AI Synthetic Data Companies

MDCLONE

Synthetic data for healthcare

datagen

Human-centric synthetic data for computer vision

FINCRIME DYNAMICS

Financial fraud prevention and compliance control

Synthesis AI

Human and environmetal synthetic data for computer vision

Synthetic data generation platforms

hazy

gretel

TONIC

Datomize

Specialized

General Purpose

parallel domain

Synthetic data for outdoor perception models

SYNTHESIZED

Diveplane

Neurolabs

Consumer packaged goods image generation for computer vision

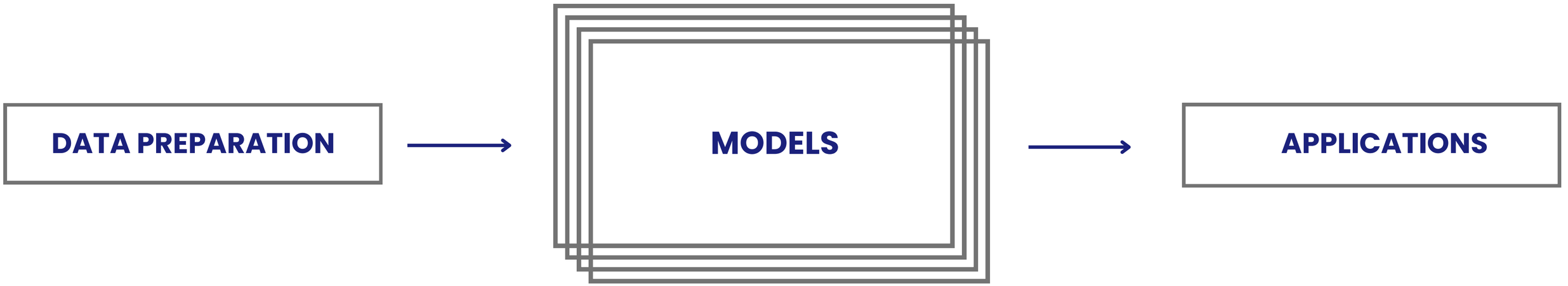
MOSTLY AI

RENDERED.AI

Physics-based synthetic data for computer vision

The Strategy Deck

Synthetic Data in the ML Value Chain



1 Identify gaps & biases in real world datasets

2 Replace sensitive information with artificial data

3 Generate new data to augment training dataset

4 Adapt model for adjacent use cases & environments

5 Test and validate quality of model output

The Strategy Deck

Computer Vision Spans Human-Centric Use Cases and Outdoor & Environmental Perception

Human-Centric Computer Vision

Synthesis AI

Human and environmetal synthetic data for computer vision

datagen

Human-centric synthetic data for computer vision

Neurolabs

Consumer packaged goods image generation for computer vision

Outdoors and Environmental Perception

parallel domain

Synthetic data for outdoor perception models

RENDERED.AI

Physics-based synthetic data for computer vision

Company

datagen

- Full body humans in context
- **US$50M Series B** - Mar 2022
- HQ in **Tel Aviv, Israel**
- https://datagen.tech/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Generate **human-centric** synthetic data for **computer vision**
- **Synthetic data management platform** with generation, versioning and sharing capabilities
- Generate **fully annotated, full body humans** and realistic scenes of **humans interacting with objects** in various environments

Value Proposition & Use Cases

- Use cases: **in-cabin automative, home and building security, fitness, cosmetics, smart office**
- Accurate representations of the in-cabin environment for Driver and Occupant Monitoring Systems, hands-on wheel detection

Company

Neurolabs

- New Generation Image Recognition Technology for Retail
- **US$3.50M Seed** - May 2022
- HQ in **Edinburgh, UK**
- https://www.neurolabs.ai

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Vision technology for **consumer packaged goods**
- **Image recognition datasets for over 100,000 SKUs**
- High accuracy for SKU-level product recognition from day 1
- ChatCPG(beta) - **retail shelf auditing tool** for insights into in-store visits and customer behaviour

Value Proposition & Use Cases

- Use cases: store set-up for CPG brands, insights gathering by field marketing agencies, sales force automation for retail suppliers

Company

Synthesis AI

- Synthetic data for computer vision to enable more capable and ethical AI
- **US$17M Series A** - Apr 2022
- HQ in **San Francisco, CA**
- https://synthesis.ai/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Synthesis Humans: generate **fully labeled humans in near-infinite combinations of environments** / lighting and sensor types
- Synthesis Scenarios: craft sophisticated **scenarios for multi-human ML models in diverse environments** and engaging in various activities

Value Proposition & Use Cases

- Use cases: **avatar creation, pose understanding, ID verification, AR/VR/XR, driver monitoring, pedestrian detection, virtual try-on, teleconferencing, security**
- Rich labels, photorealistic images, diversity and control, variable camera locations

Company

RENDERED.AI

- The Platform as a Service for Creating Synthetic Data
- **US$6M Seed** - Oct 2021
- HQ in **Bellevue,WA**
- https://www.rendered.ai/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- PaaS for generating physics-based synthetic data for visible and non-visible computer vision appications
- Procedurally modeled **landscapes, vegetation, buildings, water bodies, and cities**
- Reality capture meshes and models
- **User and community contributed 3D assets**
- Sample sensor models to generate simulated optical, radar, IR and other types of imagery output

Value Proposition & Use Cases

- Use cases: **automotive, medical AI, robotics, satellite imagery, security**
- Enable collaboration between data scientists and data engineers with a no-code configuration experience

Company

- Accelerating machine perception with synthetic data
- **US$30M Series B** - Nov 2022
- HQ in **Vancouver, Canada**
- https://paralleldomain.com/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Generate synthetic data to **train and test perception models**
- Rich **scenario diversity**, **content fidelity** for a variety of regions, agents and environmental conditions, custom annotations, sensor simulation for cameras, LiDAR, radar
- Vizualizer: **visually explore datasets** and annotations to identify patterns and outliers

Value Proposition & Use Cases

- Use cases: **Advanced Driver Assistance, L4/L5 Automotive Autonomy, Drones / Aerial**
- Worlds contain deep synthetic complexity while maintaining the realism of the details needed to accurately model the imperfections of the real world

Specialized Synthetic Data Platforms (Healthcare, Finance) Compete with General Purpose Tools

Specialized

MDCLONE

Synthetic data for healthcare

FINCRIME DYNAMICS

Financial fraud prevention and compliance control

General Purpose

hazy

gretel

TONIC

Datomize

SYNTHESIZED

Diveplane

MOSTLY AI

Company

MDCLONE

- Synthetic Data for Healthcare
- **US$50M Series B** - Mar 2022
- HQ in **Tel Aviv, Israel**
- https://www.mdclone.com/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- MDClone ADAMS Platform is a **self-service data analytics environmen**t enabling healthcare collaboration, research, and innovation.
- Synthetic data protects patient privacy while preserving maximum data utility for **faster research,** positive impact on **operational processes**, and improvement of patient outcomes while saving costs and resources.

Value Proposition & Use Cases

- Use cases: access and exploration of medical data, cross-organizational sharing and collaboration

Company

- Financial Crime Testing and ML Training
- HQ in **Cambridge, UK**
- **https://fincrimedynamics.com**

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Synthetic data generation for **financial crime prevention**
- We enable financial institutions to create tailored synthetic data with the latest **financial crime simulations** to build high performance machine learning crime detection
- Synthetizor: financial crime and **compliance control simulations and scoring**

Value Proposition & Use Cases

- Use cases: **banking, payments, financial and crypto exchanges**

Company

MOSTLY·AI

- Mission: We enable organizations to thrive ethically and responsibly with smart and safe synthetic data
- **US$25M Series B** - Jan 2022
- HQ in **Vienna, Austria**
- https://mostly.ai/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Synthetic data platform for: **smart data imputation, data augmentation, rebalancing, automated data privacy and diversity, data exploration**

Value Proposition & Use Cases

- Use cases: **data democratization, cross-border and enterprise data sharing, data anonymization, testing and product development, model bias removal**
- Banking, insurance, telecommunications, healthcare

Company

gretel

- We help developers build with data, together.
- **US$50M Series B** - Oct 2021
- HQ in **San Diego, CA**
- https://gretel.ai/

Partnerships & GTM

- Google Cloud - compute
- Illumina - genomics
- B2B SaaS

Tech Stack & Product

- **Generate synthetic datasets** with the same characteristics as real data, to develop and test AI models without compromising quality and privacy
- Supported data types: **relational, tabular, unstructured text, time-series, images**
- Validate models and use cases with **quality and privacy scores**

Value Proposition & Use Cases

- General purpose platform
- **Gretel's Transform API automates transformations that anonymize your data to preserve privacy, and use Synthetics for additional privacy guarantees.**

Company

SYNTHESIZED

- The fastest way to create trusted data
- **US$2.8M Seed** - Mar 2020
- HQ in **London, UK**
- https://www.synthesized.io/

Partnerships & GTM

- Accenture, Nextwave, Bigid
- B2B SaaS

Tech Stack & Product

- Create high-quality data for machine learning, application development and testing with **easy-to-use config files.**

Value Proposition & Use Cases

Use cases: **customer data analytics, customer churn projection, data monetization in financial services, bias mitigation in recruitment processes, data sharing with 3rd-parties, fraud prevention, clinical trials, insurance quote conversion, ecommerce trend projection**

Company

- Unparalleled Privacy. Unbeatable Accuracy.
- **US$25M Series A** - Sep 2022
- HQ in **Raleigh, NC**
- https://diveplane.com/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Create a **safe, accurate synthetic twin of your data**. All the value from your data without the risk to data privacy

Value Proposition & Use Cases

- Use cases: **financial services, healthcare, defense**
- **Buying habits** and trends, population and **census information, credit worthiness and financial transactions, travel and vacation** preferences, **medical records and diagnosis trends**

Company

Datomize

- Limited data. Unlimited insights.
- **US$6M Seed** - Feb 2021
- HQ in **Tel Aviv, Israel**
- https://www.datomize.com/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Increase the amount of data, reduce the number of errors, and balance the model. Datomize generates optimal synthetic training data with the **lowest bias**.
- Data Health Report - **validation of source data,** equipped with **quality and balance scores,** and insights into **data gaps.**

Value Proposition & Use Cases

- Simulate scenarios and outcomes, complete data gaps, **reduce inherent bias, examine previously unseen environments**

Company

hazy

- Set your enterprise data free
- **US$9M Series A** - Mar 2023
- HQ in **London, UK**
- **https://hazy.com/**

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Hazy **re-engineers your enterprise data** to make it faster, easier and safer to use.
- We equip and enable businesses to get the most out of their data in a safe and ethical way. We create synthetic data that is completely artificial and **statistically representative to your raw data**.

Value Proposition & Use Cases

- Use cases: **financial services, telecommunications, government, research**
- **Privacy compliance, data sharing, model testing**

Company

- At Tonic.ai, we fake things seriously
- Mission: to unlock data for the efficiency of organizations while guaranteeing privacy for entitites and individuals.
- **US$35M Series B** - Sep 2021
- HQ in **San Francisco, CA**
- https://www.tonic.ai/

Partnerships & GTM

- B2B SaaS

Tech Stack & Product

- Automating data synthesis to advance data privacy.
- Create **realistic, targeted, representative test data** based on your real world data, preserving critical relationships and maintaining **input-to-output consistency across tables and databases.**
- **Model, shape, and size the data** to specific requirements.

Value Proposition & Use Cases

- Run **pre-production environments** that look, act, and feel like production.
- **De-identify data, achieve privacy compliance**

AI Synthetic Data Companies

MDCLONE

Synthetic data for healthcare

datagen

Human-centric synthetic data for computer vision

FINCRIME DYNAMICS

Financial fraud prevention and compliance control

Synthesis AI

Human and environmetal synthetic data for computer vision

Synthetic data generation platforms

hazy

gretel

TONIC

Datomize

Specialized

General Purpose

parallel domain

Synthetic data for outdoor perception models

SYNTHESIZED

Diveplane

Neurolabs

Consumer packaged goods image generation for computer vision

MOSTLY AI

RENDERED.AI

Physics-based synthetic data for computer vision

The Strategy Deck

July 2023; List not exhaustive

Need Help with Product Strategy for your AI app?

- Alex Irina Sandu

- **alex@thestrategydeck.com**

- www.thestrategydeck.com

The Strategy Deck