

futuresource✱



BRAGI

Device-to-Service Integration in Audio: From Fragmentation to Platform Infrastructure

September 2026

Whitepaper co-authored by Bragi and Futuresource

Nikolaj Hviid, CEO, Bragi

Saranraj Mathivanan, Senior Analyst, Futuresource



Contents

System Overview: The Six-Layer Architecture	3
Structural Conclusion: Why a Coordination Layer is Inevitable.....	4
PART I — Market Analysis	5
1.1 A Market in Transition.....	5
1.2 The Silicon Intelligence Race.....	6
1.3 From Playback to Interaction: What the Evidence Shows.....	8
1.4 The Fragmentation Architecture	9
1.5 The Fragmentation, Measured.....	10
1.6 The Case for Inevitability: Counterfactual Analysis.....	12
1.7 The Closing Window	12
1.8 How Value Will Concentrate: Platform Economics	13
PART II — Competitive Landscape	14
2.1 The Structural Constraint Map.....	14
2.2 Big Tech: Capability Without Neutrality.....	15
2.3 SoC Vendors: Distribution Without a Service Layer	15
2.4 Chinese AI Platforms: Vertical Integration and Its Limits	16
2.5 Long-Tail Software Platforms: App-First Architecture and Its Ceiling.....	16
2.6 Where Competitors Demonstrably Lead.....	17
PART III — Bragi: Strategy Analysis	18
3.1 What Bragi is Doing: Descriptive.....	18
3.2 Strategic Alignment: Where the Strategy Fits the Market Structure	19
3.3 Execution Risks	19
Conclusion: The Standard for Evaluation.....	20
Sources and Data Notes.....	21

This document presents a structural analysis of the global audio wearables market. Its central question is straightforward: under what conditions does a software coordination layer emerge as an economic necessity in this market, rather than a company aspiration?

The argument proceeds in three parts. Part I establishes the market structure, the nature of the fragmentation problem, and why a coordination layer is a structural inevitability. Part II analyses the competitive landscape — the players attempting to address this market and the structural constraints that prevent any of them from fully solving the coordination problem. Part III applies this framework to the strategy of one company, Bragi, as an analytical case study in the same register.

A note on terminology. In the audio hardware supply chain, SoC vendors design the chips that power headphones. ODMs (original design manufacturers) build physical devices using those chips. Brands market and sell finished products. Service providers — Spotify, Fireflies, AI platforms — integrate above all three. And the platform layer — the subject of this analysis — sits between them all, coordinating identity, monetisation, and service routing across brands and chipset families. This document uses these terms precisely

System Overview: The Six-Layer Architecture

Before examining the fragmentation problem, it helps to establish how the audio supply chain functions as a system. Six participant types interact in this market. Each has distinct economics, distinct incentives, and a distinct structural constraint that prevents it from solving the coordination problem alone.

Participant	Role	Revenue Model	Primary Constraint
SoC Vendor	Designs the chip: radio, audio processor, NPU, power management	Chip margin per unit shipped (one-time)	Building the application layer above the chip creates channel conflict with the OEM brands it sells silicon to
ODM	Manufactures the physical device on top of the chipset reference design	Assembly margin (9–12%); project fees for bespoke work	Margins too thin to absorb AI integration risk; no engineering capacity to build independently
Brand	Markets and sells the finished product under its own identity	Hardware margin; minimal recurring revenue	Cannot build AI in-house at viable cost; companion apps aside, loses the meaningful user relationship as app engagement drops sharply after setup
Service Provider	Delivers content, AI, or productivity functionality via the audio interface	Subscription, advertising, or affiliate revenue from end users	Bilateral integration with each brand is economically infeasible at the scale this market requires
Platform Layer	Coordinates identity, service routing, subscription, and memory across the whole ecosystem	Per-device licensing; subscription revenue share; marketplace take-rate	Must be adopted by the SoC, ODM, and brand before it can reach the end user
End User	Uses the device for audio, AI, and connected services	Subscription for AI services; one-time hardware purchase	Limited by what the device and platform make available; loses accumulated data on every device upgrade

The implication is structural. No single participant has both the economic incentive and the technical capability to build the coordination layer. The SoC vendor cannot do it without creating channel conflict. The ODM cannot afford it. The brand lacks the engineering scale. The service provider faces a bilateral cost that is mathematically unviable. The platform layer, if it exists, resolves all of these simultaneously — but it must first be adopted by participants who have no precedent for paying for it.

This is the core tension the document examines.

Structural Conclusion: Why a Coordination Layer is Inevitable

The emergence of a software coordination layer in the audio wearables market is not contingent on any particular company succeeding. It is a structural outcome of the economic forces this document examines. While only a few major players have the resources to build it themselves, the majority of brands and services are left behind without an independent coordination layer.

This is a strong claim. It requires precise grounding.

The argument rests on three observations. Each one, standing alone, is sufficient to drive platform emergence. Together, they make the outcome close to certain.

First: hardware capability has outrun the software layer above it. Five Bluetooth Audio SoC companies have begun to embed on-chip AI accelerators (micro NPUs or equivalent) into earphone chips. This capability is so far confined to flagship models shipping in relatively low volumes; the large majority of earphones shipped to date carry no on-device AI accelerator. These devices are technically capable of running real-time AI-enabled interaction. What is missing sits above the silicon: a shared software layer that would make that capability economically meaningful across brands at scale. This is not a shortcoming of any one chip or vendor. It is a system-level coordination gap.

Second: fragmentation makes bilateral integration impossible. Reaching up to 3,000 brands active on Amazon US alone through bilateral service integration would require \$1.05 billion to \$3.96 billion in upfront engineering investment — for a single service, in a single country. Most brands selling fewer than 1,000 units per year have no engineering team to receive an integration. They are unreachable by any mechanism except chipset-level default embedding. This is arithmetic, not opinion.

Third: every existing player is prevented from serving as the neutral layer. Apple captures value through ecosystem exclusivity. Google's revenue model is advertising. Amazon routes users toward commerce. SoC vendors cannot build application software without channel conflict. Chinese AI platforms serve non-neutral ecosystem interests. None of these players can create the aligned incentives the market requires.

Taken together, these three observations establish that a neutral coordination layer — one embedded at the chipset reference design level, revenue-neutral with respect to brand identity, and structured to create financial upside for every participant — is not a business strategy. It is a market structure requirement.

The sections that follow build the evidentiary case for each component of this argument.

PART I — Market Analysis

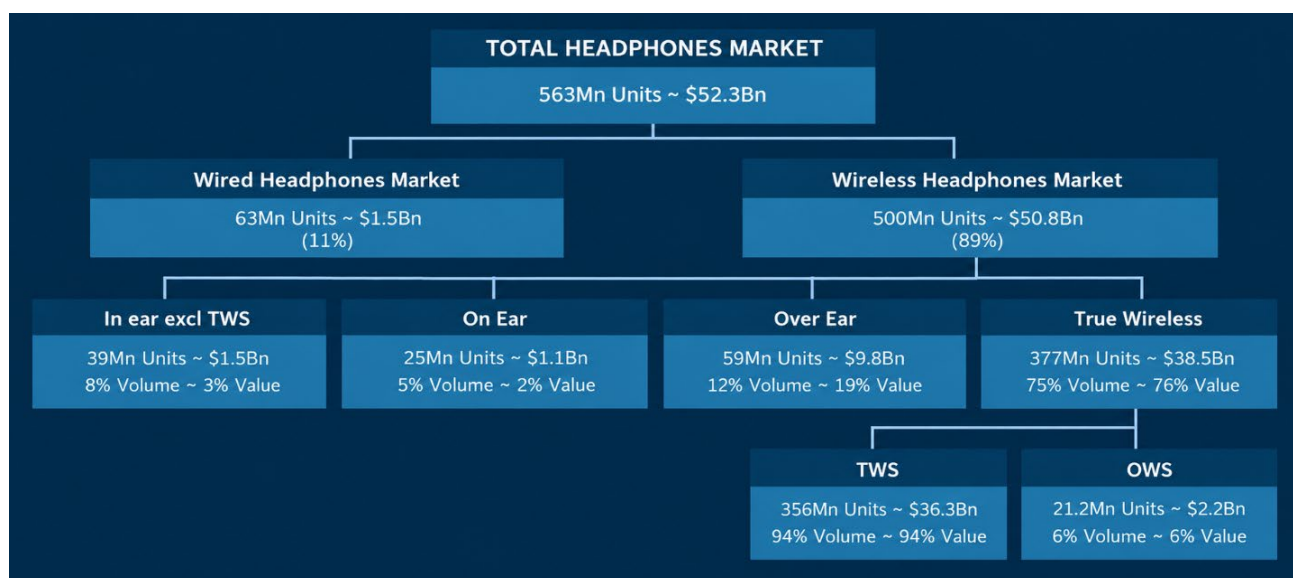
The forces shaping audio wearables — independently of any specific market participant

1.1 A Market in Transition

Audio devices have historically been evaluated as accessories — hardware endpoints differentiated by acoustic performance, price-to-performance ratio, and incremental feature additions. This framing is no longer adequate.

What is unfolding in the global hearables market is not the evolution of a product category. It is the reorganisation of a computing interface.

*2025 Headphones Market Size

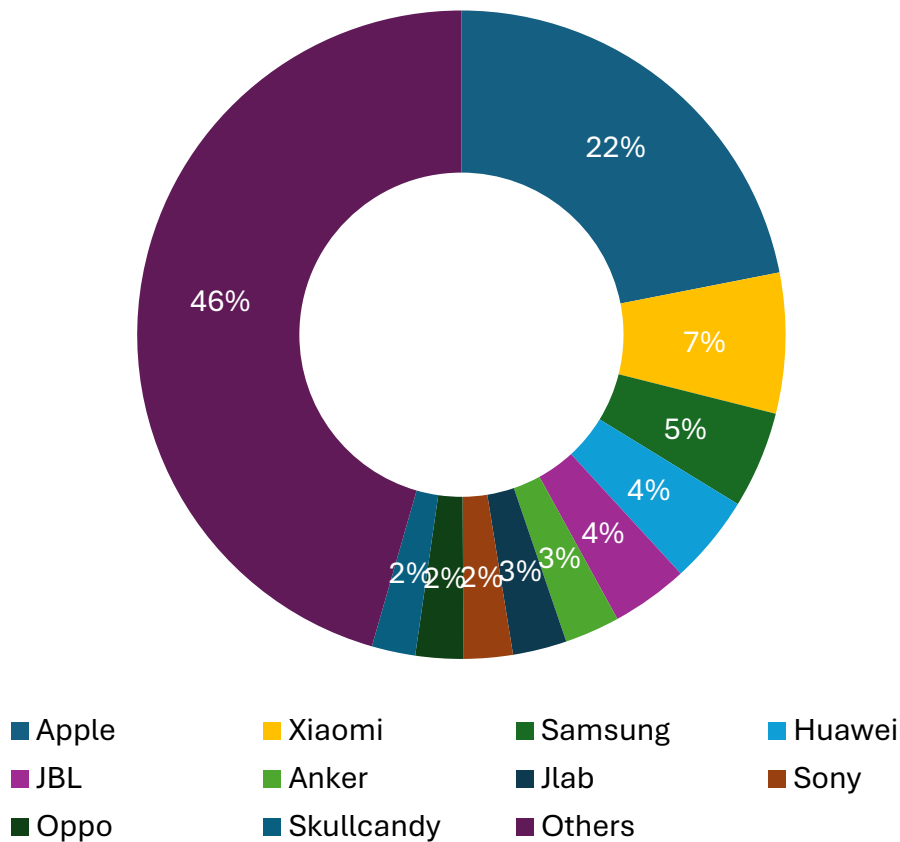


(Chart above excludes White Label Products)

The global branded headphones market reached 563m units in 2025, with the industry now overwhelmingly dominated by wireless devices. Wireless headphones accounted for 500m units, generating US\$50.8bn in revenue, while wired headphones represented 63m units and only US\$1.5bn in market value. This highlights the extent to which value creation has shifted towards connected devices, with wireless products accounting for the vast majority of both market volume and revenue.

Within the wireless segment, the market has consolidated around true wireless stereo (TWS) and open wireless stereo (OWS) form factors. Together, these categories generated 377m units and US\$38.5bn in revenue during 2025, reinforcing their position as the dominant segment within the global headphones market. The combined TWS and OWS market is projected to grow to 448m units by 2030, supported by continued consumer demand for cable-free audio solutions and ongoing innovation in wireless technologies.

TWS & OWS Market : Top 10 Brand Share (2025)



The TWS and OWS market also exhibits a relatively concentrated competitive landscape. Apple remained the clear market leader, accounting for 22% of combined TWS and OWS shipments while capturing 44% of category value, reflecting its premium positioning and strong ecosystem advantages. It was followed by Xiaomi, Samsung, Huawei, JBL, Anker, JLab, Sony, Oppo and Skullcandy. Collectively, these ten manufacturers accounted for approximately 54% of global TWS and OWS shipments, indicating that although a small group of leading brands dominates the market, nearly half of global shipments remain fragmented across numerous regional and specialist manufacturers.

Fragmented markets exist across every consumer electronics category. That, on its own, is not the observation that matters.

The critical observation is the combination of fragmentation with advancing silicon AI capability that has arrived faster than any software layer to coordinate it. The earphone market has acquired hardware capable of running real-time ambient AI interaction — and has done so without the platform infrastructure to deploy that capability consistently across brands, services, and users. The hardware investment cycle has run 18 to 24 months ahead of the platform layer.

That gap is the structural opportunity this document examines.

1.2 The Silicon Intelligence Race

Every pair of wireless earbuds and headphones is powered by a Bluetooth audio system on chip (SoC), which integrates wireless connectivity, audio processing, memory and power management. The SoC market is relatively concentrated, with Bestechnic, Apple, MediaTek's Airoha, Qualcomm, Jieli and

HiSilicon collectively accounting for over 90% of branded wireless audio SoC shipments in 2025 (Futuresource tracked market, excluding white label production).

Bestechnic is the largest supplier to this branded market by shipment volume, at nearly 39%; across the total market including white label, Jieli leads on units. Its strong position has been driven by widespread adoption across Chinese brands and recent premium design wins, including Samsung's Galaxy Buds3 Pro. Its latest BES2800 platform integrates a neural processing unit (NPU) capable of GOPS-level AI performance.

Apple is the second largest SoC vendor, accounting for 18.5% of global shipments through its in-house H series chips, which power the entire AirPods portfolio.

MediaTek's Airoha, with a 14.1% shipment share, has established a strong position across leading consumer audio brands, including Sony, JBL and Jabra, while Qualcomm continues to serve the premium Android ecosystem with its latest S7 Pro platform, which incorporates a dedicated AI processor delivering better AI performance than its predecessor.

Jieli remains a primary supplier for entry level wireless audio devices, while HiSilicon continues to support Huawei's expanding audio portfolio.

The latest generation of wireless audio SoCs is introducing dedicated AI hardware to support on-device AI processing. Qualcomm, Bestechnic and Apple have already integrated AI processing capabilities into their flagship chipsets, laying the foundation for features such as real-time voice processing, contextual awareness and personalised audio. As more SoCs incorporate these capabilities, value creation is expected to shift increasingly towards the software platforms and services that orchestrate the user experience. Bluetrum has also demonstrated one of the earliest publicly announced integrations between an audio SoC platform and a large language model, with its BT895x platform integrated with ByteDance's Doubao LLM through the Volcano Engine cloud platform. As AI capabilities continue to evolve, the industry's focus is expected to shift from silicon performance towards the software platforms and services that can fully leverage these capabilities.

Reading these trajectories together reveals a conclusion that none of the vendors would state publicly: the hardware they are building has no software platform above it to coordinate it into a unified user experience. Every vendor is racing toward on-device neural processing from different directions. The silicon is ready. The coordination layer is not.

This is the central paradox of the audio wearables market in 2026. Semiconductor companies are collectively spending hundreds of millions of dollars embedding TOPS-level AI compute into earphone chips. Every one of those chips is technically capable of running embedded AI interaction. None of them is connected to a software platform that manages user identity, routes service interactions across brands, accumulates interaction memory, or creates the economic infrastructure for subscription monetisation.

The hardware investment cycle has outrun the platform investment cycle. The silicon waits for the layer that will make it useful.

1.3 From Playback to Interaction: What the Evidence Shows

The transition from passive playback endpoint to active AI interaction surface is real. But it requires careful characterisation. It is not yet a mass-market reality in most markets. It is an emerging behaviour with strong directional evidence, accelerating fastest in specific geographies and user segments.

Three sources of evidence are worth examining in sequence.

Adjacent category validation. Smart glasses saw a substantial increase in 2025 with a year-on-year volume increase of 150%. This led to shipments of 6.1 million in 2025 with a CAGR of 25.1% between 2026-2030. Revenue reached \$2.1 billion in 2025 and is expected to reach \$7.1 billion by 2030.

Demand is driven by lightweight, hands-free functionality enabled by miniaturised components, improved sensors, and AI-enhanced voice interfaces without relying on visual overlays. The USA leads global adoption, while ongoing innovation in materials, battery life, audio, and on-device AI will continue to boost both volume and value through 2030. Furthermore, Futuresource sees this segment converge with AR which would lead to a volume CAGR of 26.8%.

The segment is overwhelmingly led by Meta with 73% of global demand. In June 2026, Meta introduced a new line-up of smart glasses, without the Ray-Ban branding, although the glasses are still manufactured in partnership with EssilorLuxottica. The competitive landscape in the space is also heating up with recent introductions from Alibaba earlier this year, and notable collaborations from Samsung and Google with fashion brands (Warby Parker and Gentle Monster).

Hardware as health platform. Apple's AirPods Pro 2 received FDA clearance in September 2024 as an over-the-counter hearing aid — the first such clearance for a consumer earphone — enabling clinical-grade audiometry for mild to moderate hearing loss. AirPods Pro 3, announced in September 2025, added heart rate sensing, live AI-powered translation, and dedicated fitness coaching. Supply chain analysts project AirPods with infrared cameras for 2026. Apple's services revenue reached \$109 billion in fiscal year 2025 across 2.35 billion active devices — approximately \$46 per device per year.

The implication is clear: Apple is building AirPods as an AI wearable platform, not an audio accessory.

China as leading indicator. In China, where LLM-native earphone products are twelve to eighteen months ahead of Western markets, the trajectory is steeper. ByteDance launched Ola Friend earbuds in October 2024 with its Doubao large language model — China's most widely used generative AI application, with 27.5 million monthly active users — integrated directly into the earphone firmware. Sales of AI-classified earphones on Chinese e-commerce platforms grew 261 percent year-over-year in 2024, then reached a further 382,000 pairs in Q1 2025 alone. That is roughly a tenfold increase from the equivalent quarter two years prior.

Alibaba has followed a similar path with its Qwen-powered Clip earbuds, which integrate Alibaba's Qwen large language model into the earbud platform through an AI-enabled SoC. The integration enables more natural, conversational interactions and AI functions directly through the earbuds, illustrating how LLM capabilities are moving deeper into the device architecture. Together with Ola Friend, the example points to a broader Chinese trend of embedding generative AI into the earbud hardware stack. The model is now spreading beyond China. In India, boAt has partnered with Google Cloud to bring Gemini-powered AI to its next generation of smart audio devices, enabling conversational interactions and capabilities such as real-time translation and hands-free access to services. This

suggests that the LLM-earbud opportunity is beginning to move from a China-led experiment toward a broader ecosystem involving global AI providers and established audio brands.

The *Audio Tech Lifestyle 2024* yearly consumer research study conducted by Futuresource Consulting to understand consumer usage and attitudes towards audio equipment surveyed 10,000 respondents across the US, UK, China, Japan and Germany. The study found that 45% of consumers own headphones capable of triggering a voice assistant, and among these users, 65% frequently use the functionality, underscoring the growing importance of voice-enabled experiences in the connected audio ecosystem.

GPT-4-class voice models reduce AI response latency to approximately 320 milliseconds, enabling natural real-time conversation rather than command-and-response. Speech recognition error rates have dropped 30 percent with current-generation models. The AI interaction that failed to achieve mass adoption between 2017 and 2022 failed not because consumers rejected the concept. It failed because the technology was not good enough. That condition no longer holds.

The subscription evidence is equally instructive.

Subscription intent. Research conducted by Bragi with 500 AI-paying TWS users in the United States in August 2025 (Bragi proprietary research, 2025) found 75 percent stated likelihood to subscribe to a Chat AI application on headphones at \$3 to \$10 per month. Users who are voice-first in their AI behaviour — preferring voice input to typing — are 16 times more likely to subscribe. Approximately 50 percent of AI-engaged TWS users in the United States and Germany already pay for at least one AI subscription.

What drives purchase intent. A follow-on Bragi survey in November 2025, covering 1,500 respondents across the United States, Germany, and India, found that those willing to subscribe consistently rank AI model quality and tool and service integration depth above audio quality — which inverts the traditional earphone purchase hierarchy. Users evaluating an AI subscription are not primarily evaluating acoustic performance. They already made that decision when they bought the device. They are evaluating whether the AI layer delivers genuine utility in voice-first workflows.

The platform creates the conditions for engagement to accumulate; subscription converts the engaged.

1.4 The Fragmentation Architecture

Unlike smartphones, which consolidated around two operating systems and a small number of chipset vendors, the audio market evolved through a distributed supply chain with no natural coordination point. Each participant handles its own layer. No participant is responsible for the whole.

The fragmentation exists at multiple levels simultaneously. Hardware capabilities vary by chipset family — a Qualcomm QCC device and a Jieli budget device have fundamentally different firmware architectures, SDK conventions, and software development toolchains, even though both are Bluetooth earphones. Software stacks differ across brands, and sometimes across product lines within the same brand. Service integrations are inconsistent and shallow. The same action — initiating a voice interaction, connecting to a service, triggering a feature — can behave differently depending on the device, the companion app, and the service layer involved.

For years, this fragmentation was economically tolerable. The primary function of audio devices was playback. Integration complexity did not constrain the user experience. The device played audio, the phone managed interaction, and the service layer remained external.

This equilibrium breaks down once the device becomes an interface. A device that exists primarily to deliver audio requires no coordination between its layers. A device that exists to serve as an AI interface does. The distinction is structural, not a matter of degree.

The cost of building that coordination from scratch is substantial.

The cost of a single integration. A full bespoke AI voice assistant integration on a Qualcomm QCC earphone platform involves five distinct engineering workstreams: adaptation of embedded firmware to the SoC's proprietary audio chain and Bluetooth stack; wake-word detection and acoustic noise-cancellation tuning across real-world environments; a cloud AI pipeline (speech-to-text at approximately \$0.0043–\$0.006 per audio minute; LLM backend at \$2.50 per million tokens, enabling per-conversation costs of \$0.05–\$0.15 per interaction minute; neural text-to-speech synthesis); iOS and Android companion application development; and vendor certification. The engineering team required is five to ten specialists — embedded firmware engineers, audio DSP specialists, cloud integration engineers, and mobile developers — representing a thin global talent pool. At US contract rates of \$150–\$250 per hour for qualified embedded audio engineers, a six-to-twelve-month project costs \$400,000 to \$1.5 million in engineering fees alone, before ongoing per-device cloud service costs of \$0.50 to \$2.00 annually per active device.

Why it does not scale. This investment is not portable across chipset families. Firmware written for Qualcomm's QCC audio chain does not run on Jiel's proprietary SDK, which does not run on Bestechnic's architecture. Each of the SoC families requires a substantially fresh integration project. Most mid-tier ODMs do not have the engineering headcount to execute even one such project. They ship what the chipset reference design contains.

This is the mechanism that makes the coordination problem self-perpetuating. The ODM lacks the engineering capacity to integrate AI independently. The brand customer lacks the leverage to demand it. The chipset vendor has the reference design but focuses on silicon differentiation, not application-layer innovation. The service provider faces the steepest barrier of all.

Integrating across every chipset family costs \$400,000 to \$1.5 million per firmware integration. Across the thousands of brands active on Amazon US alone, that produces a bilateral investment of \$1.05 billion to \$3.96 billion. Reaching the 6,000-plus brands believed active globally adds an estimated \$1.3 billion to \$5 billion more — conservatively assuming a single chipset integration per brand. In practice brands ship across multiple chipset families, so the true figure is higher.

The result: the majority of earphones ship with unmodified chipset vendor reference firmware, despite the existence of AI technology that consumers demonstrably want.

1.5 The Fragmentation, Measured

The structural fragmentation described above is not an industry estimate. It can be measured directly.

On Amazon US alone, thousands of distinct headphone brands sell across tens of thousands of SKUs each year. Globally, 500 million wireless headphones are shipped annually, creating a large addressable market for AI platform providers.

The power law. The global branded headphone market comprises an estimated 800+ manufacturers (excluding white-label brands) and around 12,000 active SKUs yet remains highly concentrated. Apple, JBL, Sony, Xiaomi, and Samsung account for 37.3% of global branded headphone shipments, while the top 10 brands represent 48.8%. Beyond these leaders, market share fragments rapidly across hundreds of smaller brands with limited shipment volumes. This reflects a classic power law, where a small group of global brands captures disproportionate commercial success through strong product portfolios, extensive distribution, brand equity, and ecosystem integration.

SKU concentration. Concentration is even more pronounced at the product level. Despite an estimated 12,000 active SKUs, the top 100 models account for approximately 42% of global branded headphone shipments. The remaining SKUs primarily serve niche segments, regional markets, or specific retail channels, each contributing only modest shipment volumes. This demonstrates that commercial success depends far more on developing a handful of globally successful products than on maintaining an extensive product catalogue.

Beyond the branded market sits a large base of white-label SKUs. (The 800-plus branded manufacturers counted here, and the several thousand brands observed in the Amazon US dataset are different populations: the retail dataset includes white-label and marketplace-native sellers excluded from the branded-manufacturer count.) This is not a competitive market with many viable participants. It is a market with a small number of commercially meaningful products surrounded by an enormous catalogue that generates no material revenue and is built by ODM factories that ship whatever the chipset reference design contains.

The revenue concentration. Revenue is even more concentrated than shipments. Despite an estimated 12,000 active headphone SKUs, the top 100 models account for 42.0% of global shipments but 64.2% of industry revenue, reflecting the premium pricing of flagship products. Conversely, the remaining thousands of SKUs contribute just over one-third of market revenue.

This concentration highlights that value creation is driven by a limited portfolio of globally successful products rather than catalogue breadth. While 64% of active SKUs are priced below US\$50, they account for only 40.4% of shipments and primarily compete on affordability. In comparison, products priced above US\$100 also account for around 40% of shipment volume, representing a sizeable premium user segment with greater willingness to pay. This makes the premium tier the most attractive opportunity for AI subscription services.

Three figures define the coordination problem with a precision that abstract market descriptions cannot match: thousands of brands active on Amazon US over twelve months; \$400,000 to \$1.5 million per bilateral service integration; and \$1.05 billion to \$3.96 billion as the total bilateral investment required to reach a single country on a single retail channel. The 1,800 brands selling fewer than 1,000 units per year are structurally unreachable regardless of investment. A coordination layer that reaches these brands through chipset reference design inclusion is the only economically viable mechanism at scale by which they are reachable at all.

1.6 The Case for Inevitability: Counterfactual Analysis

The inevitability claim is best understood by examining what happens under each alternative architecture. There are three structural alternatives to a neutral platform layer. Each fails on identifiable grounds.

Alternative 1: Bilateral integration at scale. Every service provider integrates with every brand independently, and every brand integrates with every service independently. This is the current state of the market. The cost calculation above establishes why it fails: \$1.05 billion to \$3.96 billion to reach Amazon US brands alone for a single service, before accounting for the 1,800 brands with no engineering team to receive any integration. The cost problem, by itself, is sufficient. But beyond cost, bilateral integration produces an incoherent user experience. The same user's earphone behaves differently on different devices. Services cannot create consistent interaction patterns across brands. No participant accumulates the cross-brand user intelligence that would make the system valuable over time. Bilateral integration is not merely expensive. It is architecturally incapable of producing the platform-grade experience that the market trajectory requires.

Alternative 2: Vertical integration by a major platform. A large technology company — Apple, Google, Amazon, or a Chinese equivalent — builds the coordination layer and deploys it as a proprietary platform. This resolves the coordination problem for devices within its ecosystem. It does not solve it for the market. Apple serves 22 percent of the TWS and OWS market by volume. A Google or Amazon audio platform that routes users toward its own services will be resisted by brands that do not want their customers converted into a competitor's commercial traffic. A Chinese platform cannot serve Western brands on neutral terms without conflicting with the commercial interests that define its existence. Vertical integration produces a coordination layer that serves the integrator's interests, not the market's. That constraint cannot be engineered around. It is structural.

Alternative 3: Status quo — no coordination layer. The market continues to ship AI-capable hardware without a platform above it. The economic consequences are predictable. Service providers cannot access the installed base at viable cost. Brands cannot differentiate after the point of sale. Consumers receive AI capability that resets with each device upgrade. SoC vendors compete on silicon specifications with no recurring software revenue. The market's aggregate economic potential — approximately \$1.5 billion in gross annual recurring subscription revenue at 5 percent attach and \$5 per month across 500 million global wireless devices — remains uncaptured, accruing to no participant. The market does not tolerate this condition indefinitely. The uncaptured value is too large, and the coordination mechanism is too tractable.

The logical conclusion follows. A neutral platform layer is not one possible outcome. It is the only architecture that is simultaneously economically viable for all participants, technically feasible at scale, and sustainable against competitive pressure from adjacent players.

The question is not whether this layer emerges. The question is who builds it, whether it is neutral, and whether it achieves default embedding before the window closes.

1.7 The Closing Window

The opportunity to establish a neutral coordination layer is time-bound by semiconductor product cycles. This is the argument the market has not yet fully priced.

Bluetrum's BT895x chip integrating with ByteDance's Doubao LLM in 2024 is not a feature announcement. It is a directional signal. Chinese SoC vendors are beginning to build upward toward the application layer — not through neutral platform architecture, but through vertical integration with China's dominant AI ecosystem. If this model extends to Jieli Technology — which has shipped over 5.736 billion audio chips cumulatively from 2023 to 2025 and dominates the global sub-\$30 segment — the Chinese market, representing approximately 55 to 60 percent of global TWS production, could be captured by vertically integrated Chinese AI ecosystems before a neutral coordination layer achieves default embedding.

The question is not whether a coordination layer will emerge in the Chinese market. It is whether it will be neutral when it does.

The Western market faces equivalent time pressure from a different direction. Airoha grew revenue 41 percent in 2024, and its leadership has publicly stated the ambition to become the world's top TWS Bluetooth chip supplier (Taipei Times, December 2024). If Airoha achieves the number one TWS chip position by volume, it would control a majority of premium and mid-tier Android earphone production globally.

Chipset generation cycles run approximately 18 to 24 months. Each cycle that passes without a coordination layer achieving default embedding in the reference design is a cycle in which the application layer could be claimed by a chipset vendor, a Chinese AI platform, or a major technology company with the structural incentives to build it from their current position.

The opportunity is real. The timeline is not indefinite.

1.8 How Value Will Concentrate: Platform Economics

Platform value in this model is not a linear function of device count. It is the product of four variables that compound with each other: Installed Base, Attach Rate, Conversion Rate, and Device Lifetime Value. Each variable amplifies the others.

Installed Base is the output of the SoC partnership strategy. It scales through chipset reference design embedding, not through direct brand sales — the Dolby distribution model, embedded in the manufacturing default rather than sold to each downstream customer. The measured wireless audio market provides the scale reference: 500 million wireless audio devices ship globally per year. The SoC partnership model is what makes this non-linear: a single reference design inclusion reaches every ODM building on that chipset and every brand that comes to that ODM, without requiring a separate sales engagement for each.

Attach Rate is the share of shipped devices where a user activates the platform's AI features. A device that ships with the coordination layer firmware but is never activated contributes to distribution but not to platform economics. Attach Rate is driven by the quality of the first interaction and — critically — by a premium partner brand's financial incentive to surface the platform in its companion app onboarding. Because premium partners share directly in subscription profit, they have a structural reason to optimise their onboarding flow for platform activation. This self-reinforcing incentive is what distinguishes a profit-share structure from simple licensing: it turns a premium partner's product team into a platform conversion optimisation team.

Conversion Rate — the share of the installed base that becomes monetising users — converts distribution into revenue. Subscription conversion concentrates among users who have already

adopted voice as their primary AI interaction modality — disproportionately owners of premium earphones in markets where paid AI subscription has normalised. At 5 percent Conversion Rate and \$5 per month across 500 million global wireless devices, the gross subscription ARR is \$1.5 billion. At 10 percent, it is \$3.0 billion. These are not projections. They are arithmetic applied to a measured market.

Device Lifetime Value closes the model. Current managed voice AI inference costs approximately \$0.05 to \$0.15 per minute. At a \$5 monthly subscription and moderate usage, monthly inference costs are \$2 to \$4, leaving \$1 to \$3 gross margin per subscriber per month. Inference costs are declining at approximately tenfold per year at the frontier model level, meaning that by 2027 a \$5 monthly subscription at moderate usage will carry software-as-a-service grade gross margins. LTV improves as users age on the platform: longer-tenured users generate richer personalisation, higher service integration revenue, and stronger retention.

When these four variables compound, the system exhibits the defining characteristic of a platform. Roku grew from negligible software revenue to \$4.1 billion annually — 85 percent from its platform — by licensing its operating system to television manufacturers at near-zero cost and monetising through advertising and subscription revenue shares once the installed base existed. Below approximately 20 million active devices, Roku was a hardware company with modest software revenue. Above that threshold, the platform economics became self-reinforcing: content providers needed Roku because users were there; users chose Roku devices because content was there. The coordination layer for earphones faces an equivalent threshold — the point at which the installed base is large enough that service providers cannot afford to be absent and brands cannot afford to be excluded.

Proximity as the structural asset. The earphone is worn for hours of the day — across work, commute, fitness, and rest. Smartwatches are worn longer still, but across a smaller installed base and a lower-bandwidth interaction surface; no other device combines this wear time with voice-first interaction at this scale. A layer embedded at the firmware level occupies the closest position to the user that exists in consumer technology: closer than the phone, closer than any app. That position carries a forward implication. It is the natural place for a user's context and preferences to persist — across devices, across brands, under the user's explicit control, rather than resetting with every hardware upgrade. What that persistence ultimately enables is beyond the scope of this analysis. The structural point is narrower: the embedded, neutral position is the only place in the architecture from which it can be offered at all.

PART II — Competitive Landscape

Why existing players cannot fill the coordination gap — and where each is genuinely ahead

2.1 The Structural Constraint Map

Every major technology company active in audio wearables faces a structural conflict that prevents it from serving as the neutral coordination layer the market requires. These conflicts are not strategic choices that can be reversed. They are features of each company's business model.

Mapping them is essential to understanding why the coordination gap persists despite the technical capability and commercial interest of well-resourced players.

The competitive landscape can be mapped along two axes: SoC distribution coverage — the ability to reach brands through chipset-level embedding — and revenue model neutrality — the absence of

commercial interests that would direct users toward a specific company's services, commerce, or AI pipeline. The coordination gap exists at the intersection of broad SoC coverage and genuine neutrality. No existing player occupies that intersection.

2.2 Big Tech: Capability Without Neutrality

Apple. Apple has built the most advanced earphone AI platform in existence. AirPods Pro 3 is a health-monitoring, translation-capable, AI-connected device with no peer in the consumer market. But every capability requires Apple Intelligence on an iPhone. The 78 percent of the TWS market outside Apple's ecosystem is structurally excluded, by design. Apple generates approximately \$46 per active device per year from its services layer. That economics requires the wall to remain intact.

Apple's ecosystem exclusivity is not a limitation to be fixed. It is the model.

Google. Google has made the most expansive non-proprietary move, deploying Gemini capabilities to any Android-connected headphones in over 70 languages, albeit via hybrid AI compute with connectivity to the cloud. But this integration is phone-mediated, not firmware-embedded. It provides no platform economics for ODMs, no mechanism for earphone brands to differentiate their user relationship. Google's business model is data and advertising. Its incentive is to make earphones transparent conduits to its services infrastructure — not to create an independent platform with its own economics. A January 2026 academic security disclosure identified a widespread implementation flaw in Google's Fast Pair protocol across consumer audio accessories — since addressed by Google and device manufacturers — illustrating the complexity of the security surface that deep phone-mediated integration creates.

Android demonstrates that Google can operate a neutral layer when device makers are its distribution channel for services; an audio coordination layer is different, because its value lies in routing users between services — a market in which Google itself competes. This constraint defines the system: a platform whose revenue model depends on advertising cannot serve as a neutral coordination layer for brands whose users it would direct toward its own advertising inventory.

Amazon. Amazon's Alexa earphone strategy has been commercially limited because brands resist a platform that converts their customers into Amazon shoppers. Alexa's incentive is to route users toward Amazon Prime, shopping, and the Alexa commerce ecosystem. Amazon's Echo Buds repositioning to the budget \$49.99 tier confirms that earphones as a primary Alexa surface are not a strategic priority.

OpenAI. OpenAI's reported AI earphone ambitions represent a product play competing with AirPods for market share, not a coordination layer serving the existing production base. A neutral platform operator does not compete with its own brand customers for those same end users. OpenAI's revenue model comes from ChatGPT subscriptions — a conflict of interest that is structural, not correctable.

2.3 SoC Vendors: Distribution Without a Service Layer

SoC vendors have the most direct path to earphone distribution of any non-brand player. They supply the silicon that goes into every device. Embedding software at the reference design level is technically straightforward.

But building the application layer above the chip creates an insolvable channel conflict. The OEM brands that are a leading SoC vendor's primary silicon customers would object to a chipset vendor competing with them on software, AI experiences, and subscription revenue.

Qualcomm's Snapdragon Sound programme provides a certification standard and audio quality brand that creates some of the conditions for a coordination layer — but it stops at the hardware capability boundary. It does not route service integrations, manage user identity, accumulate memory, or create subscription economics. The same constraint applies to Airoha, Bestechnic, Bluetrum, and Jieli. Their business is silicon. Building the software layer above it conflicts with the interests of every OEM they supply.

SoC vendors are the optimal distribution partners for a coordination layer built by someone else. A neutral layer is therefore most naturally realised in partnership with them, rather than in competition with them.

2.4 Chinese AI Platforms: Vertical Integration and Its Limits

ByteDance's Doubao LLM integration into Bluetrum's BT895x chip represents the most significant competitive move in the audio wearables market in 2024. It demonstrates that Chinese AI platforms are building upward from the silicon layer into the application layer — not through neutral platform architecture, but through vertical integration with China's dominant AI ecosystem.

The limitation of this approach is not technical. ByteDance can build a capable AI audio platform. The limitation is neutrality. It is also why a coordination layer is more plausibly a partner than a rival: for ByteDance, reach into a fragmented installed base and per-device revenue through a neutral layer is a stronger motivator than building proprietary embedding brand by brand.

ByteDance monetises through content engagement, advertising, and ecosystem lock-in. The Doubao/Ola Friend integration routes users into TikTok and ByteDance's content ecosystem. Western brands wanting ByteDance AI features face developer access barriers. Chinese brands wanting OpenAI features face mirror barriers. The platform that emerges from Chinese AI ecosystem vertical integration will be deeply capable — and non-neutral by design.

Brands will not voluntarily route their customers into a competitor's content ecosystem without control and compensation. The neutrality requirement cannot be satisfied by any player whose primary revenue model requires directing users toward its own content or commerce.

The risk to the market is not that ByteDance builds a bad platform. The risk is that it builds a dominant one for the Chinese market — which represents 55 to 60 percent of global TWS production — before a neutral layer achieves default embedding. A market where 55 to 60 percent of production is embedded in a non-neutral AI ecosystem and the remaining 40 to 45 percent operates independently is a fragmented market with different fragmentation than today, but no better coordination.

2.5 Long-Tail Software Platforms: App-First Architecture and Its Ceiling

A distinct group of competitors addresses the long-tail brand market directly: NebulaBuds, Original Sound, Tuya, Dana AI, and DUI (AISpeech). Each has meaningful commercial presence, particularly in China. Their limitation is not commercial scale. It is architectural position.

Every one of them treats the earphone as a rendering pipe for intelligence that lives on the phone screen.

The app-first vs audio-first distinction. Every existing long-tail software platform follows an app-first architecture: the intelligence lives on the phone, the trigger requires screen interaction, and the headphone functions as an audio output peripheral for an application running elsewhere. The user must open an app, tap a button, or interact with a phone screen before the AI responds. An audio-first

architecture inverts this entirely: the trigger is the headphone itself — a programmable button, a custom wake word, a contextual gesture — and the AI responds in-ear without requiring the phone. The device is the product, not a peripheral. This distinction is commercially significant. It determines whether the brand owns the interaction moment with its users. And device-level triggers are the prerequisite for the ambient AI use cases — hands-free productivity, real-time translation, meeting intelligence, fitness coaching — that drive subscription conversion. App-first platforms can add AI features indefinitely without closing this gap, because the gap is architectural rather than algorithmic.

NebulaBuds is the most visible long-tail competitor, with disclosed brand partnerships including Realme, Huawei, Lenovo, Baseus, and MINISO. It offers 200 or more AI capabilities across multiple LLMs and leads the market in translation depth: offline dual-ear translation, WeChat video integration, and real-time multilingual modes that are genuinely ahead of the field. But the interaction model requires the phone screen. The device trigger is limited to a single proprietary wake word. NebulaBuds charges approximately \$1.40 per device for activation, with subscription packages sold by NebulaBuds rather than shared with the brand. The brand is a cost centre rather than a profit participant.

Original Sound charges \$4,000 to \$15,000 per integration project plus \$0.87 per device for ongoing feature access. It has extensive SoC coverage — Bluetrum, Jieli, BEStech, Airoha, Realtek — making it accessible to ODMs across a wide price range. It does not offer a device-level trigger or a brand subscription revenue profit-share model.

Tuya is an IoT platform, not an audio AI platform. Its integration model is firmware-level, with no voice interaction layer and no subscription model. **Dana AI** charges \$0.87 per device on a lifetime licence model. **DUI (AISpeech)** operates on SDK and API integration fees with no disclosed per-device pricing.

In every case, the brand is a cost centre rather than a profit participant. None of these platforms offers a device-level trigger or a brand subscription revenue or profit-share model. Their competitive advantage is speed, price, and China market presence — not platform depth.

2.6 Where Competitors Demonstrably Lead

Analytical rigour requires acknowledging the dimensions on which the competitive field currently holds genuine advantages. These are real, not trivial, and they define the areas where a coordination layer must close the gap to compete effectively.

Translation depth. NebulaBuds leads in offline translation depth. Its dual-ear real-time translation and WeChat video integration are mature, commercially deployed features without a direct equivalent at platform scale.

China local engineering presence. NebulaBuds and Original Sound both maintain 20 to 30 field application engineers physically located in China, enabling faster ODM support and shorter issue resolution cycles than remote teams can achieve.

Low-end SoC breadth. Original Sound supports a broader range of low-cost Chinese chipsets, including Realtek variants, making it accessible to ODMs operating at the very lowest price tiers.

Proprietary hardware performance. JBL's Personi-Fi 3.0, within the Samsung-Harman group, delivers per-ear hearing profiling that goes beyond standard EQ personalisation. Sony's QN3 processor with twelve-microphone arrays in the WH-1000XM6 and WF-1000XM6 sets the hardware benchmark for noise cancellation that no neutral platform has replicated at comparable specification.

These advantages are real. They are also feature-specific and non-transferable. Sony's ANC hardware does not license to third parties. NebulaBuds' translation depth does not address the brand revenue model or device-level trigger problem.

The structural advantages of a neutral coordination layer — cross-chipset distribution, premium-partner profit participation, firmware-embedded interaction — are architectural rather than feature-based. A competitor who closes the translation gap still does not have the platform. A competitor who hires FAE engineers in Shenzhen still charges brands rather than paying them.

PART III — Bragi: Strategy Analysis

How Bragi is positioned to fill the coordination gap

3.1 What Bragi is Doing: Descriptive

Platform architecture. Bragi operates a three-layer infrastructure stack designed to sit between earphone firmware and service providers. The firmware layer is an embedded service module — present by default in chipset reference designs, activatable by a brand's product decision, transparent to the consumer as a branded experience. The brand layer provides a white-label companion application that brands can configure with their own identity, through which platform services are surfaced to end users. The cloud layer manages user accounts, subscriptions, service routing, and user context across all devices and brands simultaneously.

This architecture is structurally analogous to Android in mobile: an open, freely embeddable firmware layer that creates distribution at scale, with a proprietary service and monetisation layer above it that captures economic value once distribution is established. The audio coordination layer follows the same mechanic with one important difference: it does not require owning the operating system. The earphone firmware remains the brand's or ODM's. The platform's proprietary value is the identity infrastructure, service routing layer, subscription management system, and Audio App marketplace that brands access through that module.

SoC integrations. As of March 2026, Bragi has completed AI integration across multiple major SoC vendors including Airoha, Qualcomm, Jieli and Bluetrum, with more to come. No other neutral coordination layer company holds simultaneous completed integrations across a comparable combination of chipset families and market tiers.

Brand and ODM network. Bragi's brand relationship base spans Tier 1, Tier 2, and Tier 3 brands, alongside an established ODM, IDH, and service partner ecosystem built through years of project work. Each relationship represents technical, commercial, and operational integration depth. This institutional knowledge — which hardware architectures specific ODMs prefer, which brands have which product roadmap constraints, which chipset combinations are used across which price tiers — cannot be replicated by a product launch or accelerated through capital alone.

Incentive architecture. The profit-share architecture is the most structurally distinctive element of Bragi's model: it is designed so that its partners across the supply chain — SoC vendor, ODM, premium brand, and service provider — can share in the platform's success rather than only paying for access.

No other platform in this market creates this structure of aligned incentives. Every competing platform charges brands. NebulaBuds, Original Sound, and Tuya — examples, not an exhaustive list — all charge brands. Bragi shares profit with its premium partners. Apple, Sony, and JBL do not license at all.

Precedent. The precedent that matters most is Dolby Laboratories. Dolby generated \$1.35 billion in fiscal year 2025 revenue at approximately 92 percent gross margins. Its implementation licences to chipmakers embed Dolby processing in silicon reference designs. Its system licences sell to manufacturers who ship finished products. The Dolby brand mark on the product box became a consumer-recognised quality signal that manufacturers were reluctant to abandon once they had earned it. Dolby's certification process created both a quality standard and a switching cost simultaneously. The coordination layer model follows this mechanic — but with the addition of a direct consumer subscription layer and a service marketplace that Dolby, as a pure audio processing licensor, never required.

3.2 Strategic Alignment: Where the Strategy Fits the Market Structure

Alignment with the coordination layer requirement. Bragi's architecture directly addresses the three structural observations that make a coordination layer inevitable. On the fragmentation observation: the SoC-level default embedding strategy is the only mechanism that reaches the 1,800 brands with no engineering team. On the bilateral integration impossibility: the single SDK integration model that reaches all brands through chipset partnerships directly solves the $M \times N$ problem. On the neutrality requirement: a profit-sharing structure that flows value to premium brands and SoC vendors is the mechanism that makes Bragi's commercial interests aligned with, rather than opposed to, the interests of every participant in the supply chain.

Revenue model differentiation. The premium-partner profit-share model has no equivalent in the market. Every competing platform charges brands. Bragi shares profit with its premium partners. This is not a pricing decision — it is a commercial architecture. It makes brand adoption a net positive rather than a cost centre, which fundamentally changes the buying decision. It also creates a self-reinforcing flywheel: a premium partner that shares in subscription profit has a financial incentive to activate users on the platform, which increases subscription conversion, which increases that partner's profit, which reinforces adoption.

Data sovereignty by architecture. The privacy architecture for this system is designed to align with GDPR principles. User data persists under the user's explicit control: on-device first, end-to-end encrypted where cloud processing is involved, consent-gated per data category, with full export and deletion rights. The platform holds structured preference models, not conversation recordings. It is a combination that purely cloud-based AI assistants cannot achieve without device-level firmware presence.

3.3 Execution Risks

Failure mode 1: A non-neutral player reaches default embedding first. The window for a neutral layer is defined by chipset generation cycles. Vertical integration between major AI ecosystems and SoC vendors is already observable in the Chinese market, and well-resourced Western players retain the option of deepening surface-level integrations into the firmware. If a non-neutral player reaches reference design adoption at scale before a neutral layer does, the addressable market for neutral coordination shrinks materially — and for the affected segment, the neutrality question becomes moot.

A fast-following long-tail platform closing its architectural gaps represents the equivalent risk from below.

Failure mode 2: The subscription model does not convert at the required scale. If consumer-facing applications and conversion design do not deliver, the Conversion Rate stays at the low end of the modelled range and the platform's economics revert to per-device licensing rather than compounding subscription revenue.

Conclusion: The Standard for Evaluation

The audio market is not lacking in technology. Devices are capable, services are abundant, and consumer willingness to pay for embedded AI on earphones is measurable and high. What has been missing is a layer that coordinates these elements across a fragmented ecosystem — one that is neutral enough for brands to trust with their users, financially structured to create aligned incentives at every layer of the supply chain, and structurally durable because the value of its embedded position compounds over time.

The brands with volume need a neutral platform to compete against Apple. The brands without volume have no path to AI capability at all without chipset-level default embedding. Both conditions point to the same structural requirement.

As established in the opening sections, SoC companies are collectively spending hundreds of millions of dollars embedding TOPS-level AI compute into earphone chips. None of those chips are connected to a software platform that makes that interaction economically significant, personally persistent, or commercially scalable across brands. That coordination gap is closing — not because a neutral layer has been built, but because chipset vendors and Chinese AI platforms are beginning to build application-layer solutions from their own positions. Those solutions are not neutral. They serve their own commercial interests. They will close the window for a neutral layer if they reach reference design adoption before the neutral layer does.

The standard against which the coordination layer opportunity — and any company attempting to fill it — should be evaluated is not the size of the market or the quality of the structural argument. Both are compelling. The standard is whether the execution is sequenced correctly and moves quickly enough for the structural position to convert into demonstrated platform economics before the window closes. That standard is falsifiable: default SoC embedding achieved, Audio Apps at market quality, subscription conversion demonstrated empirically, and brand network activated through a functioning channel program — all within a timeline defined by the semiconductor product cycles this document has mapped.

The central thesis is simple and structural: a neutral audio coordination layer is inevitable because no economically viable alternative architecture exists. The market will produce it, or it will remain fragmented in ways that prevent the enormous economic potential of AI-capable earphones from being captured by any participant. The question the market is answering in 2026 is not whether this layer exists. It is who builds it, whether it is neutral, and whether the window defined by chipset generation cycles remains open long enough for that layer to achieve the default embedding that makes its position self-reinforcing.

Sources and Data Notes

Primary data. Amazon US retail transaction data Sep 2024–Sep 2025: proprietary transaction dataset used as the primary empirical basis for the fragmentation analysis and market extrapolation. Consumer AI subscription data: proprietary two-wave market research across the United States, Germany, and India (2025). ODM assembly margin: Triangulated from gross margins reported in Goertek and Luxshare annual reports

Market data. Wireless headphone shipments, brand share, and SoC share are sourced from Futuresource Consulting's Headphones Quarterly Tracker, Q4 2025. Market data and brand shares for smart glasses are sourced from Futuresource Consulting's Global Wearables Market Outlook, March 2026. China AI earphone sales: 36Kr; Pandaily; Gizmochina (2024–2025). AI voice quality benchmarks: GfK Consumer Electronics Study 2023; OpenAI latency benchmarks 2024. "Fast Pair security disclosure: S. Duttgupta, S. Wynn, N. Antonijević, D. Singelée and B. Preneel, 'One Tap to Hijack Them All: A Security Analysis of the Google Fast Pair Protocol,' KU Leuven (COSIC/DistriNet), IEEE Symposium on Security and Privacy 2026; CVE-2025-36911; disclosed January 2026."

Consumer survey data. Futuresource Consulting's Audio Tech Lifestyle study conducted across USA, China, UK, Germany and Japan among 10k respondents

Company financials. Qualcomm: SEC FY2024 10-K; FY2025 public reporting. MediaTek / Airoha: Yahoo Finance (TPE:2454, TPE:6526); Taipei Times December 2024. Bestechnic (688608): Shanghai Star Market disclosures; Futunn Financial; Yicai Global (2025). Jieli: PitchBook (private); World Electronics Achievement Awards 2024. Bluetrum (688332): PitchBook; company disclosures. Apple: press release September 2024; FDA OTC hearing aid clearance; FY2025 earnings. ODM margins: Luxshare Precision H1 2024 interim report; GoerTek financial filings. Dolby: FY2025 earnings release. Roku: Q4 2024 shareholder letter; SEC filings.

Methodology. Platform valuation multiples: public technology company data. AI inference cost trends: Stanford HAI AI Index 2024. Currency conversions at approximate 2024 annual average rates: NT\$32/USD, CNY 7.1/USD, EUR 1.085/USD. Reasoned inferences are identified where they appear in the text. Claims confirmed by disclosed information are stated as confirmed; claims that are targets or work in progress are stated as such.