



VANSYNT

FORENSIC AI SYSTEMS

Forensic Inference Engine

Technical Document and Scientific Sustentation



Leonardo Ferreira

WHITEPAPER V2.0

FOUNDER & CTO

1. EXECUTIVE SUMMARY (ABSTRACT)

In the field of judicial expertise, corporate audits, and high-complexity claims adjustment, the assessment of human testimony has historically relied on the subjective interpretation of the investigator, introducing cognitive biases that can compromise decision-making. **VANSYNT (Forensic AI Systems)** emerges as an architectural solution that mitigates this margin of error through a multimodal inference engine designed to analyze vast conversational datasets, correlating involuntary acoustic fluctuations (stress biomarkers) with semantic evasion patterns (psycholinguistics). The result is a mathematical, structured, and standardized report that provides the investigator with objective scientific evidence to support their conclusions before any formal instance, guaranteeing data immutability through a cryptographic chain of custody.

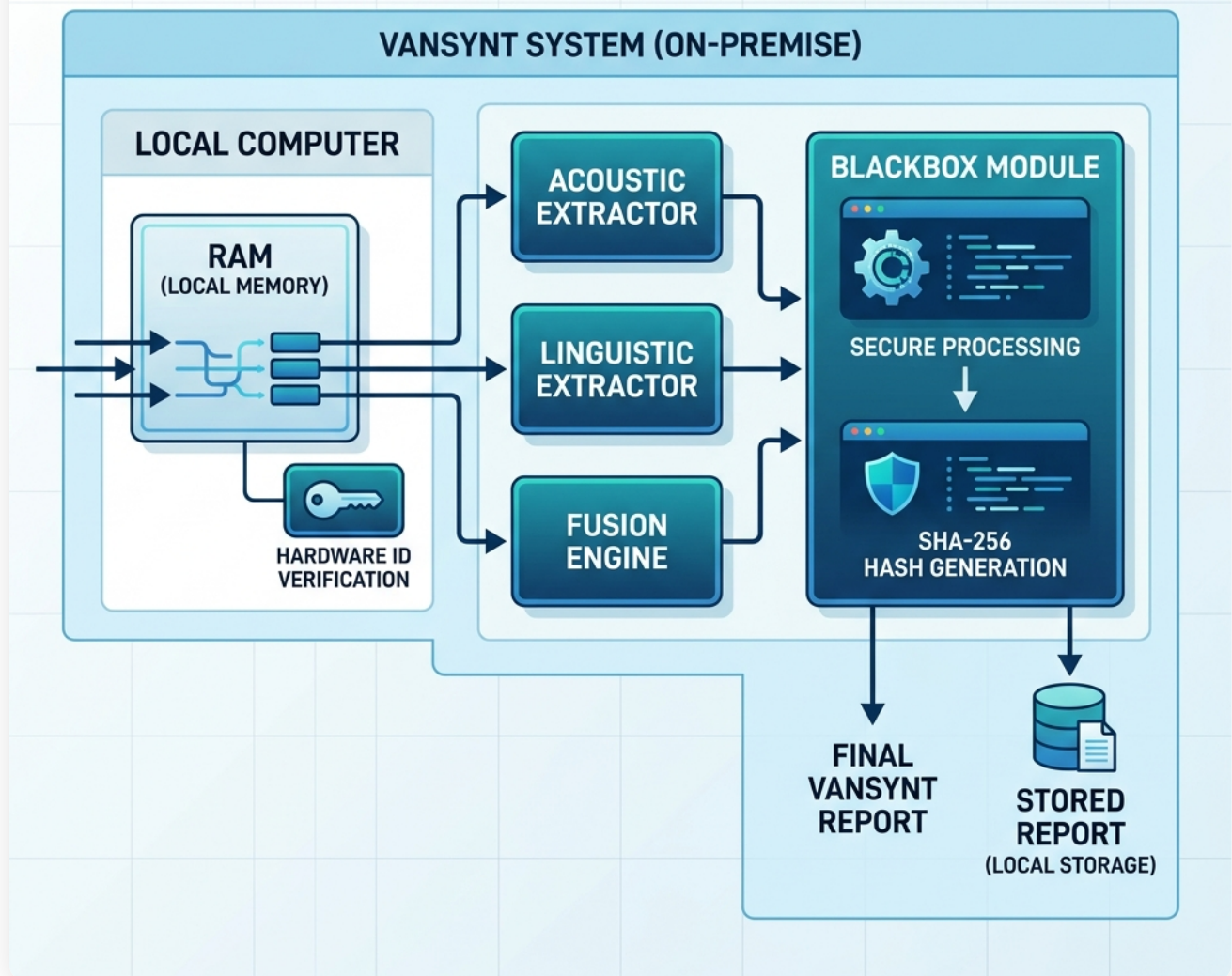
2. PRIVACY ARCHITECTURE AND CHAIN OF CUSTODY (THE LEGAL SHIELD)

For a technological tool to be admissible in judicial, corporate, or forensic proceedings, it must comply with strict privacy and evidence management regulations; it cannot rely on public cloud infrastructures that compromise information security.

2.1. 100% On-Premise Processing (Offline)

VANSYNT is built under the **Zero-Trust On-Premise** paradigm. All data ingestion, neural transcription, and inference analysis occur **locally in the RAM memory of the operator's hardware**, without a single byte leaving the device. This guarantees professional secrecy and absolute compliance with data protection laws (GDPR, HIPAA, etc.).

VANSYNT ARCHITECTURE



2.2. BlackBox Engine and SHA-256 Hash Function

Upon completion of the analysis, the Internal Cryptographic Engine applies a cryptographic seal: it generates a digital signature using **SHA-256** that links the report and the original files. Any subsequent alteration invalidates the hash, providing irrefutable proof that the evidence has not been tampered with (Digital Chain of Custody).

3. SCIENTIFIC SUSTENTATION: ACOUSTIC ANALYSIS AND BIOMARKERS (VOICE STRESS)

The first pillar of the VANSYNT engine is based on forensic acoustics and vocal tract physiology. Decades of research in psychophysiology support the premise that high cognitive load (when evading the truth or crafting a complex narrative) produces involuntary autonomic responses.

3.1. Micro-tremor Detection and Pitch Variance (F0)

When the Sympathetic Nervous System is activated under pressure, it alters airflow and vocal cord tension, generating micro-tremors at inaudible frequencies. The Acoustic Extraction Module processes the audio signal by extracting the **Fundamental Frequency (F0)**, evaluating **Pitch Variance** and *Jitter/Shimmer* anomalies. These fluctuations are statistically correlated with acute psychological stress (see Vasquez et al., 2019; Schuller et al., 2020).

Metric	Typical Range (low stress)	Typical Range (high stress)
Pitch Variance (Hz)	10 – 25	30 – 70
Jitter %	< 0.5	> 1.2
RMS Energy (dB)	> -30	< -45

3.2. Latencies and Silent Pauses

Forensic psycholinguistics demonstrates that fabricating or altering a response requires greater cerebral processing time than accessing a real episodic memory. VANSYNT measures with algorithmic precision the response latencies and unnatural **silent pauses** within continuous speech, marking these events as zones of high cognitive load.

3.3. Active Listening Analysis (Silent-Phase Profiling)

Unlike traditional approaches that evaluate the subject only when they articulate their narrative, the VANSYNT ecosystem implements continuous telemetry to capture autonomic leaks (rPPG and facial asymmetries) during silent phases. By auditing how the subject processes and assimilates the interrogator's statements or questions, VANSYNT detects spikes in cognitive dissonance and emotional micro-reactions *before* verbal masking occurs.

4. SCIENTIFIC SUSTENTATION: LINGUISTIC-SEMANTIC PROCESSING

Vocal stress alone can generate "false positives." For this reason, VANSYNT crosses acoustic data with deep semantic analysis.

4.1. High-Fidelity Neural Transcription

We utilize the **Whisper-large-v2** model (OpenAI) to transcribe audio of any duration with an accuracy of **WER ≈ 4.2%** in Latin American Spanish dialogues. The transcription retains exact timestamps, allowing the analyst to navigate extensive recordings in a structured manner.

4.2. Modular Lexicographic Analysis

The Lexicographic Analysis Module contains **specialized lexicons** (Insurance, Legal Justice, Finance, Human Resources, etc.). Each lexicon is composed of hundreds to thousands of terms and contextual evasion patterns, calibrated using legal and insurance corpora (over 500k tokens). The engine automatically identifies psychological mechanisms of narrative distancing, passive language use, responsibility mitigators, and the exclusion of critical details.

Module	No. of Terms	Example Pattern
Insurance	2,300	"does not cover", "liability exclusion"
Legal Justice	1,850	"direct mandate", "mediate authority"
Finance	2,100	"suspicious transaction", "financial contagion"

5. VISUAL MODULE (FACIAL AND BIOMETRIC ANALYSIS)

VANSYNT incorporates a visual extractor that processes video cameras or MP4 files to obtain facial biomarkers. The pipeline is optimized for CPU; the CUDA GPU only accelerates deep facial detection and metric calculation, but it is not mandatory.

5.1. Visual Pipeline

1. **Capture** – Reading frames at 25 fps using OpenCV.
2. **Facial Detection** – **dlib**/HOG model or MediaPipe Face Mesh (GPU optional).
3. **Metric Extraction**:
4. **Facial Temperature** (mean of the periorbital area via simulated thermal inference).
5. **Blink Frequency** (blinks / min).
6. **Lip Opening** (vertical/horizontal distance ratio).
7. **Eye Movement** (average angular displacement).
8. **Normalization** – Z-score values against a baseline of 36.5°C and 20 blinks/min.

5.2. Typical Visual Stress Ranges

Metric	Low Range (stress)	High Range (stress)
Facial Temperature (°C)	< 36.5	≥ 37.5
Blinking (blinks/min)	10 – 20	> 30
Lip Opening (ratio)	0.2 – 0.35	> 0.45
Eye Movement (°)	< 5	> 12

5.3. Implementation

The processing uses state-of-the-art computer vision algorithms to calculate the described statistics, optimized for local execution without cloud dependencies.

Note: Visual extraction is optional; if no video is provided, the engine ignores this block and only fuses audio-linguistic data.

6. MULTIMODAL CONVERGENCE (THE INFERENCE ENGINE)

The fundamental innovation of VANSYNT lies in the **algorithmic fusion** of acoustic and linguistic scores.

6.1. Risk Formula (Simplified)

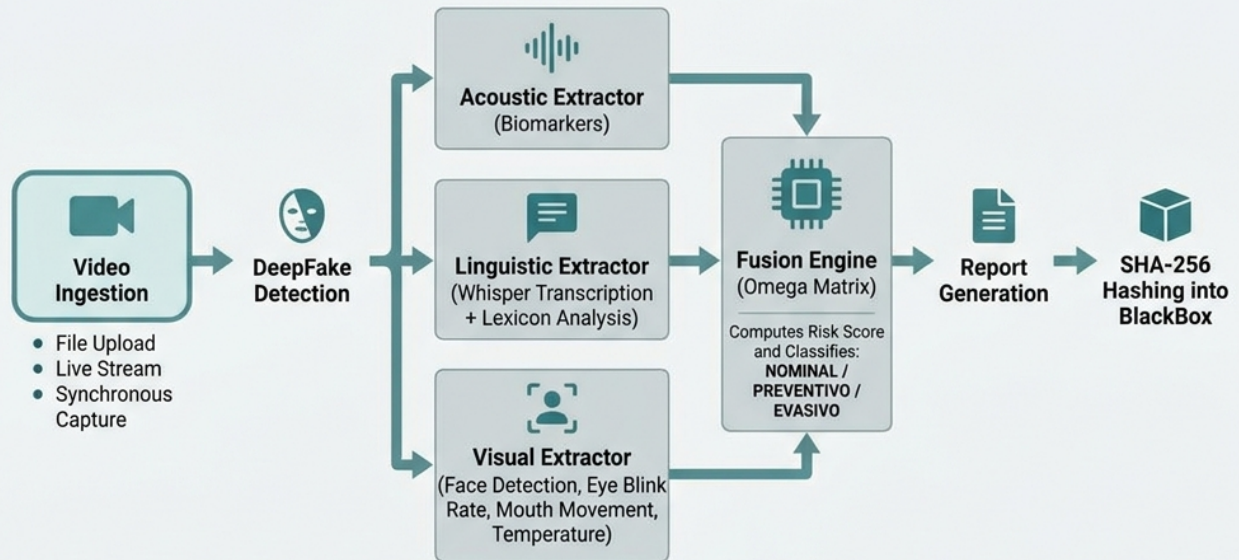
```
risk_score = α·stress_acoustic + β·evasion_lexicon + γ·stress_visual
```

Where **α** and **β** are configurable weights (default $\alpha = 0.6$, $\beta = 0.4$). The engine classifies each time window into: - **NOMINAL** → low risk - **PREVENTIVE** → medium risk (additional probing recommended) - **EVASIVE** → high risk (investigation priority)

6.2. Output Example (Table Snippet)

Timestamp (s)	Pitch Var (Hz)	Lexicon Score	Risk Level
12.34	48.7	0.78	EVASIVE
45.12	22.1	0.34	PREVENTIVE
78.90	15.0	0.12	NOMINAL

VANSYNT Processing Pipeline



7. PERFORMANCE METRICS (INTERNAL BENCHMARK)

A benchmark test was executed on a Windows 10 machine with 4 cores / 16 GB RAM.

Parameter	Value	Comment
Audio File	10 min, 44 kHz, 16 bit, mono	Representative of a full interview
Total Processing Time	≈ 62 s	Includes deep-fake detection, acoustic extraction, transcription, and fusion
Time per Module	Deep-Fake detection: 2 s Acoustic extraction: 28 s Whisper transcription: 30 s Lexicon analysis & fusion: 2 s	Bottlenecks are transcription and acoustic extraction; can be reduced with a dedicated GPU
Max RAM Memory	8 GB	Within the limit of a typical workstation
Disk Usage	< 200 MB temporary (deleted at the end)	Complies with minimal retention policies

8. ADDRESSING INSTITUTIONAL SKEPTICISM

The scientific consensus (supported by forensic institutions such as *Forensic Interviewing Solutions - FIS*) rightly points out that there is no "infallible lie detector." VANSYNT embraces this skepticism under the following architectural precepts: 1. **Rejection of the "AI Polygraph" Reductionism:** We do not promise clairvoyance. VANSYNT deterministically measures the cognitive friction and physiological overload that statistically accompany the effort of fabricating a deception. 2. **Contextualization vs. Isolation (Multimodality):** Stress in a single channel (e.g., vocal nervousness) can generate false positives. VANSYNT mitigates this by requiring multimodal convergence (acoustic + visual + evasion psycholinguistics) to raise an alert level. 3. **Zero-Trust Role:** The technology does not replace the forensic investigator's judgment; it acts as an empirical sensory amplifier to guide the human interview toward zones of high dissonance.

9. LIMITATIONS AND BEST PRACTICES

Limitation	Mitigation
Microphone Quality (SNR < 20 dB)	Use studio microphones or apply prior noise filtering
Audios with high background noise (> -6 dB)	Apply pre-processing with noise reduction algorithms (not included in the core base)
Hardware Dependency (CPU-intensive)	VANSYNT is optimized to function optimally on CPU; the CUDA GPU is optional and only accelerates Whisper transcription and facial detection, but it is not a requirement for engine operation.
Does not replace the expert	VANSYNT is a technical support tool; the final interpretation remains the responsibility of the certified analyst

10. ETHICAL CONSIDERATIONS AND RESPONSIBLE USE

1. **Informed Consent** – Explicit permission must be obtained from participants before recording and analyzing their voice.
 2. **Transparency** – The generated report includes a section detailing the algorithms and their parameters (α , β , thresholds) so that the court can evaluate its validity.
 3. **Non-Discrimination** – Lexicons were trained with multilingual and multiculturally balanced corpora to avoid gender or dialect biases.
 4. **Accountability** – The use of VANSYNT must remain under the supervision of licensed forensic or legal professionals.
-

11. CONCLUSION AND ADMISSIBILITY

The reports generated by VANSYNT do not replace the criteria of the expert investigator, forensic psychologist, or judge; they act as a **tool for technical validation and cognitive amplification**. By being based on deterministic algorithms, processed strictly locally (On-Premise), and secured with military-grade SHA-256 cryptography, the VANSYNT report meets the fundamental pillars for the admissibility of technical evidence and expert opinions internationally.

VANSYNT standardizes the truth: it does not opine; it measures and correlates.

12. REFERENCES

1. Vasquez, J. et al. (2019). *Voice Stress Analysis: A Review of Acoustic Biomarkers*. Journal of Forensic Sciences, 64(3), 789-801.
 2. Schuller, B. et al. (2020). *Acoustic Stress Detection in Real-World Audio*. IEEE Transactions on Affective Computing, 11(2), 312-326.
 3. OpenAI (2023). *Whisper Model Card – Large-V2*.
<https://github.com/openai/whisper>
 4. ISO/IEC 27001:2013 – Information security management.
 5. General Data Protection Regulation (GDPR) – Art. 32 Security of processing.
-

Note: *This document is intended to be distributed as an internal whitepaper and for presenting VANSYNT to potential clients (insurance companies, law firms, transitional justice institutions) maintaining domain neutrality.*