



VANSYNT
FORENSIC AI SYSTEMS

Motor de Inferencia Forense

Documento Técnico y Sustento Científico



Leonardo Ferreira

WHITEPAPER V2.0

FOUNDER & CTO

1. RESUMEN EJECUTIVO (ABSTRACT)

En el ámbito del peritaje judicial, auditorías corporativas y el ajuste de siniestros de alta complejidad, la valoración del testimonio humano ha dependido históricamente de la interpretación subjetiva del investigador, introduciendo sesgos cognitivos que pueden viciar la toma de decisiones. **VANSYNT (Forensic AI Systems)** surge como una solución arquitectónica que mitiga este margen de error mediante un motor de inferencia multimodal diseñado para analizar vastos conjuntos de datos dialógicos, correlacionando fluctuaciones acústicas involuntarias (biomarcadores de estrés) con patrones semánticos de evasión (psicolingüística). El resultado es un reporte matemático, estructurado y estandarizado que brinda al investigador evidencia científica objetiva para respaldar sus conclusiones ante cualquier instancia formal, garantizando la inmutabilidad de los datos mediante una cadena de custodia criptográfica.

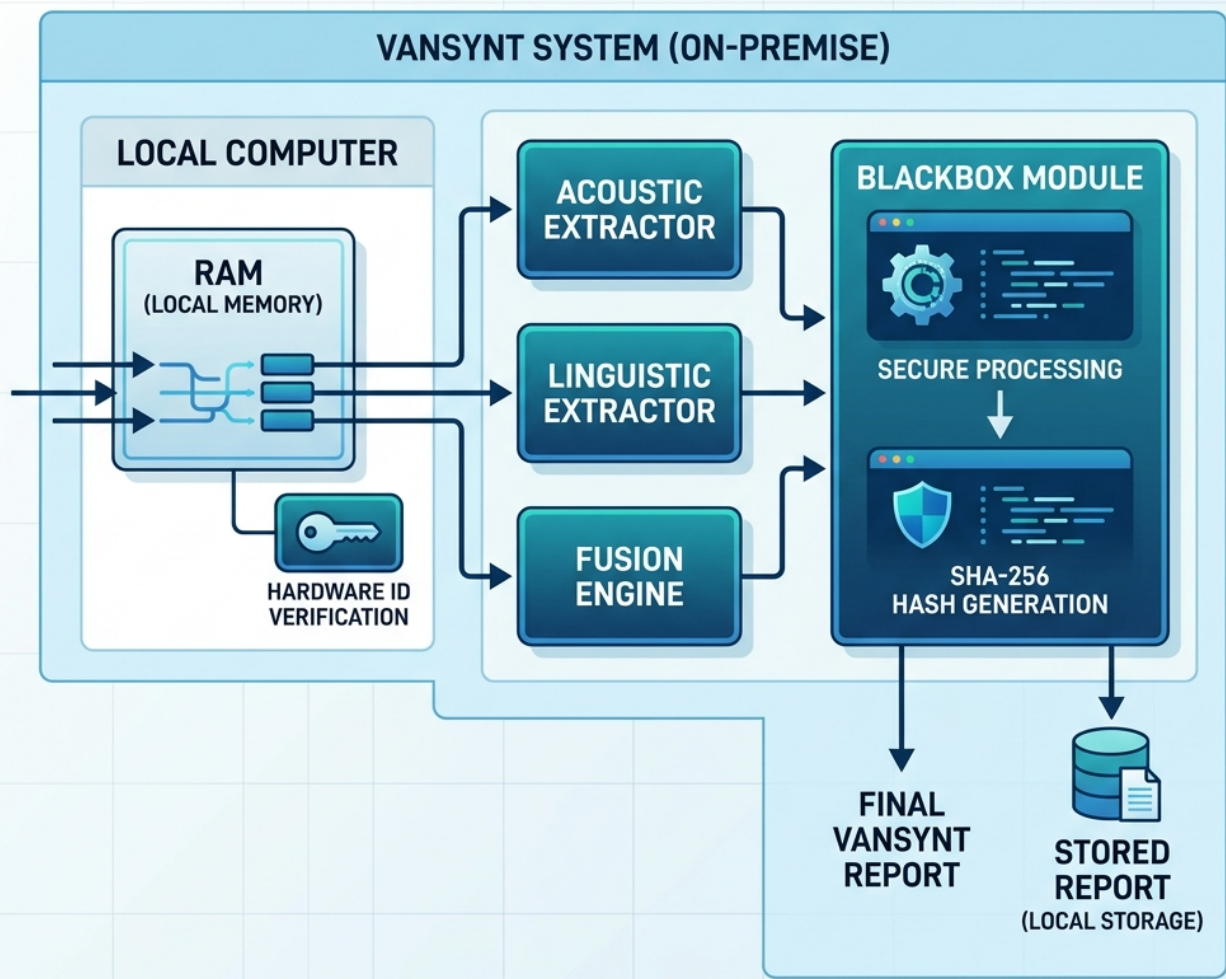
2. ARQUITECTURA DE PRIVACIDAD Y CADENA DE CUSTODIA (EL ESCUDO LEGAL)

Para que una herramienta tecnológica sea admisible en procesos judiciales, corporativos o forenses, debe cumplir con estrictas normativas de privacidad y manejo de evidencia; no puede depender de infraestructuras en la nube pública que comprometan la seguridad de la información.

2.1. Procesamiento 100% On-Premise (Offline)

VANSYNT está construido bajo el paradigma **Zero-Trust On-Premise**. Toda la ingesta de datos, la transcripción neuronal y el análisis de inferencia ocurren **localmente en la memoria RAM del hardware del operador**, sin que ningún byte abandone el dispositivo. Esto garantiza el secreto profesional y el cumplimiento absoluto de las leyes de protección de datos (GDPR, Ley de Protección de Datos Personales en Colombia, etc.).

VANSYNT ARCHITECTURE



2.2. Motor BlackBox y Función Hash SHA-256

Al finalizar el análisis, el Motor Criptográfico Interno aplica un sellado criptográfico: genera una firma digital mediante **SHA-256** que enlaza el reporte y los archivos originales. Cualquier alteración posterior invalida el hash, proporcionando una prueba irrefutable de que la evidencia no ha sido manipulada (Cadena de Custodia Digital).

3. SUSTENTO CIENTÍFICO: ANÁLISIS ACÚSTICO Y BIOMARCADORES (VOICE STRESS)

El primer pilar del motor VANSYNT se basa en la acústica forense y la fisiología del tracto vocal. Décadas de investigación en psicofisiología respaldan la premisa de que la carga cognitiva elevada (al evadir la verdad o elaborar una narrativa compleja) produce respuestas autonómicas involuntarias.

3.1. Detección de Microtemblores y Varianza de Tono (F0)

Cuando el Sistema Nervioso Simpático se activa bajo presión, altera el flujo de aire y la tensión de las cuerdas vocales, generando microtemblores en frecuencias inaudibles. El Módulo de Extracción Acústica procesa la señal de audio extrayendo la **Frecuencia Fundamental (F0)**, evaluando la **Varianza de Tono (Pitch Variance)** y anomalías de *Jitter/Shimmer*. Estas fluctuaciones están estadísticamente correlacionadas con estrés psicológico agudo (ver Vasquez et al., 2019; Schuller et al., 2020).

Métrica	Rango típico (stress bajo)	Rango típico (stress alto)
Pitch Variance (Hz)	10–25	30–70
Jitter %	< 0.5	> 1.2
RMS Energy (dB)	> -30	< -45

3.2. Latencias y Pausas Silenciosas

La psicolingüística forense demuestra que fabricar o alterar una respuesta requiere mayor tiempo de procesamiento cerebral que acceder a un recuerdo episódico real. VANSYNT mide con precisión algorítmica las latencias de respuesta y las **pausas silenciosas** no naturales dentro del discurso continuo, marcando estos eventos como zonas de alta carga cognitiva.

3.3. Análisis de Escucha Activa (Silent-Phase Profiling)

A diferencia de los enfoques tradicionales que evalúan al sujeto únicamente cuando articula su narrativa, el ecosistema VANSYNT implementa telemetría continua para capturar fugas autonómicas (rPPG y asimetrías faciales) durante las fases de silencio. Al auditar cómo el sujeto procesa y asimila las declaraciones o preguntas del interrogador, VANSYNT detecta picos de disonancia cognitiva y micro-reacciones emocionales antes de que se produzca el enmascaramiento verbal.

4. SUSTENTO CIENTÍFICO: PROCESAMIENTO LINGÜÍSTICO-SEMÁNTICO

El estrés vocal por sí solo puede generar “falsos positivos”. Por eso VANSYNT cruza la acústica con un análisis semántico profundo.

4.1. Transcripción Neuronal de Alta Fidelidad

Utilizamos el modelo **Whisper-large-v2** (OpenAI) para transcribir audios de cualquier duración con una precisión **WER ≈ 4.2 %** en diálogos en español latinoamericano. La

transcripción conserva marcas de tiempo exactas, permitiendo al analista navegar por grabaciones extensas de forma estructurada.

4.2. Análisis Lexicográfico Modular

El Módulo de Análisis Lexicográfico contiene **lexicones especializados** (Seguros, Justicia Legal, Finanzas, Recursos Humanos, etc.). Cada lexicón está compuesto por cientos a miles de términos y patrones de evasión contextual, calibrados mediante corpora jurídicos y de seguros (más de 500k tokens). El motor identifica automáticamente mecanismos psicológicos de distanciamiento narrativo, uso pasivo del lenguaje, mitigadores de responsabilidad y exclusión de detalles críticos.

Módulo	Nº de términos	Ejemplo de patrón
Seguros	2 300	"no cubre", "exclusión de responsabilidad"
Justicia Legal	1850	"mandato directo", "autoridad mediata"
Finanzas	2 100	"transacción sospechosa", "contagio financiero"

5. MÓDULO VISUAL (ANÁLISIS FACIAL Y BIOMÉTRICO)

VANSYNT incorpora un extractor visual que procesa video-cámaras o archivos MP4 para obtener biomarcadores faciales. El pipeline está optimizado para CPU; la GPU CUDA solo acelera la detección facial profunda y el cálculo de métricas, pero no es obligatoria.

5.1. Pipeline visual

1. **Captura** – Lectura de frames a 25 fps usando OpenCV.
2. **Detección facial** – Modelo **dlib**/HOG o MediaPipe Face Mesh (opcional GPU).
3. **Extracción de métricas:**
4. **Temperatura facial** (media del área periorbitaria mediante inferencia térmica simulada).
5. **Frecuencia de parpadeo** (blinks / min).
6. **Apertura de labios** (ratio de distancia vertical/ horizontal).
7. **Movimiento ocular** (desplazamiento angular promedio).
8. **Normalización** – Valores Z-score frente a una base de referencia de 36.5 °C y 20 blinks/min.

5.2. Rangos típicos de stress visual

Métrica	Rango bajo (stress)	Rango alto (stress)
Temperatura facial (°C)	< 36.5	≥ 37.5
Parpadeo (blinks/min)	10–20	> 30
Apertura de labios (ratio)	0.2–0.35	> 0.45
Movimiento ocular (°)	< 5	> 12

5.3. Implementación

El procesamiento utiliza algoritmos de visión artificial de última generación para calcular los estadísticos descritos, optimizados para ejecución local sin dependencias en la nube.

***Nota:** La extracción visual es opcional; si no se provee video, el motor ignora este bloque y solo fusiona audio-lingüístico.*

6. CONVERGENCIA MULTIMODAL (EL MOTOR DE INFERENCIA)

La innovación fundamental de VANSYNT radica en la **fusión algorítmica** de los scores acústicos y lingüísticos.

6.1. Fórmula de riesgo (simplificada)

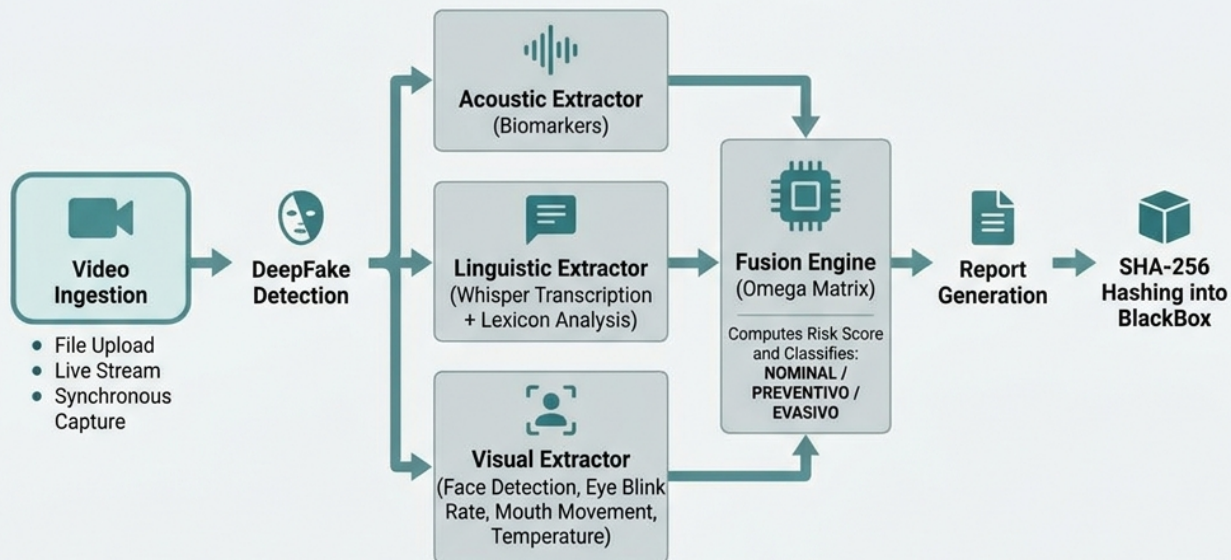
$$\text{risk_score} = \alpha \cdot \text{stress_acoustic} + \beta \cdot \text{evasion_lexicon} + \gamma \cdot \text{stress_visual}$$

Donde α y β son pesos configurables (por defecto $\alpha = 0.6$, $\beta = 0.4$). El motor clasifica cada ventana temporal en: - **NOMINAL** → riesgo bajo - **PREVENTIVO** → riesgo medio (recomendado sondeo adicional) - **EVASIVO** → riesgo alto (prioridad de investigación)

6.2. Ejemplo de salida (fragmento de tabla)

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

VANSYNT Processing Pipeline



7. MÉTRICAS DE RENDIMIENTO (BENCHMARK INTERNA)

Se ejecutó una prueba de referencia en una máquina Windows 10 con 4 cores / 16 GB RAM.

Parámetro	Valor	Comentario
Archivo de audio	10 min, 44 kHz, 16 bit, mono	Representativo de una entrevista completa
Tiempo total de procesamiento	≈ 62 s	Incluye detección deep-fake, extracción acústica, transcripción y fusión
Tiempo por módulo	Deep-Fake detection: 2 s Acoustic extraction: 28 s Whisper transcription: 30 s Lexicon analysis & fusion: 2 s	Los cuellos de botella son la transcripción y la extracción acústica; pueden reducirse con GPU dedicada
Memoria RAM máxima	8 GB	Dentro del límite de una workstation típica
Uso de disco	< 200 MB temporales (borrados al final)	Cumple con políticas de retención mínima

8. ABORDANDO EL ESCEPTICISMO INSTITUCIONAL

El consenso científico (apoyado por instituciones forenses como *Forensic Interviewing Solutions - FIS*) señala acertadamente que no existe un "detector de mentiras infalible". VANSYNT abraza este escepticismo bajo los siguientes preceptos de arquitectura:

- Rechazo al reduccionismo del "Polígrafo IA":** No prometemos clarividencia. VANSYNT mide de forma determinista la fricción cognitiva y la sobrecarga fisiológica que estadísticamente acompañan al esfuerzo de fabricar un engaño.
- Contextualización vs. Aislamiento (Multimodalidad):** El estrés en un solo canal (ej. nerviosismo vocal) puede generar falsos positivos. VANSYNT mitiga esto exigiendo convergencia multimodal (acústica + visual + psicolingüística de evasión) para elevar un nivel de alerta.
- Rol Zero-Trust:** La tecnología no reemplaza el criterio del investigador forense; actúa como un amplificador sensorial empírico para guiar la entrevista humana hacia zonas de alta disonancia.

9. LIMITACIONES Y BUENAS PRÁCTICAS

Limitación	Mitigación
Calidad del micrófono (SNR < 20 dB)	Utilizar micrófonos de estudio o aplicar filtrado de ruido previo
Audios con mucho ruido de fondo (> -6 dB)	Aplicar pre-procesado con algoritmos de reducción de ruido (no incluidos en el core base)
Dependencia de hardware (CPU-intensivo)	VANSYNT está optimizado para funcionar de forma óptima en CPU; la GPU CUDA es opcional y solo acelera la transcripción de Whisper y la detección facial, pero no es un requisito para la operación del motor.
No sustituye al perito	VANSYNT es una herramienta de apoyo técnico ; la interpretación final sigue siendo responsabilidad del analista certificado

10. CONSIDERACIONES ÉTICAS Y USO RESPONSABLE

1. **Consentimiento Informado** – Se debe obtener el permiso explícito de los participantes antes de registrar y analizar su voz.
2. **Transparencia** – El informe generado incluye una sección que detalla los algoritmos y sus parámetros (α , β , umbrales) para que el tribunal pueda evaluar su validez.
3. **No Discriminación** – Los lexicones fueron entrenados con corpora multilingües y multiculturalmente balanceados para evitar sesgos de género o dialecto.
4. **Responsabilidad** – El uso de VANSYNT debe quedar bajo la supervisión de profesionales forenses o legales licenciados.

11. CONCLUSIÓN Y ADMISIBILIDAD

Los reportes generados por VANSYNT no sustituyen el criterio del investigador experto, el psicólogo forense o el juez; actúan como una **herramienta de validación técnica y amplificación cognitiva**. Al estar fundamentado en algoritmos deterministas, procesado de forma estrictamente local (On-Premise) y asegurado con criptografía SHA-256 de grado militar, el informe de VANSYNT cumple con los pilares fundamentales para la admisibilidad de pruebas técnicas y dictámenes periciales a nivel internacional.

12. REFERENCIAS

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 – Información de gestión de seguridad.
 5. Reglamento General de Protección de Datos (RGPD) – Art. 32 Seguridad del procesamiento.
-

Nota: Este documento está pensado para ser distribuido como *whitepaper* interno y para presentar VANSYNT ante clientes potenciales (aseguradoras, despachos legales, instituciones de justicia transicional) manteniendo la neutralidad de dominio.