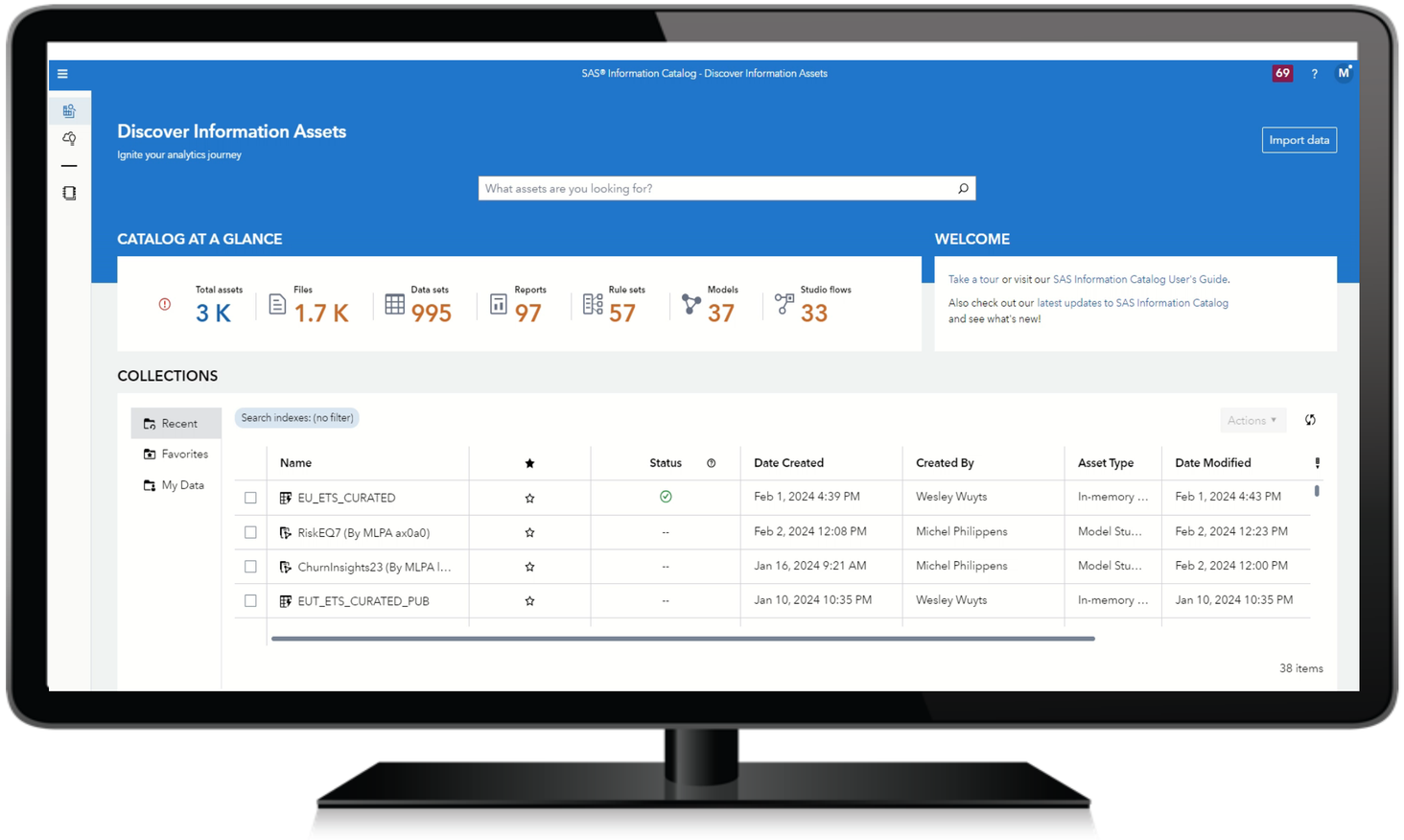
Mitigation
- Bias Mitigation
- Bias Prevention
- Synthetic Data Generation

Model Ops
- Model Cards
- Decisioning
- Lifecycle Management
- Metric Monitoring
- Model Robustness
- Model Oversight

SAS Viya supports FAIR data management principles

SAS Information Catalog facilitates and automates metadata capture

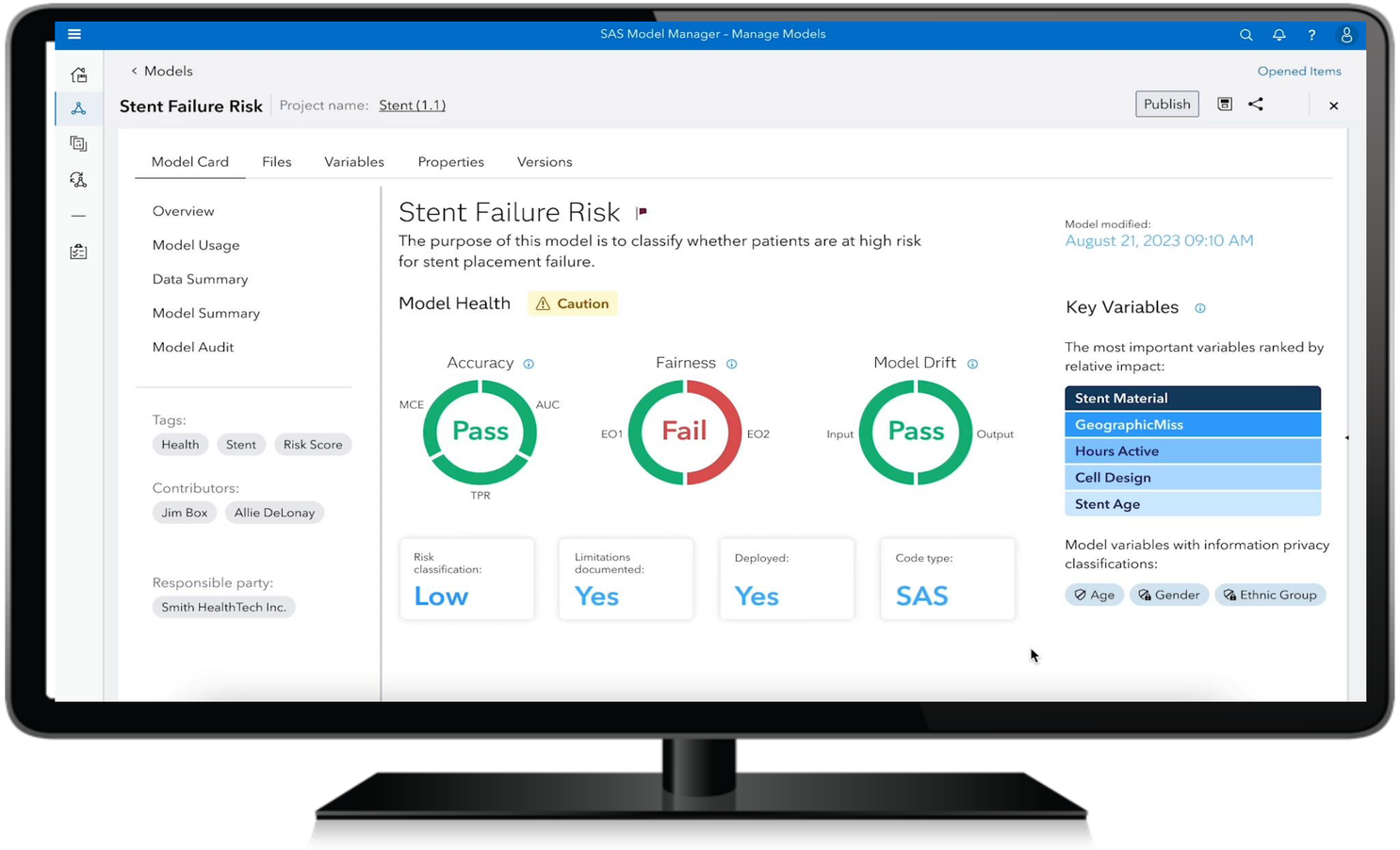
1. Automatic extraction of metadata for datasets
2. Automatic Semantic/Privacy Classification
3. Natural Language search for data sources (powered by Elastic Search)
4. Lineage Viewer
5. Semantic enrichment/glossary
6. Graph-based/Cypher-based Catalog API



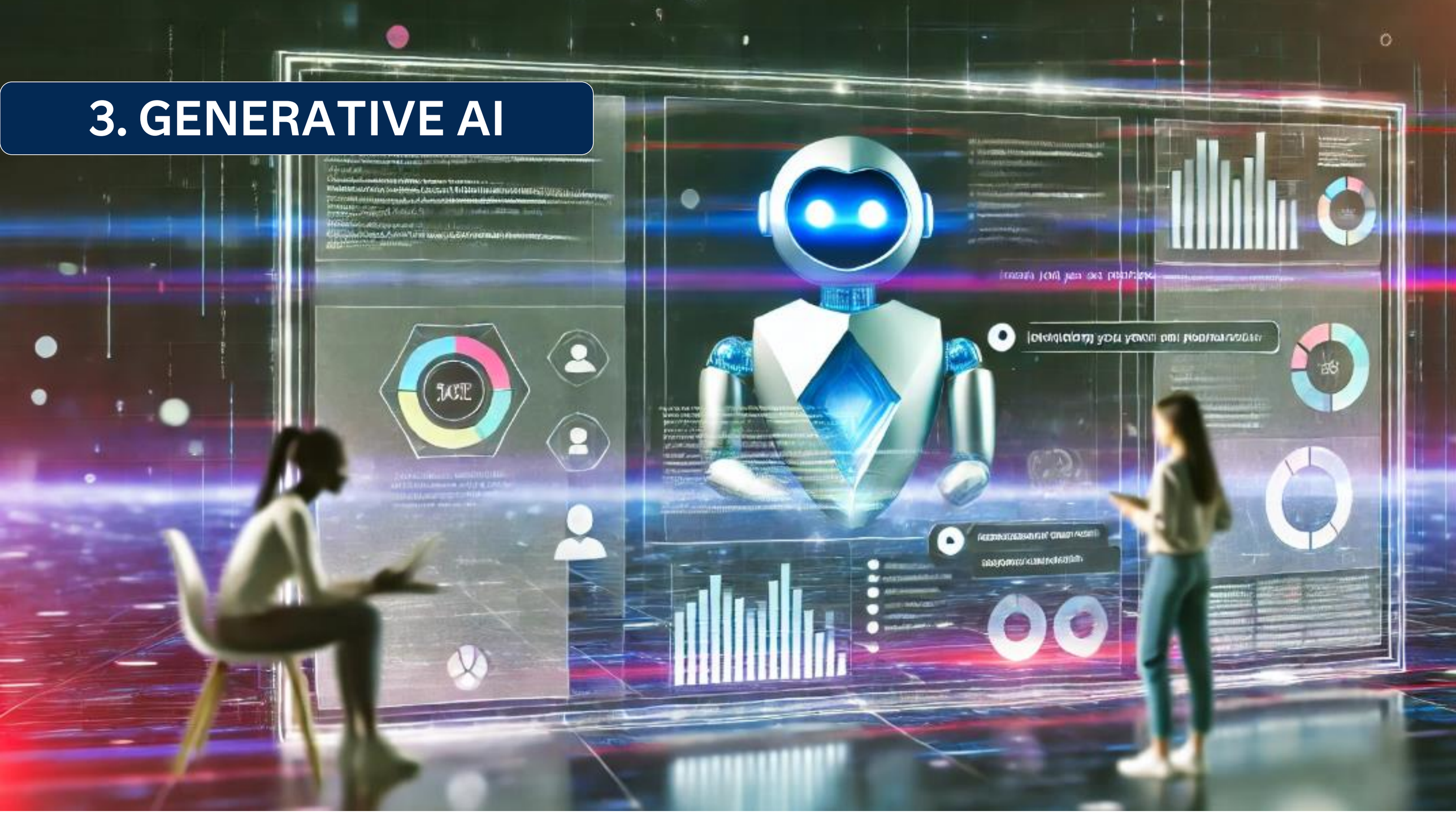
Trustworthy AI and ModelOps

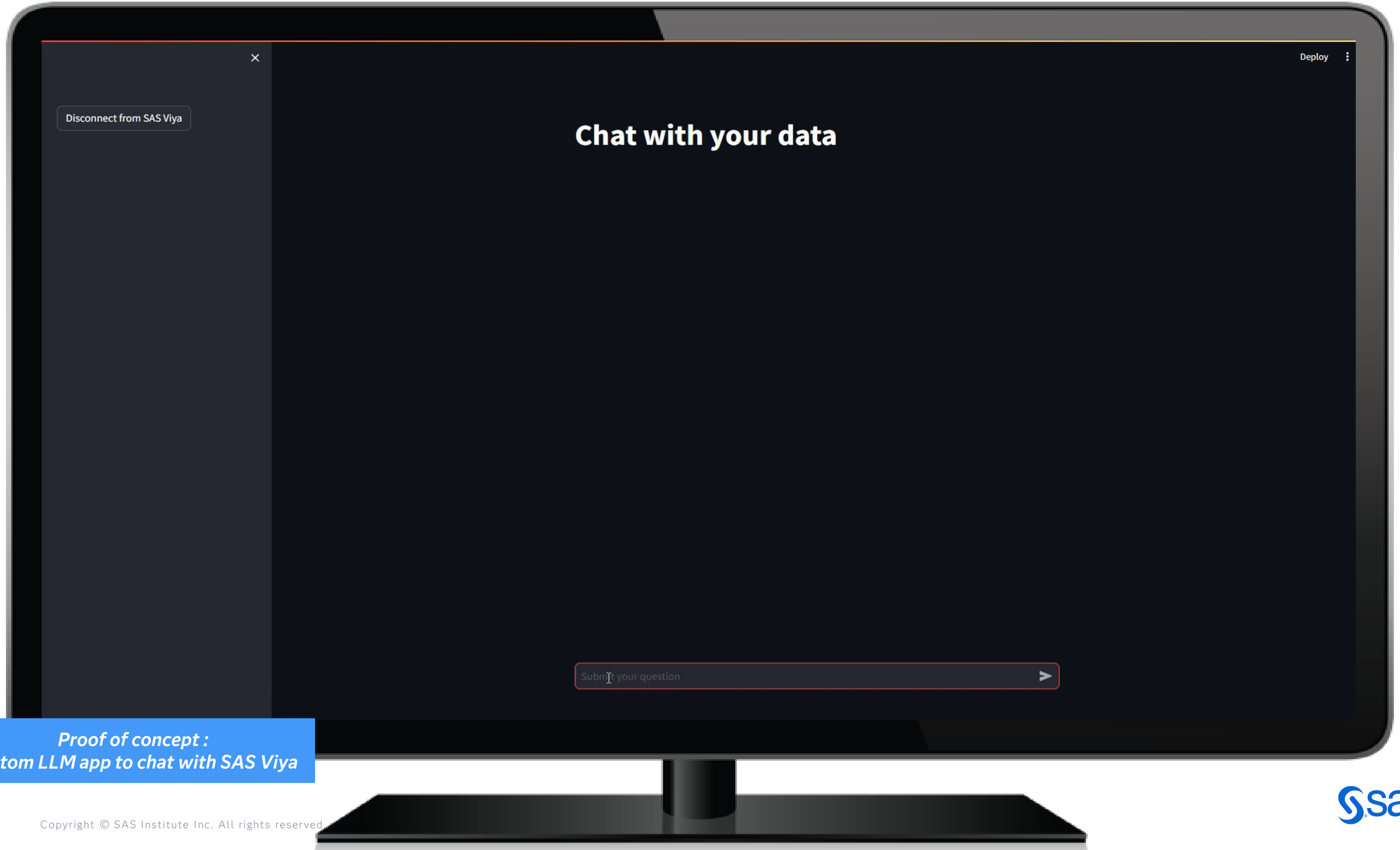
Model Cards : A nutrition label for your AI/ML models

1. Model Repository for SAS and Open-Source Models
2. Model Cards
3. Governance Workflow based on EU AI ACT principles
4. Deployment to SAS-independent scoring containers



3. GENERATIVE AI

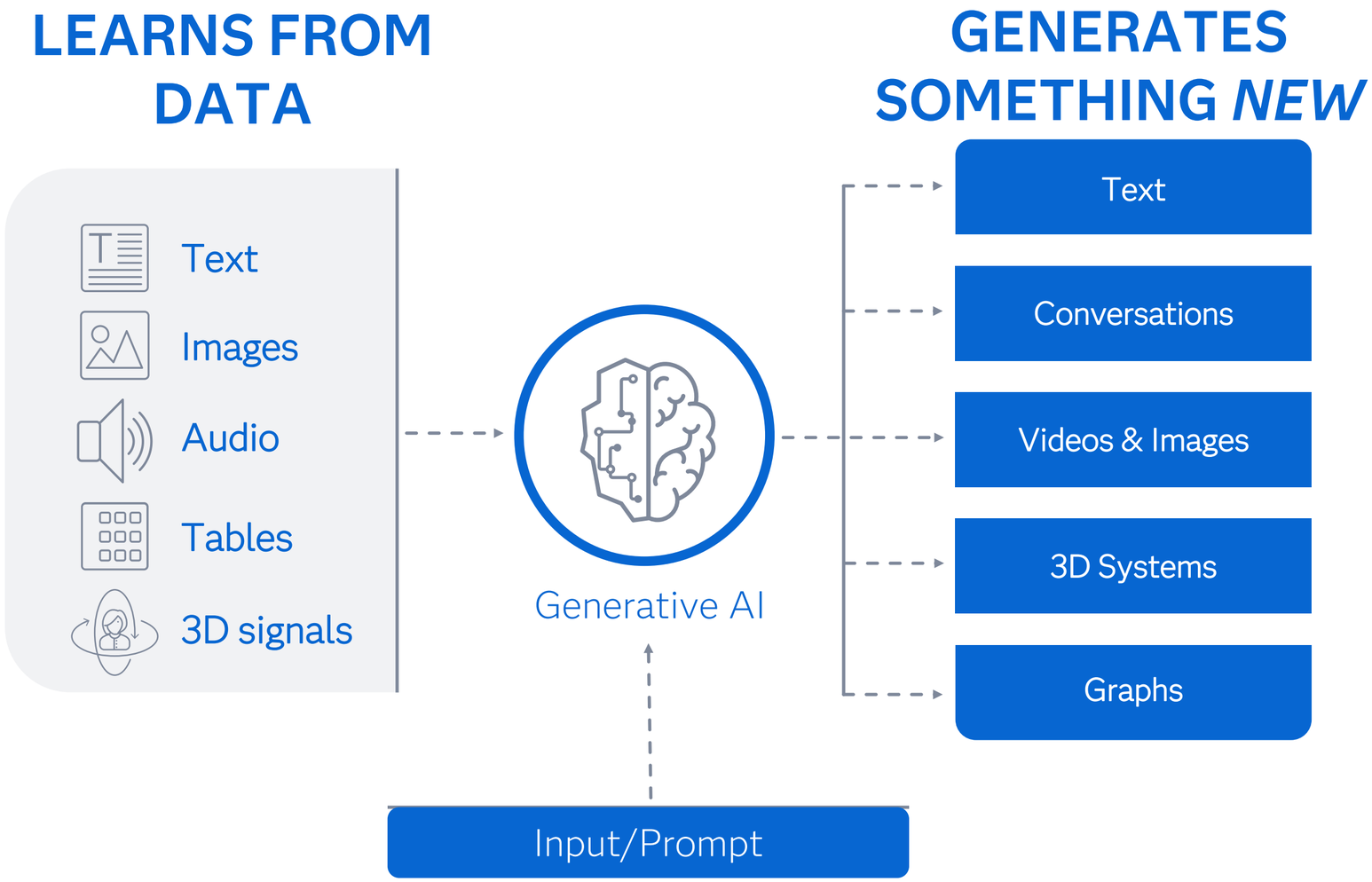




Proof of concept :
Custom LLM app to chat with SAS Viya

Creating New Realities

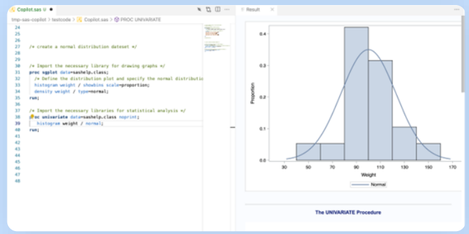
Generative AI



SAS and Generative AI (GenAI)

SAS Viya Copilot
to accelerate user productivity

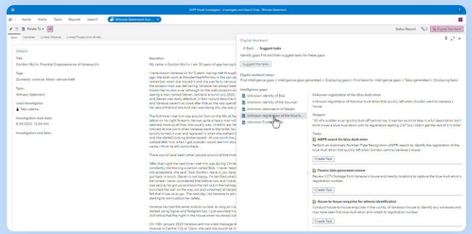
- Conversationally query SAS Viya to increase user productivity, further democratizing analytics



example: SAS Code Generation

Specialized
SAS Viya Copilots
to accelerate industry workflows and processes

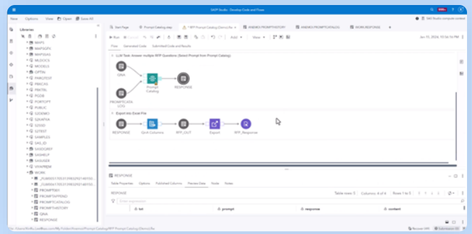
- GenAI driven assistants to accelerate domain-specific workflows



example. Law Enforcement Investigations

SAS Viya to help build your LLM application stack to unlock value in your proprietary business assets

- Flexible SAS Viya GenAI building blocks to help with development and orchestration of your LLM application stack



example. Build Prompt Library using existing SAS Viya APIs and SAS Studio steps

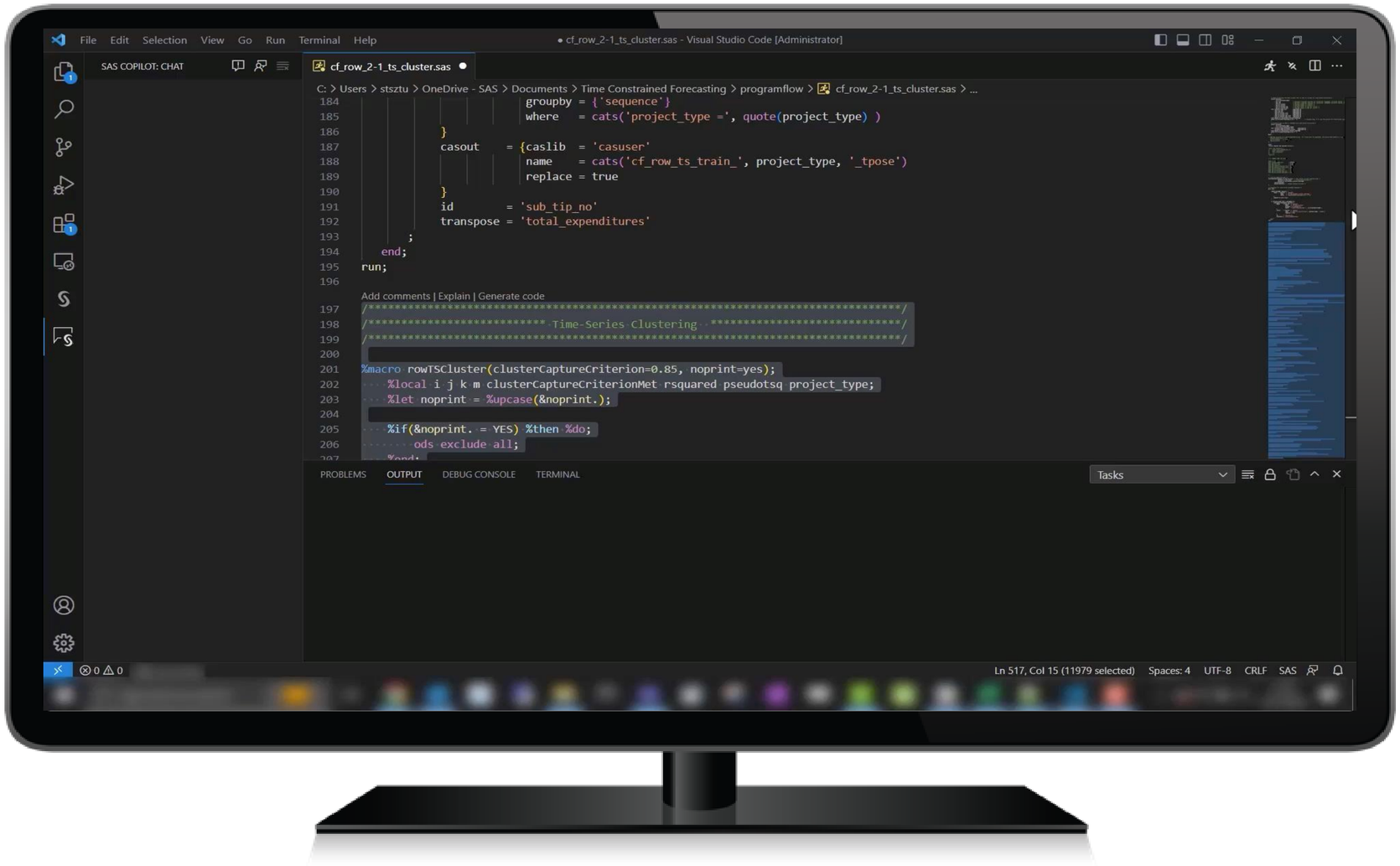
Standalone GenAI services
to enable your toolchains, like a **synthetic data generation service**

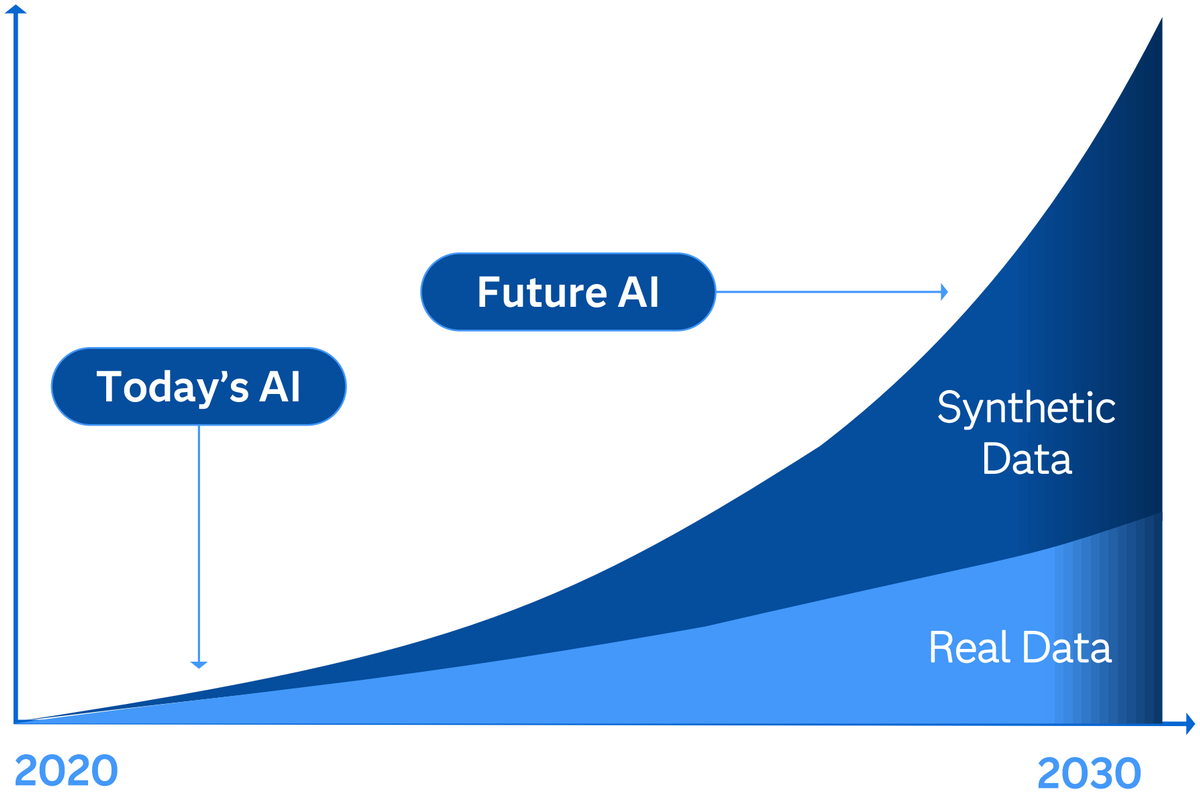
- API-first low-code/no-code synthetic data generation and validation experience

example. Synthetic data generation API

SAS Viya Copilot for Code (Customer Preview)

A finetuned LLM for Explaining and Generating SAS Code



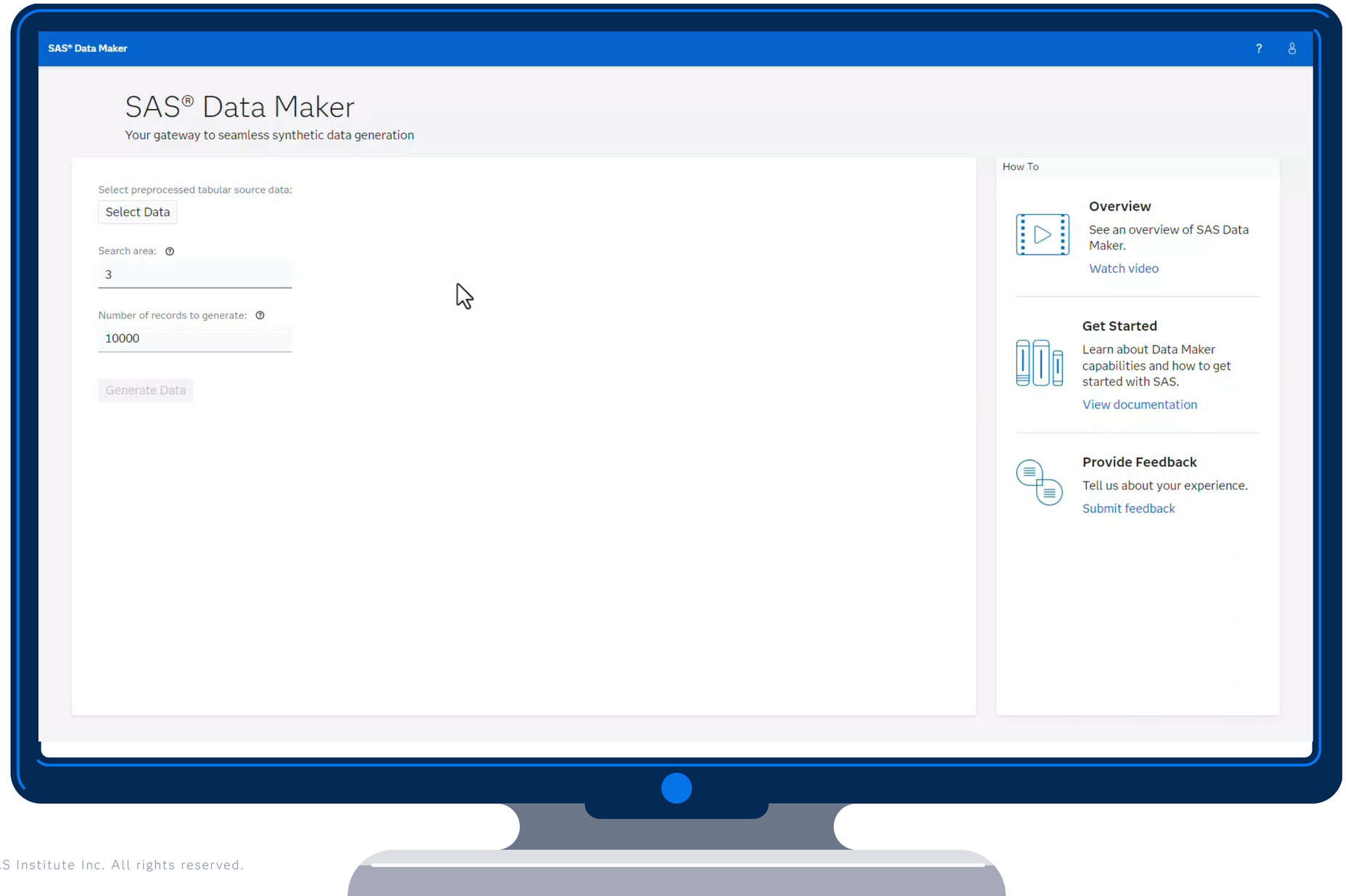


BY 2030, SYNTHETIC DATA WILL OVERSHADOW REAL DATA IN AI MODELS

Source: Maverick Research: Forget About Your Real Data — Synthetic Data Is the Future of AI, Gartner, June 24, 2021

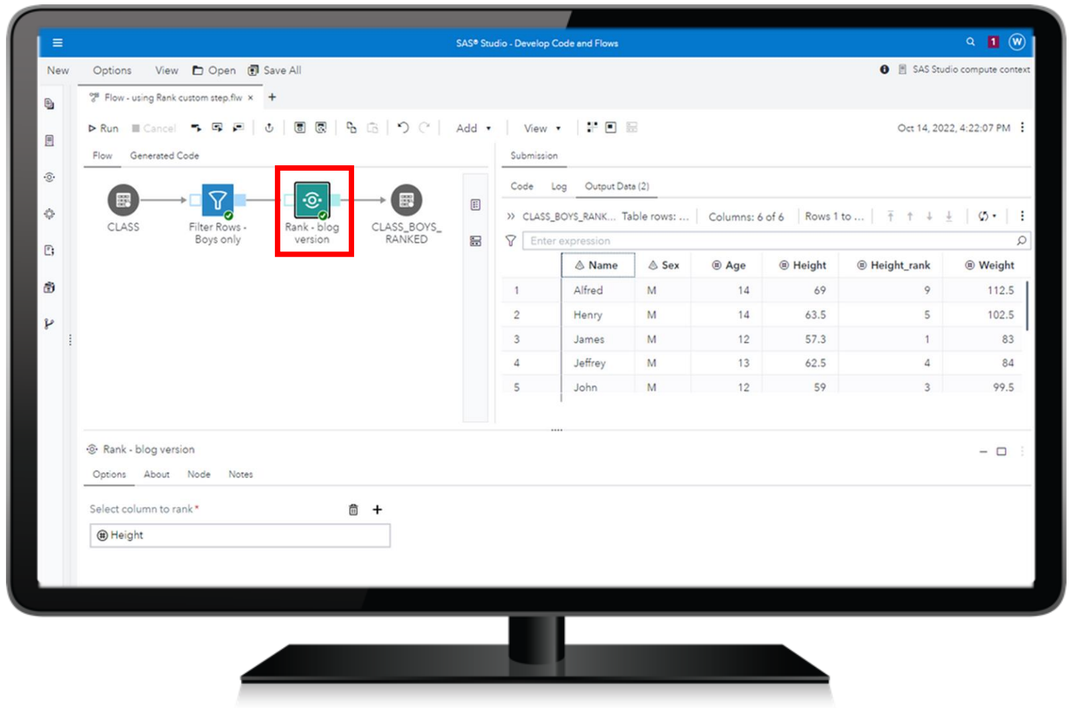
SAS Data Maker Demo

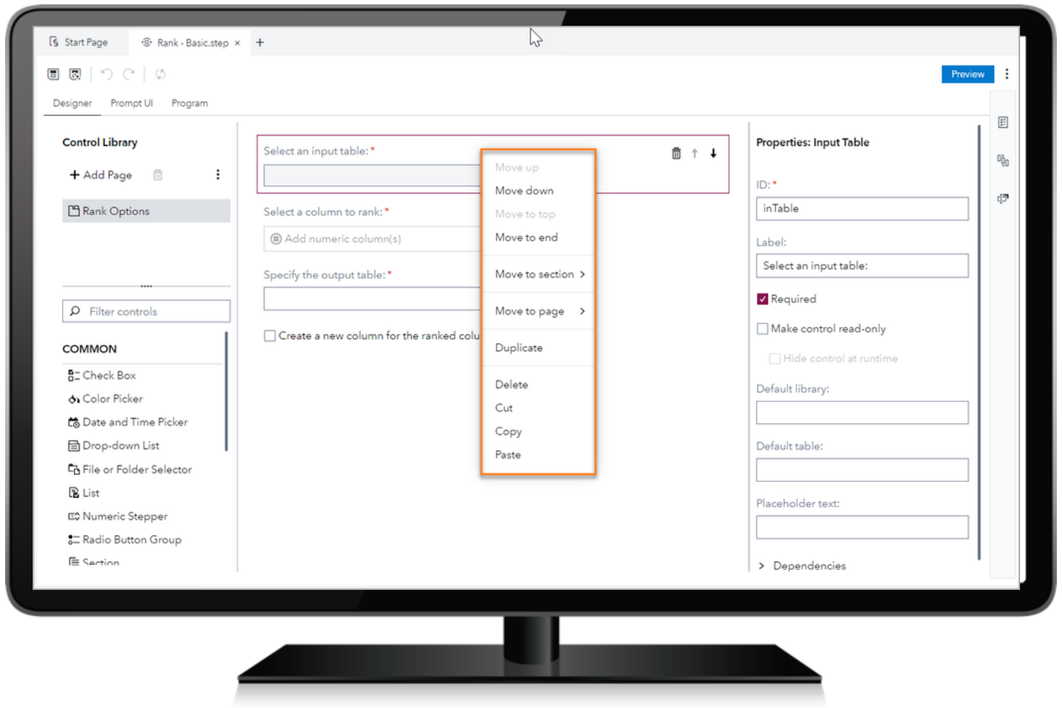
Now in Private Preview



SAS Studio Custom Steps

Custom Steps are Apps (UI) that complete a specific data/statistics/AI task





Custom Steps enable you to create a user interface to complete a specific task. They can then be used in multiple SAS Studio Flows to ensure a common, repeatable process. Custom Steps can also be exported as a JSON file and shared with other sites or be part of an internal DevOps process

Modernizing Statistical Production

eurostat

Eurostat Viya Modernization Project

PROJECT GOALS

- Modernize Existing (SAS-based) framework for statistical production
- Move to Cloud-Native Infrastructure and Data Storage (S3) on a private cloud
- Adopt DevOps-CI/CD methodologies and technologies (Github, Bamboo) for automation and better versioning
- Pilot new approach on EUROFARM

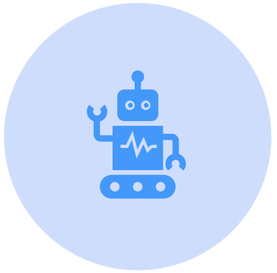
SAS Viya as a catalyst for innovation

1. Statistical Offices Modernization
2. Overview SAS Viya
3. Conclusion

Conclusion

How to navigate the myriad of challenges in Stat Office transformation?

FROM “PROGRAMMING LANGUAGE” TO PERFORMANT AND CLOUD EFFICIENT DATA AND AI PLATFORM



FROM DATA PROCESSING TO METADATA PROCESSING ON DATA AND AI ALGORITHMS USING DATA AND AI CATALOGS

LEVERAGE GENERATIVE AI TO BOOST PRODUCTIVITY

START WITH MODERNIZING YOUR CLASSICAL STATISTICAL PRODUCTION PROCESSES AND ADOPT NEW WAYS OF WORKING