

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,
Petitioner,

v.

DIALECT, LLC,
Patent Owner.

IPR2024-00749
Patent 7,693,720 B2

Before KRISTEN L. DROESCH, NEIL T. POWELL, and
GEORGE R. HOSKINS, *Administrative Patent Judges*.

POWELL, *Administrative Patent Judge*.

DECISION
Denying Institution of *Inter Partes* Review
35 U.S.C. § 314

I. INTRODUCTION

A. BACKGROUND

Google LLC (“Petitioner”) filed a Petition for *inter partes* review of claims 1–16, 18–25, 27, 29–43, 52, and 55 of U.S. Patent No. 7,693,720 B2 (Ex. 1001, “the ’720 patent”). Paper 2 (“Pet.”). Dialect, LLC (“Patent Owner”) filed a Preliminary Response. Paper 8 (“Prelim. Resp.”).

An *inter partes* review may not be instituted “unless . . . the information presented in the petition . . . shows that there is a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition.” 35 U.S.C. § 314(a). Having considered the arguments and the associated evidence, we determine that Petitioner has not shown a reasonable likelihood that it will prevail in establishing the unpatentability of any of the challenged claims. Accordingly, we do not institute *inter partes* review.

B. RELATED PROCEEDINGS

The parties indicate that the ’720 patent is involved in *Dialect, LLC v. Google LLC*, No. 1:23-cv-00378 (D. Del.); and *Dialect, LLC v. Amazon.com, Inc.*, No. 1:23-cv-00581 (E.D. Va.). Pet. 17; Paper 5, 1. Petitioner also identifies IPR2024-00746, IPR2024-00747, IPR2024-00748, IPR2024-00750, IPR2024-00751, IPR2024-00752, and IPR2024-00753 as involving the same parties and patents related to the ’720 patent. Pet. 17–18.

C. THE ’720 PATENT

The ’720 patent relates to retrieving online information or processing commands through a speech interface in a vehicle environment. Ex. 1001, 1:12–14. The ’720 patent is directed to “a fully integrated environment allowing mobile users to ask natural language speech questions or give

“[Figure] 5 is an overall diagrammatic view of the interactive natural language speech processing system according to an embodiment of the invention.” Ex. 1001, 15:44–46.

In particular, Figure 5 shows “[a]gents 106 [that] may include packages of both generic and domain specific behavior for the system.” Ex. 1001, 20:12–13. “User specific data, parameters, session information, location data and history information determining the behavior of agents are stored in one or more user profiles 110.” *Id.* at 20:12–19. Agents may organize “generic and domain specific behavior and information,” and the “system agent provides default functionality and basic services” while the “domain specific agents may provide complete, convenient and re-distributable packages or modules for each application area.” *Id.* at 21:38–43.

D. ILLUSTRATIVE CLAIMS

Of the challenged claims, claims 1 and 31 are independent. Each of claims 2–16, 18–25, 27, 29, 30, 32–43, 52, and 55 depends, directly or indirectly, from claim 1 or 31. Independent claim 1 is reproduced below:

1. A mobile system responsive to a user generated natural language speech utterance, comprising:

a speech unit connected to a computer device on a vehicle, wherein the speech unit receives a natural language speech utterance from a user and converts the received natural language speech utterance into an electronic signal; and

a natural language speech processing system connected to the computer device on the vehicle, wherein the natural language speech processing system receives, processes, and responds to the electronic signal using data received from a plurality of domain agents, wherein the natural language speech processing system includes:

a speech recognition engine that recognizes at least one of words or phrases from the electronic signal using at least the data received from the plurality of domain agents, wherein the data used by the speech recognition engine includes a plurality of dictionary and phrase entries that are dynamically updated based on at least a history of a current dialog and one or more prior dialogs associated with the user;

a parser that interprets the recognized words or phrases, wherein the parser uses at least the data received from the plurality of domain agents to interpret the recognized words or phrases, wherein the parser interprets the recognized words or phrases by:

determining a context for the natural language speech utterance;

selecting at least one of the plurality of domain agents based on the determined context; and

transforming the recognized words or phrases into at least one of a question or a command, wherein the at least one question or command is formulated in a grammar that the selected domain agent uses to process the formulated question or command; and

an agent architecture that communicatively couples services of each of an agent manager, a system agent, the plurality of domain agents, and an agent library that includes one or more utilities that can be used by the system agent and the plurality of domain agents, wherein the selected domain agent uses the communicatively coupled services to create a response to the formulated question or command and format the response for presentation to the user.

Ex. 1001, 32:30–33:6.

E. ASSERTED GROUNDS OF UNPATENTABILITY

Petitioner challenges the patentability of claims 1–16, 18–25, 27, 29–43, 52, and 55 of the '720 patent on the following grounds (Pet. 19–20):

Claims Challenged	35 U.S.C. §	References
1–16, 22–25, 27, 29–31, 35–43, 52, 55	103(a) ¹	Coffman ² , Kanevsky ³ , Kneser ⁴ , Zadrozny ⁵
18, 20, 21, 32, 33	103(a)	Coffman, Kanevsky, Kneser, Zadrozny, DeLine ⁶
19, 34	103(a)	Coffman, Kanevsky, Kneser, Zadrozny, Gerson ⁷

In support of its challenges, Petitioner also relies on the Declaration of Christopher M. Schmandt. Ex. 1002.

¹ The Leahy-Smith America Invents Act (“AIA”), Pub. L. No. 112-29, 125 Stat. 284, 287–88 (2011), revised 35 U.S.C. § 103, effective March 16, 2013. The '720 patent's filing date is before March 16, 2013, so we apply the pre-AIA version of § 103. Ex. 1001, code (22).

² Coffman et al., International Patent Application Publication No. WO 00/20962, published Apr. 13, 2000 (Ex. 1003, “Coffman”).

³ Kanevsky et al., U.S. Patent No. 5,897,616, issued Apr. 27, 1999 (Ex. 1004, “Kanevsky”).

⁴ Kneser et al., U.S. Patent No. 6,157,912, issued Dec. 5, 2000 (Ex. 1005, “Kneser”).

⁵ Zadrozny et al., *NL Understanding with a Grammar of Constructions*, Coling 94: The 15th International Conference on Computational Linguistics, 1994 (Ex. 1006, “Zadrozny”).

⁶ DeLine et al., U.S. Patent No. 6,420,975 B1, issued July 16, 2002 (Ex. 1010, “DeLine”).

⁷ Gerson, U.S. Patent No. 6,937,977 B2, issued Aug. 30, 2005 (Ex. 1011, “Gerson”).

II. ANALYSIS

A. LEVEL OF ORDINARY SKILL IN THE ART

Petitioner contends that an ordinarily skilled artisan for the '720 patent “would have possessed a bachelor’s degree in electrical engineering or computer science, and two years of work experience in the field of speech- and natural language-interfaces for interacting with and directing computer systems and software, including speech recognition and processing (and more experience could substitute for less formal education, and vice versa).” Pet. 21 (citing Ex. 1002 ¶¶ 11–15). At this stage, Patent Owner does not dispute Petitioner’s proposed level of skill in the art. *See, generally*, Prelim. Resp.

Based on this record, we adopt Petitioner’s articulation of the level of ordinary skill in the art, which appears to be consistent with the '720 patent and the asserted prior art. *See Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001).

B. CLAIM INTERPRETATION

We apply the same claim construction standard used in district courts, namely that articulated in *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005) (en banc). *See* 37 C.F.R. § 42.100(b) (2022). In applying that standard, claim terms generally are given their ordinary and customary meaning as would have been understood by a person of ordinary skill in the art at the time of the invention and in the context of the entire patent disclosure. *Phillips*, 415 F.3d at 1312–13.

Both Parties express that we do not need to construe any claim language explicitly. Pet. 23; Prelim. Resp. 13. We agree that no explicit claim construction is necessary for this Decision. *See Nidec Motor Corp. v.*

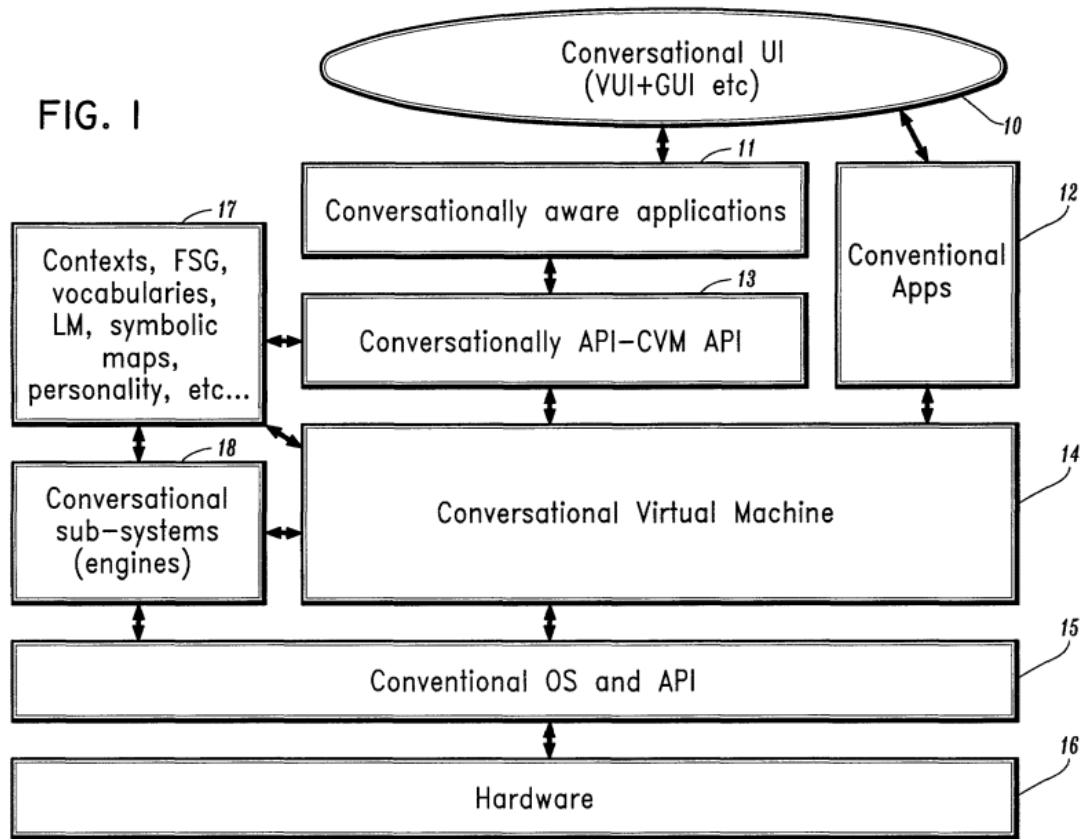
Zhongshan Broad Ocean Motor Co., 868 F.3d 1013, 1017 (Fed. Cir. 2017) (holding that only claim terms in controversy need to be construed, and only to the extent necessary to resolve the controversy (citing *Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999))).

C. PRIOR ART REFERENCES

1. *Coffman*

Coffman relates to conversational computing. Ex. 1003, 1:7–8.⁸ Coffman discusses “a system that provides conversational computing across multiple platforms, devices and applications through a universal conversational user interface, which goes far beyond adding speech I/O or conversational capabilities to existing applications, building conventional conversational applications or superficially integrating ‘speech’ in conventional operating systems.” *Id.* at 3:29–4:2. Coffman discusses an example embodiment in connection with Figure 1, which is reproduced below.

⁸ When addressing Coffman, we cite its native pagination.



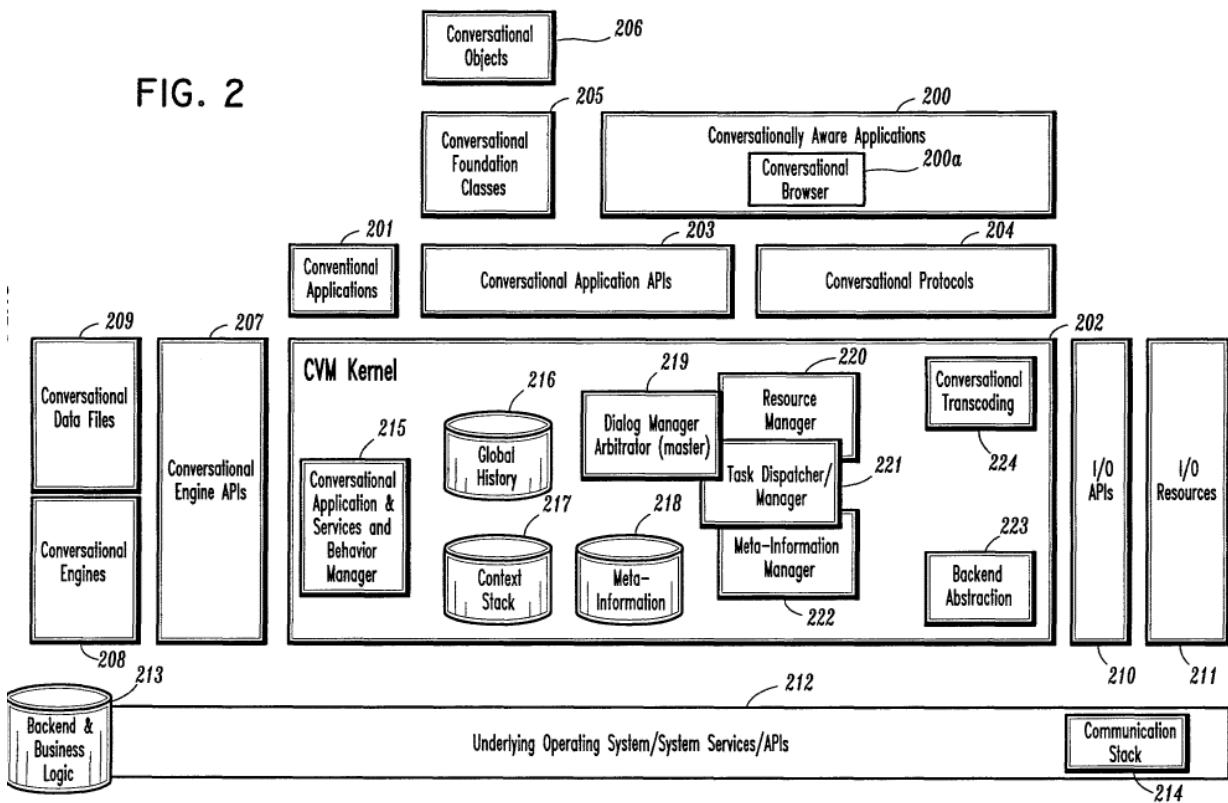
“[Figure] 1 is a block diagram of a conversational computing system according to an embodiment of the present invention.” Ex. 1003, 9:27–28.

Coffman discloses conversational virtual machine (CVM) 14 as a kernel that runs on top of conventional operating system (OS) 16 to provide backward compatibility for conventional platforms by managing dialog and context, conversational sub-systems (engines) and resources, and conversational protocols/communications across platforms and devices with different capabilities. Ex. 1003, 1:12–17, 4:17–26, 13:1–3. “The CVM kernel is the core layer that controls the dialog across applications and devices on the basis of their registered conversational capabilities and requirements.” *Id.* at 4:27–28. “[C]onventional applications 12 are managed by the CVM kernel layer 14 which is responsible for accessing, via

the OS APIs, GUI menus and commands of the conventional applications as well as the underlying OS commands.” *Id.* at 13:3–5. Conversational subsystems 18, e.g., speech recognition, natural language understanding (NLU) and natural dialog generation (NLG) engines, are accessed through the CVM. *Id.* at 5:1–7.

Coffman also discloses conversationally aware applications 11, which “are specifically programmed for operating with a CVM core layer (or kernel) 14 via conversation applications APIs 13,” based on “their registered conversational capabilities and requirements.” Ex. 1003, 12:28–32. CVM kernel 14 controls the dialog across applications and devices based on their registered conversational capabilities and provides a unified conversational user interface (CUI). *Id.* A conversationally aware application 11 provides its data file to the CVM for use by the conversational subsystems 18 by “registering all its conversational capabilities and requirements,” including active vocabulary, symbolic mapping (if NLU is supported or in a context state list), associated meta-information to allow categorization of the output, conversational I/O information, and CVM capabilities/state (master, slave, or partner CVM). *Id.* at 57:22–58:2.

Figure 2 of Coffman, reproduced below, illustrates abstract layers of an embodiment of Coffman’s conversational computing system.
Ex. 1003, 9:29–30.



Ex. 1003, Fig. 2. Conversationally aware applications 200 interact with CVM kernel layer 202 via conversationally aware application API layer 203. *Id.* at 13:16–19. CVM kernel layer (or CVM controller) 202 includes a meta-information manager 222, resource manager 220, context stack 217, global history 216, dialog manager 219, and task dispatcher 221. *Id.* at Fig. 2, 19:26–29, 21:3–4, 21:23, 21:28–29. Within kernel 202, context stack 217 “is associated with the organized/sorted context corresponding to each active dialog (or deferred dialog- agents/memorization)” and “includes all the variable, states, input, output and queries to the backend that are performed in the context of the dialog and any extraneous event that occurs during the dialog.” *Id.* at 21:23–28.

2. *Kanevsky*

Kanevsky discloses a “method and apparatus for securing access to a service or facility employing automatic speech recognition, text-independent speaker identification, natural language understanding techniques and additional dynamic and static features.” Ex. 1004, code (57). Kanevsky explains employing “a unique combination of random questions, automatic speech recognition (ASR) and text-independent speaker recognition to provide a significant improvement in secure access to services and/or facilities . . . requiring security measures.” *Id.* at 5:34–39.

3. *Kneser*

Kneser discloses a method “for adapting the unigram language model values which automatically carries out a subdivision of the words into semantic classes.” Ex. 1005, code (57). As part of this system and method, Kneser explains that “unigram language model values obtained from a large quantity of texts are adapted on the basis of an updated speech signal, and adapted M-gram language model values are obtained by means of said adapted unigram language model values.” *Id.* at 2:1–6. “This renders it possible to adapt the M-gram values already on the basis of a small quantity of spoken text data in which a major portion of the word sequences on which the M-gram values are based has not yet occurred.” *Id.* at 2:9–14. “[C]lassification can be achieved automatically when different training speech signals are available which originate, for example, from different subject fields and comprise different styles of speech or language.” *Id.* at 2:34–38.

4. *Zadrozny*

Zadrozny discloses “an approach to a natural language understanding based on a *computable grammar of constructions*,” each of which “consists of a set of features of form and a description of meaning in a context.”

Ex. 1006, 19. Zadrozny explains that “[a] grammar is a set of constructions.” *Id.* Zadrozny further discloses that “grammars of constructions can be parsed like ‘standard’ grammars, except that the set of features is richer. Given a string (representing a sentence, a fragment of a discourse or a paragraph), the parser assigns it a construction.” *Id.* Specifically, “the program uses application-specific knowledge to translate the action and its parameters into a form that can be executed by the application.” *Id.* at 20.

5. *DeLine*

DeLine relates to “[a] digital sound processor [that] is provided to enhance the vocal to non-vocal noise ratio of the signal processed by a vehicle audio system such as a cellular telephone, emergency communication device, or other audio device.” Ex. 1010, code (57). DeLine explains “an indicator is provided for use with the vehicular audio system in order to provide a user of the audio system with a status signal relating to a reception quality of a vocal signal from the user.” *Id.*

6. *Gerson*

Gerson relates to “accurately establish[ing] a context of an input speech signal relative to an output audio signal regardless of the delay characteristics of the underlying communication system.” Ex. 1011, code (57). Gerson explains “[i]nformation signals comprising data and/or control signals are provided in response to at least the contextual information

provided, i.e., the input start time and/or the identification of the output audio signal.” *Id.*

D. PETITIONER’S OBVIOUSNESS CHALLENGES

Petitioner asserts that its “grounds rely primarily on Coffman.”

Pet. 23. Petitioner elaborates that “[a]ll grounds combine Coffman with Kanevsky (EX1004), Kneser (EX1005) and Zadrozny (EX1006), which disclose known techniques related to speech recognition and natural language processing.” *Id.* Petitioner adds that “[g]rounds 2-3 combine these references with DeLine (EX1010) and Gerson (EX1011) for features recited in dependent claims 18-21 and 32-34.” *Id.* at 24. Petitioner then elaborates on its bases for asserting that the challenged claims would have been obvious over the prior art, addressing each limitation of claims 1–16, 18–25, 27, 29–43, 52, and 55 in detail. *Id.* at 24–96.

Patent Owner argues that Petitioner fails to address persuasively “every claim requirement of a ‘domain agent,’” as recited in the challenged claims. Prelim. Resp. 4. For example, Patent Owner contends that Petitioner does not address sufficiently persuasively the challenged claims’ limitations related to a parser selecting “at least one of the plurality of domain agents based on the determined context” (which we also refer to as the “selecting limitations”), as recited in independent claims 1 and 31. *Id.* at 27–33. Patent Owner also argues that Petitioner does not address sufficiently persuasively the challenged claims’ limitation “the selected domain agent uses the communicatively coupled services to create a response to the formulated question or command and format the response for presentation to the user,” as recited in independent claims 1 and 31. *Id.* at 34–40.

In view of all of the Parties’ arguments and evidence, the record does not demonstrate a reasonable likelihood of Petitioner prevailing with respect to any challenged claim. For example, Petitioner and Mr. Schmandt do not address the challenged claims’ selecting limitations in a sufficiently persuasive manner. Coffman’s disclosure is high-level, which Petitioner and Mr. Schmandt recognize. *See, e.g.*, Pet. 48; Ex. 1002 ¶ 125. Coffman’s high-level discussion lacks clear disclosures of certain key things that Petitioner and Mr. Schmandt perceive in Coffman’s disclosure. And Petitioner and Mr. Schmandt do not provide sufficiently clear, persuasive explanation to substantiate their position that Coffman teaches the challenged claims’ selecting limitations.

Petitioner contends that “[a] conversationally-aware application in Coffman—the executable software and associated data, such as vocabulary, grammar, language models and NLU symbolic mapping—is a ‘**domain agent**,’” as recited in the ’720 patent’s challenged claims. Pet. 33 (citing Ex. 1003, 13:26–29; Ex. 1002 ¶ 89). Addressing the selecting limitations, Petitioner and Mr. Schmandt state that in Coffman’s system, “[a] conversationally-aware application (‘**domain agent**’) is ‘**select[ed]...based on the determined context**’ because activation of an application for processing the user’s query is based on the application’s association with the determined context.” Pet. 43–44 (citing Ex. 1002 ¶ 116; Ex. 1003, 42:1–7, 37:2–4). This alleged teaching is not clearly present in the cited passages of Coffman. And Petitioner and Mr. Schmandt do not provide a sufficiently cohesive, persuasive explanation of why they perceive such a teaching in Coffman.

The first cited passage discusses searching for an active context in Coffman's context stack:

Completion/search for the active context is performed from context to context down the stack. That is, new queries or arguments are compared by the dialog engine by going down the stack until an acceptable match is obtained and optionally confirmation is obtained from the user. As soon [as] a context is found that fits the utterance at the level of the NLU symbolic language, the context becomes active and the corresponding process becomes active. Until the active command is completed, or until a new command is provided, the selected context is marked active, and pushed to the top of the context stack 405.

Ex. 1003, 42:1–7. This passage is not a clear disclosure of selecting an application based on its association with a determined context. Indeed, the passage never mentions an “application.” And Petitioner and Mr. Schmandt do not explain persuasively why this passage discloses or teaches selection of an application based on a determined context. *See, e.g.*, Pet. 43–44; Ex. 1002 ¶ 116.

We note that the passage discloses that “the corresponding *process* becomes active.” Ex. 1003, 42:5 (emphasis added). And Petitioner and Mr. Schmandt seem to suggest that Coffman uses the word “process” to refer to a conversationally-aware application. Pet. 43 (citing Ex. 1003, 29:18–30:10; Ex. 1002 ¶ 114) (“Coffman explains that the conversational system includes a ‘context stack’ where the ‘context,’ including the state and history of each active process, such as an application process, is tracked.”); Ex. 1002 ¶ 114 (citing Ex. 1003, 29:18–30:10). But Petitioner and Mr. Schmandt suggest this in a conclusory manner, ad-libbing that “an application process” constitutes an example of “each active process” referred to in Coffman's passage at page 29, line 18 through page 30, line 10, which

lacks a clear disclosure that “process” refers to an application, let alone a conversationally-aware application. Pet. 43 (citing Ex. 1003, 29:18–30:10; Ex. 1002 ¶ 114); Ex. 1002 ¶ 114 (citing Ex. 1003, 29:18–30:10).

Petitioner and Mr. Schmandt also cite Coffman’s disclosure that “[t]he first task to have its dialog completed by producing a non-ambiguous query is executed and the corresponding query as interpreted by the other application is stored to activate if the recognized query is rejected by the user.” Ex. 1003, 37:2–4; Pet. 43–44; Ex. 1002 ¶ 116. This passage, like the passages discussed above, lacks a clear teaching of selecting a conversationally-aware application based on its association with a determined context. And no persuasive explanation of how such a teaching could be gleaned from this passage is provided by either Petitioner or Mr. Schmandt. *See, e.g.*, Pet. 43–44; Ex. 1002 ¶ 116.

When addressing claim limitations other than the selecting limitations, Petitioner and Mr. Schmandt provide other citations to various disclosures in Coffman and the other prior art references. *E.g.*, Pet. 24–43. In doing so, Petitioner and Mr. Schmandt also fail to provide a cohesive, persuasive explanation in support of their position that Coffman discloses selecting a conversationally-aware application based on its association with a determined context. For example, in support of their view that “Coffman teaches that its NLU processing uses data provided by conversationally-aware applications (such as grammars and symbolic query maps) to interpret the words or phrases in the user’s utterance in order to identify an appropriate application and to formulate an application call/query to perform,” Petitioner and Mr. Schmandt rely on a string cite to numerous Coffman passages. Pet. 42–43 (citing Ex. 1002 ¶ 112; Ex. 1003, 6:32–7:3,

13:8–11, 30:8–10, 37:20–26, 41:2–5, 42:1–5, 51:12–14, 53:25–26); Ex. 1002 ¶ 112 (citing Ex. 1003, 6:32–7:3, 13:8–11, 30:8–10, 37:20–26, 41:2–5, 42:1–5, 51:12–14, 53:25–26). Mr. Schmandt testifies that he “explain[s]” how the cited passages allegedly teach this, but we do not find he does so in a cohesive, persuasive way. Thus, Petitioner and Mr. Schmandt do not address sufficiently persuasively the challenged claims’ requirement of selecting “at least one of the plurality of domain agents based on the determined context.”

In addition to this deficiency, Petitioner’s arguments and evidence have another, related shortcoming. In the challenged claims, the selection of at least one domain agent must be done by a parser, not just any system component. Ex. 1001, 32:54–59, 35:42–44. Even if Petitioner’s arguments and evidence did sufficiently show that Coffman teaches selecting a conversationally-aware application, they do not show that Coffman teaches doing so with a parser. Coffman’s natural language understanding and natural language parsing system (NLU) 410 is a parser, as Petitioner and Mr. Schmandt note. Pet. 42 (citing Ex. 1002 ¶ 111; Ex. 1003, 29:1–4, 40:12–15, 51:12–14, Fig. 4); Ex. 1002 ¶ 111 (citing Ex. 1003, 29:1–4, 40:12–15, 51:12–14, Fig. 4). But we agree with Patent Owner that neither Petitioner nor Mr. Schmandt clearly alleges, let alone substantiates, that Coffman’s alleged selection of a conversationally-aware application based on a determined context is performed by Coffman’s parser. Prelim. Resp. 32–33. Instead, Petitioner and Mr. Schmandt merely allege that “[a] conversationally-aware application (‘**domain agent**’) is ‘**select[ed]...based on the determined context,**’” without specifying a component of

Coffman's system that does the alleged selecting. Pet. 43–44; Ex. 1002 ¶ 116.

For the foregoing reasons, the record does not demonstrate a reasonable likelihood of Petitioner prevailing with respect to any of the challenged claims.

III. CONCLUSION

For the above reasons, we conclude the record does not demonstrate a reasonable likelihood that Petitioner would prevail with respect to at least one challenged claim, so we deny institution of the petitioned *inter partes* review per 35 U.S.C. § 314(a)

IV. ORDER

In consideration of the foregoing, it is:

ORDERED that, pursuant to 35 U.S.C. § 314(a), the petitioned institution of *inter partes* review of the '720 patent is *denied*.

IPR2024-00749
Patent 7,693,720 B2

FOR PETITIONER:

Heidi L. Keefe
Andrew C. Mace
COOLEY LLP
hkeefe@cooley.com
amace@cooley.com

FOR PATENT OWNER:

Wayne M. Helge
James T. Wilson
BUNSOW DE MORY, LLP
whelge@bdiplaw.com
jwilson@bdiplaw.com