

Deck

humans && machines



Founding team

Shalini Kurapati, PhD



CEO

Data Privacy, Data
Management, and ethics.
CIPP/E.

Luca Gilli, PhD



CTO

Generative AI, MLOps.

Matteo Giovannetti



COO

Project management,
Operations.



**We provide
Synthetic Data
solutions and
services.**

[Link](#): what's Synthetic Data

Synthetic data

Definition

Synthetic data is **fictitious data** artificially generated that incorporates the **statistical properties** and distributions of the original data, thus resulting **realistic**.

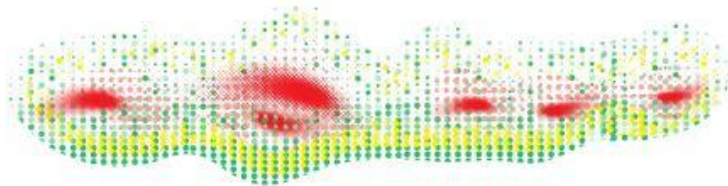
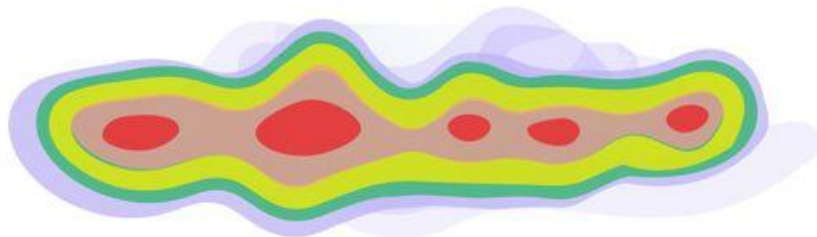
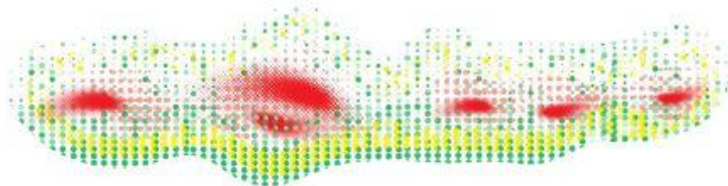
ORIGINAL DATA



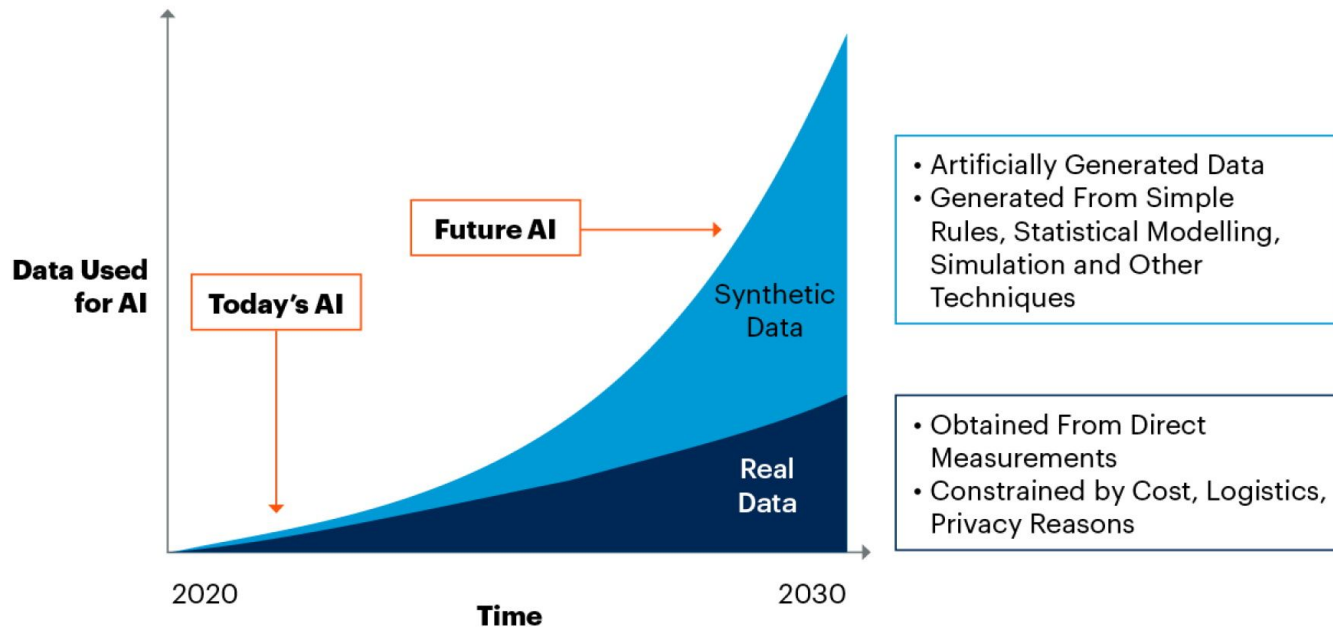
ESTIMATED
DISTRIBUTION



SYNTHETIC DATA

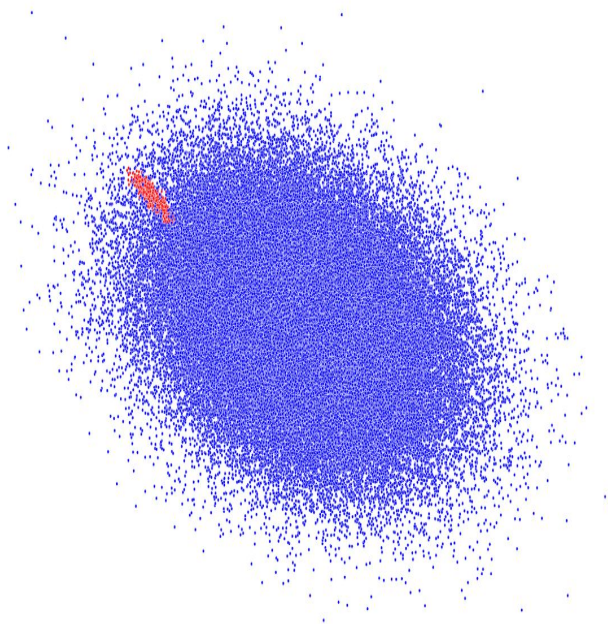


By 2030, Synthetic Data Will Completely Overshadow Real Data in AI Models



Source: Gartner
750175_C

Why Synthetic Data

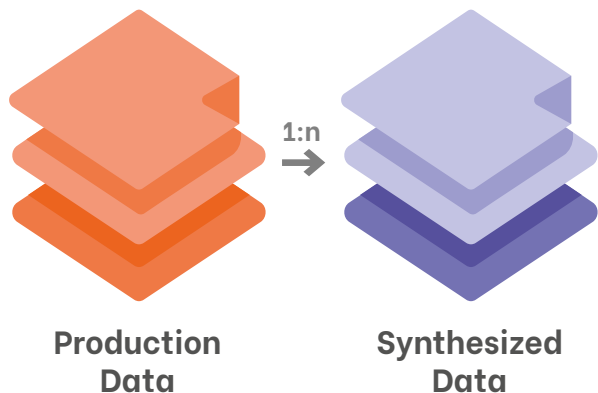


➤ To solve **Robustness problems** affecting AI Projects

- DATA SCARCITY
- CLASS IMBALANCE
- BIAS MITIGATION

**+15% MODEL
PERFORMANCES**

Why Synthetic Data



➤ *To scale and facilitate
Testing procedures*

- FUNCTIONAL TESTING
- PERFORMANCE TESTING
- CLOUD MIGRATION
- NEARSHORING

Regulators love Synthetic Data

- **GDPR:**
Data Minimisation principle
- **EDPS**
TechSonar
- **AI ACT**
It mentions Synthetic Data as a PET (Privacy-Enhancing Technologies) to facilitate Data Sharing and AI Model training



Clients and Partners



How

to work with Clearbox AI

01 Directly using our Enterprise Solution

- Target users: Data / Data Science / ML experts
- Poc + timed license
- On-prem or private cloud or managed/hosted deployment

02 On-Demand Generation Service

- Dedicated data generation made by Clearbox AI, based on a specific use case and coordinated generation strategy
- Data generation from an original dataset, from scratch or mixing different data sources

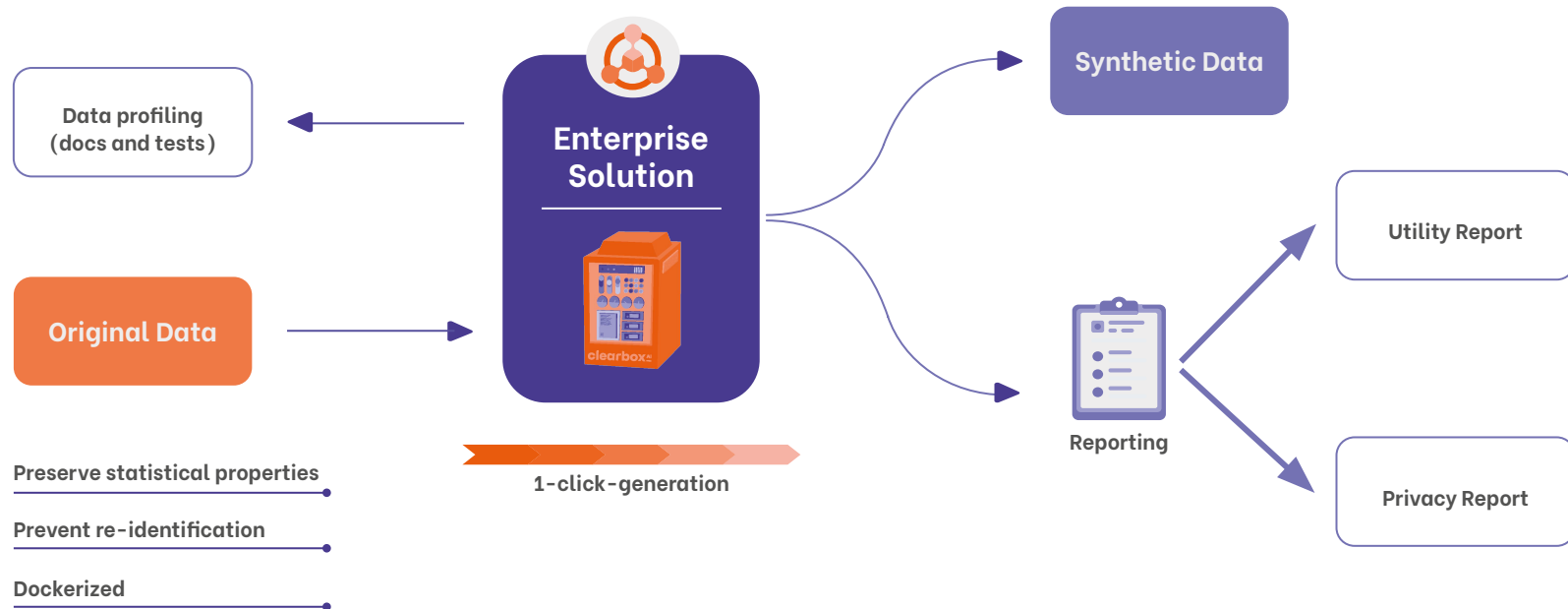
03 Dedicated Project

- Based on our core competencies: Generative AI, Synthetic Data and MLOps

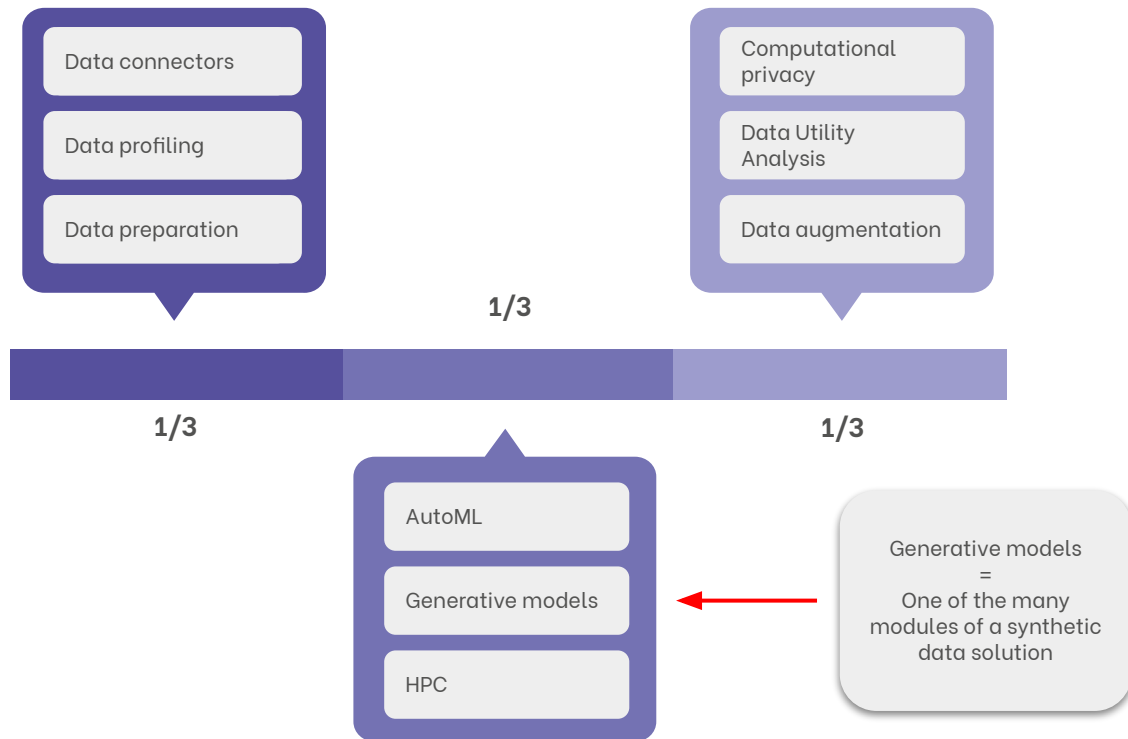
**Clearbox AI has developed
a product based on
Generative AI that
automates synthetic data
generation.**

Our Enterprise Solution

The technology has been packed in a scalable product



Synthetic Data as a Product



Some use cases

links

01 Improving ML models with synthetic data: [link](#)

02 A Synthetic Data Sandbox: [link](#)

03 Automated and more secure software testing: [link](#)

04 Overcome Data Retention periods with Synthetic Data: [link](#)

05 Augmenting historical data with synthetic time series: [link](#)





Thanks for Reading

Feel free to contact us:



www.clearbox.ai



info@clearbox.ai



[@ClearboxAI](https://twitter.com/ClearboxAI)