hana

The State of Voice AI in Healthcare - 2026

247 companies. Three layers. One map.

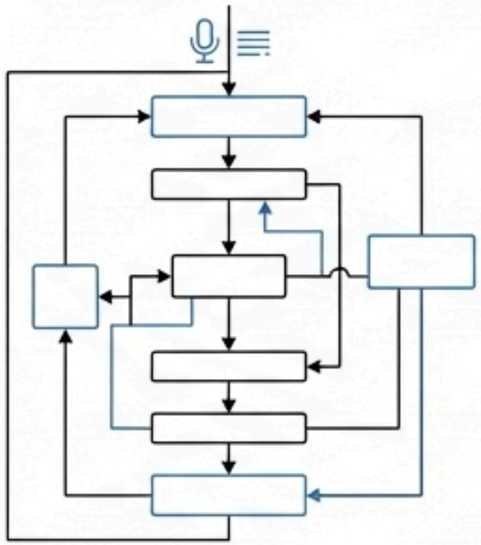
"Bringing clarity to a noisy space"

Voice AI exploded in 2025, but most healthcare buyers still struggle to separate demos from deployable systems. This report maps the market for clinics, investors, and builders.

Openaccess report. No gate.

70%
Pilots Fail to Scale

VOICE AI SYSTEM:
CLINICAL INTERFACE

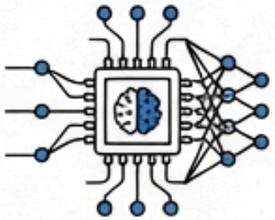


THE MARKET ANATOMY

SIGNAL vs. NOISE

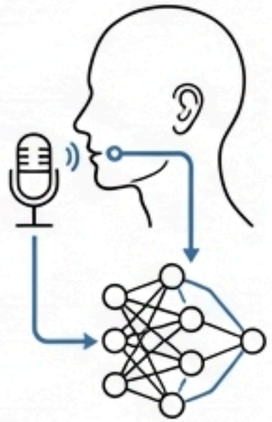
DEMOS ARE LOUD.
SYSTEMS ARE SILENT.
Most noise comes from marketing driven pilots. True signal is found in seamless, background integration with clinical workflows.

CORE FUNCTIONS



COMPUTATIONAL PROCESSING
Neural routonal processing comnonisces data computational, dieta, and processed integration with clinical workflows.

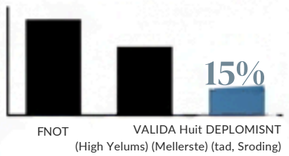
DEPLOYABLE SYSTEMS MAP



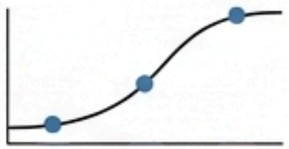
CLINICS		FOCUS: Workflow & ROI Need: Proven EMR integration, reduced documentation time, and guaranteed accuracy.
NVESTORS		FOCUS: Scalability & Retention Need: Defensible IP, clear path to revenue, and strong team.
BUILDERS		FOCUS: Tech & Ethics Need: Robust data privacy, explainable AI models, and interoperability.

85%
of Deployable Systems Offer EMR APIs

DEPLOYMENT STAGES



MARKET MATURITY



The path to clarity lies in understanding the underlying mechanics of sustainable, clinically-validated Voice AI solutions, not just the surface-level hype.

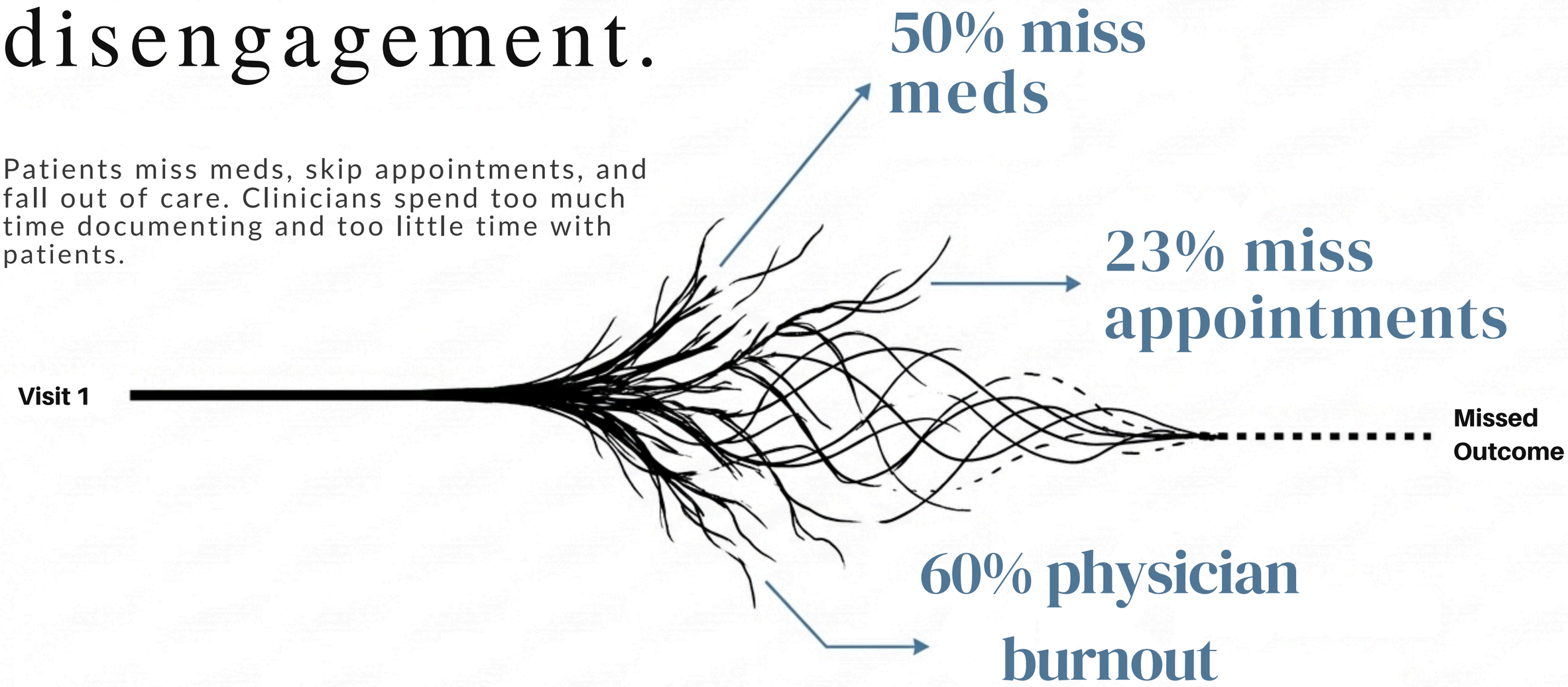
hana

How we built this map.

We analyzed 247 companies across funding data, websites, product materials, and hiring signals. Each was classified by layer, function, and healthcare relevance.

$300 billion lost to disengagement.

Patients miss meds, skip appointments, and fall out of care. Clinicians spend too much time documenting and too little time with patients.

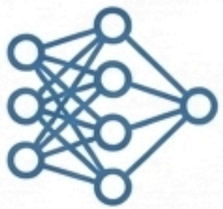


hana

Three things changed in 2025.

STT reached clinical-grade performance.

Sub-700ms latency, multilingual, and noise-robust. Real-world clinical audio accuracy pushed past 95%.



LLMs got better at real conversations.

Multi-turn intent recognition, tool use, and structured output transformed agents into executors of workflows.

Healthcare demand accelerated.

1 in 3 patients now use AI for health. 54% of AI funding went to health tech to fill nursing vacancies.

247 companies. Three layers. Twelve buckets.

The market breaks into Infrastructure, Platforms, and Applications.
The structure matters more than any individual logo.

Applications
128 companies

Platforms
44 companies

Infrastructure
75 companies

Layer 1: Core Voice Infrastucture

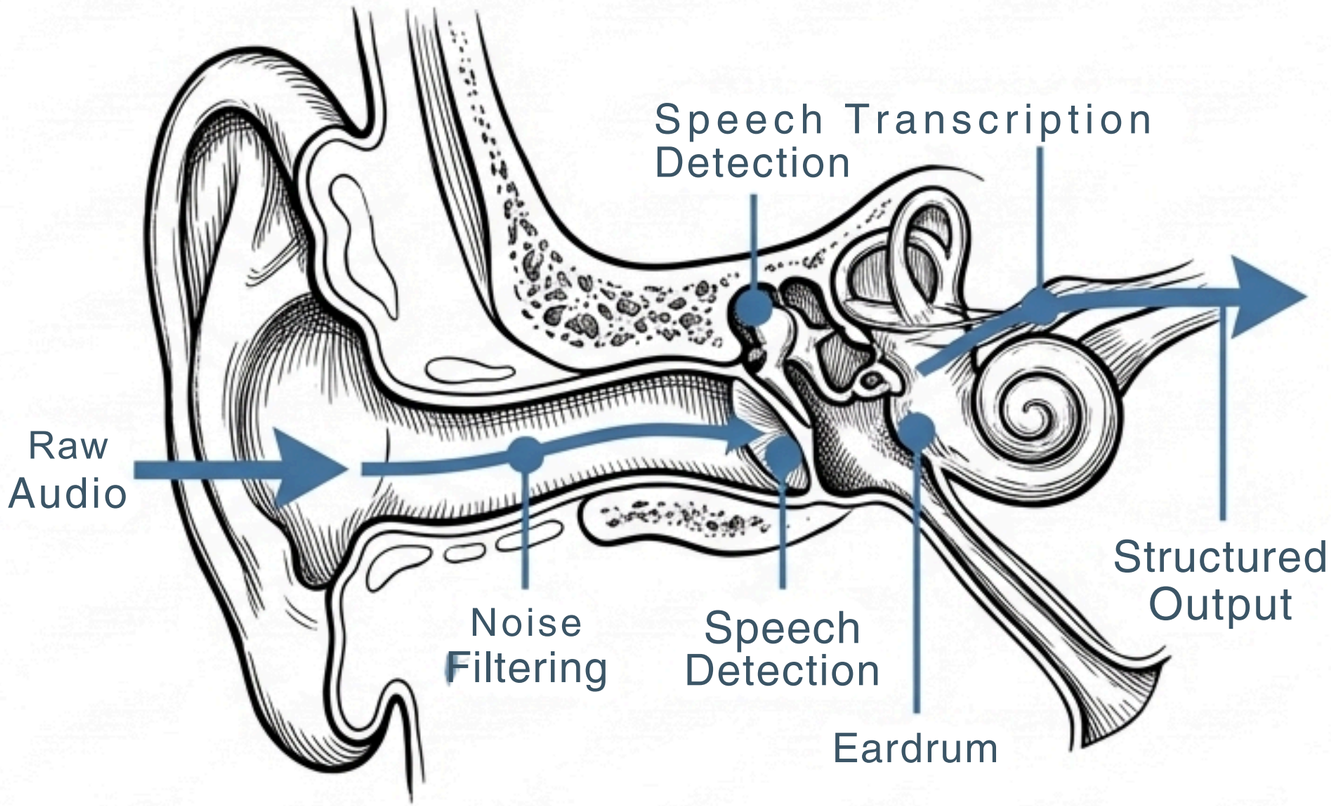
75 companies

STT, TTS, audio intelligence, and voice-native models.

hana

The Ears: Speech-to-Text

Transcription is now fast, accurate, multilingual, and real-time. The real challenge is messy healthcare audio, not clean demos.

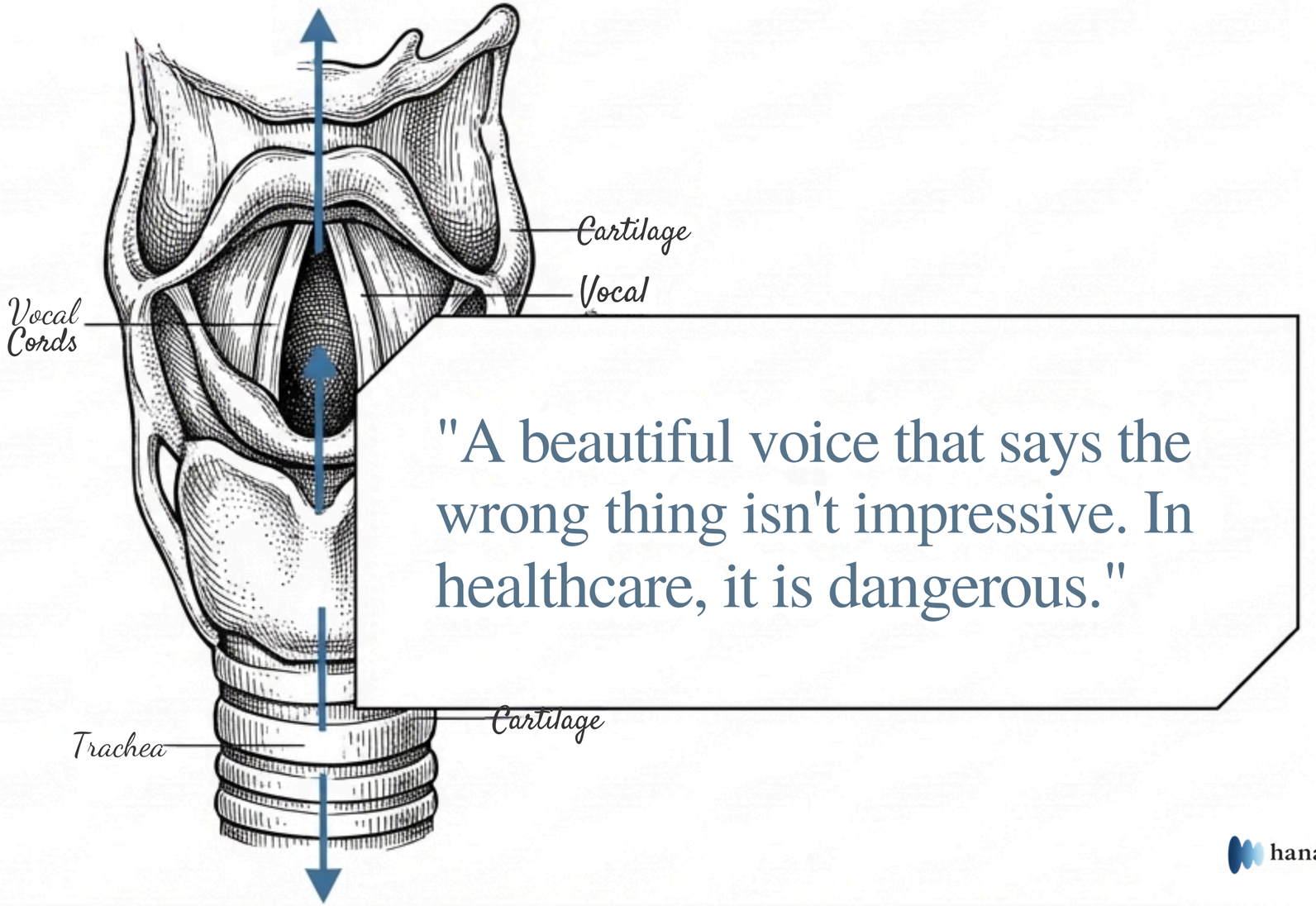


Deepgram Assembly AI Otter | Gladia Verbit | Ultravox

hana

The Mouth: Text-to-Speech

Human-like voice quality is becoming a commodity. Inhealthcare, the risk is not sounding robotic it is sounding convincing while being wrong.

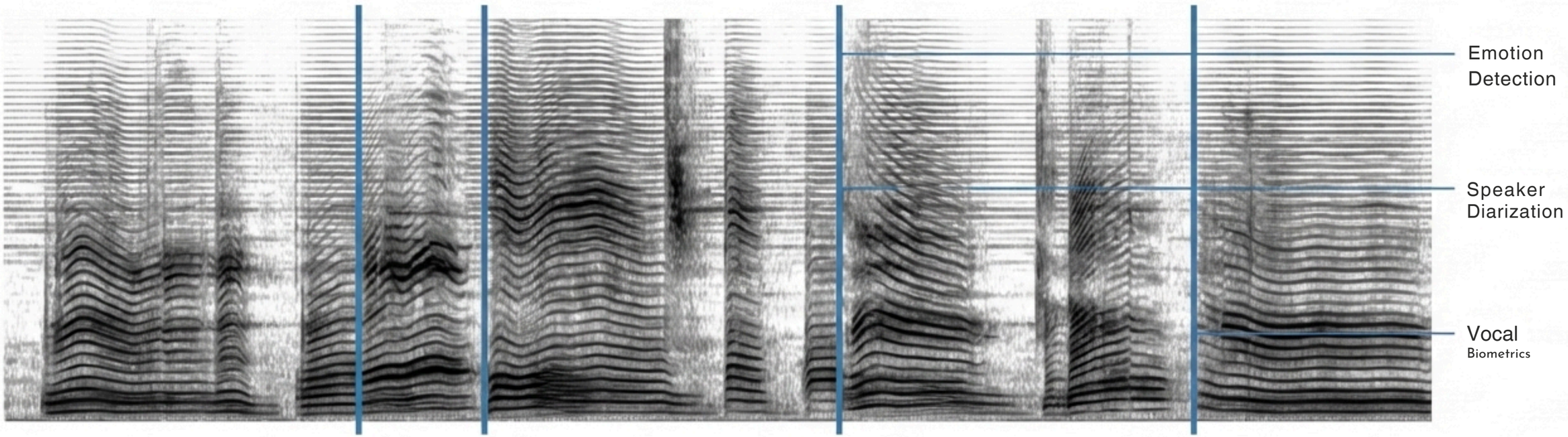


ElevenLabs | WellSaid | Inworld | Respeecher | PlayAI | Murf

hana

The Nervous System: Audio Intelligence

Emotion, hesitation, diarization, and biometrics sit in this layer. It is underappreciated and potentially high-value in care settings.

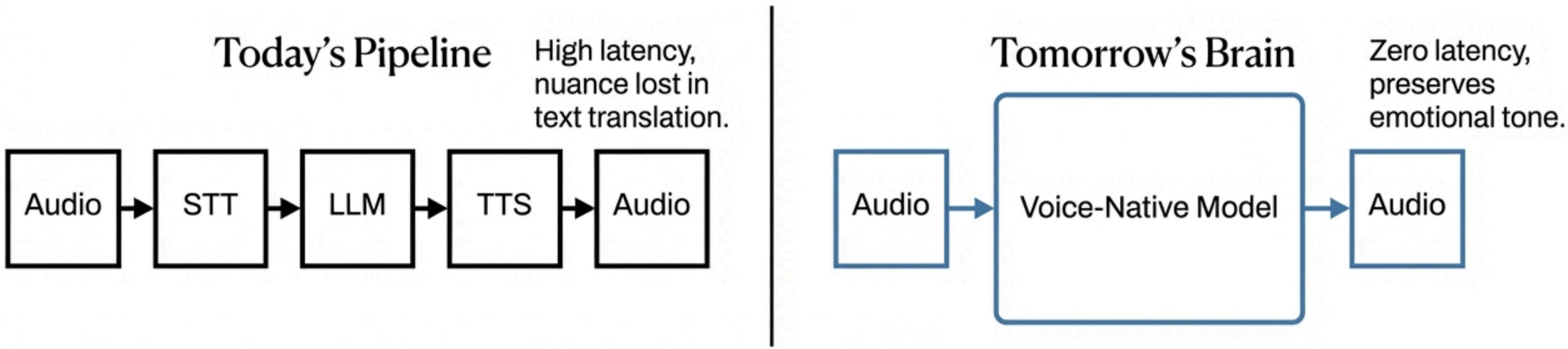


| Hume AI Behavioral Signals | pyannoteAI | IRIS Audio | audEERING

hana

The Emerging Brain

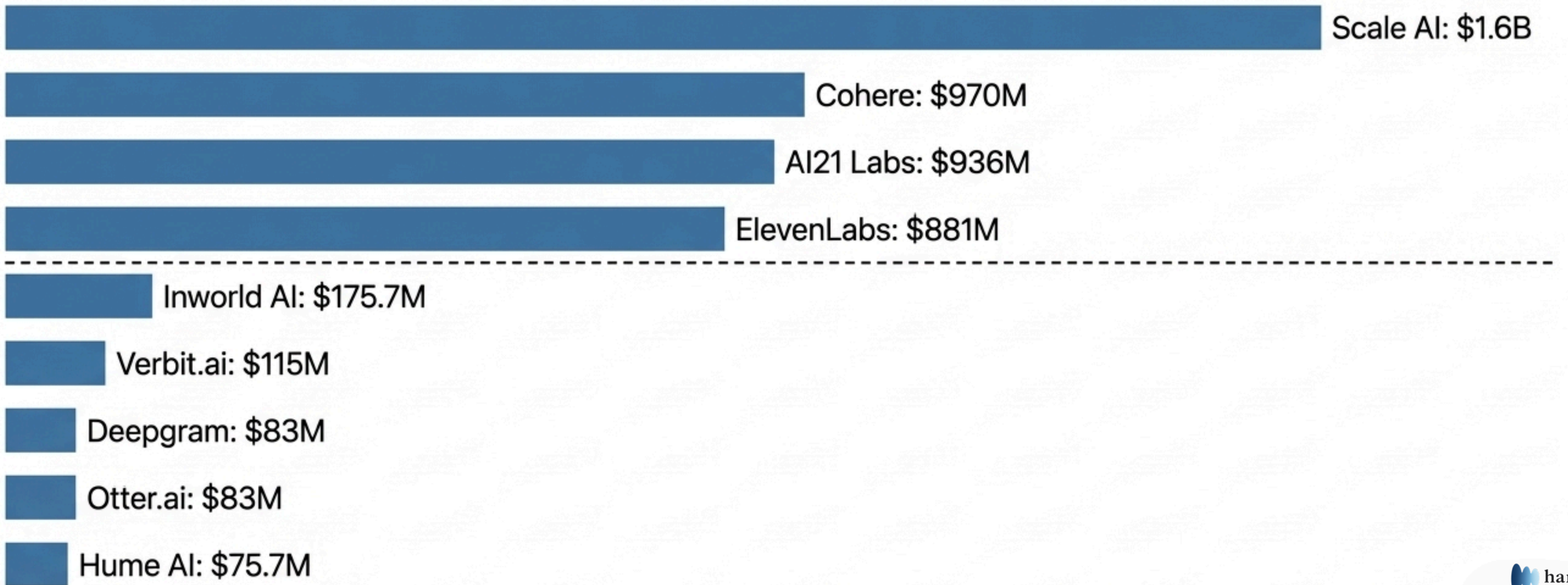
A small group is building models that reason directly from speech. That could reduce latency and preserve nuance, but it also makes auditing harder.



Scale AI | Cohere | A121 | Blue Machines AI | Deepslate

Where the infrastructure money went.

Infrastructure captured major rounds, but signs of commoditization are clear.
Scale will be won through distribution more than technical novelty.

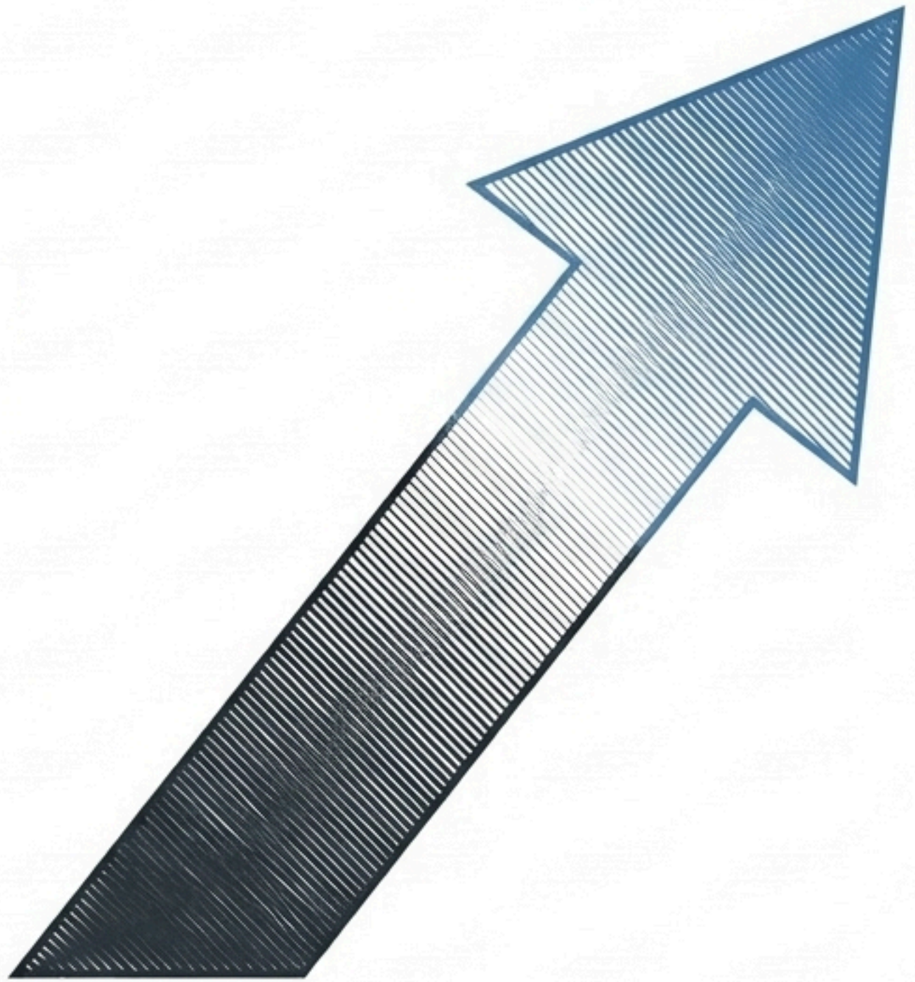


hana

Infrastructure is commoditizing.

The ears and mouth are increasingly interchangeable. Value is shifting toward reasoning, memory, safety, and workflow control.

Applications: Defensible Clinical Workflow & Memory



Infrastructure:

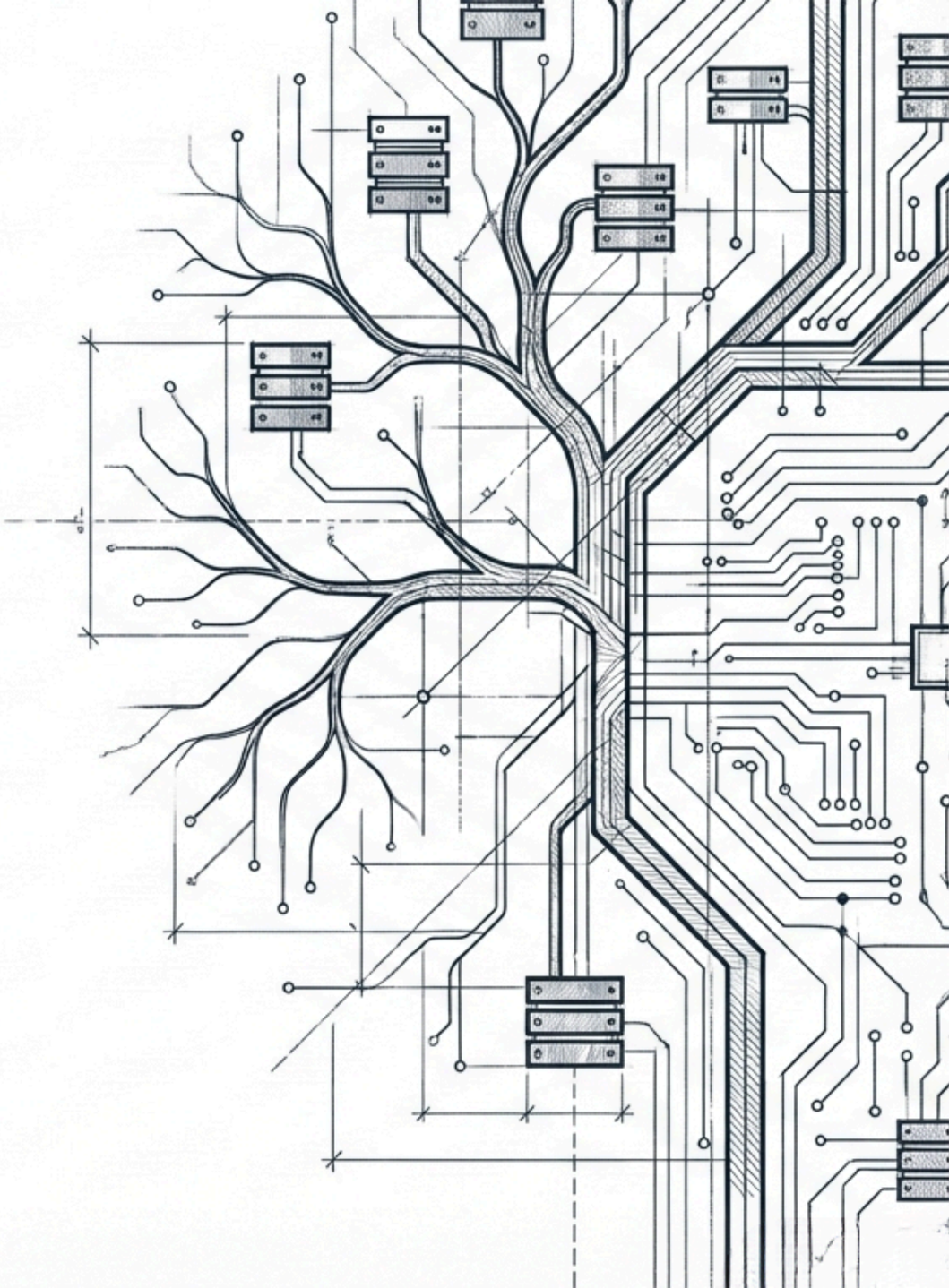
Commoditized & Interchangeable

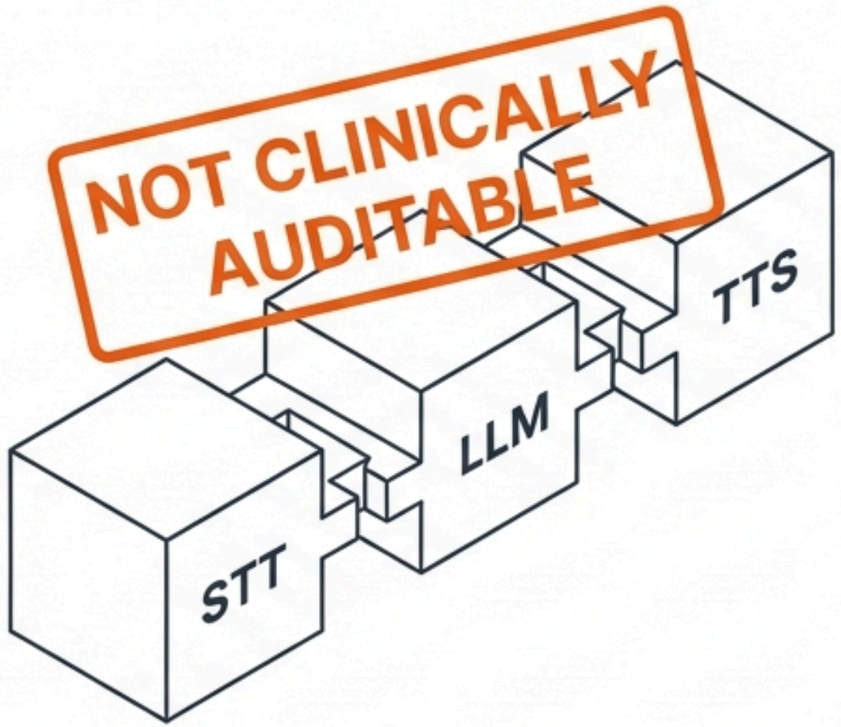
hana

Layer 2: Voice Agent Platforms

Orchestration, telephony, and deployment tooling.

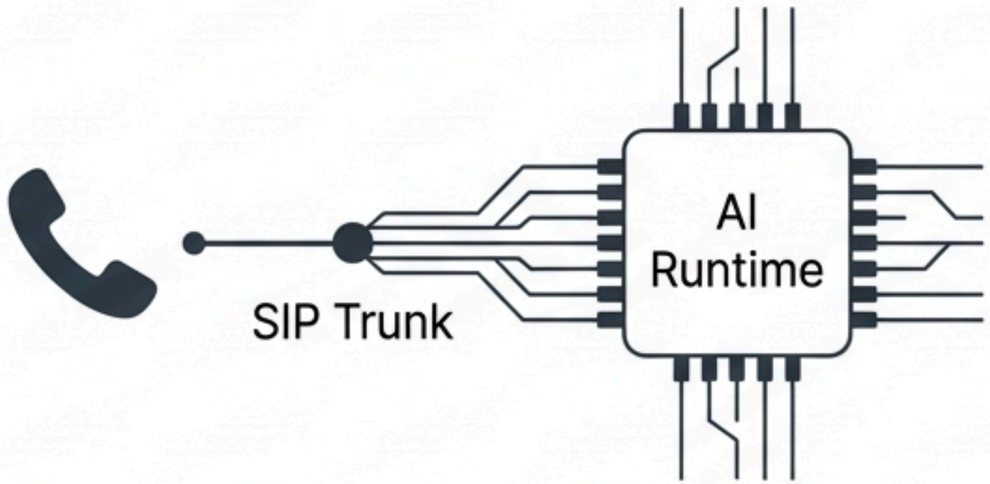
44 Companies





The Builders (Orchestration)

Platforms like Voiceflow ($78.5M funded), Synthflow AI ($49M), and Retell AI made it possible to assemble agents from modular parts in days. They solved assembly, not clinical readiness.



The Plumbing (Telephony)

This layer handles routing, SIP, WebRTC, and runtime latency. It is essential, mostly invisible, and handled by infrastructure players like Telnyx ($120M), or Twilio

The Missing Layer

Out of 247 companies mapping the Voice AI ecosystem, only 3 (Cyara, Langtail, ViClinic) focus explicitly on testing and observability.

In a safety-critical industry where a hallucinated medication dosage is an existential liability, this is the ecosystem's most dangerous blind spot.

Easy to build. Hard to trust.

A talking agent is commoditized. A clinically deployable one is not.

	The Demo	The Healthcare Reality
State	Stateless session API.	Persistent longitudinal memory across encounters.
Logic	Answers generic LLM questions.	Executes rigid clinical protocol reasoning.
Safety	Standard prompt guardrails.	Hard-coded hallucination interlocks & medical triage handoffs.
Security	Standard API keys.	HIPAA/BAA compliance, E2E encryption, EHR read-write.

hana

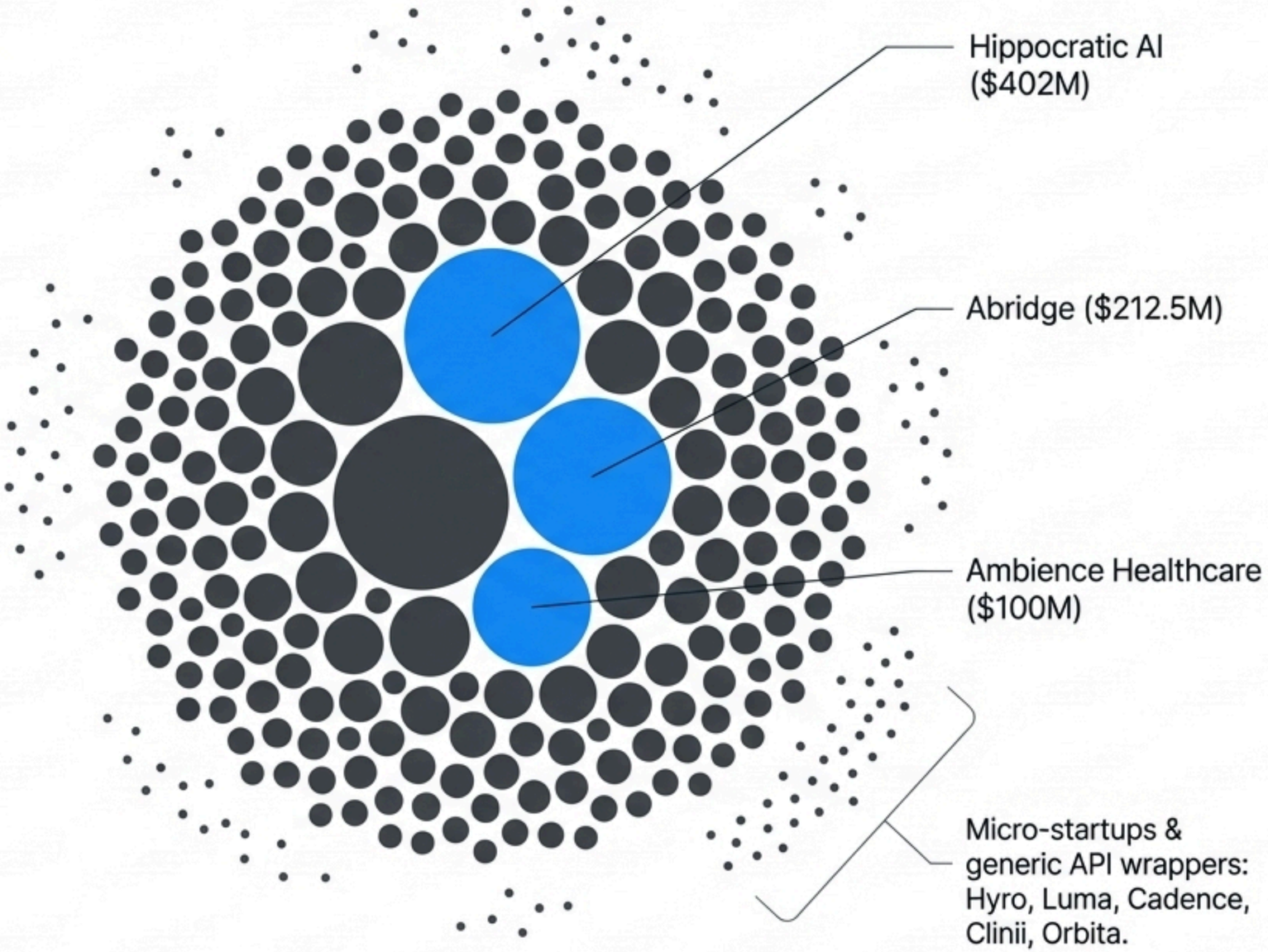
Layer 3: Engagement Applications

Where voice AI meets patients, clinicians, and care workflows.

128 Companies

The Biggest Bucket: Healthcare Engagement

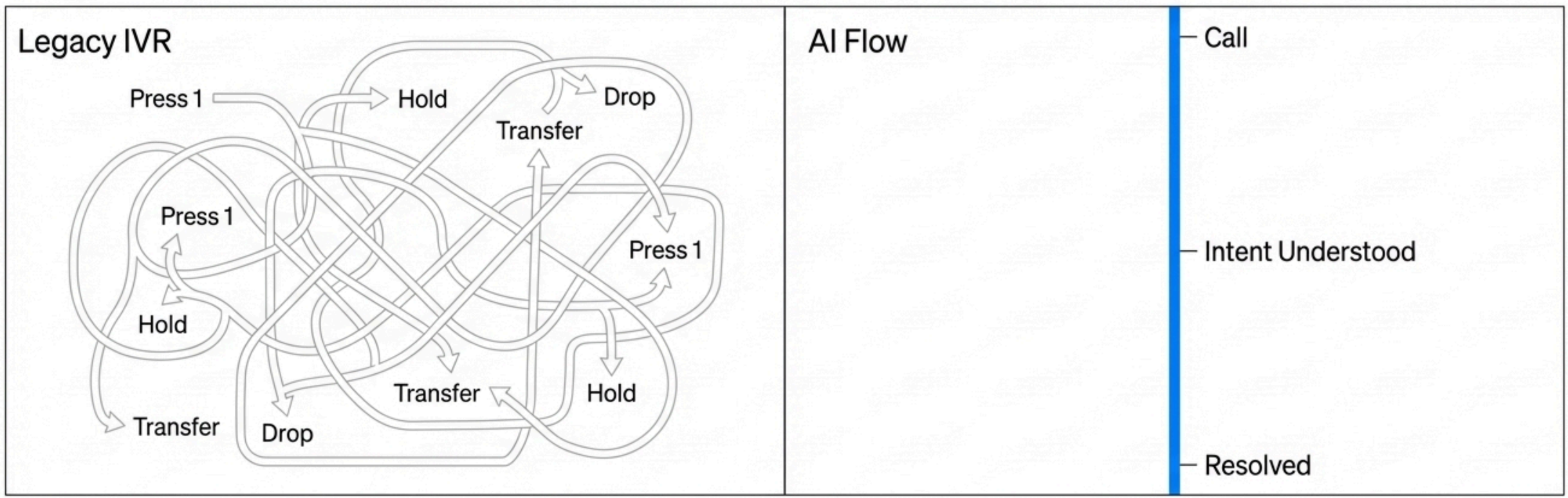
Patient engagement is the largest category, but highly fragmented. Many entrants are simply repackaged generic voice APIs wrapped in healthcare messaging, lacking deep clinical moats.



The Front Door: Access & Scheduling

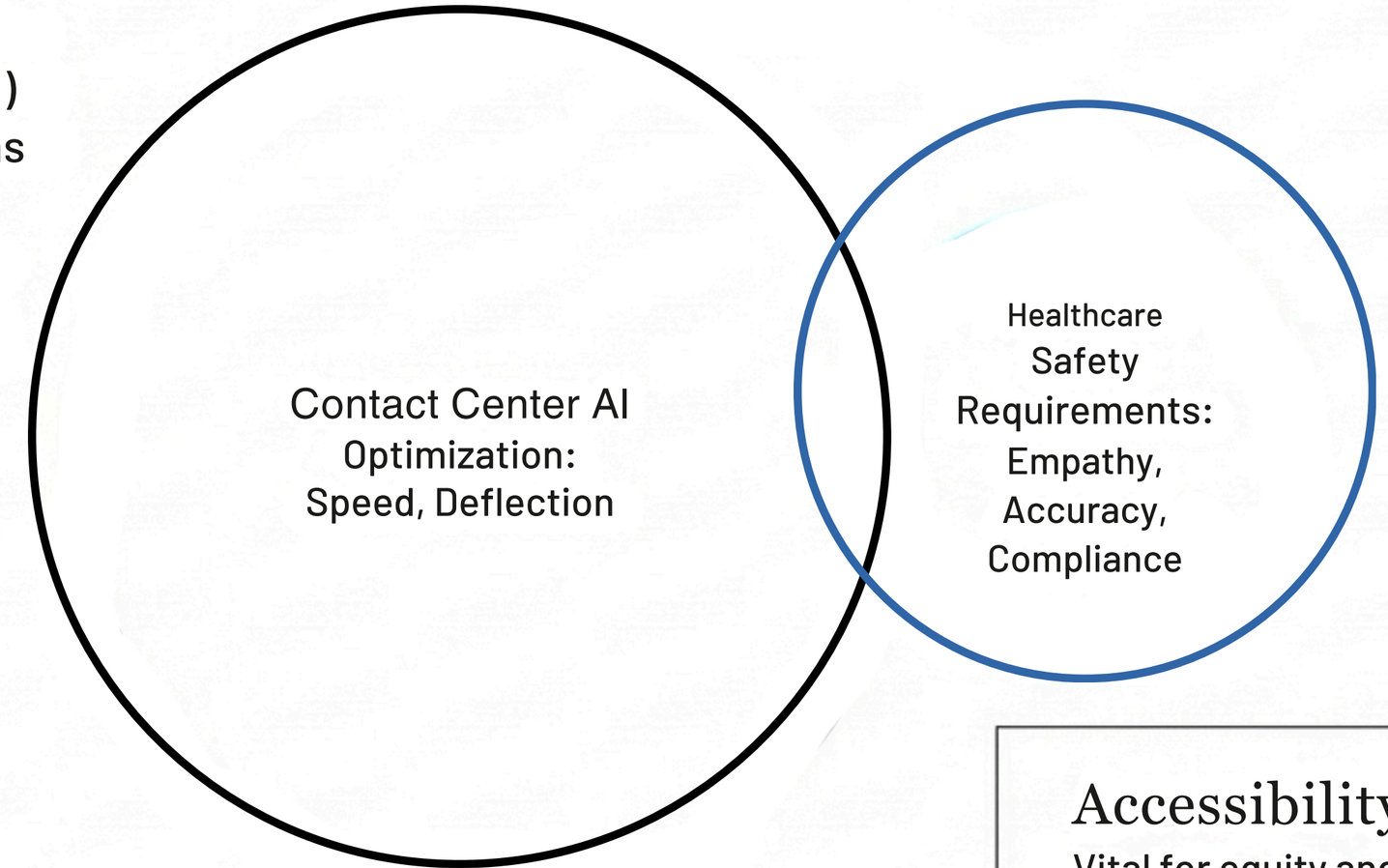
Scheduling, insurance verification, and inbound call handling offer the clearest, most immediate ROI in healthcare. Every minute spent on hold with a payer is a minute not spent on care.

Key Players: Syllable AI ($122M), Infinitus ($102.9M), SuperDial, Hello Patient,Talkie.ai.



The Crossover Problem & Adjacencies

Massive contact center companies (PolyAI, ASAPP, Cresta, Observe.AI) are shaping consumer expectations for Voice AI. However, their However, their optimization goals actively conflict with healthcare's strict safety requirements.



Sales Intelligence

Minor market impact, high revenue efficiency (Gong, Dialo).

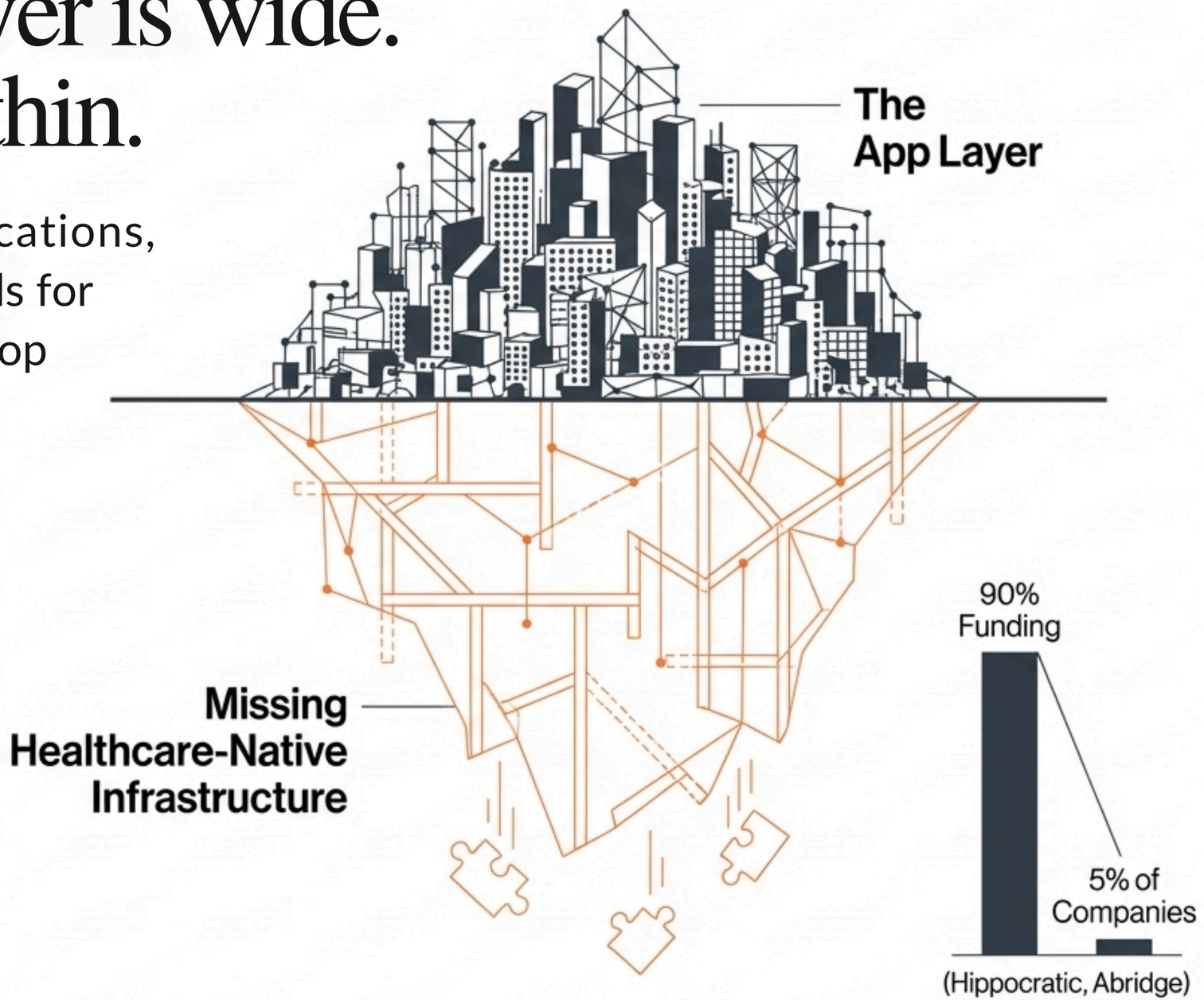
Accessibility

Vital for equity and inclusion; live translation and captioning (Sanas, (Sanas, DeepL, Nagish).

The application layer is wide. The foundation is thin.

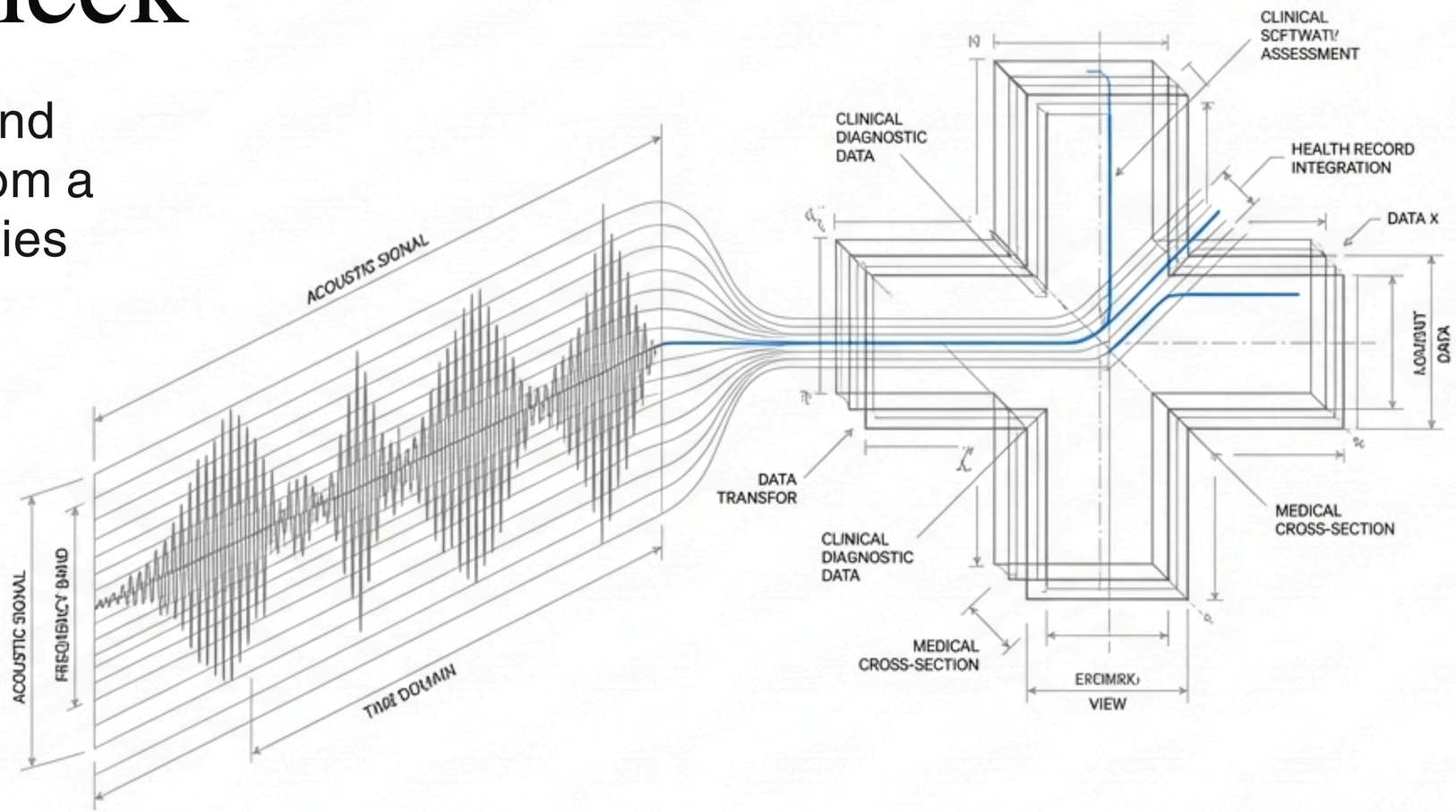
128 companies are building applications, yet there are zero common standards for clinical-grade Voice AI. Most sit on top op of generic, non-medical stacks.

The real venture opportunity is not in launching another engagement app—it is in building the missing healthcare-native infrastructure beneath them.



The State of Clinical Voice AI: A Reality Check

Four structural insights and five predictions drawn from a teardown of 247 companies



⚠ SAFETY NOTE: HALLUCINATION RISK ASSESSMENT INCLUDED

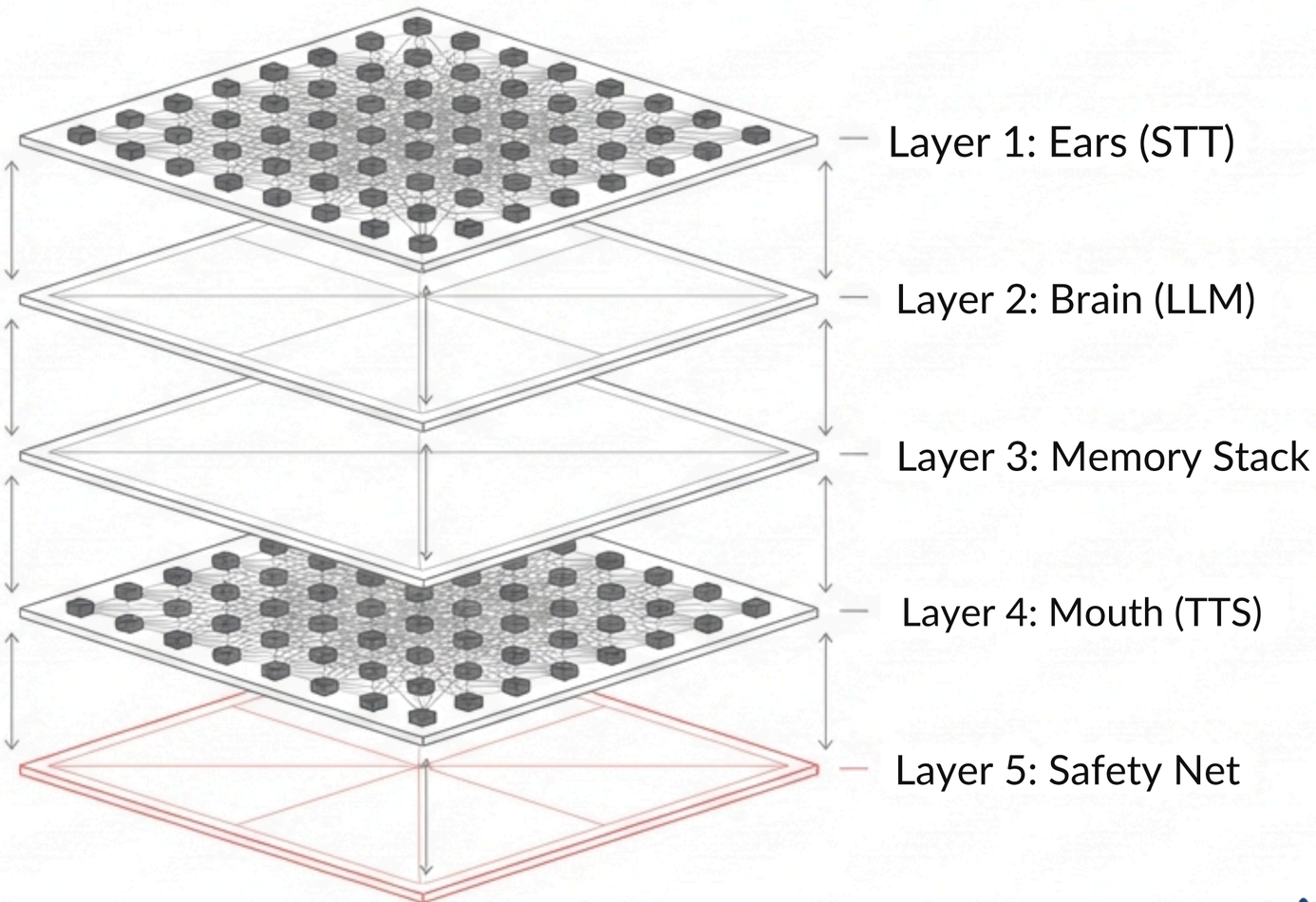
FIGURE 1: ACOUSTIC SIGNAL TRANSFORMATION TO CLINICAL DATA BLUEPRINT

STRUCTURAL INSIGHTS & PREDICTIONS FOLLOW

hana

Most companies are building organs, not organisms.

Many companies build one component well, but few connect ears, reasoning, memory, speech, and safety into one coherent system.

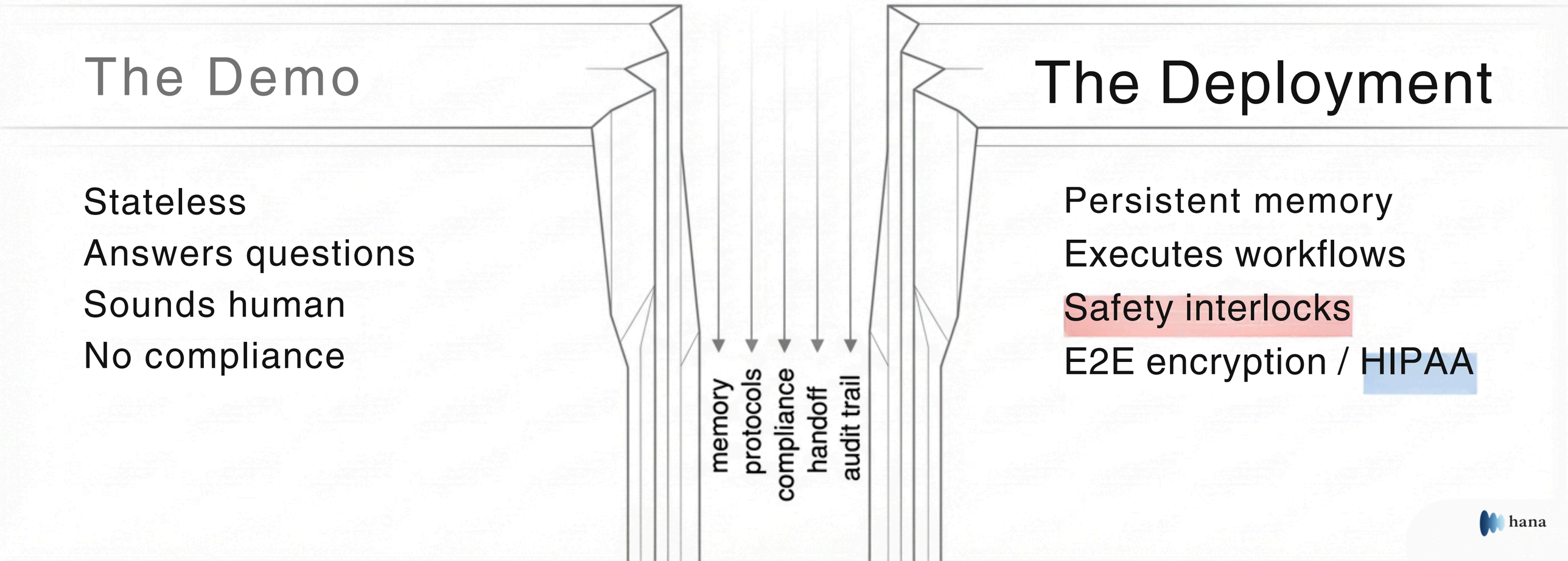


The demo-to-deployment gap is the product.

The difference between a compelling 60-second demo and a real healthcare deployment is everything that happens behind the scenes.

The Demo

Stateless
Answers questions
Sounds human
No compliance



The Deployment

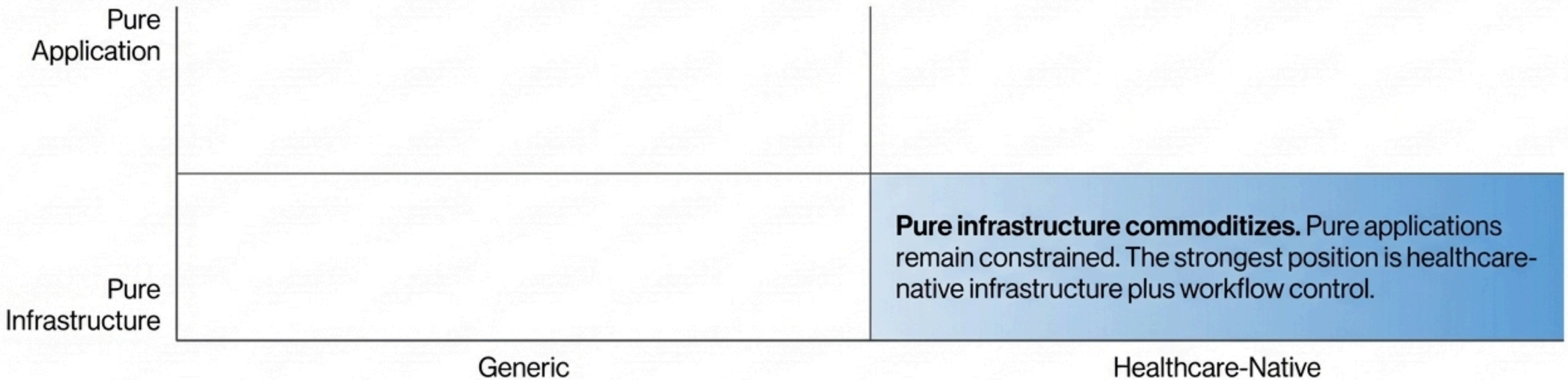
Persistent memory
Executes workflows
Safety interlocks
E2E encryption / HIPAA

hana

Healthcare breaks generic infrastructure.

The Stakes Spectrum

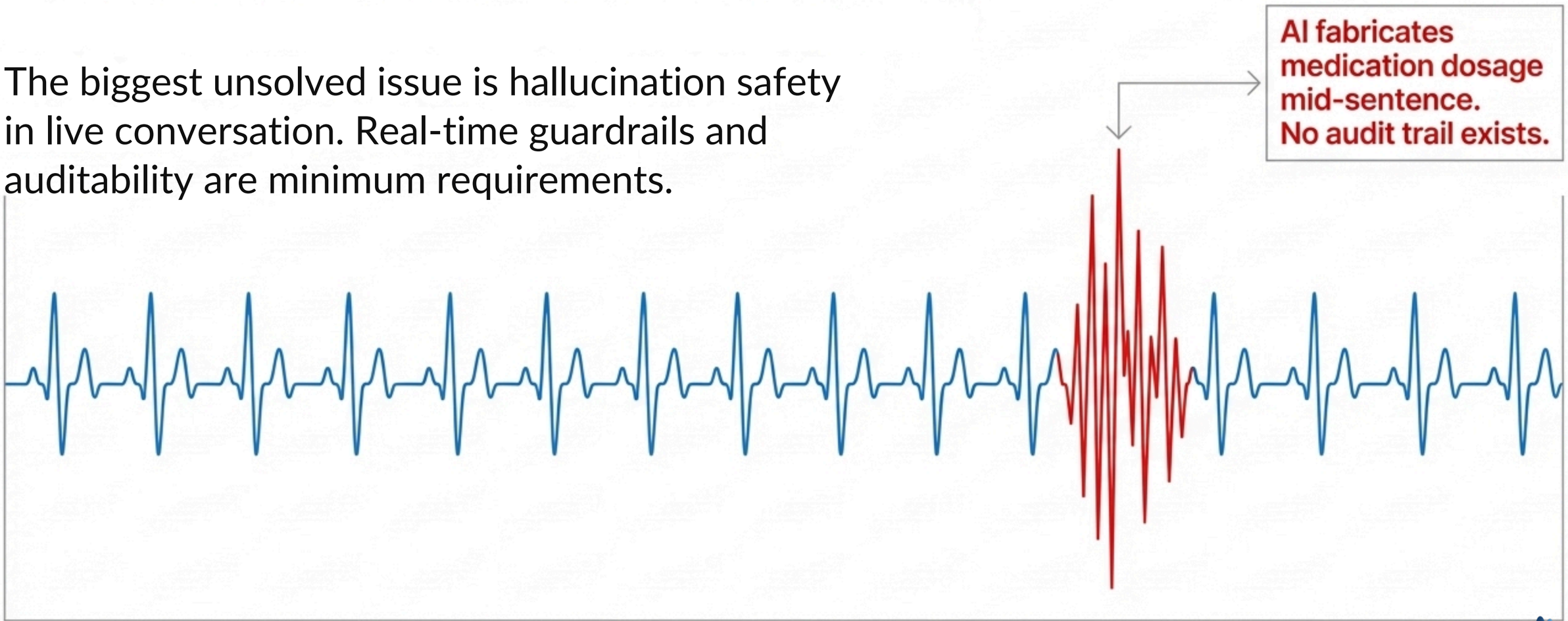
Winners will own both application and infrastructure.



hana

In healthcare, sounding human is not enough.

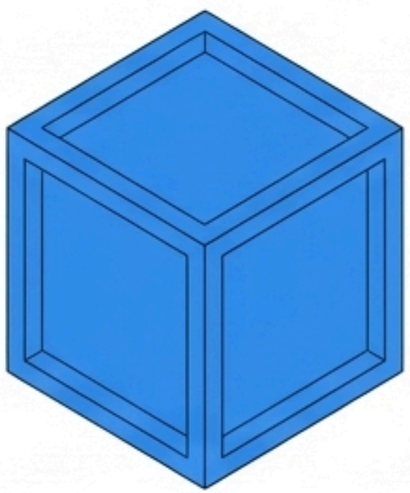
The biggest unsolved issue is hallucination safety in live conversation. Real-time guardrails and auditability are minimum requirements.



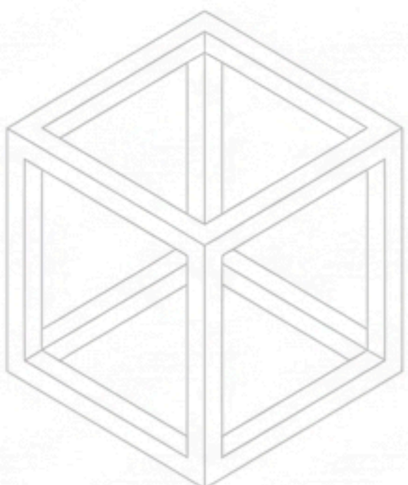
hana

Agents can talk.
Almost none can remember.

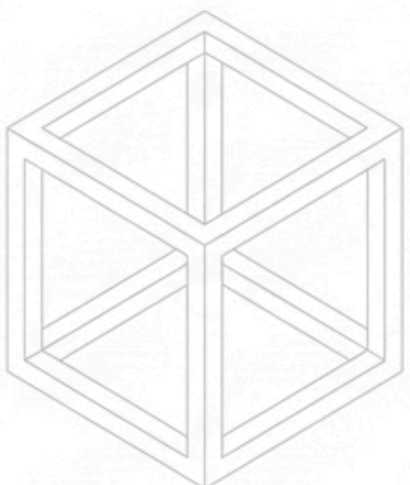
Persistent longitudinal memory is missing across most of the market. In healthcare, patient continuity depends entirely on it.



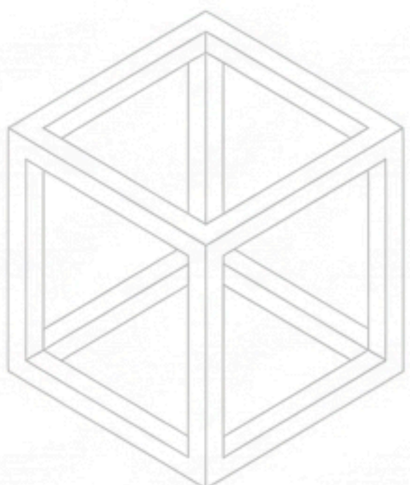
Short-term session memory



Semantic memory



Episodic memory



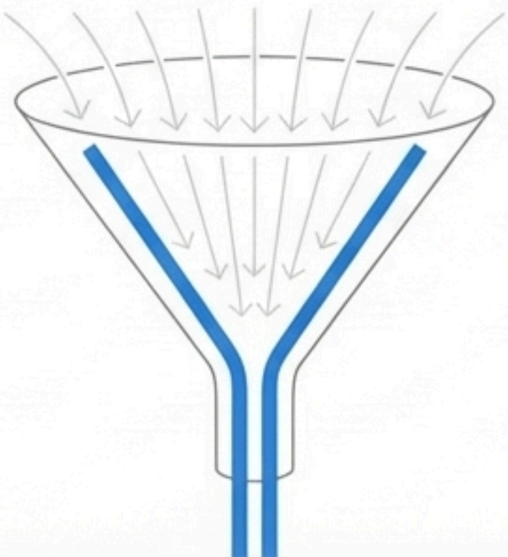
Longitudinal memory

hana

The 2026-2027 Horizon: Consolidation & Compliance.

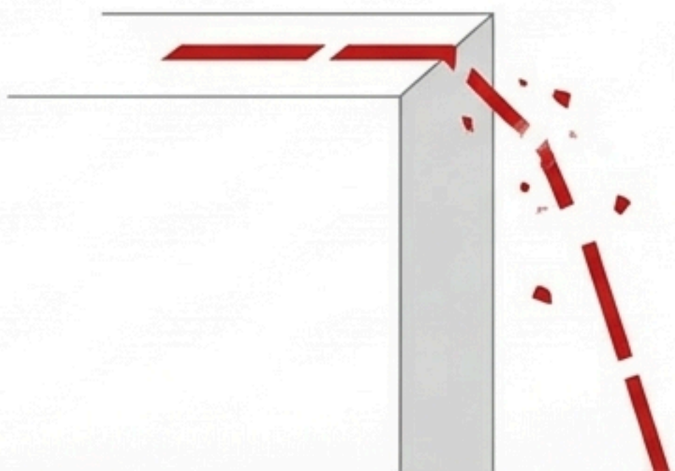
1 Infrastructure will consolidate.

STT and TTS will narrow to a few winners. The long tail of generic providers will struggle to differentiate or survive.



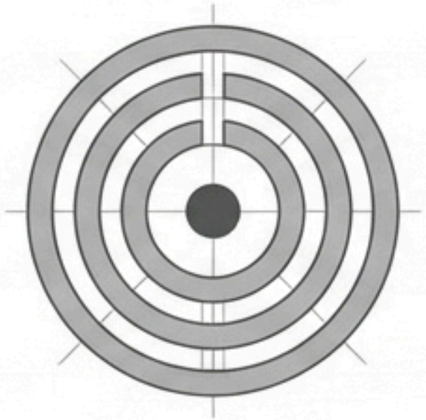
2 Safety failures will trigger a compliance cliff.

A high-profile clinical mistake will bring scrutiny. Enterprise buyers will rapidly shift toward vendors with strong audit and safety systems.

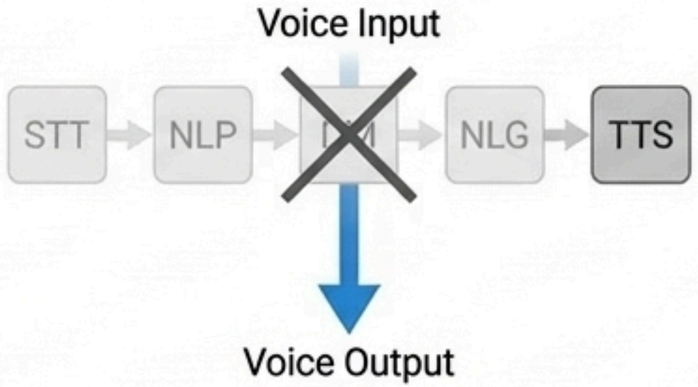


hana

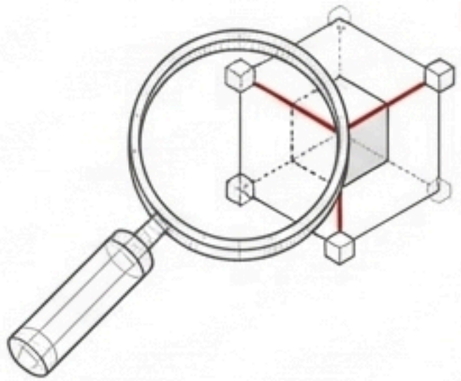
The 2026-2027 Horizon: New Models & Moats.



Memory becomes the moat.
The next competitive advantage in healthcare voice AI will come from continuity, personalization, and longitudinal context.



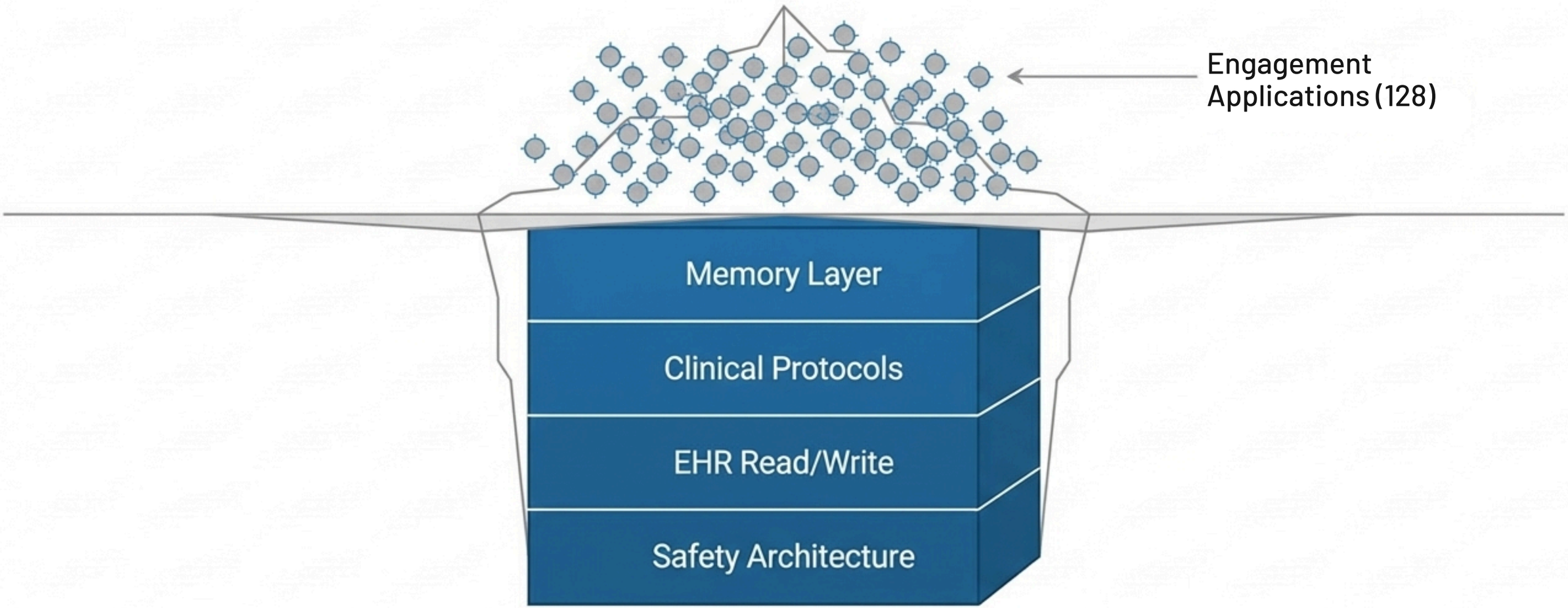
Voice-native models collapse the stack.
Models that bypass text translation will shatter latency barriers.



The sophisticated buyer.
Hospitals will buy based on verifiable safety architecture, not slick demos.

128 applications. Zero shared standards.

The application layer is wide and noisy. The structural layer beneath it is dangerously thin.



hana

The opportunity is not another voice app.

The market has plenty of interfaces and not enough infrastructure. The winners will build the memory, safety, and clinical reasoning layers that make voice AI usable in care.