

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC

Petitioner,

v.

CELLULAR SOUTH, INC.

Patent Owner.

Case No. IPR2025-00875

U.S. Patent No. 9,940,972

PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I. INTRODUCTION1

II. BACKGROUND3

III. THE '972 PATENT5

IV. OVERVIEW OF THE CITED PRIOR ART14

 A. OVERVIEW OF FONTANA14

 B. OVERVIEW OF LAU.....18

V. LEVEL OF ORDINARY SKILL IN THE ART21

VI. CLAIM CONSTRUCTION21

VII. THE PETITION FAILS TO SHOW THAT THERE IS A REASONABLE
LIKELIHOOD THAT ANY CLAIM IS UNPATENTABLE UNDER
GROUND 122

 A. FONTANA IN VIEW OF LAU FAILS TO RENDER OBVIOUS “GENERATING A
TOPICAL META-DATA THAT DESCRIBES CONTENT OF THE VIDEO BY
DERIVING SEMANTIC INFORMATION FROM THE IDENTIFICATION OF THE
ONE OR MORE OBJECTS AND SEMANTIC INFORMATION FROM THE AUDIO
FILES” (CLAIM 1[F] AND CLAIM 17[G])22

 B. FONTANA IN VIEW OF LAU FAILS TO RENDER OBVIOUS “CROSS-
REFERENCING THE TEXT AND THE VIDEO DATA BASED ON THE GENERATED
TOPICAL META-DATA TO DETERMINE TOPICS” (CLAIM 1[H] AND CLAIM
17[I])26

 C. FONTANA IN VIEW OF LAU FAILS TO RENDER OBVIOUS “GENERATING
VIDEO TEXT BASED ON THE CROSS-REFERENCING, WHEREIN THE VIDEO
TEXT DESCRIBES CONTENT OF THE VIDEO” (CLAIM 1[I])32

 D. FONTANA IN VIEW OF LAU FAILS TO RENDER OBVIOUS “GENERATING A
TEXT, IMAGE, OR ANIMATION BASED ON THE VIDEO TEXT” (CLAIM 1[J]) 35

 E. FONTANA IN VIEW OF LAU FAILS TO RENDER OBVIOUS “PLACING THE
TEXT, IMAGE, OR ANIMATION IN THE VIDEO.” (CLAIM 1[K] AND CLAIM
17[K])37

 F. A PERSON OF ORDINARY SKILL IN THE ART WOULD NOT BE MOTIVATED
TO COMBINE FONTANA AND LAU38

 G. HOUH’S EXPERT TESTIMONY SHOULD BE ENTITLED TO LITTLE OR NO
WEIGHT40

VIII. GROUND 2 FAILS FOR THE SAME REASONS AS GROUND 141

IX. CONCLUSION.....42

X. CERTIFICATE OF SERVICE43
XI. CERTIFICATE OF COMPLIANCE WITH 37 C.F.R. § 42.2444

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Cellular South, Inc. v. Google, LLC</i> , Case No. 4:25-cv-01487-YGR (N.D. Cal. Aug. 5, 2025)	iv
<i>Cellular South, Inc. v. Google, LLC</i> , Case No. 6:24-cv-00245 (W.D. Tex. May 9, 2024)	iv
<i>Kinetic Techs., Inc., Petitioner, v. Skyworks Sols., Inc.</i> , IPR2014-00529, 2014 WL 4787238, at *10 (P.T.A.B. Sept. 23, 2014)	25, 41
<i>Nautilus Hyosung Inc. v. Diebold Self-service Systems division of Diebold</i> , IPR2016-00633, Paper 9 at 21 (P.T.A.B. Aug. 22, 2016).....	31
<i>TQ Delta, LLC v. Cisco Sys.</i> , 942 F.3d 1352, 1360 (Fed. Cir. 2019)	32
<i>Upjohn Co. v. Mova Pharm. Corp.</i> , 225 F.3d 1306, 1311 (Fed. Cir. 2000)	25, 41
<i>Wellman, Inc. v. Eastman Chem. Co.</i> , 642 F.3d 1355, 1361 (Fed. Cir. 2011)	21, 41
<i>Xerox Corp. v. Bytemark, Inc.</i> , IPR2022-00624, 2022 WL 3648989, at *6 (P.T.A.B. Aug. 24, 2022).....	26, 41
Statutes	
37 C.F.R. § 42.65(a).....	25, 40

EXHIBIT LIST

Exhibit No.	Document
2001	U.S. Patent No. 9,940,972
2002	U.S. Patent No. 10,218,954
2003	U.S. Patent No. 11,126,853
2004	Article, Vū Digital Wins Coveted TMC Communications Solutions 2015 Product of the Year Award
2005	Article, NAB Show Unveils Final 10 Participants Selected for SPROCKIT 2015
2006	Notice of References Cited in U.S. Patent Application No. 15/444,054 (Oct. 30, 2019)
2007	Information Disclosure Statement by Applicant in U.S. Patent Application No. 17/423,623 (July 20, 2021)
2008	CSI's Complaint for Patent Infringement against Google in <i>Cellular South, Inc. v. Google, LLC</i> , Case No. 6:24-cv-00245 (W.D. Tex. May 9, 2024)
2009	Proof of Service of Complaint
2010	Joint Case Management and Pretrial Order as Modified in <i>Cellular South, Inc. v. Google, LLC</i> , Case No. 4:25-cv-01487-YGR (N.D. Cal. Aug. 5, 2025)
2011	Transcript of Conference held July 7, 2025 in <i>Cellular South, Inc. v. Google, LLC</i> , Case No. 4:25-cv-01487-YGR (N.D. Cal. Aug. 5, 2025)
2012	Preliminary Infringement Contentions in <i>Cellular South, Inc. v. Google, LLC</i> , Case No. 6:24-cv-00245 (W.D. Tex.) (served Dec. 13, 2024)

Exhibit No.	Document
2013	Comparison of portions of Petition with Ex. 1002

I. INTRODUCTION

Pursuant to 35 U.S.C. § 313 and 37 C.F.R. § 42.107, Patent Owner Cellular South, Inc. (“Patent Owner” or “CSI”) submits this Preliminary Response to the Petition for *Inter Partes* Review (“Petition”) filed by Google LLC (“Petitioner” or “Google”) and challenging claims 1-11 of U.S. Patent No. 9,940,972 (the “’972 Patent”).

Petitioner's case fails at the threshold. The asserted prior art does not teach, suggest, or even contemplate the '972 Patent's cross-referencing semantic information from both audio and video streams of the same video using topical meta-data. This feature—cross-referencing audio text and video data from the same video to determine topics using topical meta-data generated from semantic information in both streams—is absent from both references. The combination posited by the Petition using Fontana (Ex. 1003) and Lau (Ex. 1004) does not merely fail to teach this claim element—it sets the stage for compounding deficiencies. With each successive limitation, the disconnect between the prior art and the '972 Patent grows, as the references lack not only the foundational features but also the advanced functionalities that build upon them. The result is a cascade of failures, each more pronounced than the last. The Petition's attempt to manufacture an obviousness theory is nothing more than hindsight-driven speculation, not a legitimate motivation to combine.

The '972 Patent claims a specific, innovative solution to the longstanding problem of extracting meaningful, semantically-rich data from video content. It does so by generating topical meta-data from both audio text (converted from audio files) and video data (converted from image files) of the same video, and then cross-referencing these to determine the actual topics depicted. This process enables a context description of the video that neither Fontana nor Lau, alone or in combination, comes close to achieving. Fontana merely processes video into text and object metadata, while Lau matches keywords between unrelated content and advertisements. Neither reference, and certainly not their combination, provides for cross-referencing audio text and video data from the same video using topical meta-data.

To bridge this gap, Petitioner must twist and misrepresent the prior art, stitching together disparate teachings that do not—and cannot—address the claimed invention. This is classic hindsight, not a legitimate rationale for obviousness.

Institution is not warranted where the Petition relies on speculation and hindsight, rather than a clear teaching or motivation in the prior art. Petitioner has not met its burden under 35 U.S.C. § 314(a): there is no reasonable likelihood that Petitioner would prevail on any challenged claim. The Board should deny institution.

II. BACKGROUND

Patent Owner is a diversified telecommunications and technology company that focuses on delivering connectivity and technology solutions to advance the field and its community. While Patent Owner is a relatively small company compared to its competitors in the telecommunications industry, it has always been an innovator. For example, Patent Owner has worked to diversify its product offerings to remain innovative and competitive in the telecommunications and technology marketplace. Through one of these efforts, Patent Owner, through its subsidiary Vu Digital, created and offered “Video-to-Data” (or “V2D”), a cloud-based digital profiling and analytics system that could scan and turn video into searchable data faster than the real-time viewing of the video by using multiple servers that run simultaneously and in parallel.

V2D breaks down a video frame-by-frame and identifies the objects—for example, text, audio, images, faces, locations, logos—within each frame. V2D then creates a chronological record of all the objects identified within the video. V2D was a transformative product that made video content as easy to search as text, thereby providing users with unprecedented video classification and clustering capabilities, as well as significantly enhanced search engine indexing, content personalization, and targeted advertising capabilities.

The applications for V2D are far reaching. V2D could be an incredibly useful technology for video data processing and content personalization. For example, V2D could be used by law enforcement to process body camera footage which was becoming more widely adopted.

Patent Owner launched its V2D product in May 2015. Mere months later, Patent Owner's V2D product was recognized by TCM—a global integrated media conglomerate and leading source of news and information for the communications and technology industries worldwide—as the “best-of-the-best technology solutions available on the market today.” V2D received TCM's coveted Communications Solutions 2015 Product of the Year Award that recognizes exceptional voice, data, and video communications products and services. (Ex. 2004).

Patent Owner was also selected in 2015 to be part of the National Association of Broadcasters' year-long SPROCKIT accelerator program for innovative startups in media and entertainment. (Ex. 2005). SPROCKIT is a global innovation platform created to help large media, entertainment, and technology companies meet emerging tech start-ups to fast-track investment, acquisition, and partnerships between the two groups on new products and services. Petitioner is a corporate partner of SPROCKIT. *Id.*

SPROCKIT Sync is an invitation-only event put on three times a year by SPROCKIT that includes a series of private meetings between selected startups and

corporate players in the media and entertainment industry. Following the launch of its V2D product, Patent Owner was invited to attend the SPROCKIT Sync conference that was held on June 18, 2015 at Google Tech Corners in Sunnyvale, California. There, Patent Owner presented its V2D technology at the conference. Google was one of the corporate partners for this event.

Patent Owner protects its intellectual property. As is relevant here, Patent Owner is the assignee of three patents related to its V2D technology including the '972 patent at issue in this IPR, as well as U.S. Patent Nos. 10,218,954 ("the '954 patent") (Ex. 2002) and 11,126,853 ("the '853 patent") (Ex. 2003), each entitled "Video to data." The '954 patent and '853 patents are related and Google filed IPRs challenging these other two patents as well. *See* IPR2025-00876 (challenging the '954 patent) and IPR2025-00877 (challenging the '853 patent). At a high level, these patents are directed to an improvement in the functionality of machine learned video recognition and classification systems, particularly with regard to specific techniques for improving the accuracy of predictions made using image, audio, text, and other video data.

III. THE '972 PATENT

The '972 patent addresses technological challenges in extracting meaningful metadata from video content, which is a persistent problem in video classification

systems. Prior to the claimed invention in the '972 patent, video content was largely opaque to search engines and analytics platforms. While reverse image search and audio-to-text transcription existed, they operated as independent silos and lacked semantic, contextual understanding. The '972 patent addresses these limitations by integrating distributed processing, advanced image and audio analysis, and natural language processing to create a unified, and context-rich video data extraction engine.

The process described in the '972 Patent first receives a video file and extracts both its audio track and individual image frames. These audio and image components are then distributed across multiple processors, enabling parallel analysis for efficient, near real-time processing. The audio is segmented, and each segment is transcribed into text using speech recognition techniques. Simultaneously, the image frames are processed to identify objects, symbols, brands, faces, and other features through various image analysis methods. Both the transcribed audio and the descriptive text generated from images are further refined using natural language processing, which helps to extract context, topics, and semantic meaning, and can assign weights to identified elements based on their relevance. The information from both audio and image analyses are combined and cross-referenced, giving additional emphasis to topics or keywords that appear in both sources. This integrated data is used to generate a comprehensive textual

description of the video's content, along with metadata such as timestamps, frame references, and topical tags. The final output can include text, metadata, images, or animations that describe the video. The key is the ability to extract semantic meaning from both the audio and visual components of the same video and then cross-reference them to develop a holistic understanding of the video's content.

Figure 1 of the '972 Patent, shown below, illustrates an embodiment of a video-to-content¹ engine:

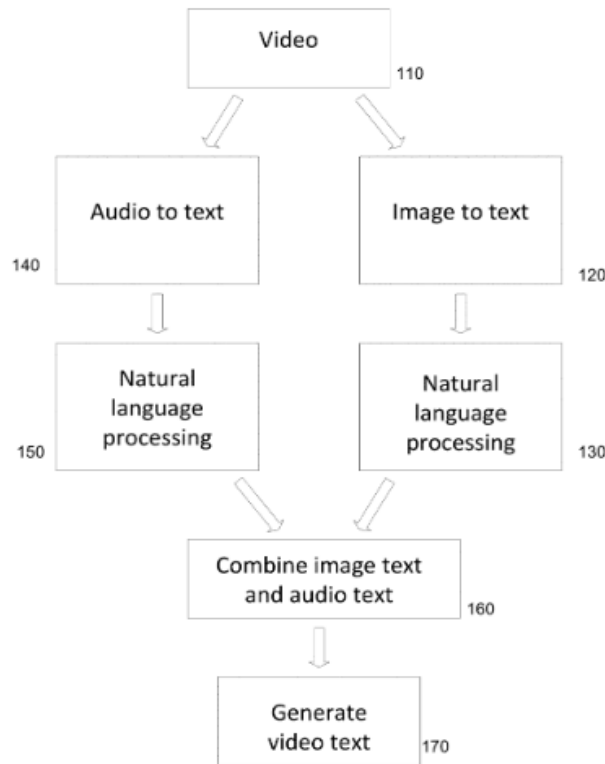


Figure 1

¹ The '972 Patent also refers to this as the “video-to-text engine”. *See, e.g.*, Ex. 1001 at 3:51; 3:64; and 5:4.

Ex. 1001 at FIG. 1.

Starting with step 110, the video-to-content engine embodied on a computer, a server, or servers, extracts audio data and image data (e.g. images or frames forming the video) from the video stream in parallel processes. Ex. 1001 at 3:51-63. At step 120, the video-to-content engine performs image data processing on the image data extracted from the video stream by distributing the image files across a plurality of processors for parallel processing. *Id.* at 3:63-4:2. An example of the distributed image data processing is illustrated in Figure 2, shown below:

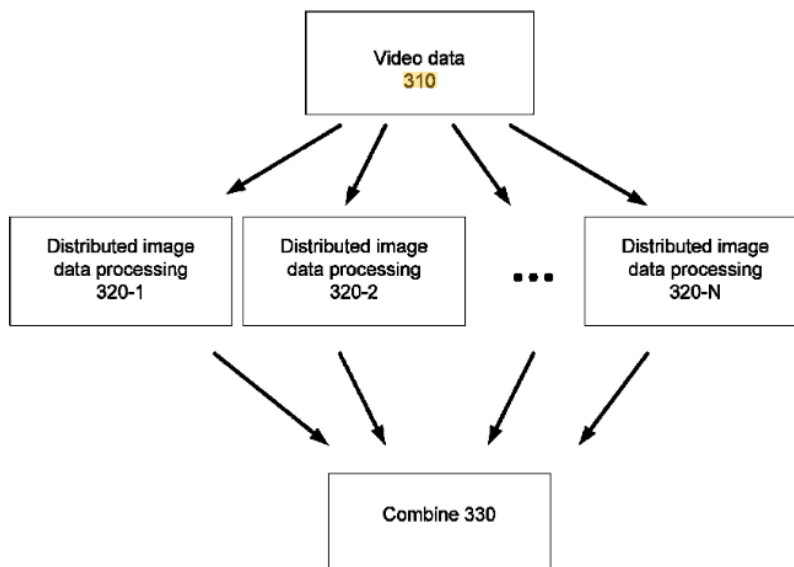


Figure 2

Ex. 1001 at FIG. 2.

At step 120 and as shown in Figure 2, the image data is segmented into image files and each image file is processed in parallel (e.g., by distributing image files to the plurality of processors for distributed processing 320-1 to 320-N). *Id.* at 3:66-

4:2. Processing the image files comprises extracting one or more objects and identifying the one or more objects. For example, using image processing techniques (such as color edge detection, image extraction, high-level vision and symbol detection, figure-ground separation, depth and motion perception, etc.) and object template database, an object or symbol of an image can be extracted, isolated, and identified in the form of text describing the object. *Id.* at 4:3-31.

At step 130, natural language processing is applied to the image text to derive semantic information from the identification of the objects. *Id.* at 4:61-5:3. For example, the image text can be modified as the video-to-text engine selects primary semantics from a plurality of possible semantics. *Id.* at 4:60-66. The video-to-content engine generates topical meta-data or topical descriptors that describe the content of the image files and applies this topical meta-data to the original video. *Id.* at 4:32-44. For example, if the object identified in the image is a dog, when compared to the topic database, the dog may be further described as running or jumping. *Id.* at 4:45-49. Importantly, because this topical meta-data generation process—using the distribute/parallel processing architecture described above—is repeated for additional symbols/objects within an image frame, the process develops granular contextual details about what's happening in each image of the video. As a result, the dog initially identified from the video frame, when compared to the topics database, might further associate the dog with the topical meta-data of

“running” and “jumping.” *Id.* at 4:42-44. The image topical metadata are stored for the entire video or each image segment resulting in several topic descriptors. From this topical meta-data, it is possible to generate a more detailed contextual representation of the video as showing a dog, which is running, on the beach, with people, by a resort. *Id.* at 4:57-5:2. In other words, the contextual topics are not just a series of images but rather explains what's happening in the image.

At step 140, the audio files generated from step 110 are processed and converted into text. *Id.* at 5:4-7. At step 150, and as with the image processing of step 130, natural language processing can be applied to the text generated from step 110 to derive semantic information from the audio files. *Id.* at 5:41-45. For example, the video-to-content engine generates topical meta-data or topical descriptors that describe the content of the audio files. *Id.* Continuing with the above example concerning topical meta-data of a dog object with actions of “running” and “jumping” on a beach identified from the image files, suppose the audio text from the video contains the following dialogue:

Person 1: “Look at Buddy go! He's chasing that frisbee like a pro.”

Person 2: “I know, he loves playing fetch on the beach. The kids are having so much fun with him today.”

Person 1: “Yeah, the waves crashing in the background make it the perfect summer day.”

In this example, the video-to-content engine could derive additional topical meta-data from the semantic information, such as: subject: dog (named Buddy); action: chasing frisbee; location: beach; activity: playing fetch; participants: kids, people (2 individuals); and environment: waves, summer.

At step 160, the text, video data, and topical meta-data are combined and cross-referenced to determine topics common to both the text of the audio files and the video data. *Id.* at 5:46-49. When the topical meta-data generated and derived from the semantic information of an image, and the topical meta-data generated and derived from the semantic information of the audio are combined, the text can be cross-referenced and topics common to both the image and the audio are determined and weighted. *Id.* By cross-referencing the topical meta-data derived from the image analysis and the audio text, the video-to-content engine can identify the common topics and themes that describe the overall content of the video, such as: common topics: dog, beach, playing/activity. The cross-referencing allows the video-to-content engine to validate and strengthen the topical meta-data, since the visual and audio components both *independently* point to a video centered around a family's dog playing at the beach. In this example, the audio text provides additional context and details that complement the visual information (e.g., identifying the dog's name, the presence of kids, and the summer beach setting) through cross-referencing.

At step 170, the video-to-content engine generates video text describing the content of the video using the cross-reference of the combined text and video data. *Id.* at 5:49-54. For example, key words indicating topic and semantic that appear in both texts can be selected or emphasized. *Id.* The video-to-content engine generates text, image, or animation based on the video text and places the generated text, image, or animation in the video stream. *Id.* at 6:21-29.

The output data from the combined image and audio generated topics is then applied in various ways to create content-rich video, enrich user experience, or user defined outputs including “search engine optimization, video categorization, recommendation engines, advertisement targeting, content personalization, analytics, etc.” *Id.* at 8:15-28.

In summary, the '972 Patent claims a method and system for distributed, parallel processing of video content, extracting and synthesizing data from both audio and image components, leveraging contextual information to improve the accuracy of object and topic identification, and generating rich metadata and content representations that make video as searchable and actionable as text. The '972 Patent's focus lies in its use of context—derived from both audio and visual data—to enable more reliable classification, identification, and personalization of video content.

Claim 1 of the '972 Patent is recited below:

A method to generate video data from a video, comprising:

generating audio files and image files from the video;

distributing the image files across a plurality of processors and processing the image files in parallel, wherein processing the image files comprises extracting one or more objects and identifying the one or more objects;

processing the audio files;

converting audio files associated with the video to text;

converting the image files associated with the video to video data;

generating a topical meta-data that describes content of the video by deriving semantic information from the identification of the one or more objects and semantic information from the audio files;

adding the topical meta-data to the video; and

cross-referencing the text and the video data based on the generated topical meta-data to determine topics;

generating video text based on the cross-referencing, wherein the video text describes content of the video;

generating a text, image, or animation based on the video text;
and

placing the text, image, or animation in the video.

Id. at Claim 1.

IV. OVERVIEW OF THE CITED PRIOR ART

A. Overview of Fontana

Fontana is directed to systems and methods for processing and delivery of multimedia content, with a particular focus on applying a container to multimedia content that enables metadata external to the content to be linked with specific portions of the content. *See* Ex. 1003 at Abstract.

The core of the claimed system is about receiving, processing, and delivering multimedia content by associating external metadata with the content, primarily to enhance playback, searchability, and interactivity for users. *See* Ex. 1003 at [0041].

Claim 1 of Fontana recites:

applying a container to multimedia content using one or more computing systems, the container defining an interface through which metadata external to the multimedia content is linked with one or more portions of the multimedia content;

constructing a set of text metadata describing an audio portion of the multimedia content;

generating a set of object metadata describing at least a portion of one or more objects appearing in the multimedia content; and

upon receiving a request for playback of the multimedia content, providing the text metadata and the object metadata associated with the container for synchronized use during playback of the multimedia content via the container.

Ex. 1003 at Claim 1.

The Fontana patent repeatedly emphasizes that the metadata is “external to the multimedia content” and is linked via a container interface, rather than being

embedded or used to generate new content from the video itself. *See, e.g.*, Ex. 1003 at Summary, [0008]-[0010], and Claims 1, 14, 23.

Fontana details the construction of text metadata from audio and object metadata from video, but the purpose is to enable enhanced search, playback, and interactivity through the containerized interface. *See* Ex. 1003 at [0087]-[0095].

FIG. 13 illustrates the claimed process of associating metadata with content and providing enhanced playback via the container. Ex. 1003 at FIG. 13; *see also* corresponding description.

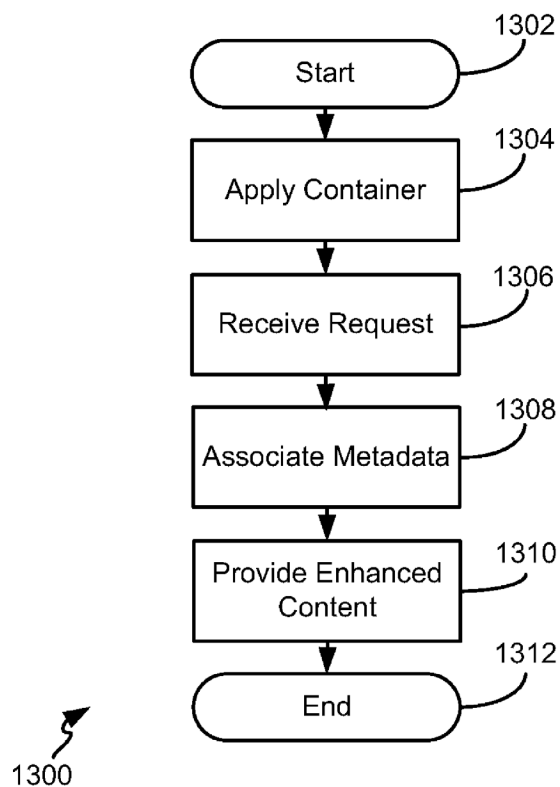


FIG. 13

Ex. 1003 at FIG. 13.

As illustrated in FIG. 13, the system applies a container to video content, receives a request and delivers metadata based on user preferences, but does not generate new content from the video data. *Id.*

Fontana’s system supports interactive features such as search, click-through events, overlays, and customized advertisements, all enabled by the external metadata and delivered in real-time or on request during playback. *See Id.* at Claims 10, 12, 18, 19; FIGS. 6, 7A-7M, 14. FIGS. 7A, 7B, 7L, and 7M are reproduced below:

Event Data 628					
Content ID	Timeline	Duration	Event	Target	Function
...					

FIG. 7A

Analytics Data - Detailed 714						
Content ID Viewed	Viewer ID Mute/Unmute	Views	Click-Through Collapse/Expand	Completed Play Pause	Time Resume	Search
Accept	Invite	Minimize	Close	Overlay View	Duration	

FIG. 7L

Text Index 630				
Content ID	Transcription Text	Start Time	End Time	Text
...				

FIG. 7B

Advertisement Data 716		
Advertiser ID	Advertisement Definition	Associated Topic/Keyword/Content
...		

FIG. 7M

Id. at FIGS. 7A, 7B, 7L, and 7M

Notably, these figures illustrate various metadata structures (e.g. event data; text index; analytics), which are used to define interactive enhancements and search capabilities. However, such enhancements do not appear to involve any semantic analysis or content generation from the video itself.

Fontana also claims methods for delivering advertisements related to objects defined in the object metadata and for capturing analytics about content usage, tightly integrated with the metadata system. *See* Ex. 1003 at Claims 10, 18, 19; and FIGS. 7L, 7M, 8.

The technical focus of Fontana is on external metadata management, distributed processing, and container-based delivery for multimedia content. Fontana's claimed systems teach linking and delivering external metadata for playback enhancement. As explained above, however, the resulting metadata in Fontana is "external to the multimedia content" and is linked to that content via a container. *See* Ex. 1003 at Claims 1, 14, 23. Fontana does **not** teach or claim the extraction of semantic context from the combination of the image files and the audio files of the same video for the purpose of generating new content or recommendations, as is central to the '972 Patent.

Fontana's focus is on metadata management and delivery, not on video-to-data transformation and contextual content generation. In the '972 Patent, the extracted data (text, objects, context) is used to generate *new* content and contextual understanding of the video itself, not merely to link external metadata for playback enhancement.

B. Overview of Lau

Lau is directed to systems and methods for matching subject matter in a portion of rich media content to an advertisement. *See* Ex. 1004 at Abstract.

Claim 1 of Lau recites:

determining when an advertisement correlates to subject matter associated with a portion of rich media content;

determining rendering preferences associated with the advertisement; and

serving the advertisement to a device, the advertisement being served such that it can be rendered relative to a time the portion of the rich media content is being played on the device based on the rendering preferences.

Ex. 1004 at Claim 1.

Lau discloses a system for serving advertisements with rich media content. *See* Ex. 1004 at [0019]. As shown in FIG. 1 of Lau (reproduced below), engine 102 provides advertisements from advertiser system 106 to user device 104 by correlating an advertisement from the advertiser system 106 with rich media content from content owner system 108.

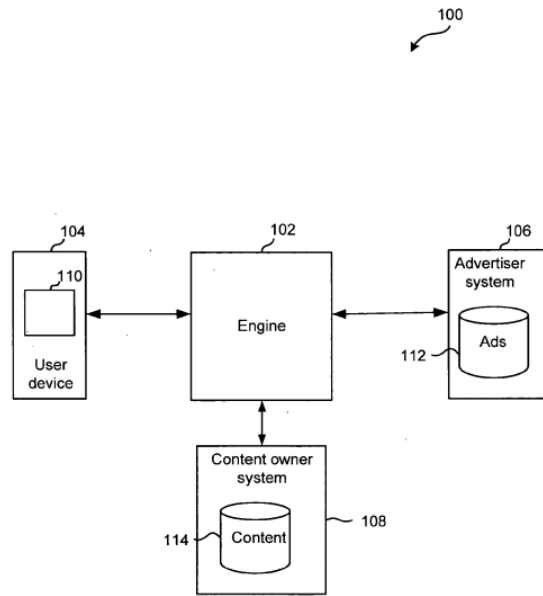


Fig. 1

Ex. 1004 at FIG. 1.

Lau describes a correlation engine that determines whether an advertisement is correlated to a portion of rich media content. *Id.* at [0040]. The rich media content is processed by generating keywords using audio recognition for audio portions of the rich media content, and image recognition is used to recognize objects. *Id.* As shown in the examples of FIGS. 4A-4B (reproduced below), advertisements are correlated to subject matter in the video across time, and, separately, advertisements are correlated to subject matter in the audio across time. *Id.* at [0102] and [0105].

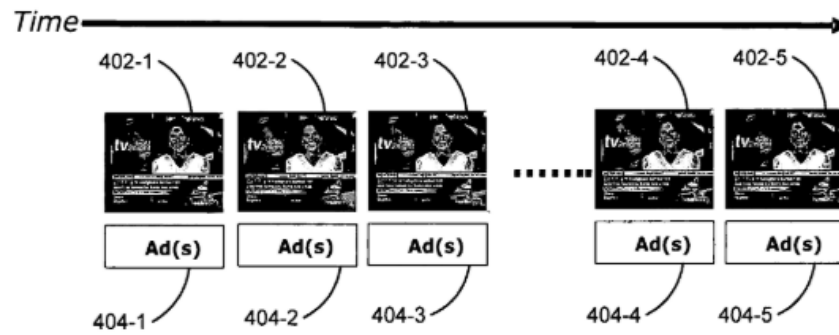


Fig. 4A

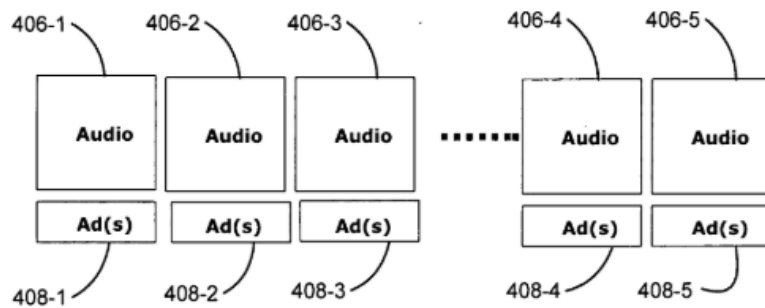


Fig. 4B

Ex. 1004 at FIGS. 4A-4B.

However, such correlations do not appear to involve any semantic analysis or cross-referencing of audio text and video data *from the same video*.

The advertisements in Lau are correlated without any reference to whether portions of the rich media content correlate *to other portions* of the rich media content. *See Id.* at [0098]-[0105]. Lau does **not** teach or claim the cross-referencing of semantic context from the combination of the image files and the audio files of the same video for the purpose of generating new content, as is central to the '972 Patent.

Lau's focus is on matching an *external* advertisement to rich media content, not on video-to-data transformation and contextual content generation using the audio and image of the same video. *See* Ex. 1004 at [0019]-[0021] and FIG. 1. In the '972 Patent, semantic meaning, extracted from both the audio and visual components of the same video, is cross-referenced internal to the components of the same video and used to automatically generate new content, not from the identification of a separate, external advertisement that correlates to rich media content based on keyword similarity as taught by Lau.

V. LEVEL OF ORDINARY SKILL IN THE ART

For the limited purpose of this Preliminary Response, Patent Owner does not contest Petitioner's definition of a person of ordinary skill in the art, but it reserves the right to do so if trial is instituted.

VI. CLAIM CONSTRUCTION

Similarly, for the limited purpose of this Preliminary Response, Patent Owner does not contend that any claim terms need construction to arrive at the conclusion that the Petition is substantively deficient. *See Wellman, Inc. v. Eastman Chem. Co.*, 642 F.3d 1355, 1361 (Fed. Cir. 2011) (terms "need only be construed 'to the extent

necessary to resolve the controversy”). Patent Owner reserves the right to construe terms if trial is instituted.

VII. THE PETITION FAILS TO SHOW THAT THERE IS A REASONABLE LIKELIHOOD THAT ANY CLAIM IS UNPATENTABLE UNDER GROUND 1

A. Fontana in View of Lau Fails to Render Obvious “generating a topical meta-data that describes content of the video by deriving semantic information from the identification of the one or more objects and semantic information from the audio files” (Claim 1[f] and Claim 17[g])

Petitioner fails to show that Fontana in view of Lau renders element 1[f] obvious. The Petition fails to adequately explain how a person of ordinary skill in the art could combine Fontana and Lau in the manner asserted in the Petition.

Claim 1[f] requires an “generating a topical meta-data that describes content of the video by deriving semantic information from the identification of the one or more objects and semantic information from the audio files.” Ex. 1001 at Claim 1.

With respect to generating topical meta-data, the '972 Patent explains, first concerning video data, “the text extracted from video images can be modified as the video-to-text engine selects primary semantics from a plurality of possible semantics.” *Id.* at 4:62-66. The '972 Patent further explains that this process is “repeated for all identifiable symbols in a video in a distributed process” and that the resulting topical meta-data includes “several topical descriptors of the content within a video.” Ex. 1001 at 4:39-41. In one example of topical meta-data generated for

the video data from the objects, the '972 Patent describes "an example of the aggregate information that would be derived using the above example would be understanding that the video presented a dog, which was jumping, on the beach, with people, by a resort." *Id.* at 4:41-44.

With respect to generating topical meta-data for the audio data, the '972 Patent similarly explains that, with natural language processing, "the text extract[ed] from the audio stream of a video can be given context, an applied sentiment, and topical weightings." *Id.* at 5:43-45. For example, the natural language processing of the audio data "can select semantics or topics, following certain rules, from a plurality of possible semantics or topics." *Id.* at 4:47-50. As such, the topical meta-data includes *semantic information* that is *derived from* the audio data.

Petitioner argues that the "contextual information"; "indexed transcript"; keywords"; and "other meta-data" derived from audio and video analyses disclosed in Fontana meets this claim limitation. *See* Petition at 20-23. In particular, Petitioner references Fontana's text metadata operation which defines text metadata associated with the multimedia content. *See Id.* at 21. Fontana describes the text metadata as the transcript and other external information (i.e., "additional contextual information", "advertisements, or hyperlinks to other websites or content", and "descriptive information related to the multimedia content"). *See* Ex. 1003 at [0121].

Petitioner argues that “the text metadata can include additional text-based descriptive and contextual information about the video content, beyond the transcript of the audio.” Petition at 22. To support this assertion, the Petitioner switches focus to keyword searchability, “Fontana discloses identifying keywords in the transcript, which enables searching of the multimedia content.” *Id.* Petitioner then asserts that “A POSA would have appreciated that Fontana’s above-described object metadata and text metadata (**including both transcript and keywords**) are all “topical meta-data that describes content of the video.” *Id.*

However, claim 1[f] requires “**generating a topical meta-data that describes content of the video by deriving** semantic information from the identification of the one or more objects and **semantic information from the audio files**”. Petitioner does not attempt to assert that the generation of a transcript by Fontana represents deriving semantic information. *See* Petition at 23. In other words, the process of converting speech to text described in Fontana does not derive semantic information describing the content of the video.

Concerning keywords as topical meta-data, Fontana is silent as to *how the keywords are generated*. It logically follows that the failure to describe how keywords are generated, also fails to teach how keywords are generated *by deriving semantic information from the audio files*. By contrast, Fontana states that “**additional information can be added to the stored information** that is available

to be searched, either within the video or **as made available for indexing by larger search engines external to the multimedia processing system.**" *See* Ex. 1003 at [0184] (emphasis added). Petitioner, relying on expert testimony, merely concludes that "a POSA would recognize keywords identified from a text-to-speech transcript as semantic information because they convey meaning about the audio." Petition at 23. Petitioner's expert does not cite to any additional supporting evidence or provide any technical reasoning to support his statement. *See* Ex. 1002 at [0097]-[0098]. Indeed, a comparison of the Petition to the declaration shows that they are nearly identical. *See* Ex. 2013 at [0097]-[0098]. Thus, the cited declaration testimony is conclusory and unsupported, adds little to the conclusory assertion for which it is offered to support, and is entitled to little weight. *See* 37 C.F.R. § 42.65(a) ("Expert testimony that does not disclose the underlying facts or data on which the opinion is based is entitled to little or no weight."); *Upjohn Co. v. Mova Pharm. Corp.*, 225 F.3d 1306, 1311 (Fed. Cir. 2000) ("Lack of factual support for expert opinion to factual determinations, however, may render the testimony of little probative value in a validity determination."); *Kinetic Techs., Inc., Petitioner, v. Skyworks Sols., Inc.*, IPR2014-00529, 2014 WL 4787238, at *10 (P.T.A.B. Sept. 23, 2014) ("Merely repeating an argument from the Petition in the declaration of a proposed expert does not give that argument enhanced probative value."); *Xerox Corp. v. Bytemark, Inc.*, IPR2022-00624, 2022 WL 3648989, at *6 (P.T.A.B. Aug. 24, 2022) ("We have

reviewed this excerpt from Dr. Jones' declaration and note that it merely repeats, *verbatim*, the conclusory assertion for which it is offered to support. . . . Thus, the cited declaration testimony is conclusory and unsupported, adds little to the conclusory assertion for which it is offered to support, and is entitled to little weight.”).

Independent claim 17[g] similarly recites “generating data topics, from the text and the video data, which describe content of the video by deriving semantic information from the identification of the one or more objects and semantic information from the audio.” For the same reasons discussed *supra* concerning deriving semantic information from the audio files, Petitioner fails to show that Fontana in view of Lau renders element 17[g] obvious and the Petition fails to adequately explain how a person of ordinary skill in the art could combine Fontana and Lau in the manner asserted in the Petition.

B. Fontana in View of Lau Fails to Render Obvious “cross-referencing the text and the video data based on the generated topical meta-data to determine topics” (Claim 1[h] and Claim 17[i])

Petitioner fails to show that Fontana in view of Lau renders element 1[h] obvious. The Petition fails to adequately explain why a person of ordinary skill in the art would be motivated to combine Fontana and Lau in the manner asserted in the Petition.

Claim 1[h] requires “cross-referencing the text and the video data based on the generated topical meta-data to determine topics.” Ex. 1001 at Claim 1.

With respect to cross-referencing the text and the video data, the '972 Patent explains that the topical meta-data “generated from an image or a frame and the topics extracted from audio can be combined.” *Id.* at 5:46-47. The '972 patent further explains that the text and video data are “cross-referenced, and topics common to both texts would be given additional weights.” *Id.* at 47-49.

Petitioner argues that “**topical meta-data** generated by Fontana would be searched using the method of Lau to determine a relevant context for the video and ultimately identify advertisements for display at specific times during a video.” Petition at 26. However, Fontana's object metadata and transcript generated from multimedia content would not be combined with Lau's correlation engine and search functionality to result in the *cross-referencing of the text and the video data based on the generated topical meta-data to determine topics*.

The Petition does not expressly identify Fontana's deficiency in teaching the limitations of claim 1[h]. *See* Petition at 26-32. Rather, the Petition references how advertisement data from Fontana “can be used in association with multimedia content, to link one or more advertisements with multimedia content during playback” and how “the matching of advertisements and content occurs based on a decision process separate from the content delivery system of the present

disclosure.” *See*, Petition at 31; Ex. 1003 at [0117]. As such, Lau is introduced in combination with Fontana as an “advertisement-matching method” to “practice Fontana’s disclosures of linking advertisements with multimedia content.” *See* Petition at 31.

Petitioner argues that “[i]t would have been obvious to a POSA that the text and object meta-data generated by Fontana could be used as the keywords/content to search for matching advertisements to video content according to Lau.” Petition at 32. Petitioner makes reference to multiple portions of Fontana that disclose *keyword searching* and concludes “a POSA would have appreciated that the **topical meta-data** generated by Fontana could be used as the keyword/content information for the video in Lau’s search method, and Fontana indeed encourages such combinations.” *See, Id.*; Ex. 1003 at [0042], [0107], [0114], [0121], and [0126-0127].

Accordingly, Petitioner’s combination of Fontana and Lau requires the linking of keywords described in Fontana to keywords, categories, or concepts described in Lau. However, as discussed *supra*, Fontana’s keywords **are not generated by deriving semantic information from the audio files**. Even assuming *arguendo* that Fontana’s keywords were topical meta-data as required by claim 1[f], the combination of Fontana and Lau proposed by the Petitioner uses an *external*

advertisement to perform the claimed cross-referencing of text and video data from the same video.

To address this logical gap requiring access to external advertisements from an advertisement system as described in Lau to somehow determine a topic from the cross-referencing of text and video data from the same video, the Petition broadly defines the determination of a “topic” as recited in claim 1[h] as searching for and matching “each of the ads, ad units, and context or subject matter of the video.” Petition at 27.

Notwithstanding this definition of “topic”, the determination of “an ad, an ad unit, or context or subject matter”, provides correlation of an ad, ad unit, or context or subject matter to object metadata and/or, **separately**, correlation of an ad, ad unit, or context or subject matter of an ad to keywords. *See* Ex. 1004 at FIGS. 4A-4B and [0102] and [0105]. Regardless of whether or not semantic information is derived to generate the keywords for the transcript, or whether a topic can be an external ad, ad unit, or context or subject matter of an ad, **the text and video data are not cross-referenced as required by claim 1[h]**. Instead, an ad, ad unit, or context or subject matter of an ad is correlated to text **independently** of object metadata, **and vice versa**.

Recognizing this, Petitioner argues that “Lau teaches matching keywords and concepts in a video to ads based on context or subject matter. (Lau ¶0051.) This

matching aggregates together the different metadata types and uses the resulting aggregated information to match the content to an ad.” Petition at 28. However, the *aggregation* of different metadata types in Lau does not *cross-reference* the text and the video data to determine a topic for the text and video data as required by claim 1[h].

Therefore, to combine Fontana and Lau as the Petition suggests, a person of ordinary skill in the art desiring to cross-reference text and video data from the same video to determine a topic would depart from the object metadata and the transcript for the video generated with the teachings of Fontana in order to access *external* advertisements with associated keywords. The person of ordinary skill in the art would then 1) compare the keywords of the external advertisements to the object metadata, 2) compare the keywords of the external advertisements to the keywords for the transcript, and then 3) determine any common advertisement, ad unit, context, or subject matter for the matching advertisement as the topic common to the text and the video data. Given the significant effort in accessing and comparing external advertisements, a person of ordinary skill in the art would not be motivated to combine Fontana and Lau, especially not in the way alleged by the Petition. Accordingly, a person of ordinary skill in the art would have no reason to look to Lau's correlation of advertisements to rich content media in order to cross-reference

Fontana's transcript keywords and object metadata to determine a topic for the transcript keywords and object metadata from the same video.

Further, Petitioner argues that "a POSA would have found it obvious that correlating advertisements to the video could cross-reference the text (speech-to-text transcript and any derived keywords) and the video data (the identified objects of interest) based on the generated topical meta-data (additional contextual information). Petition at 28-29. To support this argument, the Petition cites to the teachings of the '972 Patent itself and Lau, stating "the contextually aware keyword and topic searching aligns with the '972 specification's teaching regarding cross-referencing". *See, Id.* at 29; Ex. 1001 at 5:46-54; Ex. 1004 at 0057-0059. Petitioner is arguing that the elements of claim 1[h] would have been obvious to a person of ordinary skill in the art by citing to the '972 Patent itself as primary evidence to support what a person of ordinary skill in the art would understand. This type of circular reasoning (i.e., relying on the patent itself as a basis for obviousness) is a paradigm of legally impermissible hindsight. Accordingly, there is not a reasonable likelihood that Petitioners will prevail on Ground 1. *See Nautilus Hyosung Inc. v. Diebold Self-service Systems division of Diebold*, IPR2016-00633, Paper 9 at 21 (P.T.A.B. Aug. 22, 2016) ("[a]n assertion that something could be done does not articulate a reason why something would be done by one of ordinary skill in the art at the time of the invention and, therefore, raises a specter of impermissible hindsight

bias in a obviousness analysis.”) (emphasis added). Rather, Petitioner improperly uses the claims of the '972 patent as a roadmap for its obviousness combination. *See, e.g., TQ Delta, LLC v. Cisco Sys.*, 942 F.3d 1352, 1360 (Fed. Cir. 2019) (“we concluded that the expert appeared to have improperly relied on the challenged patent itself as her roadmap for putting what she referred to as pieces of a jigsaw puzzle together”) (internal quotations omitted).

Independent claim 17[i] recites “cross-referencing the text, the video data, and the topics with the video based on the generated data topics.” For the same reasons discussed *supra* concerning cross-referencing the text and video data, Petitioner fails to show that Fontana in view of Lau renders element 17[i] obvious and the Petition fails to adequately explain how a person of ordinary skill in the art could combine Fontana and Lau in the manner asserted in the Petition. Thus, for the same reasons provided as to claim 1 for Ground 1, the Petition should also be dismissed for claim 17.

C. Fontana in View of Lau Fails to Render Obvious “generating video text based on the cross-referencing, wherein the video text describes content of the video” (Claim 1[i])

Petitioner fails to show that Fontana in view of Lau renders element 1[i] obvious. The Petition fails to adequately explain why a person of ordinary skill in the art would be motivated to combine Fontana and Lau in the manner asserted in the Petition.

Claim 1[i] requires “generating video text based on the cross-referencing, wherein the video text describes content of the video.” Ex. 1001 at Claim 1.

As noted above, Lau does not teach “cross-referencing the text and the video data based on the generated topical meta-data to determine topics” per claim 1[h]. This failure in the combination of Fontana and Lau is fatal to Petitioner's arguments. Unfortunately, this deficiency only grows as the Petition attempts to make something more of Lau's advertisement correlation. For example, the Petition states “Lau discloses matching ads to multimedia content through cross-referencing by determining a context or subject matter (for simplicity, “context”) (“video text”) that “describes content of the video” and using that context to determine “ad units” from which the ads are generated.” Petition at 32-33. However, and as noted above, Petitioner simultaneously argues that “Each of the ads, ad units, and context or subject matter of the video is a ‘topic.’” *Id.* at 27. As such, Petitioner's asserted combination of Fontana and Lau requires subject matter or context to simultaneously render obvious the claimed “topic” determined by cross-referencing and, separately, the claimed “video text” that describes the content of the video.

To support the obviousness of assigning multiple claim elements to the same context or subject matter, the Petition relies on an example from Lau concerning a correlating a BMW advertisement to rich media content:

Correlation engine 202, when determining the advertisement, may determine one or more ad units that correlate to the subject matter. For example, based on one or more keywords, ad units from the ad matrix are determined. The ad units are then combined into an advertisement that is correlated to the subject matter. One example of this is BMW may provide a general ad unit for their logo and have a different ad unit for different models, such as the 330 model, 530 model, etc. The logo unit and each of the model units can be combined at runtime based on the context of the content. If the content talks about the 330 model then the logo and the 330 ad units may be combined and presented to the user.

Ex. 1004 at [0038]

Petitioner argues that “Lau determines that the ‘subject matter’ of the video is the BMW 330 model.” Petition at 34. However, nothing in Lau *generates* video text that describes the content of the video using the cross-referencing. Lau merely correlates a general BMW advertisement (i.e., BMW logo) and a 330 model-specific advertisement (i.e., 330 ad unit) to content that “talks about the 330 model”. No video text describing the content is generated by Lau’s simple matching of keywords from the advertisements to the rich media content.

The above example and its analysis also demonstrate the nonsensical overlapping of the proposed combination of Fontana and Lau’s alleged “topic” and “video text” that ignores the ‘972 Patent’s claimed determination of topics from cross-referencing **and** generation of video text describing the content of the video. For example, the ‘972 Patent describes the generation of video text as “text describing the content of the video, using the result of the combined texts and cross

reference". Ex. 1001 at 5:49-51. By contrast, the combination of Fontana and Lau merely correlates an advertisement to a portion of rich media content and provides that advertisement with the rich media content without generating the claimed "video text" that describes the content of the video using the cross-referencing. *See* Ex. 1004 at [0038].

For these reasons, Petitioner's arguments fail to show that Fontana in view of Lau renders element 1[i] obvious.

D. Fontana in View of Lau Fails to Render Obvious "generating a text, image, or animation based on the video text" (Claim 1[j])

Petitioner fails to show that Fontana in view of Lau renders element 1[j] obvious. The Petition fails to adequately explain why a person of ordinary skill in the art would be motivated to combine Fontana and Lau in the manner asserted in the Petition.

Claim 1[j] requires "generating a text, image, or animation based on the video text." Ex. 1001 at Claim 1.

As noted above, the Petition struggles to expand Lau's process to render the additional, specific elements and content generation in claims 1[h-j] obvious because Lau merely discloses the correlation of advertisements to portions of rich media content and the rendering of those correlated advertisements with the associated portions of the rich media content. *See* Ex. 1004 at [0006] and [0037]-[0038].

Petitioner conflates the claimed “topic”; “video text”, and now “text, image, or animation” by asserting that the selecting and combining of Lau’s ad units (previously “topics”) which are selected by matching to context (previously a “topic”, now “video text”), into an advertisement (previously a “topic”, now asserted as “a text, image, or animation”) generates something new. *See* Petition at 35.

However, Lau does not *generate* text, image, or animation *from the video text that describes the content of the video*. Rather, Lau simply renders the advertisement with the correlated rich media content. No change in granularity (i.e., “ad” or “ad unit”) or combination of advertisements as described by Lau resolves this deficiency because no “text, image, or animation” in Lau is generated *from the video text that describes the content of the video*. *See* Petition at 35. Lau’s correlation of advertisements to rich media content cannot extend as far as Petitioner wishes it to. Despite Petitioner’s efforts to differentiate an advertisement from “an advertiser-specified message”, or a logo, or a “video that may serve as pre/mid/post-roll”, all of these are just the correlated advertisement that is presented with the rich content media, lacking *any* reference to the claimed “video text” that describes the content of the video. *Id.*

For these reasons, Petitioner’s arguments fail to show that Fontana in view of Lau renders element 1[j] obvious.

Independent claim 17[j] recites “generating a text, image, or animation based on the data topics.” For the same reasons discussed *supra* concerning generating a text, image, or animation based on the video text, Petitioner fails to show that Fontana in view of Lau renders element 17[j] obvious and the Petition fails to adequately explain how a person of ordinary skill in the art could combine Fontana and Lau in the manner asserted in the Petition. Thus, for the same reasons provided as to claim 1 for Ground 1, the Petition should also be dismissed for claim 17.

E. Fontana in View of Lau Fails to Render Obvious “placing the text, image, or animation in the video.” (Claim 1[k] and Claim 17[k])

Petitioner fails to show that Fontana in view of Lau renders element 1[k] obvious. The Petition fails to adequately explain why a person of ordinary skill in the art would be motivated to combine Fontana and Lau in the manner asserted in the Petition.

Claim 1[k] requires “placing the text, image, or animation in the video.” Ex. 1001 at Claim 1.

The combination of Fontana and Lau does not place the text, image, or animation *generated* as described in claim 1[j] based on the video text *that describes the content of the video* as *generated* in claim 1[i] using the cross-referencing of text and video data described in claim 1[h]. Petitioner and its expert merely conclude that Lau’s “injecting the advertisement into the rich media content (i.e., video),

plac[es] the text, image, or animation in the video.” Petition at 36. However and as noted above, Lau does not teach anything beyond correlating an advertisement to a portion of rich media content and providing that advertisement with the rich media content. *See* Ex. 1004 at [0038].

For these reasons, Petitioner's arguments fail to show that Fontana in view of Lau renders element 1[k] obvious.

Independent claim 17[k] recites “placing the text, image, or animation in the video.” For the same reasons discussed *supra*, Petitioner fails to show that Fontana in view of Lau renders element 17[k] obvious and the Petition fails to adequately explain how a person of ordinary skill in the art could combine Fontana and Lau in the manner asserted in the Petition. Thus, for the same reasons provided as to claim 1 for Ground 1, the Petition should also be dismissed for claim 17.

F. A Person of Ordinary Skill in the Art Would Not be Motivated to Combine Fontana and Lau

A person of ordinary skill in the art would never have the motivation to combine Fontana and Lau as posited by the Petitioner. The Petition alleges that “[a] POSA would have been motivated to combine Fontana and Lau in order to use Lau's method for placing advertisements within a video to carry out Fontana's disclosures of linking advertisements to video content, and would have had a reasonable

expectation of success in the combination.” Petition at 30. However, the Petition's motivation to combine rests on the following flawed statement:

A POSA would have found it obvious to combine Lau and Fontana by using the **topical meta-data** generated by Fontana (*see* claims 1[f] and 1[g]) as well as **the text and the video data** (*see* claims 1[d] and 1[e]) as keywords and concepts searched by Lau for purposes of identifying ads or ad units to match to the video based on context or subject matter. Each of the ads, ad units, and context or subject matter of the video is a “**topic.**”

Petition at 27.

This motivation is nonsensical and loses all hope for a reasonable expectation of success when the Petition further states:

A POSA would have appreciated that this includes determining “context,” which is a way of describing what a video is about beyond just keywords. In other words, a POSA would have appreciated that “context” as used in Lau can be a form of **video text based on the cross-referencing** in claim 1[h] and that **describes content of the video.**

Id. at 33.

To complete the logical chain required by the combination of Fontana and Lau, a person of ordinary skill in the art generates video text that describes the content of the video based on cross-referencing of text and video data by: comparing keywords of an advertisement external to a video as taught by Lau against the transcript and object metadata generated by Fontana to determine the context of the advertisement and then using that searching of text and video data to identify

keywords that can be matched to the advertisement which actually describe the context of the video. The combination of Fontana and Lau as argued by the Petitioner begins where it ends – with keywords of an advertisement that can be correlated with subject matter in a video, and, separately, that can be correlated to subject matter in the audio for determining whether to and/or where to insert the advertisement into the video. As such, no person of ordinary skill in the art would ever combine Fontana and Lau as argued with the slightest hope of achieving what Petitioner concludes is obvious.

G. Houh's Expert Testimony Should be Entitled to Little or No Weight

As shown previously, Petitioner's expert testimony does not cite to any additional supporting evidence or provide any technical reasoning to support his statements concerning the elements of claim 1. *See* Ex. 1002 at [0097]-[0098]; and [0107]-[0130]. In fact, a comparison of the Petition to the declaration shows that they are nearly identical. *See* Ex. 2013 at [0097]-[0098]; and [0107]-[0130]. Thus, the cited expert testimony is conclusory and unsupported, adds little to the bare assertions for which it is offered to support, and is entitled to little or no weight. *See* 37 C.F.R. § 42.65(a) ("Expert testimony that does not disclose the underlying facts or data on which the opinion is based is entitled to little or no weight."); *Upjohn Co. v. Mova Pharm. Corp.*, 225 F.3d 1306, 1311 (Fed. Cir. 2000) ("Lack of factual

support for expert opinion to factual determinations, however, may render the testimony of little probative value in a validity determination.”); *Kinetic Techs., Inc., Petitioner, v. Skyworks Sols., Inc.*, IPR2014-00529, 2014 WL 4787238, at *10 (P.T.A.B. Sept. 23, 2014) (“Merely repeating an argument from the Petition in the declaration of a proposed expert does not give that argument enhanced probative value.”); *Xerox Corp. v. Bytemark, Inc.*, IPR2022-00624, 2022 WL 3648989, at *6 (P.T.A.B. Aug. 24, 2022) (“We have reviewed this excerpt from Dr. Jones’ declaration and note that it merely repeats, *verbatim*, the conclusory assertion for which it is offered to support. . . . Thus, the cited declaration testimony is conclusory and unsupported, adds little to the conclusory assertion for which it is offered to support, and is entitled to little weight.”).

VIII. GROUND 2 FAILS FOR THE SAME REASONS AS GROUND 1

For Ground 2, Petitioner argues that certain dependent claims are obvious over Ground 1 in light of additional prior art. *See* Petition at 57-65. All of these claims depend from claim 1 or claim 17. Petitioners do not make any new arguments about the elements of claim 1 or claim 17 for Ground 2, but instead argue that the additional prior art provides disclosure as to additional elements of the dependent claims. *Id.* Thus, for the same reasons provided as to Ground 1, the Petition should also be dismissed for Ground 2.

IX. CONCLUSION

For the reasons above, Petitioner has not met its burden of showing a reasonable likelihood of prevailing. Patent Owner respectfully submits that the Board should deny institution.

Dated: September 12, 2025

Respectfully submitted,

Holland & Knight LLP

/s/ Amy Simpson

Amy Simpson (Reg. No. 54,688)

Amy.simpson@hklaw.com

HOLLAND & KNIGHT LLP

1901 Avenue of the Stars, Suite 1200

Los Angeles, CA 90067

Telephone: (310) 201-8925

Facsimile: (310) 201-8922

Counsel for Cellular South, Inc.

X. CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. 42.6(e), 42.8(b)(4) and 42.105, the undersigned certifies that a complete and entire copy of this **PATENT OWNER'S PRELIMINARY RESPONSE** was served in their entirety via filing through the Patent Trial and Appeal Case Tracking System (P-TACTS), as well as electronic mail on the attorneys of record for Petitioners:

Heidi L. Keefe (Reg. No. 40,673)
hkeefe@cooley.com
Andrew C. Mace (Reg. No. 63,342)
amace@cooley.com
Mark R. Weinstein
mweinstein@cooley.com
Reuben Chen
rchen@cooley.com
Alexandra D. Leeper
aleeper@cooley.com
COOLEY LLP
ATTN: Patent Group
1299 Pennsylvania Ave. NW, Suite 700
Washington, DC 20004
Tel: (650) 843-5001
Fax: (650) 849-7400

Dated: September 12, 2025

/s/ Amy Simpson
Amy Simpson (Reg. No. 54,688)

Counsel for Cellular South, Inc.

XI. CERTIFICATE OF COMPLIANCE WITH 37 C.F.R. § 42.24

Pursuant to 37 C.F.R. § 42.24 et seq., the undersigned certifies that this Response complies with the type-volume limitations excluding the cover page and the parts exempted by 37 C.F.R. § 42.24(a)(1). This document contains 8,617 words as calculated by the "Word Count" feature of Microsoft Word, the word processing program used to create it.

Dated: September 12, 2025

/s/ Amy Simpson
Amy Simpson (Reg. No. 54,688)
amy.simpson@hklaw.com
HOLLAND & KNIGHT LLP
1901 Avenue of the Stars, Suite 1200
Los Angeles, California 90067
Email: amy.simpson@hklaw.com
Telephone: (310) 201-8925
Facsimile: (310) 201-8922

Counsel for Cellular South, Inc.