

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

KRISP TECHNOLOGIES, INC.,
Petitioner,

v.

SANAS.AI INC.,
Patent Owner.

Case IPR2026-00274
Patent 12,131,745

PATENT OWNER’S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I. INTRODUCTION	1
II. THE '745 PATENT.....	5
III. CLAIM CONSTRUCTION	9
IV. THE PETITION FAILS TO DEMONSTRATE MOTIVATION, REASONABLE EXPECTATION OF SUCCESS, AND PARTICULARIZED MAPPING OF CLAIM ELEMENTS	11
A. The Petition Does Not Demonstrate a Reasoned Motivation to Combine Prior Art References.....	12
B. The Petition Does Not Demonstrate a Reasonable Expectation of Success	13
C. The Petition Does Not Make A Particularized Showing Of Where Each Claim Element Is Disclosed In The Asserted Prior Art.....	14
D. Institution Should Be Denied as an Inefficient use of Office Resources	15
V. PETITIONER’S CHALLENGES FAIL ON THE MERITS.....	17
A. Ground 1 Fails on the Merits	18
1. Ground 1 Independently Fails Because Keshet, King, and Cuturi Are Not Analogous Art	22
2. The Ground 1 Combination Is Mathematically and Architecturally Incoherent and Still Does Not Yield the Claimed Operation.....	27
3. Ground 1 Also Fails Because Petitioner’s Motivation-to- Combine Theory Is Hindsight-Driven and Lacks a Reasonable Expectation of Success.....	30
4. Even If the Board Does Not Treat Ground 1 as Dispositive, Grounds 2-5 Independently Fail Because the Additional References Do Not Cure the Base-Combination Defect	32
VI. CONCLUSION.....	35

TABLE OF AUTHORITIES

Page(s)

CASES

<i>ActiveVideo Networks, Inc. v. Verizon Commc'ns, Inc.</i> , 694 F.3d 1312 (Fed. Cir. 2012)	12, 13
<i>Apple Inc. v. Ginko LLC</i> , IPR2025-01388, Paper 14 (March 2, 2026)	15
<i>In re Bigio</i> , 381 F.3d 1320 (Fed. Cir. 2004)	23
<i>Chevron Oronite Co. v. Infineum USA</i> , IPR2018-00923, Paper 9 (November 7, 2018)	16
<i>CyberSecure IPS, LLC v. Network Integrity Systems, Inc.</i> , IPR2025-01441, Paper 18 (May 26, 2026)	16
<i>Honeywell Int'l Inc. v. 3G Licensing, S.A.</i> , 124 F.4th 1345 (Fed. Cir. Jan. 2, 2025)	14
<i>Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.</i> , 821 F.3d 1359 (Fed. Cir. 2016)	13
<i>In re Kahn</i> , 441 F.3d 977 (Fed. Cir. 2006)	12, 13
<i>KSR Int'l Co. v. Teleflex Inc.</i> , 550 U.S. 398 (2007)	11, 12, 13
<i>PAR Pharm., Inc. v. TWI Pharms., Inc.</i> , 773 F.3d 1186 (Fed. Cir. 2014)	14
<i>Xerox Corp. v. Bytemark, Inc.</i> , IPR2022-00624, Paper 9 (PTAB Aug. 24, 2022)	13

STATUTES

35 U.S.C. § 312(a)(3)	4, 11, 14
-----------------------------	-----------

37 C.F.R. § 42.104(b)(4).....	14
37 C.F.R. § 42.107	1
37 C.F.R. § 42.120(a).....	5

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”)
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”)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
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.”)
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> ”)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
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)
Ex2022	U.S. Publication No. 2024/0363135 A1 (published Oct. 31, 2024)
Ex2023	U.S. Patent No. 12,125,496 (issued Oct. 22, 2024)
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”)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2027	Arto Minasyan, <i>Krisp Voice AI Lab: Revolutionizing Voice Technology</i> , <i>LinkedIn</i> (posted Mar. 2026), https://www.linkedin.com/posts/artominasyan_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).
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*	Prosecution History of U.S. Patent No. 12,131,745
Ex2033*	<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),</i> https://www.uspto.gov/sites/default/files/documents/aapa_memo_final_signed.pdf. (the “Enforcement, Non Waiver, and Enforcement and Non Waiver Memo”)

*** New Exhibit(s) 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. 12,131,745. Institution should be denied because Petitioner fails to show that the combination of asserted prior art discloses the challenged claims.

The Petition fails at the threshold because it does not identify any error by the Office or show that the asserted art cures the gap the Office found.¹

In particular, the Office allowed the claims after determining that the prior art did not disclose the alignment limitations recited in the independent claims, including the corresponding limitations exemplified by claim elements **[1(d)]** and **[1(e)]**², which require:

¹ The petition's error appears to stem in part from the Petition's submission of Ex1002 as the prosecution history for the '745 Patent, which is incorrect as it is the prosecution history for U.S. Patent No. 11,715,457. *See e.g.* Ex1002 at 126. Patent Owner submits Ex2032 herewith as the accurate prosecution history for the '745 Patent.

² Although the Notice of Allowance discussed the alignment limitation using the language of claim 8, Petitioner treats the corresponding limitations of the

[1(d)] determine a differentiable alignment by jointly maximizing a cosine distance between the first phonetic embedding vectors and the second phonetic embedding vectors; and **[1(e)]** align the speech data to the phonetic content based on the differentiable alignment to generate and provide output audio data corresponding to the aligned speech data and representing the target accent.

Cf. Ex2032, at 31, 32 (color coding supplied); with Pet. at 18, 25.

Rather than show that any reference discloses those allowance-dependent limitations, the Petition sidesteps the Office’s reasoning and seeks to cobble together at least three disparate references in an illogical manner:

(1) Keshet (applied to the black text) addresses localized speech-therapy correction (rather than accent correction like the ’745 Patent) by replacing individual mispronounced phoneme segments, not by aligning first and second phonetic embedding vectors. Ex1004 at Abstract, ¶¶[0018],[0087], [0092]-[0093]. The output is then “spliced and/or joined into the original input speech” to “replace a corresponding wrongly pronounced speech segment.” *Id.* ¶[0108].

independent claims together: the Petition substantively maps claim 1 in detail and then cites back to its claim 1 analysis for claim 8. Pet. at 28–29. Accordingly, claim elements **[1(d)]** and **[1(e)]** are representative of the independent-claim alignment limitations for purposes of addressing Petitioner’s Ground 1 theory.

- (2) **King (applied to the blue bold text)** relates to aligning sound to replace dialog in a different language (rather than accent correction) and uses cosine distance only as a frame-level acoustic similarity measure for conventional DTW alignment of recordings, *not as the basis for a differentiable phonetic-embedding-vector alignment*. Ex1005 ¶¶[0019], [0030], [0032], [0038], [0041]-[0043].
- (3) Cuturi (applied to the green underscore text) supplies a differentiable DTW-style loss for generic time series, *not the claimed differentiable alignment determined by jointly maximizing cosine distance between phonetic embedding vectors*. Ex1006 at 1-2. Cuturi does not callout accent correction as an intended use. *Id* at 1-10.

That color-coded patchwork of this non-analogous Keshet, **King**, and Cuturi is **not** a disclosure of elements **[1(d)]** and **[1(e)]**; it is an admission that **no single reference—or any coherent combination—teaches the claimed operation**.

Each addresses different data, objectives, and architectures, and the Petition does not explain how they would be combined to perform a differentiable alignment that jointly maximizes cosine distance and then uses that alignment to generate output audio. None of these references are actually in the same field of use as the '745 Patent, which is real-time accent correction, and thus are not analogous art to the claimed invention.

The logic used to support this combination of references is like using a thermometer to measure distance because it has numbers on it: a thermometer is designed to capture temperature differences, not spatial relationships. Even if you repurpose its numbers, it cannot perform the function required—because the underlying concept and data it operates on are fundamentally different from what distance measurement requires.

Because Petitioner does not identify any prior-art teaching of the claimed alignment and alignment-based output—either in a single reference or a coherent combination—it does not cure the deficiency in the art identified by the Patent Office in the Notice of Allowance.

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 purportedly reconstructing the claims from disconnected fragments.

Grounds 2–5 fare no better and do not even try to cure Ground 1’s defect. They rely on the same deficient Keshet-[King-Cuturi](#) combination for the independent claims and merely add Kameoka, Fan, Chen, or Vyas for dependent-claim details. Pet. at 31-42. Those additional references do not supply the missing differentiable alignment and merely add more disparate snippets to further clutter the chaos proposed for Ground 1. Grounds 2–5 therefore fail with Ground 1—or, at

minimum, independently fail because Petitioner never shows that any added reference cures the defect in the base combination.

Institution should be denied because Petitioner fails to show the asserted prior art discloses the subject matter of the challenged claims.³

II. THE '745 PATENT

The '745 Patent is entitled "*System and Method for Automatic Alignment of Phonetic Content for Real-Time Accent Conversion.*" It is directed to real-time accent conversion by transforming phonetic characteristics of speech in a source accent to more closely match a target accent in a very specific and inventive manner.

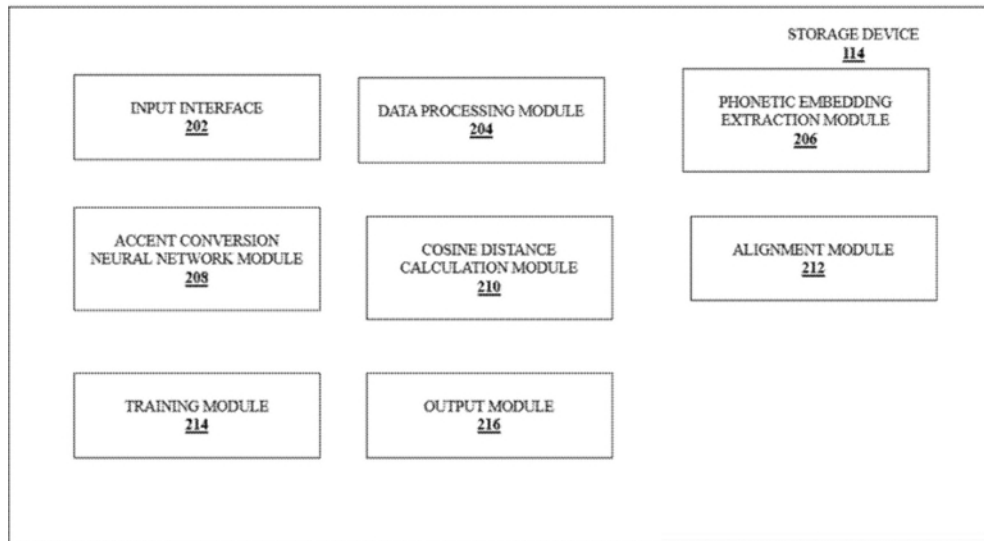
The Abstract provides an overview of the invention:

An alignment is determined by maximizing a cosine distance between the set of phonetic embedding vectors and the set of transformed phonetic embedding vectors. The speech data is then aligned to the phonetic content based on the determined alignment to generate output audio data representing the target accent. The disclosed technology transforms phonetic characteristics of a source accent to match the target accent more closely for efficient and seamless accent conversion in real-time applications.

Ex1001 at Abstract.

Fig. 2 shows this same architecture in greater detail:

³ Patent Owner does not waive any arguments and reserves the right to present a full response if inter partes review is instituted. See 37 C.F.R. § 42.120(a).



As pertinent here, storage device 114 containing the accent conversion system 100, shown in Fig. 2, includes, among other components, a phonetic embedding extraction module 206, an accent conversion neural network module 208, a cosine distance calculation module 210, and an alignment module 212. Ex1001 at Fig. 2, 4:34-43.

The phonetic embedding extraction module 206 extracts phonetic embedding vectors that “capture and represent the phonetic characteristics of the input speech or audio data in a numerical format” and may encode phonemes, speech sounds, and other phonetic units present in the source-accent input. Ex1001 at 5:20-37; *see also* 5:38-60 (elaborating on further details of how the module works).

The accent conversion neural network 208 “may predict the transformed phonetic embedding vectors that represent the target accent based on the source accent.” Ex1001 at 5:62-64; *see also* 5:64-6:43 (elaborating on the details of how this is performed).

The cosine distance calculation module 210 calculates cosine distance between the phonetic embedding vectors and the transformed phonetic embedding vectors, and the specification explains that “[m]aximizing the cosine distance may jointly align the phonetic embedding vectors in a way that minimizes their dissimilarity and maximizes their similarity.” Ex1001 at 6:44-7:5.

The alignment module 212 then uses that information to align the phonetic content of the source accent with the target accent. Ex1001 at 7:7-30. “By performing automatic alignment, the alignment module 212 may facilitate transformation of the phonetic characteristics of the input speech data in the source accent to closely match those of the target accent. This alignment process may ensure that important phonetic features are preserved while adapting the phonetic content to the desired target accent.” Ex1001 at 7:40-46. “Thus, the alignment module 212 allows for seamless and efficient conversion of accents during speech processing and may ensure that the converted speech maintains the natural flow and rhythm while accurately reflecting the desired target accent.” Ex1001 at 7:46-50.

Based on the disclosure, the Examiner found Claim 1 allowable for the following reasons as set forth in the Notice of Allowance:

The closest prior art of record, Fan et al. (US PGPub 2023/0223006) discloses obtaining a set of phonetic embedding vectors for phonetic content representing a source accent and obtained from input audio data (“*The terminal inputs the first voice into a voice feature extraction model, and extracts a PPG feature of the first voice by using the voice feature extraction model*”, [0084], see [0120] for accent); and applying a trained machine learning model to the set of phonetic embedding vectors to generate a set of transformed phonetic embedding vectors corresponding to phonetic characteristics of speech data in a target accent (“*The terminal inputs the style feature and the PPG feature into a voice fusion model for fusion, to obtain a second voice.*”, [0096], see [0120] for accent).

Fan, alone or in combination with the prior art of record, ***does not disclose determining an [8(d)] alignment by maximizing a cosine distance between the set of phonetic embedding vectors and the set of transformed phonetic embedding vectors; and [8(e)] aligning the speech data to the phonetic content based on the determined alignment to generate output audio data representing the target accent.***

Ex2032 at 31-32 (emphasis supplied).⁴ Thus, like the current petition which offers a base reference (Keshet, Ex1004) with comparable teachings to Fan et al.

⁴ Fan et al. cited in the reasons for allowance is the same prior art relied upon in the Petition for Ground 3. The Petition offers no explanation why this previously considered prior art is being presented in a new or different light from what the Office already considered and distinguished in the Notice of Allowance.

(Ex1008), neither the Office nor the Petitioner has identified art which discloses elements 1(d) and 1(e).⁵

Significantly, the only basis proffered by the Petition to address this basis for allowance was the combination of Keshet, [King](#), and [Cuturi](#) in Ground 1, 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 found by the Office in the notice of allowance that the prior art fails to teach the cosine-distance-based alignment limitations recited in elements **[1(d)]** and **[1(e)]** and the other independent claims.

III. CLAIM CONSTRUCTION

Petitioner asserts that the claims should be interpreted under the *Phillips* standard and applies the plain and ordinary meaning of the claim terms as understood by a PHOSITA. Pet. at 3. For purposes of institution, Patent Owner agrees that no express claim construction is necessary because Petitioner's grounds fail under the plain and ordinary meaning of the claims. The institution determination on the merits turns on whether the prior art teaches elements **[1(d)]**, **[8(c)]**, and **[15(c)]** and **[1(e)]**,

⁵ The Office recognized that this distinction applied to all the pending claims, including the other independent claims 8 and 15. Ex2032 at 31-32. Likewise, the Petition applies the same defective analysis to all three independent claims.

[8(d)], and **[15(d)]**, which the Petition does not come close to demonstrating. Accordingly, there is no claim construction dispute material to the institution decision. Patent Owner reserves the right to propose and address claim constructions should trial be instituted.

IV. THE PETITION FAILS TO DEMONSTRATE MOTIVATION, REASONABLE EXPECTATION OF SUCCESS, AND PARTICULARIZED MAPPING OF CLAIM ELEMENTS

The Board may deny institution where a petitioner fails to satisfy its threshold burdens, which in the present case include 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 as they each independently show that there is not a reasonable likelihood of success that the petitioner would prevail. 35 U.S.C. § 314(a).

Each of these requirements fails for the same reason: Petitioner does not identify any prior-art teaching of the claimed alignment **[1(d)]**, **[8(c)]**, and **[15(c)]** and **[1(e)]**, **[8(d)]**, and **[15(d)]**, nor explain how the asserted references—alone or in any coherent combination—perform that operation.

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 technique[s]” (Pet. at 22, 25, 34, 37, 39, 41; Ex1003 ¶¶98, 103-105, 132-133, 138, 144-145, 148), and that any modifications would have required only “reprogramming” (Pet. at 22, 25, 37, 39, 42; Ex1003 ¶¶98, 105, 139, 145, 149). 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 Dr. 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”).

Here, as discussed in Section V.A, Petitioner does not provide a reasoned explanation for how Keshet, [King](#), and [Cuturi](#) would be combined into a coherent system that performs the claimed differentiable alignment in elements **[1(d)]** and **[1(e)]**, nor explain how their distinct architectures would be meaningfully integrated to perform that operation. 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 explain how the disparate systems of Keshet, [King](#), and [Cuturi](#) would **function together** to achieve the claimed **differentiable alignment by jointly maximizing cosine distance between phonetic embedding vectors as required by element [1(d)], and to use that alignment to generate output audio as required by element [1(e)]**. Nor does it address the technical incompatibilities between these three references. These omissions are fatal to establishing 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. Ex2033 (Enforcement, Non Waiver, and Enforcement and Non Waiver Memo) at 1.

Here, Petitioner does not identify **any** prior art reference that even purports to disclose the claimed alignment determined by jointly maximizing cosine distance between phonetic embedding vectors (element [1(d)], [8(c)], [15(c)]) or the subsequent alignment-based generation of output audio (element [1(e)], [8(d)], [15(d)]). (*Cf.* [1(d)], [1(e)], [8(c)], [8(d)], [15(c)], [15(d)], *with* Pet. at 18-26, 29, 30). Instead, it attempts to cobble together that limitation from separate and disparate references—Keshet, [King](#), and [Cuturi](#)—without demonstrating that the combined system makes any sense, let alone teaches the claimed operation as a whole. This piecemeal reconstruction does not satisfy the particularity requirement.

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. 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 alignment nor explains how the proposed combination performs that limitation. *Cf.* Pet. at 18-26, 29, 30.

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

Finally, denial is warranted where 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 (May 26, 2026) at 2–3 (citing *Chevron Oronite Co. v. Infineum USA*, IPR2018-00923, Paper 9 (November 7, 2018) at 10–11 (informative)).

Here, all asserted grounds depend on the same defective Keshet–~~King~~–Cuturi combination, and the additional references in Grounds 2–5 are not even offered to cure and do not cure these deficiencies in Ground 1 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 five obviousness grounds:

<u>Ground</u>	<u>Claims Challenged</u>	<u>References Relied Upon</u>	<u>Petition Citation</u>
Ground 1	1–2, 4, 8–9, 11, 15–17	Keshet, King , Cuturi	Pet. at 10
Ground 2	3, 10	Keshet, King , Cuturi , Kameoka	Pet. at 31
Ground 3	5, 12, 18	Keshet, King , Cuturi , Fan	Pet. at 36
Ground 4	6, 13, 19	Keshet, King , Cuturi , Chen	Pet. at 38
Ground 5	7, 14, 20	Keshet, King , Cuturi , Vyas	Pet. at 40

This Preliminary Response focuses on Ground 1 because Ground 1 supplies the base combination that every other asserted ground, and every claim, requires. Ground 1 asserts that claims 1-2, 4, 8-9, 11, and 15-17 are obvious over Keshet, [King](#), and [Cuturi](#). Pet. at 10. Grounds 2-5 merely add Kameoka, Fan, Chen, or Vyas for dependent-claim limitations, while retaining the same Keshet-[King](#)-[Cuturi](#) combination for the independent-claim limitations. Pet. at 31, 36, 38, 40. Thus, if

Ground 1 fails to show the claimed alignment in independent claims 1, 8, and 15, Grounds 2-5 fail as well.

A. Ground 1 Fails on the Merits

The novelty identified by the Patent Office is embodied in claim elements **[1(d)]** and **[1(e)]**, which the Petition relies on and simply cites back to—without further analysis—for the remaining independent claims, including claims 8 and 15.

Claim **[1(d)]** requires “**determin[ing] a differentiable alignment by jointly maximizing a cosine distance between the first phonetic embedding vectors and the second phonetic embedding vectors.**” Claim **[1(e)]** then requires aligning speech data to phonetic content “**based on the differentiable alignment.**” ’745 Patent at Cl. 1.

Petitioner’s Ground 1 combination does not satisfy these requirements. Petitioner does not identify any prior-art teaching the claimed alignment in elements **[1(d)]** and **[1(e)]**, nor explain how the asserted references cure the deficiency identified by the Patent Office in allowing the claims.

Petitioner’s theory is not that any reference teaches the claimed alignment. Instead, Petitioner divides the limitation across three references. Petitioner first relies on Keshet for the alleged first and second phonetic embedding vectors, asserting that Keshet generates vectors from input speech and applies an ML model to produce output vectors. Pet. at 12-18. Petitioner then turns to **King** for the cosine-distance

component, arguing that **King**'s time-warping module aligns features by calculating a similarity matrix using cosine distance. Pet. at 19-23. Finally, Petitioner invokes Cuturi for differentiability, contending that Cuturi supplies a differentiable alignment technique through soft-DTW. Pet. at 23-25. As Petitioner summarizes its own theory, "Keshet does not expressly provide all of the details for maximizing the cosine distance between the first PEVs and second PEVs or utilizing a differentiable alignment," so Petitioner uses **King** for the cosine-distance alignment and Cuturi for differentiability. Pet. at 19-25.

That division of labor exposes the hindsight. The Patent Office allowed the claims because the prior art failed to disclose the claimed cosine-distance alignment. Petitioner does not identify that Examiner-recognized advance in any single reference. Instead, it attempts to manufacture the limitation by assigning one piece to Keshet, another to **King**, and another to Cuturi. But obviousness requires more than locating isolated words in separate references. Petitioner had to show a coherent combined system that determines **a differentiable alignment by jointly maximizing cosine distance between first and second phonetic embedding vectors**. It did not.

The Petition's own descriptions confirm the incapability. Petitioner relies on Keshet for phonetic embeddings and output audio, but Keshet is directed to local phoneme correction. The Petition states that Keshet identifies and replaces

mispronounced phonemes and that its ML model generates corrected speech for a selected phoneme segment. Pet. at 4-5, 12-18. That is not a sequence-alignment system. Petitioner then relies on [King](#), but [King](#) is an acoustic-frame alignment reference. Petitioner quotes [King](#) as comparing “**every frame of the overdub sound data 114 to every frame of the reference sound data 112**” using a normalized inner product, such as cosine distance. Pet. at 20; Ex1005 ¶[0038]. And Petitioner admits that [King](#)’s time-warping module uses that matrix to align sound data by changing timing by stretching or compressing frames. Pet. at 20-21. That is acoustic-frame DTW alignment of recordings, not differentiable alignment of phonetic embedding vectors.

[Cuturi](#) does not cure that defect. Petitioner relies on [Cuturi](#) because it allegedly teaches “differentiable alignment.” Pet. at 23-25. But [Cuturi](#) is a soft-DTW loss-function paper. Petitioner itself describes [Cuturi](#) as teaching techniques for differentiable alignment but cites [Cuturi](#) for soft-DTW only, excluding the further limitation of maximizing cosine distance between phonetic embedding vectors. *Id.* [Cuturi](#) therefore supplies, at most, differentiability for a DTW-style time-series loss. It does not transform Keshet’s local phoneme-substitution system or [King](#)’s acoustic-frame DTW matrix into the claimed differentiable cosine-distance alignment.

Petitioner's stitched-together theory never shows that combined operation and instead confirms that no single reference—or any coherent combination—teaches the claimed alignment.

Thus, Ground 1 is an improper aggregation. Keshet allegedly supplies the vectors, **King** allegedly supplies cosine distance, and Cuturi allegedly supplies differentiability. But none of those references—alone or in any coherent combination—teaches or even suggests the claimed alignment operation. The claim requires one coherent process with three linked features: the objects being aligned are the first and second phonetic embedding vectors; the mathematical basis is joint maximization of cosine distance; and the resulting alignment is differentiable. Petitioner's stitched-together theory never shows that combined operation. Petitioner's theory does not describe a working system; it merely assigns disparate concepts from unrelated references to different words in the claim.

This is not a mere implementation-detail gap. It is the limitation the Office identified as missing from the prior art and the limitation Petitioner now attempts to cobble together with no less than three disparate references with hindsight. Keshet repairs a selected phoneme segment; **King** warps acoustic frames between reference and overdub recordings; Cuturi supplies a differentiable DTW loss. Those are different data, objectives, and architectures, and the Petition does not explain how they would be combined to perform a differentiable alignment that jointly

maximizes cosine distance and then uses that alignment to generate output audio. Ground 1 therefore fails because Petitioner never shows that the proposed Keshet-King-Cuturi combination determines the claimed differentiable alignment between first and second phonetic embedding vectors.

Indeed, Petitioner’s theory is not a combination of known components performing their known functions; it requires modifying each reference to perform a different function, on different data, for a different purpose—and still does not yield the claimed differentiable alignment by jointly maximizing cosine distance between phonetic embedding vectors.

By way of illustration, replacing Cuturi’s cost function with cosine distance would not yield the claimed method. Cuturi’s framework computes a soft-minimum over alignment paths in a DTW geometry; substituting a different pairwise cost does not convert that optimization into “jointly maximizing a cosine distance” between first and second phonetic embedding vectors. It remains a DTW-based alignment objective, not the claimed direct optimization over embedding vectors.

1. Ground 1 Independently Fails Because Keshet, King, and Cuturi Are Not Analogous Art

Petitioner’s Ground 1 fails at the threshold because its asserted references—Keshet (Ex1004), King (Ex1005), and Cuturi (Ex1006)—are not analogous art. A reference is analogous only if it is from the same field of endeavor or is reasonably

pertinent to the problem faced by the inventor. *In re Bigio*, 381 F.3d 1320, 1325 (Fed. Cir. 2004) (“Two separate tests define the scope of analogous prior art: (1) whether the art is from the same field of endeavor, regardless of the problem addressed and, (2) if the reference is not within the field of the inventor’s endeavor, whether the reference still is reasonably pertinent to the particular problem with which the inventor is involved”). Neither Keshet, King, nor Cuturi satisfies either requirement.

i. Keshet (Ex1004)

Keshet is not in the same field of endeavor. The ’745 Patent is directed to real-time accent conversion using phonetic embedding vectors and alignment between source- and target-accent representations. Ex1001 Abstract. Keshet, by contrast, addresses speech-therapy-style mispronunciation correction, specifically “machine learning-based speech production correction stimuli for training patients suffering from speech disorders”. Ex1004, [0047].

Keshet operates at the level of localized phoneme repair, not transformation of phonetic characteristics across an utterance. It processes acoustic feature vectors, such as representations where “each feature vector x_t may include or may represent voice characteristics such as frequency bins of the Mel-spectrum or mel-cepstrum of the recorded speech of the speaker” (Ex1004, [0082]), identifies a target segment by

isolating “a segment of interest comprising the k -th phoneme p_k and its immediate vicinity” (Ex1004, [0087]), and generates a corrected output for that localized region, where the model produces “a speech segment of the same duration T ” corresponding to that localized region (Ex1004, [0092]–[0093]).

Keshet does not transform an entire speech signal to match a target accent or learn mappings between accents.

Keshet is also not reasonably pertinent to the claimed problem. The ’745 Patent addresses aligning phonetic representations between source and target accents to enable real-time accent conversion. Ex1001 Abstract. Keshet instead evaluates phoneme correctness and replaces isolated mispronunciations, for example by training a model “to predict whether a reconstructed phoneme is similar to the target phoneme p^* ” and to “recover a phoneme p_k which is similar to the desired phoneme p^* ” (Ex1004, [0097]–[0099]).

That localized repair problem is fundamentally different from sequence-level alignment across accents. A PHOSITA would not look to phoneme-repair systems to solve accent-conversion alignment. Petitioner’s reliance on Keshet reflects hindsight.

Because Keshet is not analogous art, it cannot serve as a proper primary reference.

ii. **King (Ex1005)**

King likewise is not analogous art.

King is not in the same field of endeavor. The '745 Patent is directed to real-time accent conversion using phonetic embedding vectors and alignment between source- and target-accent representations. Ex1001 Abstract. King instead relates to sound-processing techniques such as “temporally align[ing] sound data, such as an overdub recording to a reference recording,” which “may be utilized to support a variety of functionality,” including replacing sound data with other sounds—for example, to “utilize a foreign overdub to replace the sound data with dialog in a different language.” Ex1005, King, [0012], [0001].

King is also not reasonably pertinent. The '745 Patent addresses aligning phonetic representations between source and target accents to enable real-time accent conversion. Ex1001 Abstract. King’s objective is timing and perceptual synchronization between recordings—addressing how conventional alignment techniques may “sound unnatural” due to “stretching and/or compressing of portions of the audio data to align the corresponding portions to each other,” resulting in differences “perceptible to a user” (Ex1005, King, [0002])—not determining an alignment between phonetic embedding vectors or transforming speech characteristics across accents.

iii. Cuturi (Ex1006)

Cuturi is also not analogous art.

Cuturi is not in the same field of endeavor. The '745 Patent is directed to real-time accent conversion using phonetic embedding vectors and alignment between source- and target-accent representations. Ex1001 Abstract. Cuturi is a general-purpose time-series learning paper that introduces a differentiable soft-DTW loss. Ex1006 at 1–3. It does not disclose accent conversion, phonetic embedding vectors tied to source and target accents, or alignment used to generate output accented speech.

Cuturi is also not reasonably pertinent. The '745 Patent addresses aligning phonetic representations between source and target accents to enable real-time accent conversion. Ex1001 Abstract. Cuturi instead addresses a different problem: defining “a differentiable learning loss between time series” that “build[s] upon” dynamic time warping. Ex1006 Abstract. Its focus is on providing a loss function for generic time-series comparison, not determining an alignment between phonetic embedding vectors or transforming speech characteristics across accents.

In sum, none of Keshet, King, or Cuturi is directed to the same field of endeavor as the '745 Patent, and none is reasonably pertinent to the specific problem

of aligning phonetic representations between source and target accents for real-time accent conversion.

2. The Ground 1 Combination Is Mathematically and Architecturally Incoherent and Still Does Not Yield the Claimed Operation

The Petition's proposed Ground 1 combination fails at a foundational level because it assumes compatibility among incompatible data, objectives, and architectures. Significantly, Petitioner's combination would change the operation of each reference and still would not produce the claimed alignment.

Keshet's relevant vectors are local speech-feature vectors and acoustic phoneme embeddings used for phoneme correction. Keshet's D-dimensional feature-vector sequence represents voice characteristics such as Mel-spectrum or mel-cepstrum bins. Ex1004 ¶¶[0081]-[0082]. Keshet's ML output is a sequence of vectors with the same representation as x_k , such as Mel-spectrum bins, and the same duration as the local input segment. *Id.* ¶¶[0092]-[0093]. Keshet's acoustic phoneme embedding, by contrast, is a fixed-size vector used to evaluate whether two acoustic inputs correspond to the same phoneme. *Id.* at [0100].

King's vectors are acoustic-frame features of reference and overdub recordings, such as spectral characteristics and temporal features. Ex1005 ¶¶[0030]-[0033]. **King's** similarity matrix compares every frame of overdub sound data to

every frame of reference sound data using spectral or spectrally-derived features. *Id.* ¶[0038].

[Cuturi](#)'s soft-DTW is based on a pairwise cost matrix between time-series samples, generally using squared Euclidean distance. Ex1006 at 2.

Petitioner does not reconcile these differences in the combined Ground 1 system. If Petitioner uses Keshet's Mel-spectrum vectors, those are not the claimed phonetic embedding vectors and [King](#)'s acoustic-frame DTW does not become phonetic-embedding alignment. If Petitioner uses Keshet's fixed-size acoustic phoneme embeddings, those embeddings are designed to compare two speech segments for phoneme identity and are explicitly independent of input length. Ex1004 ¶[0100]. Such fixed-size embeddings do not supply the time-ordered frame sequences required for [King](#)'s similarity matrix or [Cuturi](#)'s time-series alignment. If Petitioner uses [King](#)'s spectral frame features, those are not Keshet's phonetic embeddings and not the claimed first and second phonetic embedding vectors.

The mathematical operation used for vector alignment also changes across references. Keshet uses cosine similarity or cosine distance in a loss function to evaluate phoneme identity. Ex1004 ¶¶[0100]-[0101], [0147]-[0148]. [King](#) uses cosine distance as an element-wise similarity value in a frame-by-frame matrix for DTW. Ex1005 ¶¶[0036]-[0038]. [Cuturi](#) uses a soft-minimum over alignment costs in a DTW geometry, with costs typically based on squared Euclidean distance.

Ex1006 at 1-3. Petitioner provides no reasoned explanation for how the architecture which relies on these different mathematical operations are combined to achieve the claimed “jointly maximizing a cosine distance” between first and second phonetic embedding vectors. Ex1001 at Cl. 1.

This is not a routine substitution of known components. It is a purported reconstruction that changes the purpose and operation of each reference. Keshet’s purpose is to provide speech-production correction feedback by substituting a corrected phoneme segment. Ex1004 ¶¶[0057]-[0069]. [King](#)’s purpose is to align overdub and reference recordings in a way that sounds natural to a listener. Ex1005 ¶¶[0001]-[0003], [0012]-[0013]. [Cuturi](#)’s purpose is to provide a differentiable loss for time-series learning under DTW geometry. Ex1006 at 1-2. The Petition does not show that a PHOSITA would combine those systems to arrive at the claimed accent-conversion alignment.

Petitioner’s theory improperly collapses distinct representations into “vectors,” but the claimed method requires alignment of phonetic embedding vectors corresponding across source and target speech, not arbitrary feature vectors or acoustic frames. Changing the representation to enable Petitioner’s mapping would alter the operation of each reference and eliminate the very structures Petitioner relies upon.

Petitioner's theory does not describe a working system; it merely assigns disparate concepts to claim language and, even if implemented, would not determine a differentiable alignment by jointly maximizing cosine distance between first and second phonetic embedding vectors.

In sum, the proposed combination requires rewriting each reference's core operation, not merely combining them—and even then does not yield the claimed method.

3. Ground 1 Also Fails Because Petitioner's Motivation-to-Combine Theory Is Hindsight-Driven and Lacks a Reasonable Expectation of Success

Petitioner argues that a PHOSITA would have been motivated to use [King's](#) time-warping module and sound-feature alignment techniques in Keshet to improve alignment, and then use [Cuturi's](#) differentiable alignment to improve speed, efficiency, and accuracy. Pet. at 21-25. But that reasoning abstracts away the actual disclosures.

Keshet already solves a different problem than accent-correction without sequence alignment. It identifies a local mispronounced phoneme, masks the local region, uses surrounding context to generate a corrected local segment, and splices that segment into the original speech. Ex1004 ¶¶[0087]-[0093], [0106]-[0108], [0133]-[0137]. [King's](#) DTW system, by contrast, aligns entire or extended reference and overdub recordings by stretching and compressing frames of the recordings.

Ex1005 at [0034]-[0043]. There is no reasoned explanation why a PHOSITA would replace Keshet's local mask-and-fill architecture with **King**'s frame-level recording alignment.

The Petition also fails to establish reasonable expectation of success. Petitioner does not explain how the combined system would generate compatible first and second phonetic embedding vector sequences; how **King**'s similarity matrix would operate on those vectors rather than on acoustic/spectral frames; how Cuturi's soft-DTW would be incorporated while preserving the claimed "jointly maximizing a cosine distance"; or how the resulting system would align speech data to phonetic content as claimed. Simply asserting that software could be reprogrammed does not satisfy Petitioner's burden.

Further, Petitioner's proposed modifications would also defeat the purpose of the cited references. Keshet's localized phoneme-replacement architecture does not require sequence alignment; King's DTW alignment operates on acoustic frames, not phonetic embeddings; and Cuturi's objective depends on DTW structure. Modifying each to serve Petitioner's theory would discard their intended operation and replace it with a different system altogether.

In short, Petitioner has not shown a reasoned path from the Ground 1 prior art to the claimed invention. Petitioner's mapping is built from the claim language backward: Keshet is selected for "phonetic embeddings," **King** for "cosine

distance,” and Cuturi for “differentiable alignment.” That is hindsight, not obviousness.

4. Even If the Board Does Not Treat Ground 1 as Dispositive, Grounds 2-5 Independently Fail Because the Additional References Do Not Cure the Base-Combination Defect

As explained above, Grounds 2-5 should fail with Ground 1 because they depend on the same Keshet-King-Cuturi base combination. But even if the Board separately considers Grounds 2-5, those grounds still fail because Petitioner does not rely on Kameoka, Fan, Chen, or Vyas to cure the absence of the claimed cosine-distance alignment in elements [1(d)] and [1(e)], [8(c)] and [8(d)], and [15(c)] and [15(d)] as identified by the Office.

Ground 2 does not offer any fix for the failures of Ground 1 discussed above. In any event, ground 2 adds Kameoka for claims 3 and 10. *See* Pet. at 31-36. Petitioner relies on Kameoka for training details, including alleged labeled datasets and encoder/decoder layers. *Id.* But claim 3 depends from claim 1 and claim 10 depends from claim 8. Kameoka does not supply the missing limitation of determining a differentiable alignment by jointly maximizing cosine distance between first and second phonetic embedding vectors. Nor does Petitioner contend that Kameoka replaces the Keshet-King-Cuturi mapping for claim elements [1(d)], [1(e)], or the corresponding method limitations. Accordingly, Ground 2 fails and

therefore does not cure the fundamental defect in the Keshet-[King-Cuturi](#) base combination identified above.

Ground 3 does not offer any fix for the failures of Ground 1 discussed above. In any event, ground 3 adds Fan for claims 5, 12, and 18. *See* Pet. at 36-38. Petitioner relies on Fan for output-generation details such as prosody modeling, intonation adjustment, accent-specific acoustic modeling, speech rate, pitch, or gender. *Id.* Those features come after the independent-claim alignment problem. Fan does not remedy the absence of the claimed differentiable alignment in the base combination. Ground 3 therefore fails and therefore does not cure the fundamental defect in the Keshet-[King-Cuturi](#) base combination identified above.

Ground 4 does not offer any fix for the failures of Ground 1 discussed above. In any event, ground 4 adds Chen for claims 6, 13, and 19. *See* Pet. at 38-40. Petitioner relies on Chen for gradient-based optimization. *Id.* But adding a reference that teaches gradient-based optimization in another context does not establish that the Keshet-[King-Cuturi](#) base combination determines a differentiable alignment by jointly maximizing cosine distance between first and second phonetic embedding vectors. Chen does not supply the missing claimed alignment, does not make [King](#)'s acoustic-frame DTW into phonetic-embedding alignment, and does not cure Keshet's local phoneme-substitution architecture. Ground 4 therefore fails and

therefore does not cure the fundamental defect in the Keshet-**King**-Cuturi base combination identified above.

Ground 5 does not offer any fix for the failures of Ground 1 discussed above. In any event, ground 5 adds Vyas for claims 7, 14, and 20. *See* Pet. at 40-42. Petitioner relies on Vyas for dimensionality reduction. *Id.* Dimensionality reduction is irrelevant to the dispositive defect in Ground 1. Reducing the dimensionality of vectors does not show that the base combination uses the claimed first and second phonetic embedding vectors, determines the claimed differentiable alignment, or jointly maximizes cosine distance between those vectors. Ground 5 therefore fails and therefore does not cure the fundamental defect in the Keshet-**King**-Cuturi base combination identified above.

Thus, regardless of whether the Board treats Ground 1 as dispositive, the Petition does not establish a reasonable likelihood of prevailing on any challenged claim.

VI. CONCLUSION

Ground 1 fails because Petitioner does not show—consistent with the Patent Office’s allowance determination—that any prior art discloses or suggests the claimed alignment in elements **[1(d)]** and **[1(e)]**, **[8(c)]** and **[8(d)]**, and **[15(c)]** and **[15(d)]**. Further, Petitioner does not show a supported motivation to combine Keshet, [King](#), and [Cuturi](#) or a reasonable expectation of success and does not satisfy its burden to show with particularity where the claimed alignment in elements **[1(d)]** and **[1(e)]**, **[8(c)]** and **[8(d)]**, and **[15(c)]** and **[15(d)]** is disclosed.

Because Grounds 2–5 depend on Ground 1’s defective base combination, those grounds fail as well. Furthermore, Grounds 2–5 each fail for defective combinations.

For the foregoing reasons, Patent Owner respectfully requests that the Board deny institution of inter partes review of Claims 1–20 of the '745 Patent.

Respectfully submitted,

AMSTER, ROTHSTEIN & EBENSTEIN
LLP

Attorneys for Patent Owner
405 Lexington Avenue
New York, NY 10174
(212)336-8000

Dated: June 25, 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

The undersigned certifies that the foregoing PATENT OWNER'S 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 6,720 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 25, 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 25, 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 25, 2026
New York, New York

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