

WITNESS DATA FACTORY™

DATA LICENSE AND COMMERCIAL USAGE FRAMEWORK

DOCUMENT TYPE

Synthetic Dataset Commercial License

DOCUMENT VERSION

2.0

EFFECTIVE DATE

April 2026

ISSUING AUTHORITY

WITNESS DATA FACTORY™

PRINCIPAL

Hector Ortiz, Data Engineering Lead

CONTACT

WitnessDataFactory@gmail.com

Framework ID: WDF-LICENSE-2026-04-V1

Classification: Public — May be shared with auditors, regulators, and clients

This license governs all synthetic datasets and data files generated, distributed, or delivered by WITNESS DATA FACTORY™, including commercial portal deliveries, free-tier evaluation samples, and repository-distributed data artifacts.

Page 1

WITNESS DATA FACTORY™ Synthetic Dataset Commercial License v2.0

SECTION 1: NATURE OF THE DATA

1.1 Synthetic Origin

The Data is 100% synthetically generated. It is produced by the WITNESS DATA FACTORY Trinity Ensemble pipeline and is not derived from, nor intended to represent, any real patient, individual, or clinical system.

1.2 No PHI / PII

No Protected Health Information (PHI) or Personally Identifiable Information (PII) from real individuals is present in any dataset produced by WITNESS DATA FACTORY™.

1.3 QA Certificate

All datasets carry a machine-readable QA certificate (`qacertificate.json`) documenting Trinity Consensus Score, pipeline version, DP parameters, compliance flags, and a SHA-256 integrity hash.

SECTION 2: FREE-TIER EVALUATION SAMPLES

2.1 Permitted Uses

- Internal technical evaluation and pipeline validation.
- Integration testing and schema compatibility assessment.
- Demonstration, research, and non-commercial experimentation.

2.2 Prohibited Commercial Use

Free-tier data may not be used in production ML model training, commercial model development, or any revenue-generating workflow without upgrading to a commercial license.

2.3 Attribution

Dataset generated with WITNESS DATA FACTORY™ (synthetic medical data).
<https://witness-data-factory.onrender.com>

SECTION 3: COMMERCIAL DATASET LICENSE

3.1 Licensed Uses

- Production ML model training and fine-tuning.
- Benchmarking, evaluation, and validation in commercial systems.
- Internal enterprise use across the purchasing organization.
- Inclusion in derivative model artifacts, provided source datasets are not redistributed in raw form.

3.2 Entity Scope

The commercial license is granted to the purchasing entity only and is non-transferable without written consent from WITNESS DATA FACTORY™.

3.3 Volume and Domain Scope

Volume and domain scope are defined by the specific purchase (SKU). Use beyond the purchased volume or domain requires a separate purchase.

Page 2

WITNESS DATA FACTORY™ Synthetic Dataset Commercial License v2.0

SECTION 4: RESTRICTIONS (ALL TIERS)

The following are prohibited under all license tiers:

4.1

Representing the Data as real-world patient data, real EHR records, or as containing actual PHI or PII.

4.2

Using the Data as the direct basis for clinical decision-making, patient diagnosis, or any regulated clinical workflow.

4.3

Reselling, sublicensing, redistributing, or publicly hosting raw dataset files without express written authorization from WITNESS DATA FACTORY™.

4.4

Removing, modifying, or obscuring any QA certificate, license notice, SHA-256 hash, or provenance metadata included with a dataset.

4.5

Using the Data to train a model whose primary purpose is to compete with WITNESS DATA FACTORY's data generation or delivery services.

SECTION 5: INDEMNIFICATION

WITNESS DATA FACTORY indemnifies purchasers of commercial datasets against third-party PHI claims arising solely from confirmed synthetic content delivered under a valid commercial purchase, subject to the purchaser:

- Not misrepresenting the synthetic origin of the data.
- Maintaining the `qacertificate.json` and provenance documentation delivered with the dataset.

- Complying with all terms of this Data License.

This indemnification does not extend to free-tier evaluation samples.

SECTION 6: COMPLIANCE POSTURE

6.1 HIPAA Safe Harbor

Datasets contain no PHI and are exempt from HIPAA Privacy Rule requirements. No Business Associate Agreement (BAA) is required for use of WITNESS DATA FACTORY™ synthetic datasets.

6.2 GDPR

Fully synthetic data with no connection to real individuals is outside the scope of GDPR personal data obligations.

6.3 EU AI Act Article 10

Datasets are designed and documented in alignment with Article 10 data governance requirements for high-risk AI systems.

Page 3

WITNESS DATA FACTORY™ Synthetic Dataset Commercial License v2.0

6.4 CCPA

No personal information of California residents is present or processed.

6.5 FDA Guidance on AI/ML-Based Software as a Medical Device (SaMD)

Datasets are aligned with FDA guidance on the use of synthetic data for AI/ML model development and validation.

SECTION 7: NO WARRANTY

THE DATA IS PROVIDED “AS IS”. WITNESS DATA FACTORY™ MAKES NO WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO WARRANTIES OF FITNESS FOR A PARTICULAR CLINICAL PURPOSE, REGULATORY SUFFICIENCY, OR TASK-SPECIFIC ACCURACY. PURCHASERS REMAIN SOLELY RESPONSIBLE FOR DETERMINING FITNESS OF THE DATA FOR THEIR SPECIFIC USE CASE, REGULATORY SETTING, AND DEPLOYMENT CONTEXT.

SECTION 8: GOVERNING TERMS

8.1

This license supersedes all prior WITNESS DATA FACTORY™ data license versions for datasets delivered after April 2026.

8.2

Where a written commercial agreement specifies different terms for a specific dataset purchase, those terms govern for that purchase.

8.3

Inquiries regarding enterprise licensing, custom agreements, or volume arrangements:

- WitnessDataFactory@gmail.com
- <https://witness-data-factory.onrender.com>

Page 4