

The background of the slide is a dark, futuristic digital landscape. It features glowing blue and white data visualizations, including bar charts and line graphs. Scattered throughout are icons representing artificial intelligence, such as a brain with a gear and a person silhouette. A central focus is a glowing red and white sphere with a complex, crystalline structure inside, which is being touched by a human hand on the left and a white, articulated robotic hand on the right. The overall aesthetic is high-tech and data-driven.

zensar

Zensar's Data Buddy: Synthetic Test Data Generator

GenAI led synthetic data accelerator

Why do we need Synthetic Data

What are the businesses doing?

Fueled by privacy concerns and the need for diverse training data, the synthetic data market is booming. Businesses are projected to spend billions on solutions that create realistic, artificial data sets for AI development and testing.

What are the challenges?

- **Data Privacy & Ethics** Real-world data can be sensitive
- **Data Scarcity:** Limited access to real data for specific needs hinders development
- **Cost & Time:** Acquiring and labeling real-world data can be expensive and time-consuming
- **Data Bias:** Real-world data can be biased, impacting AI models
- **Corner Case Coverage:** Capturing rare scenarios in real-world data is difficult
- **Explainability & Transparency:** Understanding how models make decisions based on real-world data can be challenging

Throwing light on some facts

60%

of data used for AI projects will be synthetic by next 2 years as per *Fortune Business Insights*

\$2.1
bn

Market size by 2028 as per *MarketsandMarkets* due to its' cost-effectiveness

What is Data Buddy: Synthetic Test Data Generator

zensar

About us: Zensar is a global technology consulting and services company with a strong presence across 30+ locations and a team of over 11,500 associates. We empower over 145 leading enterprises to drive disruption, agility, and competitiveness through our expertise. Our focus lies in conceptualizing, designing, engineering, marketing, and managing digital products and experiences for high-growth companies seeking innovation and velocity.

Our offering, *Data Buddy: Synthetic Test Data Generator* is a generative AI-enabled service. It helps

- businesses to unleash the synthetic test data generation potential
- creates realistic, privacy-preserving test data for multiple test scenarios for specific models
- allows focus on fine tuning the model than generating data
- cut across multiple industries and can be repurposed across environments

With our team of experts and AI-driven efficiencies, we accelerate the time-to-market intervals, enhance innovation, and deliver high-quality solutions for our clients.

Data Buddy: Synthetic Test Data Generator Key Features

zensar

Unleash Generative AI to skyrocket engineering testing velocity & deliver high-quality results

01. Data Simulation

- Replicate real-data patterns and relations – control format
- Create diverse representative datasets, filling gaps where real data might be lacking

02. Multi-table Data Generation

- Handle complex datasets with multiple connected tables
- Relationships and dependencies between tables are preserved

03. Boundary Conditions

- Optimal representation of corner/ edge cases
- Enhance the test data coverage

04. Privacy Preservation

- Mimic real data without exposing sensitive information
- Bias reduction in data

05. Model Reusability

- Scalability: Fine-tune model for different fields as required
- Increases standardization to accelerate development workflow

Data Buddy: Synthetic Test Data Generator Benefits

Data Buddy fuels data-driven innovation with faster training and privacy-protected testing for informed decisions

01. Cost Reduction

~60% reduction

in data acquiring & labelling costs as compared to traditional method

02. Quick-model Testing

2x to 5x

Model training time acceleration due to augmentation of synthetic data to existing datasets

03. Privacy Compliance

100 %

Eliminates the risk of privacy breaches associated with real-world data

04. Model Performance Improvement

5-10% increase

in accuracy due to creation of diverse and unbiased datasets

05. Time-to-market Advantage

10-20% reduction

Faster time-to-market due to faster development cycles

* % numbers are average and based on our past experiences and could change based on scenarios and complexities

Potential Use Cases By Industry



Finance

- Assess fraud & risks
- Test trading algos
- Simulate market scenarios
- Privacy-safe Analytics.



Autonomous Vehicles

- Simulate diverse & challenging driving scenarios
- Faster, cheaper training and validation of self-driving car algorithms



Retail/ E-commerce

- Simulate customer behavior for
 - Demand forecasting
 - Inventory manage
 - Personal marketing
 - Recommendation systems
 - Dynamic pricing



Healthcare

- Validate new treatments/diagnostics
- Protect privacy while enhancing AI training



Entertainment

- Create new genres of music, unique videos etc.
- Film restoration by enhancing quality



Cybersecurity

- Train & test threat detection systems
- Strengthen defenses & enable privacy-safe security analytics

How we do it – 5 Day Plan

Zensar's structured approach to Data Driven Innovation via Synthetic Data Generation is executed with below steps

Day 1: Assess



Understand Data Challenges

- **Data Challenges:** Identify key challenges faced such as privacy, scarcity, quality, and bias
- **Business Impact:** Assess the impact of on customer's operations and decision-making processes
- **Data Requirements:** Understand the specific data requirements

Day 2: Evaluate



Explain Syndata

- **Module Demonstration**
- **Synthetic Data Concept:** Concept and benefits in addressing challenges faced
- **Use Cases:** Discuss potential use cases within the customer's industry and domain
- **Generation Process:** Provide an overview of generation process and technologies involved

Day 3: Review



Data Assessment and Requirement

- **Data Analysis:** Evaluate customer's existing data to identify potential areas for Syndata
- **Privacy and Security:** Discuss concerns and how Syndata can address them
- **Requirements:** Define the specific requirements for Syndata based on the customer's needs

Day 4: Rationalize



Syndata Generation & Benefits Analysis

- **Data Generation:** Generate requirement based synthetic data samples
- **Quality Assessment:** Evaluate quality and realism of synthetic data
- **Model Performance:** Assess the performance of AI models trained on synthetic data compared to real-world data

Day 5: Next Steps



Wrap-up & Next Steps

- **Key Findings & Value Proposition:** Summarize assessment and the potential value (quantify) our Syndata can bring
- **Address Concerns:** Answer remaining questions/ concerns
- **Steps Forward:** Discuss next steps, including potential pilots/ consulting as per customer needs

Why partner with us?

- We follow Privacy & Security by Design
- Varied skilled team of AI experts, Prompt Engineers & SMEs across geographies
- We leverage due diligence and explainable AI by design
- Train faster & shorten GTM
- Cost effective solutioning
- Right partner to create co-sell AI solutions
- Future-Proof AI by partnering with a leader in synthetic data generation for AI and stay ahead of the curve

zensar

Next Steps

zensar

Take your next step with us by scheduling a 30 min discussion with our experts.

Contact: data.buddy@zensar.com

Zensar helps businesses understand and utilize the benefits of generative AI through brainstorming and conversations with your team. By bringing together our AI specialists we will dive into your organization's unique needs and goals and help identify if our solution is suited to your specific business scenarios & if there is any other solution that we offer that can drive the most significant benefits from generative AI.

Together we can determine the next steps based on the results of the initial conversations. This may involve conducting a 5-day comprehensive assessment to evaluate the use case and the possibilities of pursuing a Proof of Concept (PoC) or Minimum Viable Product (MVP) build to demonstrate the value of the technology.

