

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

KRISP TECHNOLOGIES, INC.,
Petitioner,

v.

SANAS.AI INC.,
Patent Owner.

Case No. IPR2026-00273
Patent No. 12,125,496

PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	THE '496 PATENT	5
III.	CLAIM CONSTRUCTION.....	11
IV.	PETITIONER HAS NOT ESTABLISHED A PRIMA FACIE CASE OF OBVIOUSNESS BECAUSE IT IGNORES THE CLAIMED PIPELINE AND THE EXAMINER-IDENTIFIED INVENTIVE LIMITATION ..	13
	A. The Petition Does Not Demonstrate a Reasoned Motivation to Combine Prior Art References	14
	B. The Petition Does Not Demonstrate a Reasonable Expectation of Success	17
	C. The Petition Does Not Make a Particularized Showing Of Where Each Claim Element Is Disclosed In The Asserted Prior Art	20
	D. Institution Should Be Denied as an Inefficient Use of Office Resources.....	21
V.	PETITIONER’S CHALLENGES FAIL ON THE MERITS.....	24
	A. Ground 1 Fails on the Merits.....	25
	1. The Ground 1 Combination Is Architecturally Incoherent.....	30
	2. The Petition’s Motivation Theory Fails.....	35
	B. Even If the Board Does Not Treat Ground 1 as Dispositive, Grounds 2-3 Independently Fail Because the Additional References Do Not Cure the Base-Combination Defect	37
	C. Alternative Grounds 4-6 Likewise Fail on the Merits	39
VI.	CONCLUSION	43

TABLE OF AUTHORITIES

	Page(s)
CASES	
<i>ActiveVideo Networks, Inc. v. Verizon Commc'ns, Inc.</i> , 694 F.3d 1312 (Fed. Cir. 2012)	14
<i>Apple Inc. v. Ginko LLC</i> IPR2025-01388, Paper 14 (March 2, 2026).....	20
<i>Chevron Oronite Co. v. Infineum USA</i> , IPR2018-00923, Paper 9 (November 7, 2018)	21
<i>CyberSecure IPS, LLC v. Network Integrity Systems, Inc.</i> , IPR2025-01441, Paper 18 (May 26, 2026).....	21
<i>Honeywell Int'l Inc. v. 3G Licensing, S.A.</i> , 124 F.4th 1345 (Fed. Cir. Jan. 2, 2025).....	17
<i>Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.</i> , 821 F.3d 1359 (Fed. Cir. 2016)	16
<i>In re Kahn</i> , 441 F.3d 977 (Fed. Cir. 2006)	13, 26, 33
<i>Johns Manville Corp. v. Knauf Insulation, Inc.</i> , IPR2018-00827, Paper 9 (PTAB Oct. 16, 2018)	16
<i>KSR Int'l Co. v. Teleflex Inc.</i> , 550 U.S. 398 (2007).....	14, 15, 25
<i>Metalcraft of Mayville, Inc. v. Toro Co.</i> , 848 F.3d 1358 (Fed. Cir. 2017)	34
<i>PAR Pharm., Inc. v. TWI Pharms., Inc.</i> , 773 F.3d 1186 (Fed. Cir. 2014)	16
<i>Xerox Corp. v. Bytemark, Inc.</i> , IPR2022-00624, Paper 9 (PTAB Aug. 24, 2022).....	15

STATUTES

35 U.S.C. § 103	13
35 U.S.C. § 312(a)(3).....	4, 13, 20
35 U.S.C. § 314(a)	13
37 C.F.R. § 42.104(b)(4).....	20
37 C.F.R. § 42.107	1

EXHIBIT LIST

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2001	Complaint filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Jul. 7, 2025) (D.I. 1) (“Complaint”)
Ex2002	Answer filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Sep. 8, 2025) (D.I. 28) (“Answer and Counterclaims”)
Ex2003	Amended Complaint filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Sep. 22, 2025) (D.I. 35) (“Amended Complaint”)
Ex2004	Order on First Motion to Dismiss filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Sep. 29, 2025) (D.I. 37) (“Order on First Motion to Dismiss”)
Ex2005	Order Denying Krisp Motion to Dismiss First Amended Complaint and Granting in Part and Denying in Part Motion to Dismiss Defendant Krisp’s Counterclaims filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Dec. 1, 2025) (D.I. 58) (“Order Denying Krisp Motion to Dismiss First Amended Complaint”)
Ex2006	Answer to Amended Complaint filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Dec. 15, 2025) (D.I. 60) (“Answer to Amended Complaint”)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2007	Order Denying Motion to Dismiss filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Dec. 19, 2025) (D.I. 63) (“Order Denying Motion to Dismiss”)
Ex2008	Order Denying [Renewed] Motion for Judgment on Pleadings filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Feb. 23, 2026) (D.I. 72) (“Order Denying [Renewed] Motion for Judgment on Pleadings”)
Ex2009	Joint Case Management Statement filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Mar. 19, 2026) (D.I. 83) (“Joint Case Management Statement”)
Ex2010	Case Management Scheduling Order filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Mar. 26, 2026) (D.I. 92) (“Case Management Scheduling Order”)
Ex2011	Docket Sheet in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal.) (“Docket Sheet”)
Ex2012	Coke Morgan Stewart, <i>Memorandum re Interim Processes for PTAB Workload Management</i> (Mar. 26, 2025), https://www.uspto.gov/sites/default/files/documents/InterimProcesses-PTABWorkloadMgmt-20250326.pdf . (“2025 Memo.”)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2013	John A. Squires, <i>Memorandum re Additional Discretionary Institution Considerations—U.S. Manufacturing and Small Business Use of AIA Proceedings</i> (Mar. 11, 2026), https://www.uspto.gov/sites/default/files/documents/Additional_Discretionary_Institution_Considerations_US_Manufacturing_and_Small_Business_Use_of_AIA_Proceedings.pdf . (“2026 Memo.”)
Ex2014	<i>Magnolia Med. Techs., Inc. v. Kurin, Inc.</i> , No. IPR2026-00097, Paper 17 (PTAB May 14, 2026) (precedential) (“ <i>Magnolia</i> ”)
Ex2015	Scott R. Boalick, <i>Guidance on USPTO's Rescission of "Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation"</i> (Mar. 24, 2025)
Ex2016	[Intentionally Left Blank]
Ex2017	Sanas, LinkedIn, https://www.linkedin.com/company/sanastech/about/ (last visited May 13, 2026)
Ex2018	U.S. Patent No. 12,205,609 B1 (issued Jan. 21, 2025)
Ex2019	U.S. Patent No. 12,223,979 B1 (issued Feb. 11, 2025)
Ex2020	U.S. Publication No. 2022/0358903 A1 (published Nov. 10, 2022)
Ex2021	U.S. Publication No. 2024/0347070 A1 (published Oct. 17, 2024)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2022	U.S. Publication No. 2024/0363135 A1 (published Oct. 31, 2024)
Ex2023	[Intentionally Left Blank]
Ex2024	<i>About Sanas</i> , Sanas, https://www.sanas.ai/about-us (last visited May 16, 2026).
Ex2025	Krisp, LinkedIn, https://www.linkedin.com/company/krisphq (last visited May 13, 2026)
Ex2026	Krisp Team, <i>What’s It Like to Work at Krisp?</i> (Sep. 20, 2021), https://krisp.ai/blog/working-at-krisp/ (last visited May 13, 2026) (“What’s it like to work at Krisp”)
Ex2027	Arto Minasyan, <i>Krisp Voice AI Lab: Revolutionizing Voice Technology</i> , LinkedIn (posted Mar. 2026), https://www.linkedin.com/posts/artOmniasyan_eight-years-ago-we-started-with-a-simple-activity-7430219416912629760-8l6t (last visited May 13, 2026).
Ex2028	Krisp, <i>Introducing Accent Conversion for the Listener</i> , https://krisp.ai/blog/introducing-accent-conversion-for-the-listener/ (last visited May 20, 2026)
Ex2029	Arka Telecom, <i>YSU–Krisp AI Lab Opens at Yerevan State University</i> , https://arkatelecom.am/en/news/development/ysu_krisp_ai_lab_opens_at_yerevan_state_university/ (last visited May 20, 2026).

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2030	Careers at Krisp, Krisp, https://krisp.ai/careers/ (last visited May 14, 2026)
Ex2031	Krisp, Google Maps, https://www.google.com/maps (last visited May 13, 2026)
Ex2032*	<i>U.S. Patent & Trademark Office, Enforcement and Non Waiver of 37 C.F.R. § 42.104(b)(4) and Enforcement and Non Waiver of General Knowledge in Inter Partes Reviews (July 31, 2025), https://www.uspto.gov/sites/default/files/documents/aapa_memo_final_signed.pdf. (the “Enforcement, Non Waiver, and Enforcement and Non Waiver Memo”)</i>
Ex2033*	<i>Defendant’s Proposed Terms for Construction, <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i>, No. 3:25-cv-05666-RS (KAW) (N.D. Cal. June 9, 2026)</i>
Ex2034*	<i>Defendant’s Invalidity Contentions, <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i>, No. 3:25-cv-05666-RS (KAW) (N.D. Cal. May 26, 2026)</i>

*** New Exhibits submitted herewith.**

I. INTRODUCTION

Pursuant to 37 C.F.R. § 42.107, Patent Owner Sanas.AI Inc. submits this Preliminary Response to the Petition for inter partes review of Claims 1–20 of U.S. Patent No. 12,125,496 (“the ’496 Patent”).

Institution should be denied because Petitioner fails to show any system **that applies a second neural network to low-dimensional representations of input speech frames to generate target speech frames, as claimed.**

At its core, the Petition is based on an impermissible hindsight reconstruction that ignores the specific limitation the Examiner identified as supporting allowance. Ex1002 at 130. In particular, during prosecution, the Examiner found that the prior art did not teach or suggest “apply a second neural network to the low-dimensional representations of the input speech frames to generate target speech frames” (color coding supplied, corresponding to references cited in Petition as shown below). This claim element is “found in Claim 1” (as element **[1(c)]**; *see also* **[11(c)]** and **[16(d)]**) and the “[s]imilar features [which] are claimed in Claims 2-20.” *Id.*

The Petition never identifies any reference—or any coherent combination—that discloses or renders obvious this allowance-dependent limitation in the context of the claimed pipeline. Instead, as the color coding in the quoted language

illustrates, the Petition divides the limitation across three disparate references and assigns each reference only a fragment of the claim:

- Dantrey (shown in black) is used as the alleged basis for “apply[ing] a second neural network,” but Dantrey’s audio generator receives noise-reduced spectrograms within Dantrey’s own pipeline; *it is not shown to receive low-dimensional representations of input speech frames or to generate the claimed target speech frames.*
- The Petition then relies on **Strake (shown in blue bold text)** for frame-level output, but **Strake**’s second stage generates only “a vector representation of the second-stage clean signal estimate 213 for the current frame,” which is part of a *different* restoration/reconstruction architecture and is *not produced* by Dantrey’s neural network. Ex1005 at 11.
- Finally, the Petition turns to Chen (shown in green underlined text) for low-dimensional representations, but Chen’s generic PCA disclosure belongs to a *different* noise-classification and attenuation system and *does not bridge* Dantrey’s spectrogram-based generator with **Strake**’s frame-based reconstruction path.
- Additionally, for alternative Ground 4, the Petition swaps Chen for **Liu (green bold)** as an alternative to low-dimensional representations, but **Liu**’s “low dimensional mel spectrogram” is generated within its own two-stage VoiceFixer architecture—an analysis stage that produces intermediate features and a synthesis stage that uses a neural vocoder to generate waveforms—and is not shown to cure the prior defects or to be interchangeable with Dantrey’s noise-reduced spectrogram or to be usable by Dantrey’s neural network to generate the claimed target speech frames.

Thus, the Petition offers a hindsight-driven patchwork of disparate and unconnected parts, not a particularized showing that any combined system applies a second neural network to low-dimensional representations of input speech frames to generate target speech frames.

Each reference addresses different data, objectives, and architectures, and the Petition does not explain how they would be combined to perform applying a second neural network to low-dimensional representations of input speech frames to generate target speech frames.

The combination is like taking the compressor from a refrigerator, the transmission from a car, and the display from a thermostat, then calling the result an air conditioner because each component relates in some way to controlling environmental conditions or mechanical systems. The parts may share a broad field, but they are **not arranged, designed, or motivated to operate together for the claimed function.**

Because Petitioner does not identify this limitation in either a single reference or in a coherent combination—it does not cure the deficiency identified by the Patent Office.

This failure is dispositive: the Petition **does not** satisfy its burden under 35 U.S.C. § 312(a)(3) to show **with particularity where each claim element is found in the prior art**, instead reconstructing the claims from disconnected fragments.

Grounds 2-3 fare no better and do **not** cure the defects in Ground 1. Each relies on the same deficient Dantrey–**Strake**–**Chen** combination for the independent claims and merely adds **Quillen** and **Heckmann** for dependent-claim details. Those

additional references do **not** supply the missing application of a second neural network to low-dimensional representations of input speech, as required by elements [1(c)], [11(c)], and [16(d)]. The alternative Grounds 4-6 likewise fail because they simply swap Chen for Liu, which does **not** cure the defect and instead fails for the same reasons as Grounds 1-3. In addition to its own flaws, Liu, like Chen, does **not** bridge the mismatch between Dantrey's spectrogram-based generator and Strake's frame-based reconstruction path, nor does it provide the claimed data flow.

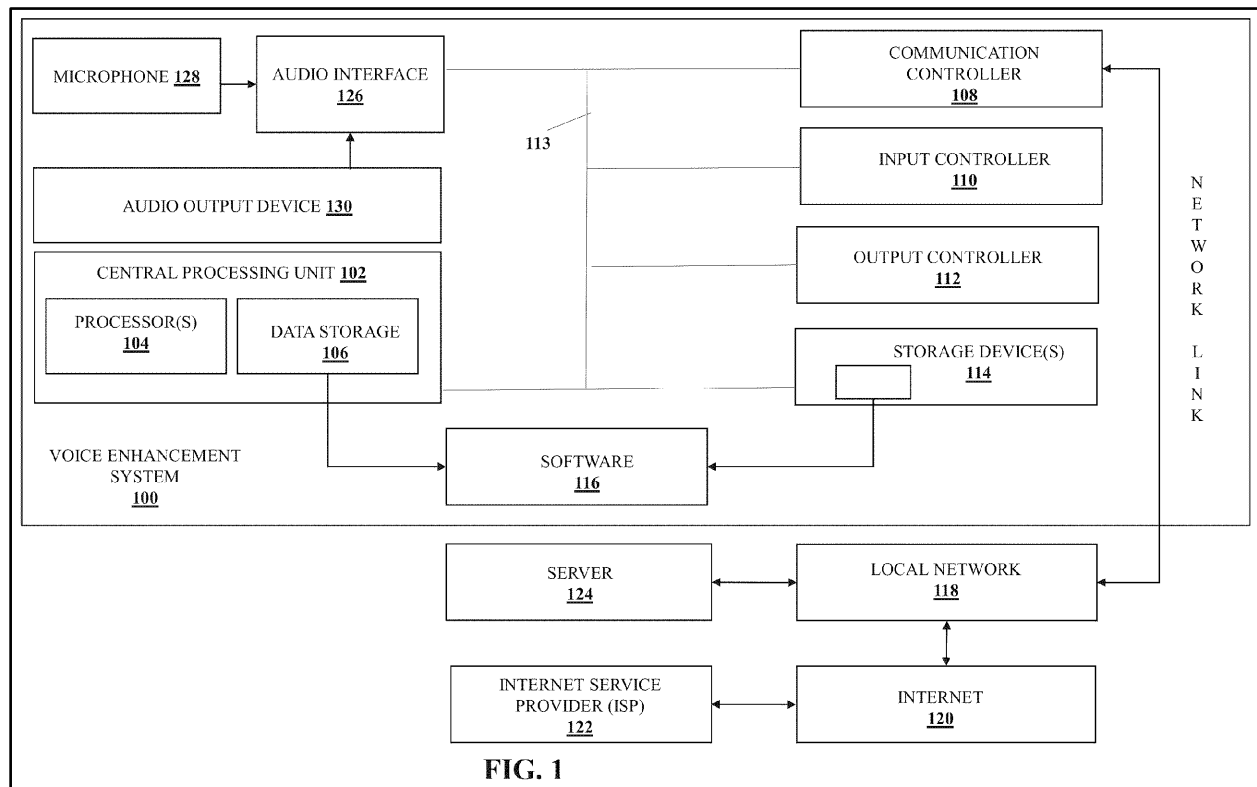
II. THE '496 PATENT

The '496 Patent is entitled “*Methods for Neural Network-Based Voice Enhancement and Systems Thereof*.” Ex1001 Title. The patent relates to “audio analysis and, more particularly, to methods and systems for voice enhancement using neural networks.” Ex1001 1:11-13. The Abstract provides an overview of this invention:

The disclosed technology relates to methods, voice enhancement systems, and non-transitory computer readable media for real-time voice enhancement. In some examples, input audio data including foreground speech content, non-content elements, and speech characteristics is fragmented into input speech frames. The **input speech frames are converted to low-dimensional representations of the input speech frames**. One or more of the fragmentation or the conversion is based on an application of a first trained neural network to the input audio data. The low-dimensional representations of the input speech frames omit one or more of the non-content elements. **A second trained neural network is applied to the low-dimensional representations of the input speech frames to generate target speech frames**. The target speech frames are combined to generate output audio data. The output audio data further includes one or more portions of the foreground speech content and one or more of the speech characteristics.

Ex1001 Abstract.

Fig. 1 shows this same architecture in greater detail:



As pertinent here, the voice enhancement system 100 includes, among other components, a voice enhancement module 206 having a first neural network 208 and a second neural network 210. Ex1001 Fig. 2, 4:18-25. The specification explains that “the voice enhancement module 206 performs voice enhancement and/or noise suppression to convert the input audio data into the output audio data using the first and second neural networks 208 and 210, respectively.” Ex1001 at 4:43-49.

Further, “[t]he first neural network 208 receives input audio data, fragments the input audio data into frames, and converts the frames to low-dimensional representations, also referred to as a reduced-dimension representation, having lower dimensionality than that of the input audio data.” Ex1001 at 4:41-46.

Figure 4 likewise shows input audio data 402 being provided to first neural network 208, which dynamically converts the input audio data 402 into a low-dimensional input audio data representation 404. Ex1001 Fig. 4, 5:49-56.

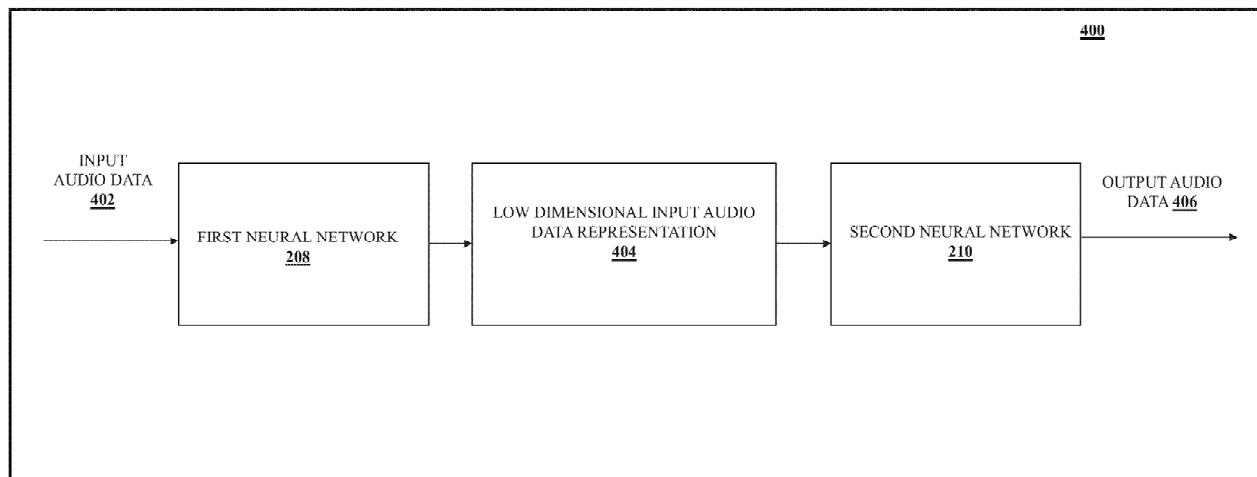


FIG. 4

The specification further explains that the input audio data may include foreground speech content, non-content elements, and speech characteristics. Ex1001 6:1-7. The speech characteristics may include, for example, pitch, intonation, melody, stress, articulation, enunciation, voice identity, and/or unintelligible speech, while the non-content elements may include background

noise, microphone pops, low-fidelity audio, and audio clipping. Ex1001 6:15-29. The first neural network 208 fragments the input audio data into input speech frames and dynamically converts each input speech frame into the low-dimensional input audio data representation 404. Ex1001 6:58-65.

The low-dimensional input audio data representation 404 may include foreground speech content and one or more speech characteristics, while omitting non-content elements such as background noise, microphone pops, low-fidelity audio, and audio clippings. Ex1001 6:66-7:7. The specification also describes that features such as pitch, intonation, and formants may be extracted and encoded into the low-dimensional representation using dimensionality-reduction techniques, such as principal component analysis (PCA) or linear discriminant analysis (LDA), so that the representation captures important speech characteristics while reducing computational complexity. Ex1001 at 7:8-29.

The second neural network 210 then receives the low-dimensional representations of the input speech frames and converts those representations into corresponding target speech frames. Ex1001 4:49-52. Figure 4 illustrates this same sequence: the low-dimensional input audio data representation 404 is fed to second neural network 210, which dynamically generates output audio data 406. Ex1001 Fig. 4, 5:52-55. Figure 5 further identifies the claimed operations as receiving, by the second neural network, the low-dimensional representations of the input speech

frames, converting each low-dimensional representation to a corresponding target speech frame, and generating output audio data by combining the target speech frames. Ex1001 Fig. 5, 8:1-14, 9:1-24.

The training disclosure further describes the two-network pipeline. As shown in Figure 6, the first neural network 208 is trained to convert input training speech frames into low-dimensional input audio training data representations. Ex1001 Fig. 6, 6:8-57. As shown in Figure 7, the second neural network 210 is trained using a target speech sample and the low-dimensional representation of input training speech to generate target training speech frames. Ex1001 Fig. 7, 8:1-14. The specification explains that the second neural network 210 may learn the conversion between low-dimensional representations of input audio training data speech frames and corresponding target training speech frames using a loss function that minimizes the difference between predicted and actual target speech frames. Ex1001 at 8:26-36.

Thus, the patent does not merely disclose generic audio enhancement or noise reduction. It discloses and claims a **specific architecture** in which the first neural network 208 produces low-dimensional representations of input speech frames that can omit non-content elements while preserving foreground speech content and speech characteristics, and the second neural network 210 applies to those low-dimensional representations to generate corresponding target speech frames. The

target speech frames are then combined to generate output audio data 406, which preserves speech characteristics while improving the quality, clarity, comprehensibility, and/or intelligibility of the input speech. Ex1001 at 9:1-34; see also Ex1001, Claim 1. Based on the disclosure, the Examiner found Claim 1 allowable for the following reasons as set forth in the Notice of Allowance:

The prior art of record fails to teach, disclose or suggest the claimed limitations of (in combination with all other limitations in the claim) **“apply a second neural network to the low-dimensional representations of the input speech frames to generate target speech frames”**, as found in Claim 1. Similar features are claimed in Claims 2-20.

Ex1002 at 130.

Significantly, the only basis proffered by the Petition to address this basis for allowance was the combination of Dantrey, [Strake](#), [Chen](#) in Grounds 1 and 2, or [Liu](#) in Ground 4, from which all the other grounds depend. As demonstrated here, the Petition fails to demonstrate any error in the Office’s analysis or fill the wide gap in the prior art identified by the Office in the Notice of Allowance, namely that the prior art fails to teach the claimed limitation.

III. CLAIM CONSTRUCTION

Petitioner proposes construing “low-dimensional representation” as “a compressed representation of an input speech frame which is a result of a dimensionality reduction technique.” Pet. at 5. That construction is unnecessary and improperly narrows the claims by importing limitations from exemplary embodiments into the claim language. Claim 1 **does not require** that every “low-dimensional representation” be “compressed,” **nor does it require** that the representation result from any particular “dimensionality reduction technique.” Instead, Claim 1 recites “low-dimensional representations of the input speech frames” and separately requires that those representations “omit one or more of the non-content elements.” Ex1001 Cl. 1. Thus, the claim imposes requirements on the representations and their relationship to the input speech frames, but it does not limit them to representations produced by a specific compression process or dimensionality-reduction technique.

In addition, the same claim term (“low-dimensional representation”) is currently the subject of claim construction proceedings in the parallel district court litigation between the parties. Ex2033 at 2. Notwithstanding Petitioner’s assertion that their “broad *Sotera* stipulations confirms Petitioner will not bring any such challenges which could create any overlap”, Petitioner has requested that the district court construe the very same term at the *Markman* hearing scheduled for October

2026—after the expected institution decision in this case, but almost one year before any final written decision would be due. *See* Petitioner’s Response to Patent Owner’s Discretionary Denial Brief, Paper 8 at 10; Ex2033 at 2. The district court proceedings will address the scope and meaning of the asserted claims based on a more developed record, including full briefing, discovery, and expert testimony. Proceeding with claim construction in parallel here risks inconsistent interpretations, duplicative efforts, and a decision on an incomplete record, further underscoring that institution would not promote efficient resolution of the dispute.

In any event, for purposes of institution, *it does not matter which construction the Board adopts because Petitioner’s analysis fails either way.* Accordingly, Patent Owner submits that no express construction is necessary at this preliminary stage. To the extent construction is required, the Board should reject Petitioner’s attempt to import exemplary compression and dimensionality-reduction embodiments into the claims. Patent Owner reserves the right to propose express constructions should trial be instituted.

IV. PETITIONER HAS NOT ESTABLISHED A PRIMA FACIE CASE OF OBVIOUSNESS BECAUSE IT IGNORES THE CLAIMED PIPELINE AND THE EXAMINER-IDENTIFIED INVENTIVE LIMITATION

The Board may deny institution where a petitioner fails to satisfy its threshold burdens, including demonstrating:

- (1) a reasoned motivation to combine prior art references,
- (2) a reasonable expectation of success, and
- (3) a particularized showing of where each claim element is disclosed in the asserted prior art.

35 U.S.C. §§ 103, 312(a)(3), 314(a); *see also KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007).

Each of these requirements is independently dispositive, and failure to satisfy any one of them warrants denial. 35 U.S.C. §§ 103, 312(a)(3), 314(a). Here, they fail for the same reason identified in the Introduction: the Petition never shows the claimed integrated application of a second neural network to low-dimensional representations of the input speech frames to generate target speech frames.

A. The Petition Does Not Demonstrate a Reasoned Motivation to Combine Prior Art References

First, obviousness *requires more than identifying references* that could be combined; a petitioner *must articulate a reasoned motivation to combine those references, supported by a rational underpinning*. *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007); *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006).

Conclusory assertions—such as invocations of “known techniques,” “routine reprogramming,” or “predictable results”—are insufficient where, as here, the petition does not explain how the specific disclosures would be integrated into a working combination. *See ActiveVideo Networks, Inc. v. Verizon Commc’ns, Inc.*, 694 F.3d 1312, 1327–28 (Fed. Cir. 2012) (rejecting generic expert testimony that “failed to explain how specific references would be combined, which combination(s) of elements in specific references would yield a predictable result, or how any specific combination would operate or read on the asserted claims”). The Petition repeatedly relies on precisely such conclusory language, asserting that the proposed combinations merely apply “known methods” (Pet. 17, 19, 24, 28, 31, 48, 65; Ex1003 ¶¶101, 109, 127, 137, 147, 200, 244), and that any modifications would have required only “minor programming modifications” (Pet. 29, 32; *see also* 19; Ex1003 ¶¶110, 138, 148) or yield “predictable results” (Pet. 17, 19, 24, 28, 31, 49, 59, 65; Ex1003 ¶¶101, 109, 127, 137, 147, 200, 244). Yet these statements are **unaccompanied by any explanation of how** the cited references would actually be

combined, **what** specific modifications would be made, or **how** the resulting system would operate to meet the claim requirements—precisely the type of unsupported reasoning the Federal Circuit has rejected. *See KSR*, 550 U.S. at 418; *Kahn*, 441 F.3d at 988; *ActiveVideo*, 694 F.3d at 1327–28.

The Board likewise accords little weight to expert testimony like that of Mr. Schmandt (Ex1003) that merely repeats such unsupported assertions. *Xerox Corp. v. Bytemark, Inc.*, IPR2022-00624, Paper 9 at 15–17 (PTAB Aug. 24, 2022) (precedential) (holding that the expert testimony should be “entitled to little weight” because the declaration “is conclusory and unsupported” and “adds little to the conclusory assertion for which it is offered to support”).

Petitioner’s argument that it is not offering expert testimony to fill the gap in the prior art is disingenuous. Petitioner’s Response to Patent Owner’s Discretionary Denial Brief, Paper 8 at 16. Petitioner asserts that “every limitation is directly mapped” and that Patent Owner has not identified any missing limitation supplied by expert testimony. *Id.* But that is ***precisely*** what the Petition does for the Examiner identified elements [1(b)] and [1(c)]. Pet. at 24-25, 27-29. During prosecution, the Examiner noted that no cited prior-art reference teaches the pipeline as arranged in the challenged claim. Ex1002 at 130. Petitioner relies on Mr. Schmandt’s testimony to bridge the gap—asserting that Chen’s alleged “low-dimensional representations” can be inserted into Dantrey’s algorithm and then combined to generate Strake’s

output. Ex1003 ¶¶ 125–128, 134–138. However, the references themselves **do not provide that combined architecture or explain why** a PHOSITA would have stitched those disparate teachings together. That is not a teaching in the prior art; it is expert reconstruction. Besides being nonsense, this is exactly the type of expert gap filling that the PTAB cautions against. *Johns Manville Corp. v. Knauf Insulation, Inc.*, IPR2018-00827, Paper 9 at 12–14, 18–20 (PTAB Oct. 16, 2018) (informative) (denying institution and according to the expert declaration little weight where conclusory assertions and expert reliance did not sufficiently explain why the asserted prior-art elements would have been selected and combined to achieve the claimed invention with predictable results).

Here, as discussed in Section V(A), Petitioner does not provide a reasoned explanation for **how** Dantrey, **Strake**, and **Chen** (or for alternative grounds 4-6 **Liu**) would be combined into a coherent system that performs the claimed pipeline, including **how** such a combination would yield the Examiner-identified limitation of applying a second neural network to low-dimensional representations of input speech frames to generate target speech frames.

For each of these reasons independently, the Petition therefore fails to establish a legally sufficient motivation to combine.

B. The Petition Does Not Demonstrate a Reasonable Expectation of Success

Second, a petitioner must demonstrate *a reasonable expectation of success for the proposed combination, not merely a motivation to attempt it.* *Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359, 1367–68 (Fed. Cir. 2016); *PAR Pharm., Inc. v. TWI Pharms., Inc.*, 773 F.3d 1186, 1193 (Fed. Cir. 2014). While a combination need not be “the preferred, or the most desirable, combination,” the petitioner must show “there is something in the prior art as a whole to suggest the desirability, and thus the obviousness, of the claimed invention.” *Honeywell Int’l Inc. v. 3G Licensing, S.A.*, 124 F.4th 1345, 1356 (Fed. Cir. Jan. 2, 2025) (precedential) (quotation omitted).

Here, as discussed in Section V(A), Petitioner **does not provide a reasoned explanation** (as the color coding illustrates) for **how** Dantrey, **Strake**, and **Chen**, or for alternative grounds 4-6 **Liu**, would be combined into a coherent system that performs the claimed pipeline, including [1(a)], **fragmenting** input audio data into **input speech frames**, [1(b)], converting those input speech frames into **low-dimensional representations that omit non-content elements** based on application of a first neural network, [1(c)] applying a second neural network **to those low-dimensional representations** to **generate target speech frames**, and [1(d)] **combining the target speech frames** to generate output audio data that preserves **speech characteristics**.

The color coding underscores why that omission is fatal. The Petition does not identify a single prior-art system that performs the claimed sequence. Instead, it **assigns the claimed fragments to different references:**

- Dantrey for the alleged second neural network,
- [Chen](#), or [Liu](#), for the alleged low-dimensional representation, and
- [Strake](#) for frame-level output and reconstruction.

But the claim does not recite those concepts in isolation. It **requires an integrated pipeline** in which the same low-dimensional representations of input speech frames are the inputs to the second neural network and are used by that second neural network to generate target speech frames.

Independent Claims 11 and 16 recite substantively the same pipeline using similar limitations, including generating low-dimensional representations of input speech frames and applying a second neural network to those representations to produce target speech frames:

- Independent Claim 16 further recites additional limitations **[16(a)]** digitizing input analog signals and **[16(f)]** converting output audio data to analog audio output signals, which the Petition maps to the same arguments relied on for Claim 2. Ex1001 Cl.16; Pet. at 42-43.
- Independent Claim 11 further recites the limitation of **[11(a)]** training the first and second neural networks, which the Petition addresses only by citing to its

discussion of dependent Claims 4–6. Ex1001 Cl.11; Pet. at 52. The remaining elements of independent Claims 11 and 16 are addressed only by reference back to Claim 1. Pet. at 42-43, 52-53.

Thus, Petitioner’s failure to identify or explain this architecture for Claim 1 applies equally to Claims 11 and 16, which recite the same claimed data path and therefore rise or fall with the same deficient combination.

Fatally, the Petition never explains:

- **how** Dantrey’s spectrogram-based generator would receive and operate on Chen-style low-dimensional representations;
- **how** Strake’s separate restoration/reconstruction path would become the output path of Dantrey’s generator; or
- **how** the resulting hybrid system would preserve the claimed data flow from input speech frames to low-dimensional representations to target speech frames.

Thus, the color-coded mapping illustrates **not a reasoned combination but the absence of one**: Petitioner has picked disconnected pieces from disparate systems without showing how those pieces would fit together in a working architecture that performs the Examiner-identified limitation.

In sum, Petitioner never explains how the proposed combination would yield the Examiner-identified limitation of “apply[ing] a second neural network to the

low-dimensional representations of the input speech frames to generate target speech frames.” Ex1002 at 130.

The Petition therefore fails to establish a reasonable expectation of success.

C. The Petition Does Not Make a Particularized Showing Of Where Each Claim Element Is Disclosed In The Asserted Prior Art

Third, the petition must identify “**with particularity**” **where** each claim element is disclosed in the prior art and **explain how** the asserted grounds render the claims unpatentable. 35 U.S.C. § 312(a)(3); 37 C.F.R. § 42.104(b)(4). The Office has made clear that missing claim limitations cannot be supplied by general knowledge, common sense, or conclusory expert testimony. Ex2032 (Enforcement, Non Waiver, and Enforcement and Non Waiver Memo) at 1.

Here, Petitioner does **not** identify any prior-art reference that discloses the claimed application of a second neural network to low-dimensional representations of input speech frames to generate target speech frames **[1(c)], [11(c)], and [16(d)]**, or any system that performs the claimed pipeline as a whole. Instead, as discussed, Petitioner attempts to assemble that limitation from disconnected disclosures in Dantrey, **Strake**, and **Chen**—Dantrey’s spectrogram-based clean-speech generator, **Strake’s** separate frame-level restoration/reconstruction architecture, and **Chen’s** generic PCA disclosure—without showing that those components would form a

coherent pipeline or perform the claimed sequence of operations, including the allowance-dependent limitation.

Apple Inc. v. Ginko LLC underscores that, where, as here, a petition fails to map claim elements with particularity, denial at institution is warranted. IPR2025-01388, Paper 14 at 3 (March 2, 2026). There, the Director emphasized that a petitioner must, “from the onset,” provide an understandable, element-by-element explanation grounded in the prior art. Petitioner’s showing here falls short for the same reasons: it neither identifies a prior-art disclosure of the claimed **applying a second neural network to low-dimensional representations of input speech frames to generate target speech frames**, nor explains how the proposed combination performs **that requirement within the claimed pipeline**.

D. Institution Should Be Denied as an Inefficient Use of Office Resources

Finally, denial is also warranted where, as here, institution would not serve the efficient administration of the Office. The Director has confirmed that the Board may deny institution where only a subset of claims or grounds has potential merit and proceeding to trial would waste resources. *CyberSecure IPS, LLC v. Network Integrity Systems, Inc.*, IPR2025-01441, Paper 18 at 2–3 (May 26, 2026) (citing *Chevron Oronite Co. v. Infineum USA*, IPR2018-00923, Paper 9 at 10–11 (November 7, 2018) (informative)).

Petitioner’s reliance on the *Sotera*-style stipulation (Petitioner’s Response to Patent Owner’s Discretionary Denial Brief, Paper 8 at 8-9) **does not meaningfully mitigate** this inefficiency. Although Petitioner purports that it will not pursue in the district court litigation “the specific grounds raised in IPR2026-00273” (Ex1051 at 2), the submitted invalidity contentions confirm that Petitioner continues to assert broad combinations of references (including the art presented in this proceeding, along with other art), system-based theories, and expert-driven PHOSITA reasoning that fall outside the scope of any estoppel. Ex2034 at 4-12. For example, Petitioner expressly reserves the ability to mix and match prior art across references and systems, expanding the scope of its invalidity case beyond the grounds asserted in the Petition and undermining the intended narrowing effect of *Sotera*-style stipulations:

The references in the below charts may be combined with one or more references in the same chart, another chart, or another piece of prior art identified in these contentions, for at least the same exemplary reasons provided in the motivation to combine sections of the charts and **Krisp’s IPR and PGR Petitions, which are incorporated by reference.** IPR2026-00275 (’457 Patent); PGR2026-00032 (’561 Patent); IPR2026-00273 (’496 Patent); IPR2026-00272 (’550 Patent); IPR2026-00274 (’745 Patent); PGR2026-00033 (’756 Patent).

Ex2034 at 10 (emphasis added).

Thus, the same underlying factual disputes, expert testimony, and technical theories will proceed in parallel regardless of institution. The stipulation alters only the labeling of Petitioner's theories—not their substance—leaving intact duplicative proceedings that do not simplify the issues before the parties or the Board. Moreover, Petitioner continues to assert invalidity and has reserved broad rights to amend and supplement its invalidity positions. Ex2034 at 2-3. These reservations ensure that the district court litigation will proceed in full irrespective of any stipulation, further confirming that institution would not reduce the scope of the dispute and would result in inefficient use of the Board's resources.

Here, all asserted grounds depend on the same defective Dantrey, [Strake](#), and [Chen](#), or for alternative Ground 4 [Liu](#), combination, and the additional references in Grounds 2-3 and 4-6 are not even offered to cure and **do not cure these deficiencies** of Ground 1 and alternative Ground 4 with respect to the independent claims.

Instituting trial would therefore require the Board to expend resources on multiple grounds that fail to address the dispositive issue.

V. PETITIONER’S CHALLENGES FAIL ON THE MERITS

Petitioner asserts six obviousness grounds:

Ground	Claims Challenged	References Relied Upon	Petition Citation
Ground 1	1–3, 7–9, 16–17, and 20	Dantrey, Strake , Chen	Pet. at 12
Ground 2	4–6, 11–15, and 18–19	Dantrey, Strake , Chen , Quillen	Pet. at 43
Ground 3	10	Dantrey, Strake , Chen , Heckmann	Pet. at 55
Ground 4	1–3, 7–9, 16–17, and 20	Dantrey, Strake , Liu	Pet. at 63
Ground 5	4–6, 11–15, and 18–19	Dantrey, Strake , Liu , Quillen	Pet. at 67
Ground 6	10	Dantrey, Strake , Liu , Heckmann	Pet. at 69

This Preliminary Response focuses on Ground 1 because it supplies the base combination on which the remaining grounds depend. Ground 1 asserts that Claims 1–3, 7–9, 16–17, and 20 are obvious over Dantrey, [Strake](#), and [Chen](#). Pet. at 4, 12. Grounds 2 and 3 merely add [Quillen](#) and [Heckmann](#) for dependent-claim limitations, while retaining the same Dantrey-[Strake](#)-[Chen](#) combination for the independent-claim limitations. Pet. at 43-55. Grounds 4-6 replace [Chen](#) with [Liu](#) for the “low-dimensional representation” limitations, but otherwise rely on the same Dantrey-

Strake framework and incorporate the Ground 1 analysis for the remaining limitations. Pet. at 60-69. Those later grounds **do not cure the defects** in Ground 1 because **Quillen** and **Heckmann** address only additional dependent-claim features, while **Liu** merely supplies another alleged low-dimensional intermediate representation without explaining how Dantrey’s spectrogram-based generator and **Strake**’s separate restoration/reconstruction architecture would be integrated into a coherent system that performs the claimed sequence. Accordingly, if Ground 1 fails to establish the claimed pipeline—including the Examiner-identified limitation in independent Claims 1, 11, and 16—then Grounds 2-6 fail as well.

A. Ground 1 Fails on the Merits

The claim elements specifically identified by the Patent Office as not taught or suggested by the prior art are embodied in claim elements **[1(c)]**, **[11(c)]**, and **[16(d)]**. Claim **[1(c)]** requires “apply[ing] a second neural network to the low-dimensional representations of the input speech frames to generate target speech frames.” Ex1001 Cl. 1 (accord Cl. **[11(c)]** & **[16(d)]**).¹ Petitioner’s Ground 1 combination does not satisfy these requirements.

¹ For Claim 11, Ground 2 adopts the same analysis as Ground 1, except it adds another reference, **Quillen** (Ex1007), to allegedly disclose neural network training

Petitioner’s theory is not that any reference teaches the claimed data path. Instead, Petitioner divides the limitation across three references. Petitioner relies on Dantrey for the alleged two-neural-network audio enhancement pipeline, asserting that Dantrey’s noise subtractor and audio generator correspond to the claimed first and second neural networks. Pet. at 21, 25-26. Petitioner then turns to [Chen](#) for the alleged low-dimensional representation, arguing that [Chen](#)’s generic PCA disclosure is used as the dimensionality-reduction component. Pet. at 20-23. Finally, Petitioner invokes [Strake](#) for frame-based processing and reconstruction, relying on [Strake](#)’s inverse transform and overlap-add disclosures to argue that the combined system generates target speech frames. Pet. at 27-28.

Petitioner’s own mapping exposes the hindsight. Petitioner asserts that Dantrey-[Strake](#) in view of [Chen](#) renders obvious converting input speech frames to low-dimensional representations based on “Dantrey’s reduced noise spectrograms that have [Chen](#)’s PCA dimensionality reduction technique applied to them,” and then sends those “PCA-compressed representations” to “Dantrey’s audio generator 118.” Pet. at 20, 23. However, the Patent Office allowed the claims because the prior art failed to disclose applying a second neural network to low-dimensional

that does not exist in Dantrey, [Strake](#), or [Chen](#), making the combination even more of a stretch.

representations of input speech frames to generate target speech frames, and Petitioner acknowledges that allowance rationale. Pet. at 3.

That division of labor is **not** a teaching in the prior art; it is a purported reconstruction from the claim language. Dantrey allegedly supplies the second network, Chen allegedly supplies dimensionality reduction, and Strake allegedly supplies frames. But obviousness requires more than assigning pieces of a claim to different references. *See KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007); *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006). Petitioner had to show a coherent combined system in which input speech frames are converted into low-dimensional representations that omit non-content elements, and further that a second neural network is applied to those representations to generate target speech frames. Ex1001 Cl. 1. **It did not.**

Dantrey confirms the mismatch. Dantrey is a spectrogram-based audio noise-removal pipeline. It explains that “primary audio . . . can be enhanced by removing background noise using a first neural network,” and that the resulting “spectrogram 204 having background noise . . . reduced by noise subtractor 116 can be provided as input to an audio generator,” which is a “frequency-domain to audio, deep learning-based clean speech generator” used “to generate clean speech audio from noise-reduced spectrograms.” Ex1004, 4:22-5:7. Dantrey therefore teaches a spectrogram-

to-audio generator, **not** a *second* neural network that receives *low-dimensional representations of input speech frames* and generates target speech frames.

Strake does not fill that gap. **Strake** is *not* a frame-based version of Dantrey's generator. It uses a different architecture: an interference-suppression stage followed by a restoration stage. Its first stage suppresses acoustic interference, while its second stage restores speech components such as "pitch frequency, pitch harmonic structure, formant structure, spectral envelope, [and] spectral phase." Ex1005 ¶[031]. And unlike Dantrey's audio generator, **Strake**'s second stage does *not* itself produce final time-domain audio; it outputs only "a vector representation of the second-stage clean signal estimate 213 for the current frame," which must then be sent to a separate "signal reconstruction circuit 214" for reconstruction. Ex1005 ¶[042]. **Strake** therefore gives Petitioner a separate frame/reconstruction disclosure to paste onto Dantrey; it does *not* show that Dantrey's generator can receive a different intermediate representation and output the frame-level target material Petitioner needs.

Chen fares no better. **Chen** is a noise-classification and attenuation reference directed to "eliminating undesired audio artifacts, such as background noises, on an audio channel." Ex1006 Abstract. It receives "a set of audio segments," uses a first machine-learning model to identify "a first probability of unwanted background noises," conditionally applies a second model to determine "a second probability

that the one or more background features exist,” and attenuates those artifacts when “the second probability exceeds a second threshold.” Ex1006 Abstract. [Chen](#)’s dimensionality-reduction disclosure is only a generic statement that “ML models can employ a dimensionality reduction approach,” including “an Incremental Principal Component Analysis (PCA) algorithm” among other options. Ex1006 3:66-4:4. [Chen](#) therefore supplies, at most, a generic PCA label; it does *not* transform Dantrey’s spectrogram generator or [Strake](#)’s restoration/reconstruction stage **into the claimed second-neural-network operation.**²

Thus, Ground 1 is an improper aggregation. None of Dantrey, [Strake](#), or [Chen](#)—alone or in combination—teaches the claimed architecture. The claim requires one coherent data path: input speech frames are converted into low-dimensional representations that omit non-content elements, and a second neural network is applied to those representations to generate target speech frames. Ex1001 Cl. 1. Petitioner’s stitched-together theory never shows that combined operation. Ground 1 therefore fails.

² In Ground 4-6, the Petition offers in the alternative to substitute [Liu](#) for Chen in a similar defective manner. As discussed in Section V, *infra*, that substitution likewise fails.

1. The Ground 1 Combination Is Architecturally Incoherent

The Petition's proposed Ground 1 combination fails at a foundational level because it assumes compatibility among incompatible data representations and processing stages. The proposed system requires using Dantrey's second neural network, applying it to the low-dimensional representations allegedly supplied by [Chen](#), and generating output in the frame-based format allegedly supplied by [Strake](#). But the Petition never explains how those three distinct teachings operate together in one coherent system. Pet. at 13-32.

Dantrey's data path is spectrogram-based. Its noise subtractor produces an "audio spectrogram with reduced background noise," which is passed to audio generator 118. Ex1004 3:62-64. Dantrey further describes audio generator 118 as a "frequency-domain to audio, deep learning-based clean speech generator" that may "generate clean speech audio from noise-reduced spectrograms." Ex1004 5:7-16. Dantrey's figures reinforce that architecture by depicting STFT-based processing, including "Input STFT" and "Output STFT Gains." Ex1004 Fig. 3A. Thus, Dantrey's generator is built to operate on a structured time-frequency representation, *not* on an abstract reduced-dimension vector.

[Chen](#)'s alleged contribution creates a fundamental input mismatch. Petitioner relies on [Chen](#)'s PCA disclosure to supply the claimed low-dimensional representation, asserting that Dantrey's reduced-noise spectrograms would have

“[Chen](#)’s PCA dimensionality reduction technique applied to them” before being sent to Dantrey’s audio generator which “is processed in frames per Strake, particularly based on Dantrey’s express sample size, sample rate, and hop size.” Pet. at 25-26. But [Chen](#) does **not** teach an input format for Dantrey’s generator. [Chen](#) merely states, in the context of its own noise-mitigation system, that “ML models can employ a dimensionality reduction approach,” including “an Incremental Principal Component Analysis (PCA) algorithm” among other options. Ex1006 3:66-4:4. [Chen](#)’s system itself classifies audio segments using probability thresholds and attenuates background noise; it does **not** disclose using PCA-reduced speech-frame representations as the input to a spectrogram-to-audio generator as the Petition proposes. Ex1006 Abstract, 2:44-57.

Applying [Chen](#)’s PCA to Dantrey’s spectrogram stream before Dantrey’s generator would materially alter the very data structure Dantrey’s generator is designed to process. A spectrogram preserves organized time-frequency structure. PCA, by contrast, transforms input data into reduced principal-component vectors that capture variance in the input data. Petitioner does **not explain how** Dantrey’s generator—trained and disclosed to generate clean speech audio from noise-reduced spectrograms—would receive, interpret, or synthesize audio from [Chen](#)-style PCA vectors. The Petition therefore assumes, without explanation, that Dantrey’s generator can speak a different input language than the one Dantrey discloses.

That technical mismatch is independently fatal because, even if Chen's generic PCA disclosure were combined with Dantrey in the abstract, Petitioner still has not shown that the resulting system would be the claimed invention.

Petitioner never explains:

- **whether** the resulting PCA-reduced data remains a spectrogram;
- **whether** it has the same dimensional arrangement expected by Dantrey's audio generator;
- **whether** it preserves the spectral, harmonic, temporal, or other speech-synthesis-relevant information needed for Dantrey's generator to generate speech from spectrograms; or
- **whether** Dantrey's generator would have to be retrained or redesigned to accept that new representation.

Without that explanation, Petitioner simply assumes that Dantrey's generator can receive, interpret, and synthesize audio from [Chen](#)-style PCA vectors as opposed to the spectrograms Dantrey is designed to use. [Strake](#) does not cure that mismatch and if anything makes it worse. [Strake](#)'s second stage is **not the same type** of spectrogram-to-audio generator Petitioner identifies in Dantrey. [Strake](#)'s second stage outputs "a vector representation of the second-stage clean signal estimate 213 for the current frame," which must then be sent to a separate "signal reconstruction circuit 214" to generate time-domain audio, for example through IFFT and overlap-add. Ex1005 ¶[042]. Thus, [Strake](#) supplies a restoration-and-reconstruction architecture that stops at frame-level intermediate vectors, while Dantrey's generator is part of a pipeline designed to "generate clean speech audio from noise-reduced

spectrograms”. Ex1004 5:7-16. Accordingly, even if Strake were added to the proposed combination, it **would not transform Dantrey’s spectrogram-to-audio generator** into the claimed second neural network that applies low-dimensional representations of input speech frames to generate target speech frames.

Petitioner still lacks a reasoned explanation of **how** Strake’s separate frame-based restoration and reconstruction disclosures would be integrated into Dantrey’s generator, or **how** that modified generator would preserve the claimed data path from Chen’s alleged low-dimensional representations to the claimed target speech frames. At best, adding in Strake’s disclosure **would result in chaos, not the claimed invention**. If Dantrey’s generator were modified to accept [Chen](#)’s PCA-reduced representations and output [Strake](#)-style frame-level estimates, Dantrey’s disclosed pipeline would no longer operate as designed. At minimum, the Petition needed to explain how Dantrey’s audio generator would be redesigned or retrained to accept PCA-reduced inputs, preserve the information needed for speech synthesis, generate [Strake](#)-style frame estimates, and then integrate those estimates into a reconstruction path. *See In re Kahn*, 441 F.3d at 988 (“[R]ejections on obviousness grounds cannot be sustained by mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness.”).

The Petition provides no articulated reasoning with rational underpinnings. Instead, it assumes that a spectrogram-to-audio generator, a PCA-based dimensionality-reduction disclosure, and a separate frame reconstruction architecture can be stitched together without addressing the resulting input/output mismatch.

This **does not merely require minor reprogramming**. (*Cf.* Pet. at 29, 32; *see also Id.* at 19; Ex1003 ¶¶110, 138, 148). It is a hindsight purported reconstruction that changes the purpose and operation of each reference. *See Metalcraft of Mayville, Inc. v. Toro Co.*, 848 F.3d 1358, 1367 (Fed. Cir. 2017) (“we cannot allow hindsight bias to be the thread that stitches together prior art patches into something that is the claimed invention”). Dantrey generates clean speech audio from noise-reduced spectrograms; [Chen](#) classifies and attenuates noise and merely lists PCA as one possible ML technique; and [Strake](#) restores speech components before separate reconstruction from frame-level vectors. Ex1004 5:7-16; Ex1006 Abstract, 3:66-4:4; Ex1005 ¶¶[031], [042]. Petitioner provides no reasoned explanation for **how** those different functions become the claimed operation of applying a second neural network to low-dimensional representations to generate target speech frames.

The Ground 1 combination is therefore architecturally incoherent.³

Because Petitioner never reconciles the architectural mismatch among Dantrey, [Chen](#), and [Strake](#), the Petition provides no coherent basis to find a reasoned motivation to combine, a reasonable expectation of success, or a particularized mapping of the claimed pipeline.

2. The Petition's Motivation Theory Fails

Petitioner argues that a PHOSITA would have been motivated to combine Dantrey, [Strake](#), and [Chen](#) because framing, dimensionality reduction, and neural-network audio enhancement were known and allegedly advantageous. Pet. at 16-31. But that reasoning abstracts away the actual disclosures and the incompatibilities among them.

Dantrey already discloses a complete audio enhancement pipeline built around noise-reduced spectrograms and a frequency-domain-to-audio generator. Ex1004 5:7-16. [Chen](#)'s PCA disclosure belongs to a different classifier/attenuation

³ Ground 2, which addresses the third independent Claim 11, relies upon the same analysis and fails for the same reason. As discussed in Ground 2, the addition of [Quillen](#) is **not** offered to address this problem, but merely to add another reference for an additional feature that has no logical connection to the other references further exacerbating the problems with the primary combination.

system. Ex1006 Abstract, 3:66-4:4, 5:1-65. [Strake](#)’s frame-based disclosure belongs to an audio restoration architecture that outputs frame-level vectors requiring separate reconstruction. Ex1005 ¶[042]. Petitioner **never explains why** a PHOSITA would replace or modify Dantrey’s spectrogram input with [Chen](#)’s PCA output, then graft [Strake](#)’s frame-reconstruction mechanism onto Dantrey’s generator.

The Petition also **fails to establish reasonable expectation of success**. Petitioner does not explain **how** Dantrey’s audio generator would be redesigned or retrained to accept PCA-reduced representations; how those representations would preserve the information needed for Dantrey’s generator; **how** [Strake](#)’s frame-level estimates would become the output of Dantrey’s generator; or **how** the resulting system would still perform the claimed operation. Simply asserting that the individual techniques were known does not satisfy Petitioner’s burden. Nor has Petitioner shown that, even if those disparate teachings were combined, the resulting system would be the claimed invention—a system in which the second neural network is applied to the low-dimensional representations of the input speech frames to generate target speech frames.

In short, Petitioner has not shown a reasoned path from Ground 1’s prior art to the claimed invention. Petitioner’s mapping is built from the claim language backward: Dantrey is selected for “second neural network,” [Chen](#) for “low-

dimensional representation,” and [Strake](#) for “frames.” **That is hindsight, not obviousness.** Moreover, combining such disparate pieces does not mean a PHOSITA would have expected the combination to result in the claimed invention—particularly where Petitioner never explains how the combined system would preserve the claimed data path from low-dimensional representations of input speech frames to target speech frames generated by the second neural network.

Because Petitioner’s motivation theory never explains *why* a PHOSITA would have combined Dantrey, [Chen](#), and [Strake](#) into the specific claimed pipeline, the Petition offers **only hindsight**—not a reasoned motivation to combine, a reasonable expectation of success, or a particularized mapping grounded in the asserted disclosures.

B. Even If the Board Does Not Treat Ground 1 as Dispositive, Grounds 2-3 Independently Fail Because the Additional References Do Not Cure the Base-Combination Defect

Ground 2 adds a fourth reference, [Quillen](#) (Ex1007), to the Ground 1 analysis for the independent claims, for dependent Claims 4-6, 11-15, and 18-19. Petitioner relies on [Quillen](#) for training neural networks using input audio, augmentations, transcripts, and mapping concepts. Pet. at 43-54. But [Quillen](#) **does not cure the missing architecture** in Ground 1.

Independent Claim 11, like independent Claim 16 found in ground 1, recite materially similar core limitations to Claim 1. Claim 11 requires applying a trained

first neural network to convert fragmented input speech frames to low-dimensional representations that omit non-content elements, applying the trained second neural network to those low-dimensional representations to generate target speech frames, and combining target speech frames to generate output audio data. Ex1001 Cl.11.

Ground 2 does not offer any fix for the failures of Ground 1 discussed above. Quillen does **not** show that Dantrey's second neural network receives Chen-style low-dimensional representations and generates target speech frames. Quillen is used only after Petitioner assumes the base Dantrey-Strake-Chen architecture exists. But as shown above, that architecture is missing.

Nor does Quillen's use of transcripts cure the defect. The challenged claims that recite transcripts require training the first neural network using input audio training data, augmentations, and transcripts to learn a mapping between input training speech frames and low-dimensional representations. Ex1001 Cls. 4, 11, 18. Petitioner's reliance on Quillen does **not** explain how transcripts are used to train Dantrey's first neural network to produce the claimed low-dimensional representations that are then consumed by Dantrey's second neural network. Pet. at 43-54.

Ground 3 does not offer any fix for the failures of Ground 1 discussed above. Ground 3 adds Heckmann for Claim 10's hierarchical feature extraction limitation. Pet. at 55-60. But Claim 10 depends from Claim 9, which depends from

Claim 1. Ex1001 Cl. 9-10. Heckmann does not cure the missing limitation of Claims [1(c)], [11(c)], or [16(d)]. Even if hierarchical feature extraction and compression were known, that does not establish that the proposed Dantrey-Strake-Chen system applies a second neural network to low-dimensional representations to generate target speech frames.

Because Quillen and Heckmann do not cure the missing claimed pipeline in the base combination, Grounds 2 and 3 inherit the same failures of particularity, motivation to combine, and reasonable expectation of success.

Accordingly, Grounds 2 and 3 fail with Ground 1.

C. Alternative Grounds 4-6 Likewise Fail on the Merits

Grounds 4-6 replace Chen with Liu for the low-dimensional-representation limitation. Petitioner relies on Liu's disclosure of a "low dimensional mel spectrogram" as an intermediate representation. Pet. at 60-70. These grounds fail for the same reasons as Ground 1 and for additional reasons.

First, the Liu grounds still rely on Dantrey and Strake for the surrounding architecture. Pet. at 60-70. As explained above, Petitioner does not overcome the incoherent combination of Dantrey and Strake. Replacing Chen with Liu does not fix that failure. Further, the addition of Chen failed in Ground 1 because Chen's generic PCA disclosure belongs to a *different* noise-classification and attenuation

system and *does not bridge* Dantrey’s spectrogram-based generator with **Strake**’s frame-based reconstruction path. **Liu** does not fix that failure either.

Second, **Liu**’s “low dimensional mel spectrogram” is **not** shown to be interchangeable with Dantrey’s noise-reduced spectrogram. Petitioner describes **Liu** as a speech-restoration system that uses an analysis stage to generate a low-dimensional mel spectrogram and a synthesis stage (a neural vocoder) to generate restored speech. Pet. at 61-62. Dantrey’s pipeline, by contrast, uses a frequency-domain noise subtractor and audio generator in its own spectrogram-based data flow. Ex1004 Fig. 1, 3:54-65, 4:41-5:16. That substitution would **not** work because **Liu**’s mel-spectrogram representation is generated for **Liu**’s own neural-vocoder synthesis stage, whereas Dantrey’s audio generator is disclosed to generate clean speech from Dantrey’s noise-reduced spectrograms; Petitioner never explains why Dantrey’s generator would accept **Liu**’s different intermediate feature representation without redesign or retraining.

Petitioner does **not** show that **Liu**’s mel spectrogram has the same or compatible structure, statistical properties, training role, or downstream compatibility as Dantrey’s noise-reduced spectrogram. Nor does Petitioner explain **how** Dantrey’s generator would be modified or retrained to receive **Liu**’s low-dimensional mel spectrogram. Petitioner simply assumes that because both references involve spectrogram-like data, the representations are interchangeable.

That assumption is unsupported, and insufficient to establish a combination would result in the claimed invention.

Third, **Liu** does **not** cure the target-speech-frame limitation. Even if **Liu** discloses a low-dimensional mel spectrogram, Petitioner still must show that a second neural network is applied to low-dimensional representations of input speech frames to generate target speech frames. Petitioner does not make that showing. Instead, the **Liu** grounds repeat the same purported claim reconstruction: Dantrey for the second network, **Strake** for frames, and **Liu** for the intermediate representation. Pet. at 60-70. That remains an improper patchwork.

Fourth, **Liu** does **not** satisfy the claim requirement to “omit one or more of the non-content elements”. Ex1001 Cl. 1. Claim **[1(b)]**, which generates the low-dimensional representations used in Claim **[1(c)]**, requires that the low-dimensional representations of the input speech frames omit one or more non-content elements. *Id.* at Cl. 1. Petitioner asserts that **Liu**’s “low dimensional mel spectrograms” omit background noise, Pet. at 63, but it does **not** substantiate that assertion or explain how **Liu**’s mel spectrogram, when substituted into Dantrey’s pipeline, is the claimed representation that both omits non-content elements and serves as the input to the second neural network that generates target speech frames in the proposed Dantrey-Strake-**Liu** system. Pet. at 60-67.

Grounds 5 and 6 add **Quillen** and **Heckmann** to the **Liu**-based combination. Pet. at 67-70. But those added references **do not cure** the base defect. **Quillen** does **not** make **Liu**'s intermediate representation compatible with Dantrey's generator, and **Heckmann** does **not** supply the missing second-neural-network target-frame-generation operation.

Because replacing **Chen** with **Liu** does **not** resolve the same architectural mismatch between Dantrey's generator and **Strake**'s frame-level reconstruction architecture, the **Liu**-based grounds still lack a particularized mapping of the claimed pipeline, a reasoned motivation to combine the references in that manner, and a reasonable expectation that the modified system would operate successfully.

Accordingly, Grounds 4-6 also fail.

VI. CONCLUSION

Petitioner’s challenges are built on hindsight. **The Petition does not identify the claimed architecture in the prior art.** It does **not** show a second neural network applied to low-dimensional representations of input speech frames to generate target speech frames. It does **not** explain how [Chen](#)’s PCA or [Liu](#)’s mel spectrogram can be inserted into Dantrey’s spectrogram-based generator. It does **not** show that [Strake](#)’s restoration and reconstruction pipeline satisfies the claimed target-speech-frame generation. And it does **not** show that the proposed combination—even if made—would result in the claimed invention.

Combining isolated pieces labeled “second neural network,” “low-dimensional representation,” and “frames” **does not establish the claimed pipeline** unless the Petition also explains **how** those pieces operate together as a system in which the second neural network applies to the low-dimensional representations of input speech frames to generate target speech frames. The Petition **provides no such explanation** and therefore fails to establish a non-hindsight motivation to combine or reasonable expectation of success.

For the reasons presented above, the Petition should be denied, and no inter partes review should be instituted.

Respectfully submitted,

AMSTER, ROTHSTEIN & EBENSTEIN LLP
Attorneys for Patent Owner
405 Lexington Avenue, 48th Floor
New York, NY 10174
(212)336-8000

Dated: June 30, 2026
New York, New York

By: /Charles R. Macedo/
Charles R. Macedo (Reg. No. 32,781)

Of counsel:
Gary J. Gershik (Reg. No. 39,992)
Lewis Derenzis III (Reg. No. 84,132)

CERTIFICATE OF WORD COUNT

PRELIMINARY RESPONSE complies with the type-volume limitation in 37 C.F.R. § 42.24(b)(1). According to the word-processing system's word count, the brief contains approximately 8,289 words, excluding the parts exempted by 37 C.F.R. § 42.24(a).

Respectfully submitted,

AMSTER, ROTHSTEIN & EBENSTEIN LLP
Attorneys for Patent Owner
405 Lexington Avenue, 48th Floor
New York, NY 10174
(212)336-8000

Dated: June 30, 2026
New York, New York

By: /Charles R. Macedo/
Charles R. Macedo (Reg. No. 32,781)

CERTIFICATE OF SERVICE

I hereby certify that on June 30, 2026, a true and correct copy of the foregoing PATENT OWNER'S PRELIMINARY RESPONSE was served via email, by consent, to Petitioner's counsel of record at the correspondence addresses of record in this proceeding.

Eric A. Buresh (Reg. No. 50,394)
Chris R. Schmidt (Reg. No. 63,982)
Nick R. Apel (Reg. No. 84,017)
Adam M. Sandwell (Reg. No. 72,484)

Erise IP, P.A.

7015 College Blvd., Suite 700
Overland Park, Kansas 66211
Telephone: (913) 777-5600
Fax: (913) 777-5601
chris.schmidt@eriseip.com
nick.apel@eriseip.com
adam.sandwell@eriseip.com
PTAB@eriseip.com

Respectfully submitted,

AMSTER, ROTHSTEIN & EBENSTEIN
LLP
Attorneys for Patent Owner
405 Lexington Avenue, 48th Floor
New York, NY 10174
(212) 336-8000

Dated: June 30, 2026
New York, New York

By: /Charles R. Macedo/
Charles R. Macedo (Reg. No. 32,781)