What is Data Greenhouse?

An operating system where data grows on
its own and proves its own quality,
so your AI performs at its best and earns trust.

01 Autonomous Evolution

**Powered by
an Agentic AI Data Scientist**

Blockchain-backed proof of data value.

02 Proven Trust

**Quality checks aligned
with ISO/IEC 5259.**

Data integrity grounded in international standards
and regulations.

03 Physical Consistency

**Synthetic data built to
prevent physical hallucinations**

Optimized for Physical AI,
including VLMs, VLAs and world models.

Pebblous

dataclinic.ai

Visit the DataClinic website

Email contact@pebblous.ai

Tel +82-44-589-3824

Web pebblous. ai

Pebblous

Data Greenhouse

An operating system where data grows on
its own and proves its own quality,
so your AI performs at its **best and earns trust.**

* Pebblous Data Greenhouse
: An autonomous data operating system that manages
your data assets and proves their value on its own.



Agentic Data OS for Physical & Sovereign AI



Data Greenhouse

The Data Greenhouse Process

An end-to-end autonomous pipeline,
run by an Agentic AI Data Scientist

01. Diagnose & Prescribe (Data Clinic)

Diagnoses data quality
and automatically prescribes
improvement scenarios

02. Scenario Design (GenSim Manager)

Turns high-level instructions
into simulation scripts for digital twins.

03. Synthetic Data Generation (PebbloSim)

Fills in missing edge cases
with synthetic data that stays
true to real-world conditions.

04. Quality Check (PebbloScope)

Visualizes data distributions
and business meaning, so every
generated sample can be explained.

05. Governance Loop (Data Greenhouse)

Feeds data back into AI retraining
and keeps audit logs ready
for regulators.

Solving the Physical AI data bottleneck at its root

Data Greenhouse + PebbloSim

PebbloSim creates synthetic data
that obeys the laws of physics.
Data Greenhouse, our autonomous
data operating system, puts it to work.
Together they keep producing
AI-ready data for Physical AI, with proof of its quality.

Autonomous Data Bulk-up

Agents find the gaps in your data and fill them with synthetic data, so you never run short. Runs on-premise and in sovereign environments.

Neuro-Symbolic Synthetic Data

High-precision synthetic data,
designed to block physical hallucinations
at the source.

Explainable Validation

Every step, from generation to improvement,
is open and traceable, proving
that your AI model can be trusted.

Solving Physical AI challenges, industry by industry



Robotic Welding · AI OCR (Drawings) · Autonomous Navigation

Shipbuilding & Maritime

Labor shortage + tight schedules

Multimodal AI that reads drawings and sensor data together makes AI adoption work on site.



Rare-Case Data · Safety Compliance

Mobility & Manufacturing

Safety-critical + physically accurate

High-fidelity synthetic data, designed to avoid physical hallucinations, builds 10,000 accident scenarios.



Sovereign AI · On-Premise Deployment

Defense

Data sovereignty + closed networks

On-premise deployment keeps sensitive data inside your network.

Pebblous

Backed by successful pilots with leading partners,
including a leading global automaker,
across manufacturing and defense, we are ready to build
the next Physical AI success story with you.