UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO. LTD. and SAMSUNG ELECTRONICS AMERICA, INC.
Petitioners

v.

VB ASSETS, LLC,
Patent Owner

IPR2025-00870
U.S. Patent No. 10,755,699

DECLARATION OF HOUMAN HOMAYOUN, PH.D. IN SUPPORT OF PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I, Houman Homayoun, declare as follows:

I. INTRODUCTION

1. I have personal knowledge regarding the facts discussed herein, I am over the age of 18, and I am able to competently testify.

2. I understand that Samsung Electronics Co. Ltd. and Samsung Electronics America, Inc. (collectively, "Petitioners" or "Samsung") challenged the patentability of U.S. Patent No. 10,755, 699 (the "'699 Patent") in a petition for *Inter Partes* Review ("IPR").

3. I have been retained as an expert by VB Assets LLC ("Patent Owner") to provide opinions regarding the patentability of the '699 Patent.

4. My opinions regarding the patentability of the '699 Patent are set forth herein. My opinions are based on the many documents I have reviewed, including the '699 Patent, its file history, the petition (Paper No. 3, referred to as the "Petition") and the exhibits and materials cited therein (Exs. 1001 – 1023), including the Declaration of Stuart Lipoff (Ex. 1003).

5. The opinions set forth herein are based on the knowledge and information currently available to me. I reserve the right to supplement and/or amend the opinions set forth herein.

6. Further, I understand that Petitioners assert two invalidity Grounds in this IPR based on five references (i) SmartKom: Foundations of Multimodal

Dialogue Systems (2006), (ii) User Models in Dialog Systems, Kobsa & Wahlster eds. (1989), (iii) U.S. Patent App. Publication No. 2004/0101198 to Barbara ("Barbara"), (iv) U.S. Patent App. Publication No. 2002/0173960 to Ross ("Ross"), and (v) U.S. Publication No. 2002/0065651 to Kellner ("Kellner"). The two invalidity grounds and the corresponding challenged claims are set forth in the table below:

Ground		Claims	Prior Art
1	103	1-22	SmartKom+Kobsa
2	103	1-22	Barbara+Ross+Kellner

7. I understand that Patent Owner will submit a Preliminary Response to the Petition, and I have been asked to review and analyze the Petition and accompanying exhibits and provide my opinions in response. I understand that Patent Owner's Preliminary Response will be submitted prior to institution. To the extent the IPR is instituted, I reserve all rights to submit a declaration in support of Patent Owner's response.

8. For the reasons set forth below, it is my opinion that the Petition fails to demonstrate that claims 1-22 are obvious in view of the asserted grounds.

9. I am being compensated at a rate of $400 per hour for my study and other work in this matter. I am also being reimbursed for reasonable and customary

expenses associated with my work in this investigation. My compensation is not contingent on the outcome of this matter or the specifics of my testimony.

II. QUALIFICATIONS

10. My curriculum vitae is attached as Exhibit 2026. A summary of my qualifications relevant to this proceeding is provided below.

11. I am a Professor of Electrical and Computer Engineering at the University of California, Davis. Prior to my current position, I was an Associate Professor at George Mason University (GMU). I am the director of GMU's Accelerated, Secure, and Energy-Efficient Computing Laboratory (ASEEC). Prior to joining GMU, I spent two years at the University of California, San Diego, as NSF Computing Innovation (CI) Fellow awarded by the CRA-CCC.

12. I hold a Ph.D. in Computer Science from the University of California, Irvine, which I earned in 2010, an M.S. in computer engineering from the University of Victoria, which I earned in 2005, and a B.S. in electrical engineering from the Sharif University of Technology, which I earned in 2003.

13. I serve as the director of the National Science Foundation Center for Hardware and Embedded Systems Security and Trust (CHEST).

14. I worked in NOVELICS, a startup company as a design architect from January 2007 to October 2008. I designed a reconfigurable and programmable processor referred to as Built-in Self-test (BIST) to test various memory architecture.

15. My research focuses on applied machine learning and AI, large language models, hardware security and trust, data-intensive computing, and heterogeneous computing.

16. I have directed over $20M in funded research projects. My team at UC Davis and prior to that at George Mason University, has successfully completed more than 20 projects funded by DARPA, NIST, NSF and, AFLR, DHS, Intel, General Motors, and a consortium on more than 20 companies which are member of CHEST center. Several of these projects are related to using large language models and/or designing large language models.

17. From 2017 to 2021 I served as an Associate Editor of IEEE Transactions on VLSI. I served as TPC Co-Chair for GLSVLSI 2018. I also served as the general chair of GLSVLSI 2019.

18. I am the co-author of the textbook Machine Learning for Computer Scientists and Data Analysts, which covers recommender systems, inference modeling, deep learning architectures, and deployment strategies for AI across software and hardware platforms.

19. My current research focuses on applied machine learning and AI, including deep learning, neural networks, natural language processing, user-aware computing, and conversational AI. I have published more than 100 peer-reviewed papers in these areas.

20. My research contributions include: “A Fast Method to Fine-Tune Neural Networks for the Least Energy Consumption on FPGAs," enabling efficient on-device AI for edge-based speech and audio processing, "Energy-Efficient and Adversarially Robust Machine Learning with Selective Dynamic Band Filtering," improving robustness in processing sensor signals, including voice, "Addressing Stereotypes in Large Language Models: A Critical Examination and Mitigation Approach," relevant to fairness, personalization, and user modeling in conversational AI, "A Hardware Accelerator for Language-Guided Reinforcement Learning," bridging natural language instructions and control logic for interactive agents, "Automatic Detection of Respiratory Symptoms Using a Low Power Multi-Input CNN Processor," advancing speech-driven health applications.

21. I teach both undergraduate and graduate courses at UC Davis. My Embedded Systems course includes projects on speech recognition applications for microcontrollers, integrating real-time inference, user interaction, and HCI design. My Applied Machine Learning course covers NLP, transformers, LLMs, and conversational AI with emphasis on practical implementation and user modeling.

III. LEGAL PRINCIPLES

22. For my opinions in this declaration, I understand that it requires applying various legal principles. As I am not an attorney, I have been informed and understand from counsel about various legal principles that involve my analysis. I

have used my understanding of those principles in forming my opinions. I summarize those principles as I understand them below.

23. For example, I understand that Petitioners bear the burden of proving unpatentability in this proceeding by a preponderance of the evidence. I understand that this preponderance of the evidence standard means that Petitioners must show that unpatentability is more probable than not.

24. I understand that a claimed invention is not patentable under 35 U.S.C. § 103 if it is obvious. A patent claim is unpatentable if the claimed invention would have been obvious to a person of ordinary skill in the field (which I discuss below) at the time the claimed invention was made. This means that even if all of the requirements of the claim cannot be found in a single prior art reference that would anticipate the claim, a person of ordinary skill in the relevant field who knew about all this prior art would have come up with the claimed invention.

25. I have been informed and understand that the obviousness analysis requires evaluating the claim language as a whole, including comparing each and every limitation of the properly construed claim language to the prior art.

26. I further understand that the ultimate conclusion of whether a claim is obvious should be based upon several factual determinations. That is, a determination of obviousness requires inquiries into: (1) the level of ordinary skill in the field; (2) the scope and content of the prior art; (3) what difference, if any,

existed between the claimed invention and the prior art; and (4) any objective factors bearing on obviousness.

27. I understand that, in determining the level of ordinary skill in the field that someone would have had at the time the claimed invention was made, I should consider: (1) the levels of education and experience of persons working in the field; (2) the types of problems encountered in the field; and (3) the sophistication of the technology.

28. I understand that, in determining the scope and content of the prior art, in order to be considered as prior art, a reference must be analogous art. I also understand that to be analogous art, the reference must be either (1) from the same field of endeavor as the claimed subject matter, regardless of the problem addressed, or (2) if not in the same field of endeavor, reasonably pertinent to the particular problem with which the inventor is involved. I am also familiar with the premise that for a reference to be reasonably pertinent, it must have logically commended itself to an inventor's attention at the time of invention.

29. I understand that a patent claim composed of several elements is not proved obvious merely by demonstrating that each of its elements was independently known in the prior art. In evaluating whether such a claim would have been obvious, I may consider whether there is a reason that would have prompted a person of

ordinary skill in the field to combine the elements or concepts from the prior art in the same way as in the claimed invention.

30. I understand that in order to show obviousness based on a single reference or a combination of references, a particular motivation to modify the reference or combine the teachings in the references, and a reasonable expectation of success must be shown.

31. I understand that the art must evidence a motivation to combine or modify the independently known elements to arrive at the claimed invention and an explanation as to how or why the references would be combined to arrive at the claimed invention to solve the particular problem. I understand that a challenger must show that there was a reason that would have prompted a person of ordinary skill in the relevant field to combine the elements in the way the claimed new invention does in an obviousness determination. I also understand that it is not enough to prove obviousness merely by a showing that the references were capable of combination without a suggestion of the desirability of the modification.

32. I further understand that there is no single way to define the line between true inventiveness on the one hand (which is patentable) and the application of common sense and ordinary skill to solve a problem on the other hand (which is not patentable). For example, market forces or other design incentives may be what produced a change, rather than true inventiveness. I may consider whether the

change was merely the predictable result of using prior art elements according to their known functions, or whether it was the result of true inventiveness. I may also consider whether there is some teaching or suggestion in the prior art to make the modification or combination of elements claimed in the patent. I may consider whether the innovation applies a known technique that had been used to improve a similar device or method in a similar way. I may also consider whether the claimed invention would have been obvious to try, meaning that the claimed innovation was one of a relatively small number of possible approaches to the problem with a reasonable expectation of success by those skilled in the art.

33. I also understand, however, that I must be careful not to determine obviousness using the benefit of hindsight; many true inventions might seem obvious after the fact. I should put myself in the position of a person of ordinary skill in the field at the time the claimed invention was made and I should not consider what is known today or what is learned from the teaching of the patent.

34. Finally, I understand that any obviousness rationale for modifying or combining prior art must include a showing that a person of ordinary skill would have had a reasonable expectation of success. I further understand that whether a proposed modification or combination of the prior art has a reasonable expectation of success is determined at the time the invention was made.

III. LEVEL OF ORDINARY SKILL IN THE ART

35. I understand that the Petition asserts that a person of ordinary skill in the art for the '699 Patent ("POSITA") would have the following level of knowledge:

> A bachelor's degree in computer science, computer engineering, electrical engineering, or related field in computing technology, and two years of experience with automatic speech recognition and natural language understanding, or equivalent education, research experience, or knowledge.

36. I understand that the Board adopted this definition for the relevant level of ordinary skill in the art in prior IPR2020-01367 challenging the parent patent of the '699 Patent, U.S. Patent No. 8,073,681. For purposes of this declaration, I do not offer an independent definition of a POSITA, and have applied the definition set forth above. I reserve the right to offer additional opinions regarding the level of ordinary skill in the art should this IPR be instituted.

IV. CLAIM CONSTRUCTION

37. Based on the representations in the Petition, I understand that the term "speech recognition engine" which appears in claims 7 and 18 of the '699 Patent was previously construed as "software or hardware that recognizes the words or phrases in the natural language utterance" based on agreement of the parties in a prior litigation. I have considered and applied this construction in my analysis and opinions set forth herein. I have otherwise applied the plain meaning of the challenged claims. At this time, I do not offer any additional opinions regarding the

proper construction of any of the claims, but I reserve the right to further address the proper construction of the claims if IPR is instituted.

V. PROSECUTION HISTORY

38. As set forth above, I have reviewed the file history of the '699 Patent, and below I provide an overview of the file history, including aspects of the file history that are relevant to my opinions set forth herein.

39. U.S. Patent Application No. 16/417,178 (the "'178 Application"), the application that resulted in the issuance of the '699 Patent, was filed on May 20, 2019. The '178 Application is a child application of U.S. Patent Application 11/580,926, filed on October 16, 2006, which issued in U.S. Patent No. 8,073,681. Ex. 1001, at Face, Item (60).

40. During prosecution of the '699 Patent, the Patent Office issued a non-final rejection on December 18, 2019, rejecting pending claims 1-20. Among other rejections, the examiner rejected pending claims 1-4, 8-14, and 18-20 under § 103 as obvious over US 2004/0044516 ("Kennewick") in view of US 2003/0167167 ("Gong"). Ex. 1002 at 219. The examiner also rejected claims 5 and 15 as unpatentable over Kennewick in view of Gong, and in further view of Perkins. Ex. 1002 at 245. The examiner argued that "Kennewick suggests . . . a computer-implemented method of facilitating natural language system responses," based on its disclosure of, for example "presenting results in a natural manner[.]" Ex. 1002 at

221. According to the examiner, Kennewick's disclosures further suggested "determining, by the computer system, an interpretation of the natural language utterance based on the identified context" based on Kennewick's disclosure that the "user asks 'what about flight one hundred and twenty too?' … parser and domain agent use flight information . . . and network information along with context to determine the most plausible interpretation among; flight 100 and flight 20 also, flight 100 and flight 22, flight 122, and the like." Ex. 1002 at 224. Thus, the examiner stated that Kennewick "describe[s] where a parser determines context of a user's question," and therefore "suggests where the parser-determined most likely context . . . is used to interpret the user's question/command . . . at least because the parser is described in paragraph 153 as the entity that determines the contexts, where the context is 'for the user's question and/or command', and where the parser is described in paragraph 155 as the entity that uses context to determine the most plausible determination." Ex. 1002 at 226. Further, the examiner found that "Kennewick suggests where the parser uses a determined most likely context to interpret a user's utterance . . . and paragraph 185 describes where responses are generated based on a user's query and can be in a natural language form . . . and paragraph 89 describes where responses are generated using components of the main unit 'computer system'." Ex. 1002 at 228. Further, the examiner stated that "Gong . . . suggests where the natural language system responses are adapted based on a

user's manner of speaking," as required by the pending claims because "Gong, like Kennewick, teaches using user profile information to determine how to communicate with a user," and further teaches "where a user's emotion/'affective state' can be determined from user speech by determining, for example, whether speech is faster/slower than the user's 'basic pattern'/'normal' . . . and also teaches where an agent . . . matches a user's speech style, which includes 'speech rate, pitch average, pitch range, and articulation." Ex. 1002 at 229-230. Thus, the examiner found that "Gong . . . suggests a combination where Kennewick's system further determines the user's 'affective state' [based on whether the user is speaking faster/slower than normal] and/or a user's speech style [which includes 'speech rate, pitch average, pitch range, and articulation'] from the user's spoken question/command . . . and then, when generating the response based on the interpretation of the user's natural language utterance . . . adapts the response to be communicated in a particular manner [more words when happy or less words/syllables when sad[.]" Ex. 1002 at 231.

41. Despite this discussion, the examiner stated that the prior art "does not teach or suggest the combination of all limitations in claims 1, 5, and 6 together, and in claims 11, 15 and 16 together." Ex. 1002 at 253. The examiner stated that Perkins "teaches where 'short-term' knowledge includes queries within the last n days, but describes where data older than n days is obsolete and is deleted (therefore not

obvious to combine with an additional reference that teaches long-term data based on utterances received more than n days ago).” Ex. 1002 at 253.

42. The examiner further rejected claims 1-20 on the ground of nonstatutory double patenting “as being unpatentable over claims 7-9 and 17-19 of U.S. Patent No. 9,015,049 . . . (and similarly claims 10-11 of US Patent 10,297,249, and claims 1 and 10 of US Patent 8,515,765, and claims 1 and 25 of US Patent 8,073,681) in view of Kennewick and Gao.” Ex. 1002 at 255.

43. After an applicant-initiated interview between applicant and examiner to discuss the Gong reference and potentially distinguishing features, the applicant eventually submitted an amendment to the claims accompanied by arguments and remarks made in amendment on March 18, 2020. Ex. 1002 at 419-433.

44. Applicant amended the independent claims as follows in representative claim 1:

LISTING OF CLAIMS:

1. **(Currently Amended)** A computer-implemented method of ~~facilitating~~ generating natural language system responses adapted based on a user's manner of speaking in a natural language utterance, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

receiving, by the computer system, a user input comprising a natural language utterance;

recognizing, by the computer system, one or more words or phrases from the natural language utterance;

identifying, by the computer system, a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance;

determining, by the computer system, an interpretation of the natural language utterance based on the identified context;

identifying, by the computer system, a manner in which the [[first]] natural language utterance was spoken based on short-term knowledge and long-term knowledge about how a user utters a request; and

generating, by the computer system, a response to the natural language utterance based on the interpretation and adapting the response based on the identified manner in which the [[first]] natural language utterance was spoken.

Ex. 1002 at 420.

45. In the applicant's remarks accompanying this amendment in response to the December 2019 Office Action, the applicant stated that claims 1-8 and 11-18 were amended and claims 21-30 were newly added. Ex. 1002 at 429. First, with respect to the double patenting rejection, the applicant indicated that "applicant will consider filing a Terminal Disclaimer once the remaining claim rejections have been withdrawn." Ex. 1002 at 429.

46. With respect to the rejections under § 103, the applicant argued that, in light of its claim amendments, the rejections were traversed. In particular, the applicant argued that "as amended, independent claims 1 and 11 now recite 'identifying, by the computer system, a manner in which the natural language utterance was spoken based on short-term knowledge and long-term knowledge about how a user utters a request' and 'generating, by the computer system, a response to the natural language utterance based on the interpretation and adapting the response based on the identified manner in which the natural language utterance was spoken.'" Ex. 1002 at 431. Thus, the applicant argued that "[n]either Kennewick nor Gong appear to teach or describe identifying a manner in which a natural language utterance was spoken **based on short-term knowledge and long-term knowledge about how a user utters a request**." Ex. 1002 at 431 (emphasis in original). Accordingly, the applicant argued that the pending claims should be allowed. Ex. 1002 at 433.

47. The applicant noted that, in light of the Examiner's finding that "[t]he prior art of record does not teach or suggest the combination of all limitations in claims 1, 5, and 6 together, and in claims 11, 15, and 16 together," the applicant introduced newly added independent claims 27 and 29. Ex. 1002 at 432. Claim 27 "includes the features of independent claim 1 and dependent claim 6, as well as features of intervening dependent claim 5." Ex. 1002 at 432. Likewise "claim 29

includes the features of independent claim 11 and dependent claim 16, as well as the features of intervening dependent claim 15." Ex. 1002 at 432.

48. In response to applicant's amendments and accompanying remarks, the Patent Office subsequently issued a Notice of Allowance on April 22, 2020, allowing claims 1-4, 7-14, and 17-26. Ex. 1002 at 439, 443. The Notice of Allowance included an Examiner Amendment as set forth in exemplary claim 1 below:

1. (**Currently Amended**) A computer-implemented method of generating natural language system responses adapted based on a user's manner of speaking ~~in a natural language utterance~~, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:

receiving, by the computer system, a user input comprising a natural language utterance;

recognizing, by the computer system, one or more words or phrases from the natural language utterance;

identifying, by the computer system, a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance;

determining, by the computer system, an interpretation of the natural language utterance based on the identified context;

accumulating, by the computer system, short-term knowledge based on one or more natural language utterances received during a predetermined time period, wherein the one or more natural language utterances received during the predetermined time period are related to a single conversation between a user and the computer system;

accumulating, by the computer system, long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period;

identifying, by the computer system, a manner in which the natural language utterance was spoken based on the short-term knowledge and the long-term knowledge ~~about how a user utters a request~~; and

generating, by the computer system, a response to the natural language utterance based on the interpretation and ~~adapting the response based on~~ the identified manner in which the natural language utterance was spoken.

Ex. 1002 at 445.

49. In its statement of reasons for allowance, the examiner stated that the applicant "amended the allowable subject matter of claims 5-6 and 15-16 into claims 1 and 11, respectively," and "[t]he prior art of record does not teach or suggest the combination of all limitations in claims 1, 5, and 6 together, and in claims 11, 15,

and 16 together." Ex. 1002 at 450. The examiner went on to discuss the prior art. Specifically, the examiner stated that "Perkins, as discussed above, teaches where 'short-term' knowledge includes queries within the las n days, but describes where data older than n days is obsolete and is deleted (therefore not obvious to combine with an additional reference that teaches long-term data based on utterances received more than n days ago)." Ex. 1002 at 451. The examiner further identified as relevant the portions of the Non-Final Office Action mailed June 11, 2019 for Application No. 16/417,173 as relevant. Ex. 1002 at 451. Specifically, the examiner stated that "Kargman (US 2005/0015256) . . . teaches determining an intended meaning for an utterance from short-term knowledge and long-term knowledge," but "does not, however, teach where context is determined based on short-term knowledge and long-term knowledge (i.e. intended meaning is not the same as context)." Ex. 1002 at 451. Further, the examiner stated that Morin (US 5,748,841) . . . suggests a short-term/immediate context and a time-out which suggests expiring short-term context information." Ex. 1002 at 451. Further, examiner noted that "Kennewick (US 2004/0044516) and Perkins (US 7,072,888) . . . suggest various limitations in claims 1 and 11 including determining context based on knowledge (Kennewick) where knowledge is short-term knowledge and where short-term knowledge is expired after a predetermined time period (Perkins teaches deleting query information older than n/30 days)," but "neither Morin nor Perkins teach/suggest where short-term

knowledge that has been expired is included in long-term information (expired information is simply deleted in Morin and Perkins), and Kennewick does not explicitly teach that both short-term and long-term information is used to determine context[.]" Ex. 1002 at 451-52.

50. Further, although the examiner noted that "prior art further teaches where information that is stored or has not been accessed in a short-term storage for a predetermined period of time is transferred/moved into a long-term storage (i.e. the short-term information is 'expired' into long-term information)," those references did not "teach expiring short-term knowledge that is based on natural language utterances received prior to the predetermined time period being expired and added to long-term storage." Ex. 1002 at 452.

51. In response to the April 22, 2020 Notice of Allowance, applicant submitted comments on the statement of reasons for allowance. Ex. 1002 at 578. First, applicant stated that "no one element or limitation in particular should be deemed to impart to, or be required for, patentability of the claims." Ex. 1002 at 579. Further, applicant submitted that "all claims are separately patentable from each other and are patentable for the subject matter specifically recited as a whole in each of those claims." Ex. 1002 at 579.

52. With respect to the examiner's statements on the teachings of the prior art, applicant did not "necessarily agree as to any specific position the Examiner sets

forth in the Statement, and explicitly disagrees with the Examiner's positions for at least the reasons set forth in Applicant's March 18, 2020 Amendment and January 22, 2020 Interview." Ex. 1002 at 579.

53. A notification was filed on August 5, 2020 regarding the issuance of U.S. Patent No. 10,755,699. Ex. 1002 at 587.

VI. OVERVIEW OF OPINIONS

54. It is my opinion that the Petition's asserted combination of SmartKom and Kobsa does not render obvious claims 1-22.

55. It is my opinion that the Petition's asserted combination of Barbara, Ross, and Kellner does not render obvious claims 1-22.

VII. GROUND 1

56. The first ground of the Petition asserts that the Challenged Claims are obvious over SmartKom (Ex. 1005) and Kobsa (Ex. 1004). I disagree that Ground 1 of the Petition demonstrates that the Challenged Claims are obvious over SmartKom and Kobsa.

A. Overview of the Asserted References

1. SmartKom (Ex. 1005)

57. SmartKom is a textbook comprising 38 different subparts, which are authored by different contributors from a list of over 70. SmartKom describes the "present the theoretical and practical foundations of multimodal dialogue systems

using the results of [the] large-scale project SmartKom as the background for [the] discussion." Ex. 1005, IX – XIV, 5. SmartKom provides a "comprehensive overview of the broad spectrum of results of the research conducted in SmartKom." Ex. 1005, VI. SmartKom was a large-scale project funded by the German Government designed to explore multimodal dialogue systems. The project spanned four years and the contributors and researchers included a consortium of multiple industrial companies, research institutes and universities. Thus, the sub-parts of SmartKom are directed to theories and principles in multimodal dialogue systems, and describes corresponding aspects of the so-called SmartKom system or architecture that was developed by dozens of different researchers at different times.

58. The Petition relies extensively on SmartKom's interaction module disclosure, which includes "indicator values" mapped to "models," a set of which is the output provided by the interaction module. The table of indicators is set forth below.

Table 1. List of indicators

Source	Description
Mimic recognizer	Mimically conveyed anger
Prosody recognizer	Prosodically conveyed anger
Mimic recognizer	Mimically conveyed joy
Prosody recognizer	Prosodically conveyed joy
Mimic recognizer	Mimically conveyed dilatoriness
Prosody recognizer	Prosodically conveyed dilatoriness
Speech recognition	Linguistically conveyed anger
Speech understanding	Ratio of unanalyzable words
Intention analysis	Overall score of the best hypothesis
Intention analysis	Difference in score between first and second best hypotheses
Intention analysis	Number of possible hypotheses (depth of lattice)
Speech recognition	Score of the speech recognizer
Gesture recognition	Score of the gesture analyzer
Speech understanding	Score of the language analyzer
Media integration	Score of multimodal integration
Discourse history	Score of the discourse module
Domain model	Score of the domain module
Intention analysis	Final score of the intention module
Intention analysis	Number of elements in the user input
Discourse history	Number of new (not previously mentioned) elements
Speech understanding	Number of elements addressed by speech
Gesture analysis	Number of elements addressed by gesture
Media integration	Number of elements addressed by speech and gesture
Intention analysis	Importance of speech recognition score for overall score
Intention analysis	Importance of gesture analysis score for overall score
Intention analysis	Importance of domain model score for overall score
Intention analysis	Importance of language understanding score for overall score
Intention analysis	Importance of discourse model score for overall score
Speech understanding	Relative number of sentence-like units in one turn
Speech understanding	Relative number of words in one turn
Speech understanding	Relative frequency of pronouns
Speech understanding	Relative frequency of verbs
Speech understanding	Relative frequency of adverbs
Speech understanding	Relative frequency of nouns
Speech understanding	Relative frequency of content words
Speech understanding, language generation	Relative frequency of content words appearing in the system output
Speech understanding, language generation	Relative frequency of content words not appearing in the system output

Ex. 1005, 322.

59. The "module delivers four sets of models":

Set	Description
Problem	Likelihood of a problem
Problem	Likelihood of an analysis problem
Problem	Discourse progress rate
Problem	Likelihood of the user being angry
Problem	Likelihood of the user being happy
UserKnowledge	Estimation of user familiarity with task
UserKnowledge	Estimation of user familiarity with system
Modality	Ratio of spoken input content
Modality	Ratio of gestural input content
Modality	Ratio of multimodal input content
ModalityContrastive	Ratio of contrastive usage of multimodal input
ModalityRedundant	Ratio of redundant usage of multimodal input
Linguistic	Adaptivity of user's lexical choices to former system output
Linguistic	Likelihood of long turns
Linguistic	Likelihood of long sentences
Linguistic	Ratio of pronoun usage
Linguistic	Ratio of verb usage
Linguistic	Ratio of adverb usage
Linguistic	Ratio of noun and verb usage

Ex. 1005 at 323.

2. Kobsa (Ex. 1006)

60. Kobsa (Ex. 1006) is also a textbook directed to documenting the results of an international workshop on user modeling that brought together 25 researchers. The stated intention of Kobsa is to form a "collection of unintegrated conference papers" and to "provide a rather coherent survey of the field of user modeling." Ex. 1006, V. Kobsa includes several Parts, each comprising numerous subparts, including a survey of user modeling, building user models, exploiting user models, and shortcomings of current models, prospects for the future. The Petition relies on broad ranging disclosure of the various subparts of Kobsa. For example, as part of Kobsa's survey of the field of user modeling, scattered throughout are descriptions of different systems and models directed to different implementations. For example,

Kobsa discloses a system called GRUNDY as an example of a system that uses an initial individual user model. GRUNDY is a system that recommends novels to people to read. Kobsa also discloses difference general user modeling systems, including the GUMS system, the TRUMP system, and the BGP-MS system.

B. The SmartKom + Kobsa Combination Fails to Describe a Coherent System that Includes all the Elements of the '699 Patent

61. The first ground of the Petition asserts that the Challenged Claims are obvious over SmartKom (Ex. 1005) and Kobsa (Ex. 1004). I have reviewed SmartKom and Kobsa, and each of these references are broad-ranging and reflect the work of many different researchers over time. In my opinion, Petitioners rely on disparate disclosures in SmartKom and Kobsa such that the proposed combination does not disclose or render obvious a coherent system that satisfies each element of the '699 Patent claims.

62. SmartKom, on the one hand, lists over 70 different contributors who, in 38 different sub-parts, "present the theoretical and practical foundations of multimodal dialogue systems using the results of [the] large-scale project SmartKom as the background for [the] discussion." Ex. 1005, IX – XIV, 5. The SmartKom reference is self-described as providing a "comprehensive overview of the broad spectrum of results of the research conducted in SmartKom." Ex. 1005, VI. The SmartKom project was developed over a four-year period, and the final SmartKom

demonstrator was presented in June 2003. Ex. 1005, 35. The plan for the project was defined by nine subprojects assigned to twelve different research groups. Ex. 1005, 31. This project was coordinated by the German Research Center for Artificial Intelligence, and comprised a consortium of six companies, one research institute, and three universities. Ex. 1005, V. As SmartKom itself describes, there is disclosure of "theoretical and practical foundations of multimodal dialogue systems," as well as disclosure of different components of concepts for implementation based on the SmartKom project developed over time, and corresponding examples of different possible implementations. Moreover, the Petition relies on different aspects of SmartKom developed by different people that are described for different implementations. The Petition fails to tie these different aspects together to show that they form a system in accordance with the challenged claims.

63. While SmartKom describes the implementation of different aspects of a system or architecture for different applications, Kobsa, on the other hand, describes itself as a "collection of unintegrated conference papers" from a conference of 25 researchers intended to provide a survey of the field of user modeling. Ex. 1006, V-VI. Thus, Kobsa is meant to broadly describe the aims and the scope of the field of user modeling and describes different systems that have been implemented and their results to achieve a broad description of the field. Kobsa

describes the scope of the field of user modeling by reference to many different systems that have been implemented.

64. As set forth in several examples below, in my opinion, the Petition relies on disparate disclosure and fails to identify a system in accordance with the claimed invention.

65. First, because of the broad-ranging nature of SmartKom and Kobsa, each of which was developed over time by different people, neither SmartKom nor Kobsa describe a single coherent system. The Petition cherry picks pieces from throughout each reference to satisfy different claim limitations, but fails to sufficiently explain how those cherry-picked disclosures would come together in a coherent system or method that includes all the elements of the challenged claims. The challenged claims recite a method or system for "generating natural language system responses adapted based on a user's manner of speaking" comprising "one or more physical processors," executing one or more computer program instructions, which configure the processors when executed to "receive a user input comprising a natural language utterance," "recognize one or more words or phrases from the natural language utterance," "identify a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance," "determine an interpretation of the natural language utterance based on the identified context," "accumulate short-term knowledge based on one or more

natural language utterances received during a predetermined time period, wherein the one or more natural language utterances received during the predetermined time period are related to a single conversation between a user and the computer system," "accumulate long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period," "identify a manner in which the natural language utterance was spoken based on the short-term knowledge and the long-term knowledge," and "generate a response to the natural language utterance based on the interpretation and the identified manner in which the natural language utterance was spoken." Ex. 1001 at cl. 12. This combination of limitations in the challenged claims are not arbitrary or ad hoc, but instead represent a non-obvious integration of components of methods to address a specific problem.

66. The Petition, however, attempts to rebuild the claimed invention by cherry-picking unrelated pieces from SmartKom and Kobsa using the challenged claims as a guide. Using this hindsight-driven reasoning is not how design of dialogue systems works, and it is not how a POSITA would approach the problems identified in the '699 Patent at the time. Specifically, the Petition extracts individual features from the separate referenced systems, or generalized concepts applicable to all systems and retrofits them into a patchwork of the '699 Patent without

acknowledging the vast combinatorial design space inherent in dialog system architecture.

67. To illustrate this technically, Claim 12 of the '699 Patent recites "a system for generating natural language system responses adapted based on a user's manner of speaking" comprising "one or more physical processors programmed with one or more computer program instructions" configured to: (i) "receive a user input comprising a natural language utterance," (ii) "recognize one or more words or phrases from the natural language utterance," (iii) "identify a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance," (iv) "determine an interpretation of the natural language utterance based on the identified context," (v) "accumulate short-term knowledge based on one or more natural language utterances received during a predetermined time period, wherein the one or more utterances received during the predetermined time period are related to a single conversation between a user and the computer system," (vi) "accumulate long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period," (vii) "identify a manner in which the natural language utterance was spoken based on the short-term knowledge and the long-term knowledge," and (viii) "generate a response to the natural language utterance

based on the interpretation and the identified manner in which the natural language utterance was spoken." Ex. 1001, cl. 12.

68. A POSITA designing a system providing a voice user interface would understand that each of these limitations could correspond to many possible modules that could be configured in many different ways. Assuming that, for instance, each limitation can be configured at least ten different ways (via alternative algorithmic approaches, interface protocols, or processing assumptions). The resulting potential design permutations exceed 1,000,000 cases to navigate. This design complexity underscores the improbability of arriving at the exact limitations of the '699 Patent while reviewing references such as SmartKom and Kobsa with far-reaching disclosures detailing many different concepts and systems absent the hindsight of the claims as a guide. In my opinion, even if the Petition identifies the individual building blocks as existing in the literature in either SmartKom or Kobsa, the Petition fails to show how those building blocks are organized, constrained, and functionally tied to one another to reach the claimed invention recited in the '699 Patent.

69. For example, the Petition relies on SmartKom's disclosure of each of a "dialogue model," that includes "what has been said, by whom," a "situation model, which records 'time, place' of an utterance," and "user state knowledge," such as whether a human is "hesitant" as examples of the alleged "short-term knowledge" that is "accumulated" in SmartKom's disclosure. Petition at 29-32. Petitioners

identify the "prosody module" as the component in SmartKom that generates the user state lattice. Petition at 32. The prosody module is described in the chapter titled "The Prosody Module" in SmartKom, which is authored by Viktor Zeibler, Johann Adelhart, Anton Batliner, Caremn Frank, Elmar Noth, Rui Ping Shi, and Heinrich Niemann. Ex. 1005 at 139. The dialogue model and the situation model, on the other hand are described in the chapter of SmartKom titled "In Context: Integrating Domain- and Situation-Specific Knowledge, which is authored by Robert Porzel, Iryna Gurevych, and Rainer Malak. Ex. 1005, 269.

70. Further, the claims require that the short-term knowledge is accumulated "based on one or more natural language utterances received during a predetermined time period" related to a "single conversation." For these limitations, the Petition sets forth that "SmartKom's discourse state is 'based on the three-tiered context representation presented in Luperfoy (1992)," which is a separate paper referenced in SmartKom authored by Luperfoy. Petition at 33. The Petition relies on Luperfoy's disclosure of "discourse objects" "decay[ing] as a function of attentional focus" to support the assertion that SmartKom "acknowledges this decay of information." Petition at 34.

71. Next, the Petition identifies SmartKom's disclosure of a "user model" storing "properties of the interlocutors," and Kobsa's disclosure of accumulating knowledge about the user "either explicitly or implicitly" as disclosures of long-term

knowledge. Petition at 35. The Petition also identifies SmartKom's disclosure of an "interaction model that 'computes information on the user' as well as Kobsa's General User Modeling System ("GUMS") "which is designed for building long term models of individual users" as alternative "techniques for 'accumulating long-term knowledge.'" Petition at 36.

72. The claims of the '699 Patent further require that the "short-term knowledge" and "long-term knowledge" described above are used to identify a "manner in which the natural language utterance was spoken[.]" Ex. 1001 at cl. 12. Petitioners identify an "interaction module" for these limitations directed to identifying "a manner in which the natural language utterance was spoken based on the short-term knowledge and the long-term knowledge." Petition at 37. The interaction module is described in the SmartKom chapter titled "Emotional Analysis and Emotion-Handling Subdialogues," authored by Michael Stret, Anton Batliner, and Thomas Portele. Ex. 1005 at 317. Specifically, Petitioners rely on the disclosure that the interaction module collects and evaluates "indications of problematic situations and the emotional state of the user" from several sources. Petition at 37 (quoting Ex. 1005 at 320). And further that the interaction module "'operates by analyzing a set of possible indicators' (listed in Table 1(322) that 'have values between 0 and 1.'" Petition at 38 (quoting Ex. 1005 at 321). Without explaining the support for its statement, the Petition asserts that Table 1, showing the set of possible

indicators below, shows that "the indication [sic] are derived from data collected from the prosody, user model, discourse model, and domain model." Petition at 38.

Table 1. List of indicators

Source	Description
Mimic recognizer	Mimically conveyed anger
Prosody recognizer	Prosodically conveyed anger
Mimic recognizer	Mimically conveyed joy
Prosody recognizer	Prosodically conveyed joy
Mimic recognizer	Mimically conveyed dilatoriness
Prosody recognizer	Prosodically conveyed dilatoriness
Speech recognition	Linguistically conveyed anger
Speech understanding	Ratio of unanalyzable words
Intention analysis	Overall score of the best hypothesis
Intention analysis	Difference in score between first and second best hypotheses
Intention analysis	Number of possible hypotheses (depth of lattice)
Speech recognition	Score of the speech recognizer
Gesture recognition	Score of the gesture analyzer
Speech understanding	Score of the language analyzer
Media integration	Score of multimodal integration
Discourse history	Score of the discourse module
Domain model	Score of the domain module
Intention analysis	Final score of the intention module
Intention analysis	Number of elements in the user input
Discourse history	Number of new (not previously mentioned) elements
Speech understanding	Number of elements addressed by speech
Gesture analysis	Number of elements addressed by gesture
Media integration	Number of elements addressed by speech and gesture
Intention analysis	Importance of speech recognition score for overall score
Intention analysis	Importance of gesture analysis score for overall score
Intention analysis	Importance of domain model score for overall score
Intention analysis	Importance of language understanding score for overall score
Intention analysis	Importance of discourse model score for overall score
Speech understanding	Relative number of sentence-like units in one turn
Speech understanding	Relative number of words in one turn
Speech understanding	Relative frequency of pronouns
Speech understanding	Relative frequency of verbs
Speech understanding	Relative frequency of adverbs
Speech understanding	Relative frequency of nouns
Speech understanding	Relative frequency of content words
Speech understanding, language generation	Relative frequency of content words appearing in the system output
Speech understanding, language generation	Relative frequency of content words not appearing in the system output

73. I disagree that this statement is supported by the disclosure of SmartKom. There is no temporal aspect in the list of the indicators above. For example, there is no support from Table 1 above that any indicator is derived from

"user model" data collected, nor that the alleged "user model" used to derive an indicator constituted knowledge accumulated prior to a predetermined time. There is no support that the discourse model includes information accumulated during a predetermined time period.

74. The "interaction module delivers four sets of models" shown in the list of models shown below.

Table 2. List of models

Set	Description
Problem	Likelihood of a problem
Problem	Likelihood of an analysis problem
Problem	Discourse progress rate
Problem	Likelihood of the user being angry
Problem	Likelihood of the user being happy
UserKnowledge	Estimation of user familiarity with task
UserKnowledge	Estimation of user familiarity with system
Modality	Ratio of spoken input content
Modality	Ratio of gestural input content
Modality	Ratio of multimodal input content
ModalityContrastive	Ratio of contrastive usage of multimodal input
ModalityRedundant	Ratio of redundant usage of multimodal input
Linguistic	Adaptivity of user's lexical choices to former system output
Linguistic	Likelihood of long turns
Linguistic	Likelihood of long sentences
Linguistic	Ratio of pronoun usage
Linguistic	Ratio of verb usage
Linguistic	Ratio of adverb usage
Linguistic	Ratio of noun and verb usage

Petition at 40. The Petition states that the UserKnowledge set of model values, which "reflects the assumed task and paradigm knowledge of the user," "is obtained from the user model, which is 'long-term knowledge.'" Petition at 40. Likewise, the Petition argues that the Linguistic and Modality, and the problem models described

as the "likelihood of the user being angry" or "likelihood of the user being happy" "include situational knowledge, discourse state, and user state information which are each 'short-term knowledge'." Petition at 41. First, I disagree that either of these arguments are supported by any disclosure of SmartKom identified by the Petitioners. For example, the Petition fails to identify any functional link between, for example, the situational model and the "ratio of spoken input content, ratio of gestural input content, ratio of multimodal input content, adaptivity of user's lexical choices to former system output, the likelihood of long turns, the likelihood of long sentences, the ratio of pronoun usage, the ratio of verb usage, the ratio of adverb usage, and the ratio of noun and verb usage." Petition at 40. Further, the Petition identifies no functional link between the discourse state that would need to reflect "short-term knowledge" accumulated during a "predetermined time period" and any of the linguistic sets of models. In my opinion, the Petition's assertion that the interaction module is functionally linked to the prior disclosures of SmartKom that Petitioners identify as satisfying the limitations requiring "short-term knowledge" and long-term knowledge" is unsupported.

75. Further, these are examples of different aspects of SmartKom, that, while funded under the same research umbrella, were developed and described by different researchers, and described using different implementations.

76. In my opinion, the fact that the different individual components or modules disclosed in SmartKom that Petitioners rely upon for different claim limitations are authored by different researchers and directed to different implementations demonstrates that SmartKom does not disclose a single, coherent system, which serves as the backdrop for my opinions that the Petition fails to identify a coherent system that includes all the limitations of the asserted claims. Specifically, as described above, SmartKom describes generalized principles of dialogue systems (e.g., user models) as well as the SmartKom the implementation and application for discrete components. Kobsa describes a survey of the field of user modeling, which includes the description of dozens of different, independent systems that necessarily have different components and different implementations. A POSITA would not be motivated to carefully select the different pieces of SmartKom and Kobsa that the Petition selects and expect to reach the claimed invention because there would be endless possibilities for implementation when selecting individual components for the resulting system.

77. Thus, in my opinion, the Petition fails to sufficiently explain how the cherry-picked disclosures of SmartKom and Kobsa would come together in a coherent system or method that includes all the elements of the challenged claims.

78. The first example of this is the Petition's reliance on both the disclosure of SmartKom that is focused on generalized descriptions of all dialogue systems, as

well as its disclosure of SmartKom's implementation system. For example, the claims require accumulating "long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period." *See, e.g.*, cl. 12. For the limitations directed to "long-term knowledge", the Petition identifies SmartKom's disclosure of a user model that provides (i) "properties of the interlocutors". Petition at 35. The Petition also relies on Kobsa for the proposition that the knowledge stored in the user model "can be acquired 'either explicitly or implicitly.'" Petition at 35-36 (citing Ex. 1006 at 416). But SmartKom's teaching of the "properties of the interlocutors" as the content stored in the user model does not refer to any component in any system described by SmartKom. Instead, this disclosure is a generalized "broad categorization of the types of context relevant to spoken dialogue systems" and their "content and respective knowledge stores" applicable to all systems. Ex. 1005, 274. SmartKom discloses that ***in general*** in human-to-human interactions, interlocutionary context is important in the process of employing "context-dependent resources" by the addressee to extract the meaning from the utterance that the speaker wanted to convey. This disclosure therefore does not refer to the SmartKom system that the Petition relies upon to satisfy other claim limitations. Instead, the disclosure simply refers to generalized foundations for all types of speech processing, including human-to-human conversations.

79. For other limitations the Petition relies on different aspects of what is described as the "SmartKom system" or "SmartKom architecture" described in various of the 38 different sub-parts of SmartKom. For example, the independent claims recite accumulating short-term knowledge and accumulating long-term knowledge, and identifying "a manner in which the natural language utterance was spoken ***based on*** the short-term knowledge and the long-term knowledge". Ex. 1001, cl. 12 (emphasis added). Similarly, the independent claims require determining "generating . . . a response to the natural language utterance ***based on the interpretation and the identified manner of speaking*** in which the natural language utterance was spoken." Ex. 1001, cl. 12 (emphasis added). For these limitations, Petitioners rely on specific implementations described in SmartKom. Namely, the Petition relies on SmartKom's "interaction module" which the Petition argues delivers a "UserKnowledge set of model values" that "reflects the assumed task and paradigm knowledge of the user" that the Petition argues "is obtained from the user model, which is 'long-term knowledge.'" Petition at 38. Where Petitioners rely on broad, generalized principles as I described above, such as the "properties of interlocutors," rather than implementation aspects of the SmartKom system for accumulating "long-term knowledge," but then subsequently rely on specific aspects or components of the SmartKom system for the claimed "long-term knowledge" in the limitations that require determining "a manner of speaking" "based on the . . .

long-term knowledge" limitations, the Petition fails to identify a coherent system in accordance with the claims.

80. Furthermore, the Petition identifies three different possible disclosures in the SmartKom-Kobsa combination to disclose or teach "long-term knowledge": (i) the generalized disclosure of a domain model in SmartKom; (ii) the generalized disclosure of a user model in SmartKom; and (iii) Kobsa's disclosure of "long-term" user models that "describe relatively static characteristics of users." Petition, 34-36. The Petition subsequently argues that SmartKom's disclosure of an "interaction model" in the SmartKom system is a technique that meets the "accumulating" "long-term knowledge" limitation. Petition at 36. The interaction model in SmartKom operates in the SmartKom system to "compute information on the user" to be used for example, "to adapt the interaction style of the system to the interaction style of the user." Ex. 1005, 287. However, as set forth above, the claims also require that the claimed "long-term knowledge" is used to identify "a manner in which the natural language utterance was spoken." Ex. 1001, cl. 1, 12. The Petitioners fail to cite to any disclosure in SmartKom indicating that information computed on the user to adapt the interaction style (as performed by the interaction model) is used to identify a manner in which the natural language was spoken by the separately described ***interaction module*** that "collects and evaluates indications of emotions, problematic situations and other aspects of the interaction." Ex. 1005 at 321.

81. Accordingly, in my opinion, the Petition cherry picks these disclosures to satisfy different limitations, but fails to explain how each of these identified disclosures relate to one another such that the asserted combination discloses or teaches a coherent system that includes all the elements of the challenged claims.

82. Furthermore, in my opinion, the combination of SmartKom and Kobsa fails to identify a coherent system in accordance with the claimed invention because Petitioners rely on different, and in many instances disparate, disclosures to satisfy the same claim limitations, without explaining how the combination of both disclosures would meet the claims.

83. For example, Petitioners allege that both the "interaction model" disclosed in SmartKom *and* the "General User Modeling System," which "is designed for building long term models of individual users" and can be "persistent," in Kobsa satisfy the limitations requiring "accumulating long-term knowledge." Petition at 36-37; Ex. 1006, 10-11, 417.

84. On the one hand, the interaction model in SmartKom operates in the SmartKom system to "compute information on the user" to be used for example, "to adapt the interaction style of the system to the interaction style of the user." Ex. 1005, 287. On the other hand, Kobsa's General User Modeling System ("GUMS") for building long term models of individual users is aimed, not at interaction preferences of the user, but at representing the "knowledge and beliefs of

individuals." Ex. 1006, 417. Kobsa's GUMS is acquired by allowing the application for which it is used to select an initial stereotype for the user, which is another significant difference from SmartKom's interaction modeling. Ex. 1006, 418. Critically, Kobsa's GUMS system organization is such that the "user modeling system does not have access to the interaction between the application and the user." Ex. 1006, 418. The Petition fails to explain this inconsistency between the interaction modeling identified from SmartKom and the GUMS identified in Kobsa, and does not explain how Kobsa's GUMS system could be implemented with SmartKom's interaction modeling. Accordingly, it is my opinion that none of these disclosures identified as satisfying the same challenged claim limitations fit together as describing the same concept or functionality, nor does the Petition explain how, for example, the General User Modeling System would be incorporated with the interaction modeling in the SmartKom system. This is therefore an example where Petitioners have identified disparate disclosures from SmartKom and Kobsa, and failed to explain how the combination of these disclosures would satisfy the challenged claims.

85. Finally, in my opinion, the Petition does not identify a coherent system based on its reliance on the disclosure of Luperfoy (Ex. 1012). Petition at 33-34. The claims require "accumulat[ing . . .] short-term knowledge based on one or more natural language utterances received during a predetermined time period." Ex. 1001,

cl. 12. The Petition argues that the SmartKom-Kobsa combination discloses or suggests these limitations requiring a "predetermined time period" because "SmartKom's discourse state is 'based on the three-tiered context representation presented in Luperfoy (1992).'" Petition at 33 (quoting Ex. 1005 at 239). As an initial matter, I understand that Luperfoy is not a part of Petitioners' Ground 1. Moreover, I disagree with the Petition's characterization of SmartKom's discourse state, and its resulting reliance on Luperfoy's description of "decay" of discourse objects "as a function of attentional focus[.]" Petition at 33 (quoting Ex. 1012 at 25).

86. I disagree with Petitioners' characterization of SmartKom's discourse state because SmartKom's approach to discourse state representation is actually a "compilation of three [different] items," only one of which relates to Luperfoy. Ex. 1005, 239-240 (emphasis added). Specifically, SmartKom explains that its "approach for representing the discourse state is a compilation of three items:

- A unified modality-independent generic representation based on the three-tiered context representation presented in Luperfoy (1992). The three tiers are the modality layer, the discourse object layer and the domain object layer.
- The context structuring ideas of Salmon-Alt (2000). In particular, we deploy compositional information of collections.

- A simple and robust thematic-based generic focus mechanism, e.g., Carter (2000). Our focus mechanism contains local and global structure organized around different focus spaces and accessible discourse objects of the focus space, respectively." Ex. 1005 at 239.

87. Thus, the representation of the discourse state in SmartKom is much more complex than just what is disclosed in Luperfoy, and relying on Luperfoy alone is not an accurate representation of the functionality disclosed in SmartKom that Petitioners rely on. Moreover, SmartKom discloses only that the discourse representation uses the "modality-independent generic representation" based on the "modality layer, the discourse object layer and the domain object layer" of Luperfoy. Ex. 1005 at 239. In other words, SmartKom merely takes from Luperfoy its idea of a "three-tiered framework" of discourse. Luperfoy explains that (i) "the first tier holds a linguistic analysis of surface forms" such that "there is a unique object (called a linguistic object or LO) for each linguistic referring expression or non-linguistic communicative gesture"; (ii) "[t]he intermediate tier is the discourse model, a tier with one unique object corresponding to each concept or guise of a concept, being discussed in the dialogue," with objects referred to as "pegs"; and (iii) the "third tier is the knowledge base (KB) that describes the belief system of one agent in the dialogue[.]" Ex. 1012 at 22. SmartKom does not, however, disclose that its system relies on Luperfoy's disclosure of "information decay," as the Petition

suggests. Ex. 1012 at 24-25; Petition at 33-34. Luperfoy's "information decay" is one of four "requirements of discourse representation based on the cognitive limitations and pressures faced by any dialogue participant." Ex. 1012 at 23. Information decay in Luperfoy refers to "the agent participating in a dialogue experiences information decay over the course of the conversation." Ex. 1012 at 24. Further Luperfoy discloses that "information from the linguistic, discourse, and belief system tiers decay at different rates and in response to different cognitive forces/limitations" such that "LOs become old and vanish at an approximately linear rate as a function of time counted from the point ," and "discourse pegs decay as a function of attentional focus, so that as long as an individual or concept is being attended to in the dialogue, the discourse peg will remain near the top of the focus stack and available as a potential discourse sponsor for upcoming dependent referring expressions." Ex. 1012 at 24-25. SmartKom only indicates that its discourse representation is based, along with two other items from other papers, on Luperfoy's three-tiered context representation.

88. Accordingly, I disagree with Petitioners' assertion that SmartKom "acknowledges [Luperfoy's] decay of information" based on the quote that" [f]or longer dialogues (more than half an hour of discourse), the discourse memory runs out of memory" and in some cases, "even for most humans," it is "necessary to forget information." Petition at 34 (quoting Ex. 1005 at 252). First, this assertion is

inconsistent with the "manipulation" of the ontology-based discourse state representation disclosed in SmartKom, which is based on "[t]wo fundamental operations": "unification and overlay (Alexandersson et al., 2006; Alexandersson and Becker, 2003)." This manipulation mechanism is entirely different from Luperfoy's mechanism because SmartKom's manipulation mechanism uses a single operation called overlay wherein discourse objects are accumulated but never expired or forgotten. Ex. 1005, 250-251. SmartKom's overlay approach describes overwriting parts of the old information with new information while keeping the old information still consistent with the new information. Ex. 1005, 246, 250-251. Employing the overlay mechanism, SmartKom explicitly discloses that this implementation does not "forget" or "discard" contextual information, but actually ***overaccumulates*** contextual information with "no control mechanism." Ex. 1005, 250-251 (emphasis added). SmartKom explains that "forgetting" information for longer dialogues "is still an open and challenging issue" that the disclosures of SmartKom does not address and "will have to [be] address[ed] in the future." Ex. 1005 at 252.

89. Thus, based on my review of SmartKom, this disclosure explicitly teaches that this implementation of managing the discourse state is not capable of limiting its accumulation of knowledge to a predetermined period of time, which is required by the claims. Because this overlay procedure is directly at odds with

Luperfoy's teachings relied on by the Petition, I disagree that Petitioners have properly relied on Luperfoy as accurately describing the functionality implemented in SmartKom that the Petition identifies for the remaining claim limitations. At bottom, based on my review of SmartKom, I have not seen any, and the Petition does not point to any, disclosure of decaying discourse objects as a function of attentional focus as described in Luperfoy. Thus, the Petition does not identify a coherent system.

90. The Petition also relies on Kobsa's disclosure of "GRUNDY" system to satisfy these same limitations. Based on Kobsa's disclosure that, "at the end of a dialog session," the system "records ***all information about the user*** inferred from his/her dialog in a corresponding file," Petitioners allege that SmartKom-Kobsa combination discloses that "short-term knowledge" is accumulated during a predetermined time." Petition at 34 (citing Ex. 1006, 10-11) (emphasis added). First, while the Petition relies on two different disclosures from different references to satisfy the same limitations, the Petition fails to explain how or why either, or both, would be combined to reach a working system in accordance with the Challenged Claims. Indeed, the Petition states that the "short-term knowledge" could be accumulated during two alternative time periods: "the shorter of (1) the duration of the conversation/dialog or (2) the size of short term storage (e.g., 30 minutes)." Petition at 34. Second, the two alleged alternatives are at odds with one another.

Even if SmartKom's system did "acknowledge[] decay of information" by forgetting information after half an hour of discourse (which it does not), Petitioners do not make any attempt to explain how or why a POSITA would reconcile this with Kobsa's recording of "***all information about the user*** inferred from his/her dialog behavior." Ex. 1006, 10-11 (emphasis added). Accordingly, in my opinion, because the Petition identifies disparate disclosures to satisfy the same limitations, Petitioners have not identified a coherent system in accordance with the claimed invention. Further, as I explain below, I disagree that a POSITA would have been motivated to combine these references.

C. The Petition Fails to Show that the SmartKom + Kobsa Combination Discloses or Teaches "accumulating" "short-term knowledge based on one or more natural language utterances during a predetermined time period"

91. The Challenged Claims require "accumulat[e]/[ing]" "short-term knowledge based on one or more natural language utterances during a predetermined time period." Ex. 1001, cl. 1, 12. In my opinion, the Petition fails to show that SmartKom in combination with Kobsa discloses or renders obvious this limitation.

92. As I set forth above, the Petition wrongly relies on the disclosure of Luperfoy, or alternatively, the disclosure of Kobsa, to support its assertion that the SmartKom-Kobsa combination discloses that "short-term knowledge" is accumulated for a dialog between the user and system over the shorter of (1) the

duration of the conversation/dialog or (2) the size of short term storage (e.g., 30 minutes)—'a predetermined period of time.'" Petition at 34. I disagree.

93. I disagree that Petitioners' reliance on the alternative disclosures of SmartKom and Kobsa discloses or teaches accumulating short term knowledge during a predetermined time period. As I explained above, SmartKom does not disclose "forgetting information," as the Petition asserts, which is the alleged disclosure of SmartKom relied upon the Petition to demonstrate that "accumulating short-term knowledge" during a "predetermined time." Petition at 34. SmartKom explicitly discloses that its system is not capable of the alleged "forgetting" and instead "overaccumulates" information. Ex. 1005 at 252. Thus, I disagree that SmartKom discloses or teaches that knowledge is accumulated over "the size of short term storage (e.g., 30 minutes)" as suggested by the Petitions.

94. Further, I disagree that Petitioners have shown that Kobsa's disclosure of recording "all information about the user inferred from his/her dialog behavior in a corresponding file" "at the end of a dialog session" satisfies these limitations. Petitioners do not explain why "all information about the user" corresponds to "short-term knowledge" at least because Petitioners rely on the exact same disclosure to satisfy "long-term knowledge." Petition at 36 ("Kobsa confirms that the user model is persistent (long-term), teaching that, 'at the end of a dialog session', the system 'records all information about the user inferred from his/her

dialog behavior in a corresponding file."). Moreover, Petitioners relied upon "situational," "user-state knowledge" such as being "hesitant," and the discourse state in SmartKom as types of "short-term knowledge." Petition at 29-32. The fact that "all information about the user" is recorded does not demonstrate that the Kobsa system is "accumulating" what Petitioners have identified as short-term knowledge during the "dialog session." Accordingly, I disagree that the Petitions demonstrates that the SmartKom-Kobsa combination discloses or suggests these limitations.

D. A POSITA Would Not be Motivated to Combine SmartKom and Kobsa to Reach the Claimed Invention

95. In my opinion, the SmartKom and Kobsa combination fails to render obvious the challenged claims because a POSITA would not have been motivated to combine SmartKom and Kobsa to reach the claimed invention.

96. I understand the Petition asserts that, among other reasons, a POSITA would have been motivated to combine Kobsa's teachings regarding user models and user modeling with SmartKom's dialogue system. Petition at 7. Petitioners' alleged motivation to combine SmartKom and Kobsa relies on SmartKom's mention of the "use of a user model and/or profiles/preferences in its context modeling and presentation planning," and mention of "storing preferences stated by a user during a conversation," but its failure to disclose further details about these user models.

Petition at 7. I disagree that a POSITA would have been motivated to combine Kobsa with SmartKom to reach the claimed invention.

97. Specifically, as I explained above, Petitioners assert that Kobsa's disclosure of General User Modeling System which discloses a "persistent" user model and SmartKom's disclosure of a domain model, user model, or interaction model satisfies the "long-term knowledge" functionality in the claims. As set forth above, the Petition offers no explanation regarding how a POSITA would combine these disclosures in accordance with the claims. For example, there is no explanation as to how Kobsa's General User Modeling System would be integrated into SmartKom's dialogue system in the user model.

98. It is my opinion that, even if the Petition provided an explanation, due to the differences between Kobsa and SmartKom's disclosures, a POSITA would not be motivated to combine Kobsa's General User Modeling System with SmartKom's user model, or interaction model to meet the "long-term knowledge" limitations. With respect to the implementation of a user model in SmartKom, SmartKom discloses that "general user model information is supplied via external sources, e.g., via a user's *SmartCard*," while the SmartKom system only "actively monitors" "interaction preferences of the users." Ex. 1005 at 276 (emphasis in original). Accordingly, apart from interaction preferences, SmartKom discloses that there is no active monitoring of user information. An external plug-in user model

with no active monitoring of user information other than interaction preferences in SmartKom's system is completely inconsistent with the "persistent" user model in Kobsa, which "records ***all information*** about the user inferred from his/her dialog behavior" "at the end of a dialog session." Petition at 22; Ex. 1006 at 10-11.

99. Petitioners have not pointed to any disclosure in SmartKom, and I have not seen any disclosure, explaining that an external plug-in user model would or could be "persistent" or continually updated to meet the "accumulating long-term knowledge" limitations of the '699 Patent claims. Petitioners have not explained how the persistent user model of Kobsa could be incorporated into a plug-in SmartCard, for example, to be used in SmartKom's system such that the combined system would work in accordance with the claims.

100. Furthermore, as I explained above, to the extent Petitioners allege that actively monitored interaction preferences would meet these limitations, the Petition does not explain how Kobsa's user model would be incorporated with the actively monitored interaction preferences (e.g., though interaction modeling) of SmartKom in the same combined system. For the foregoing reasons, it is my opinion that the Petition fails to demonstrate that a POSITA would be motivated to combine SmartKom and Kobsa in accordance with the claimed invention.

101. In addition, as I explained above, although I disagree with Petitioners' characterization of SmartKom such that it "acknowledges [the] decay of

information," and that it is "necessary to forget information" in SmartKom's system, if SmartKom did disclose this, it would be at odds with Kobsa's disclosure relied upon the Petition for the same limitations that, "at the end of a dialog session," the system "records all information about the user." Petition at 34. Forgetting information for purposes of optimizing memory in future systems, as Petitioner allege is disclosed by SmartKom, is inconsistent with Kobsa's disclosure that "all information" about the user would be recorded to a separate file at the conclusion of a dialogue session. Samsung does not explain why a POSITA would have cobbled together these inconsistent disclosures of SmartKom and Kobsa with a reasonable expectation of success that the resulting combination would be a working system that expires one or more items of short-term knowledge. Thus, in my opinion, the Petition has not shown that a POSITA would be motivated to combine these references with an expectation of success of reaching the claim limitations including accumulating long-term and short-term knowledge in relation to a "predetermined time."

102. Finally, Petitioners fail to explain how and why a POSITA would have combined Kobsa's disclosure of GRUNDY (Ex. 1006, 10-11), and SmartKom's disclosure of the discourse memory to reach the claimed invention. I disagree that a POSITA would be so motivated. As set forth above, Petitioners rely on Kobsa's GRUNDY system and the disclosure of recording information inferred about the

user at the end of a dialog session. Petition at 34, 36. However, Kobsa explicitly states that the GRUNDY system, which is an exemplary system implementing an initial user model, suffers from "legal, social and ethical problems of security and privacy." Ex. 1006, 10-11. Thus, Kobsa further explains that systems like GRUNDY with an initial user model are rarely "developed and implemented." Ex. 1006, 11. Thus, a POSITA would have understood Kobsa as discouraging implementing such a system due to their problems, and less frequent implementation of systems with models that function by recording all information about the user at the end of dialog sessions. Ex. 1006, 10-11. Thus, if a POSITA were to combine SmartKom and Kobsa, in my opinion, the POSITA, would have been guided by Kobsa to select any of the many alternative systems described in Kobsa, such as a system that "possesses no previously acquired information about individual users at the beginning of a dialog" because they were, at the time, "developed much more often" because they do not suffer from the "legal, social, and ethical problems of security and privacy" as initial user model systems, and, at the time, Kobsa disclosed that "the probability that a user will consult the same system more than once [was] quite low." Ex. 1006 at 11.

103. Accordingly, in my opinion, a POSITA would not be motivated to combine SmartKom and Kobsa in the manner proposed by the Petition.

VIII. GROUND 2

104. Ground 2 of the Petition relies on the combination of Barbara (Ex. 1007), Ross (Ex. 1008), and Kellner (Ex. 1023). In my opinion, the Petition fails to show that the asserted prior art in Ground 2 renders obvious the claims 1-22. I provide an overview of Barbara, Ross, and Kellner below.

A. Overview of the Asserted References

1. Barbara (Ex. 1007)

105. The portions of Barbara relied upon by Petitioners are directed to systems and methods for creating a voice interface where the end system responds to spoken natural language commands. Ex. 1007, ¶[0082]. Barbara discloses that the "electronic version of the spoken word is checked for accuracy, corrected and interpreted to determine what the user wanted." Ex. 1007, ¶[0082]. In Barbara, when an utterance has been detected, "a request packet is then generated, the request comprising the best guesses of the voice recognition engine as to what was said, with some indications of their likelihood and also an audio file of the utterance." Ex. 1007, ¶[0086]. "The best guesses of the voice recognition engine as to what was said are evaluated against an information database(s)," which is "built up over time, by recording personal details or preferences from what has been said, or from information directly entered into the system by the user." (Ex. 1007, ¶[0088]). The best guesses for the possible requests are also evaluated against a knowledge base,

where "the correct utterance text and the correct intention are stored." Ex. 1007, ¶[0106].

106. Barbara discloses that "[w]hen each utterance is received at the server, it is evaluated into a set of possible requests that the user may have actually meant," and a "combined likelihood for each possible request is heuristically calculated." Ex. 1007, ¶[0090]. Part of this process includes the information database being "searched for words that are similar to or the same as those interpreted," and "if anything is identified, then a message is sent . . . to the server identifying the word from the information database and providing an indication as to its meaning and the confidence of that interpretation." Ex. 1007, ¶[0089].

107. Barbara's "automatic recognition process is to consider every sub-phrase from the sentence in order of increasing length," which begins by "mapping to intentions all sub-phrases," in increasing length which is "continued . . . until a set of possible intentions is determined for the whole phrase." Ex. 1007, ¶[0118]. Then, "[c]onfidences are heuristically calculated for each possible intention, by considering the mapping rules used and their 'success' in the history" of the system. Ex. 1007, ¶[0118]. Importantly, the knowledge based is continually used to calculate rule confidences because the rule confidences are periodically calculated "based on analysis of the knowledge base." Ex. 1007, ¶[0118].

108. Importantly, Barbara's recognition process is fundamentally based on human operators. Barbara's reliance on human operators exists in two tiers: (i) there is a backstop for speech recognition where the human operator chooses the correct text of the spoken word, and (ii) there is backstop for determining the intent of the user's utterance when the system cannot deduce it, even though the system knows was spoken. Barbara's reliance on human operators is shown below in Figure 7.

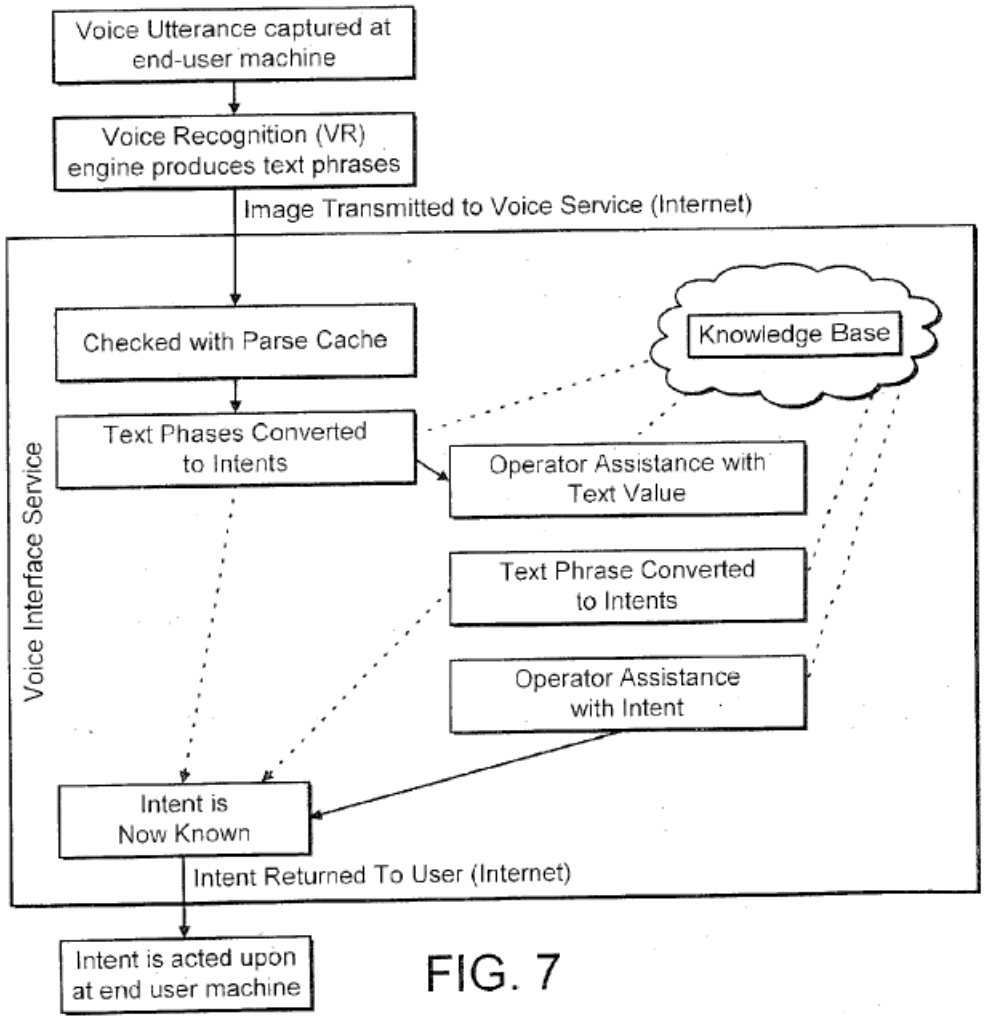

FIG. 7

Ex. 1007, Fig. 7

109. As shown in Figure 7, Barbara's system requires operator assistance with determining both the correct text value of the utterance, and the intent. However, where utterances have failed to meet a criteria for automatic acceptance, they are replayed for a human operator who chooses the correct textual interpretation. Once the textual interpretation has been deduced with the assistance of the human operator, the interpretation process begins. Barbara discloses that, in cases where it is not possible to automatically deduce what the user wanted to do despite knowing what they said, the message is sent to a second human operator who hears what was said and determines what the utterance means. Ex. 1007, ¶ [0104]-[0105]. Thus, Barbara's process is fundamentally reliant on human assistance.

2. Ross (Ex. 1008)

110. Ross (Ex. 1008) is directed to a system that uses a specific domain model for a particular application to associate a particular syntax such that, when the system interacts with a user, there is a more natural interaction and better consistency between the input and output. Ex. 1008, ¶[1005].

111. The disclosure especially relevant to my opinions below relied on by Petitioners is Ross's disclosure of a conversational record. Ex. 1008, ¶[0058]. Ross's conversation manager in Figure 3 below stores conversational record 60. Ex. 1008, ¶[0033].

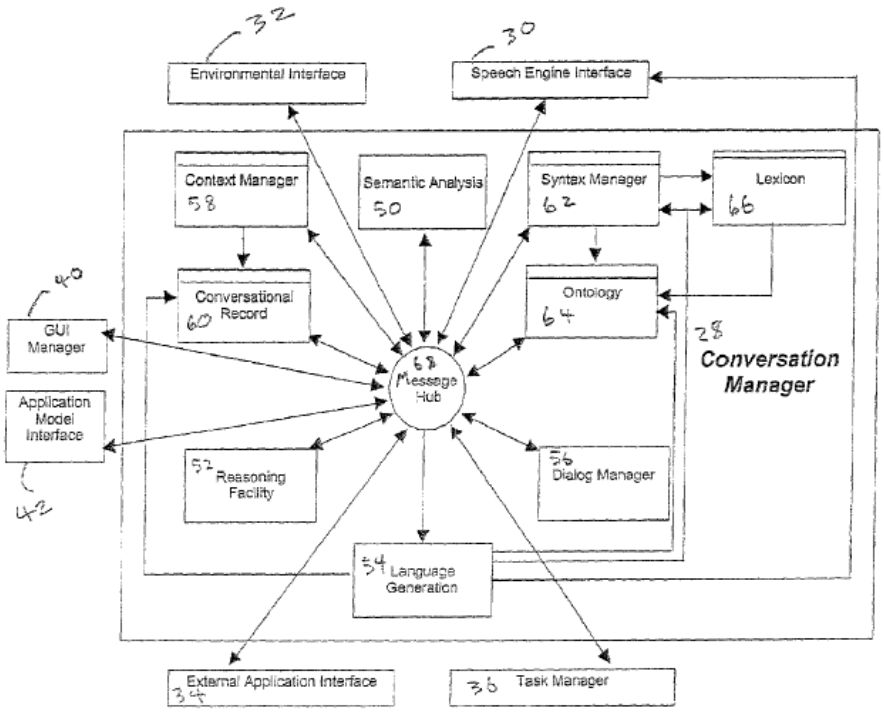

Fig. 3

Ex. 1008, Fig. 3

112. For Ground 2, I understand that Petitioners rely on Ross's architecture of the conversation manager and "data organization" including the conversational record, because of Ross's disclosure of purging information from the "conversational record when it is no longer relevant to active goals and after some predetermined time has elapsed." Ex. 1003, ¶[0057].

3. Kellner (Ex. 1023)

113. Kellner discloses as an "object" of its "invention to provide a dialog system having as comfortable and effective a dialog management for a user as

possible." Ex. 1023 at ¶[0003]. Kellner discloses that "the contents and form of the system outputs are adapted to the style of speech of user input and/or to the behavior of a user during a dialog with the dialog system." Ex. 1023 at ¶[0005]. Kellner further discloses that "[t]he details about the style of speech and dialog interactions of a user are contained in an associated user model which is evaluated by the dialogue system components." Ex. 1023 at ¶[0005]. Thus, Kellner discloses that the system is thus in a position to "generate system outputs adapted to a user's style of speech considered pleasant by the respective user." Ex. 1023 at ¶[0005].

B. The Petition Fails to Show that the Barbara, Ross, and Kellner Combination Discloses or Suggests "accumulat[ing]/[e]" "short-term knowledge" "during a predetermined time period" and "accumulate[ing]/[e]" "long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period"

114. The challenged claims require accumulat[ing]/[e]" "short-term knowledge based on one or more natural language utterances received during a predetermined time period" and "accumulate[ing]/[e]" "long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period." Ex. 1001, cl. 1, 12.

115. With respect to "short-term knowledge based on one or more natural language utterances received during a predetermined time period," the Petition

identifies Barbara's disclosure of the information base and knowledge base, which respectively store "personal details or preferences from what the user has said" and the correct utterance text and correct intention. The Petition argues, for example, that Barbara's knowledge base, which stores "[t]he correct utterance text and the correct intention" "is short term storage that uses Ross's conversational record to store . . . each utterance 'along with the results of its semantic analysis," which the Petition asserts is "eventually purged from the conversational record when it is no longer relevant to active goals and after some predefined time period has elapsed." Petition at 76 (quoting Ex. 1007 at ¶106; Ex. 1008 at ¶57). Thus, the Petition argues that these limitations are satisfied because portions of the knowledge base of Barbara are ***purged*** after some predefined period of time.

116. I disagree that the Barbara-Ross-Kellner combination discloses or teaches "accumulat[e]/[ing] "short-term knowledge" and "accumulate[ing]/[e]" "long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period" as required by the claims, because, as the Petition acknowledges, Ross discloses that the information in its conversational record is eventually "purged" "when it is no longer relevant to active goals and after some predefined period of time has elapsed." Ex. 1008 at ¶57. Thus, Ross in combination with Barbara only teaches that the alleged short-term knowledge would be purged or deleted by the

combined system, not accumulated and included in long-term storage, or later used for related functionality (i.e., to identify a context), as set forth in the claims. Accordingly, in my opinion, the Petition fails to show that Barbara and Ross disclose or render obvious "accumulat[e]/[ing] short-term knowledge.

117. Moreover, as set forth above, Petitioners assert that "[i]mplementing Ross's teachings regarding purging information from the 'conversational record' allows Barbara's dialog history to be expired and long-term data to be retained." Petition at 66-67. The Petition does not, however, demonstrate that modifying Barbara's knowledge and information databases using Ross's conversational record allows "dialog history to be expired and long-term data to be retained." Petition at 66-67. Instead, Ross explicitly discloses ***purging*** the conversational record after some predefined period of time, not expiring it to long-term storage. Because the conversational record is purged or deleted, Petitioners have not shown that the resulting combination both "accumulates" short-term knowledge based on one or more natural language utterances received during a predetermined time period" ***and*** "accumulates" "long-term knowledge" "based on one or more natural language utterances received prior to the predetermined time period," as the resulting knowledge base could not accumulate knowledge that was purged, and therefore unavailable, in order to achieve long-term knowledge based on one or more natural language utterances received prior to the predetermined time period.

118. Accordingly, in my opinion, the Petition fails to show that these claim limitations are met.

C. The Petition Fails to Show that the Barbara-Ross Combination Discloses or Suggests the Claimed "computer-implemented method" or the claimed "system for facilitating natural language system responses" "comprising one or more physical processors programmed with one or more computer program instructions"

119. Claim 1 of the '699 Patent recites a "computer-implemented method of generating natural language system responses" "the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising: . . . recognizing, by the computer system, one or more words or phrases from the natural language utterance; . . . determining, by the computer system, an interpretation of the natural language utterance based on the identified context[.]" Ex. 1001 at cl. 1. Likewise, claim 12 of the '699 Patent recites a "system for generating natural language system responses . . . comprising: one or more physical processors programmed with one or more computer program instructions which, when executed, configure the one or more physical processors to: . . . recognize one or more words or phrases from the natural language utterance;" and "determine an interpretation of the natural language utterance based on the identified context" Ex. 1001, cl. 12.

120. In my opinion, the Barbara-Ross-Kellner combination does not disclose or teach these limitations because the disclosure of Barbara identified by the Petition are fundamentally based on manual interpretation by human operators. This manual interpretation by human operators is inconsistent with the claim limitations requiring that the "recognizing" and "determining" limitations in claim 1 are done "***by the computer system***", and that the system of claim 12 uses "***one or more physical processors***" that are configured to "recognize one or more words or phrases from the natural language utterance" and "determine an interpretation of the natural language utterance based on the identified context."

121. The design of Barbara's system that performs text recognition and intention recognition is fundamentally based on human assistance and manual interpretation by the human operator, which is inconsistent with the claims. Ex. 1007, Fig. 7, ¶¶ [0100], [0104]. Petitioners rely on Barbara's voice interface system for receiving, interpreting, and correcting incoming information in combination with Ross as allegedly rendering obvious the asserted claims. Petition at 65-67. However, as I explained above in Section VIII.A.1, Barbara's voice interface system fundamentally relies on human operators, in addition to the system itself, for the utterance text recognition and the intention recognition.

122. Accordingly, in my opinion, the Barbara-Ross-Kellner combination does not render obvious the limitations set forth above.

D. A POSITA Would Not Be Motivated to Combine Barbara, Ross, and Kellner

123. I understand that the Petition alleges that a POSITA would have been motivated to combine Barbara and Ross. Specifically, the Petition alleges that "a POSITA would have been motivated to combine Ross's teachings regarding the architecture of a conversation manager with Barbara's voice interface system." Petition at 65-66. The Petition further asserts that, while Barbara discloses use of an "information database" and "knowledge base" to interpret utterances, a POSITA would have been motivated to combine Barbara's databases with Ross's conversations records for storing a dialog history because Barbara "provides limited detail regarding managing the information" in its data stores. Petition at 66. Thus, Petitioners assert that "[i]mplementing Ross's teachings regarding purging information from the 'conversational record' allows Barbara's dialog history to be expired and long-term data to be retained." Petition at 66-67. I disagree with this characterization of the references and this motivation because Ross only discloses purging the conversational record, and does not teach or suggest retaining long-term data.

124. Furthermore, Petitioners allege that a POSITA would have been motivated to combine Barbara and Ross because, while Barbara discloses use of an information database and knowledge base to interpret utterances, it does not provide

extensive details about managing the information in those data stores, whereas Ross teaches details regarding storage of dialog history because of its disclosure of purging the conversational record. Petition at 66. I disagree with the reasons provided by Petitioners that POSITA would have been motivated to combine Barbara and Ross. In my opinion, a POSITA would not have been motivated to combine the disclosures of Barbara and Ross to achieve the claimed invention.

125. I disagree with Petitioners assertion that a POSITA would modify Barbara's knowledge and information databases using Ross's organization and architecture that includes purging information from the conversational record when it is no longer relevant to active goals and after some predetermined time. This discarding of information is entirely inconsistent with the automatic recognition process in Barbara, the successful function of which is reliant on the information database and knowledge base being built up overtime and continually maintained.

126. Specifically, as I explained above, the knowledge base stores "the correct utterance text and the correct intention." Ex. 1007, ¶[0106]. Petitioners rely on Barbara's recognition process for the majority of the claim limitations, and that process works by mapping the intentions of each individual sub-phrase and word of an utterance, and calculating the corresponding confidences for accurate recognition. These confidences in Barbara are calculated based on an analysis of the knowledge base. Ex. 1008, ¶[0118]. Petitioners specifically rely on Barbara's knowledge base

and information database as allegedly disclosing or teaching the limitations of the Challenged Claims directed to short-term knowledge. Petition at 75-76.

127. In Barbara, the history of correct utterance text and correct intentions are continually used to calculate the confidences. Thus, the continual maintenance of the history of correct utterance text and correct intentions in the knowledge base is vital to the operation of Barbara's recognition process. That is, in order for Barbara's system to work, the knowledge base must be continually maintained and be as "broadly applicable as possible." Ex. 1007, ¶[0110].

128. Likewise, the information database in Barbara "is built up over time, by recording personal details from what has been said[.]" Ex. 1007, ¶[0088]. During Barbara's recognition process, the information database is searched for words that are similar to or the same. Ex. 1007, ¶[0089]. Accordingly, the functionality of Barbara's system that Petitioners rely on is dependent on the information database maintaining the history of what has been said so that future interpretations are successful.

129. In contrast, Ross's conversational record indexes each utterance with the results of its semantic analysis and eventually purges it when it is no longer relevant and after a predetermined period of elapsed time. Ex. 1008, ¶[0057].

130. I note that Petitioners appear to identify both Barbara's knowledge base and information database as satisfying the "short-term knowledge" limitations of the

claims. Petition at 75-76. However, Petitioners only allege that Barbara's knowledge based would be modified by Ross's data organization including the conversational record, and fails to make the same assertion for the information database. Petition at 75-76. In any case, I disagree that a POSITA would be motivated to modify Barbara's information database or knowledge base using Ross's conversational record.

131. Petitioners allege that a POSITA would have modified Barbara's knowledge base using Ross's data organization to make the knowledge base short-term storage that "is eventually purged from the conversational record when it is no longer relevant." Petition at 75-76. I disagree, because periodically purging the information stored in the knowledge base is entirely incompatible with the function and purpose of the knowledge base in Barbara's recognition system.

132. If Barbara's knowledge base was periodically purged, as Petitioners assert with this combination, Barbara's recognition process would not be able to successfully calculate the confidences of each interpreted word because there would be no history in the knowledge base to calculate these confidences. Barbara's knowledge base has to continually maintain all correct interpretations in order to continually be used to calculate the confidence scores for each word. Even though I understand Petitioners have not made the specific assertion to modify the information database with Ross's conversational record, that modification would not

work either because the history, "built up over time" based on what has been said is also analyzed in order to interpret utterances received by Barbara's system. Thus, a POSITA would not be motivated to modify either the information database or the knowledge base using Ross's period purge of information because Barbara's system simply would not work if modified in the way that Petitioners suggest. Thus, in my opinion, a POSITA would not be motivated to combine these references as Petitioners suggest to achieve the claimed invention.

133. Moreover, the Petition asserts that Barbara and Ross are analogous art to the '699 Patent pertaining to the same endeavor. I disagree with the Petition's assertion that a POSITA would have been motivated to combine Barbara and Ross as analogous references. Barbara, on the one hand, is directed to a communication system where the voice input from a user terminal is routed to a plurality of operator terminals manned by human operators that perform manual interpretation of the incoming information from the user terminals. Ex. 1007, ¶¶[0083]-[0085]. Ross, on the other hand, is directed to language generation method that performs its work in the context of a domain model for a particular application so that dialog between the user and the system is more natural. Ex. 1008, ¶¶[0004]-[0005]. Not only are the references directed to different domains, and solutions to different problems, but unlike Barbara, Ross's system is not reliant on manual interpretation by human

operators. In my opinion, these differences would deter a POSITA from combining these references in the manner asserted by the Petition.

134. In my opinion, a POSITA also would not have been motivated to combine Barbara and Ross with Kellner. As the Petition notes, Kellner discloses that, because its system is "in a position to generate system outputs adapted to a user's style of speech with a style of speech considered pleasant by the respective user," "the inhibition threshold of the use the dialog system can be lowered." Ex. 1023 at ¶[0005]. Thus, the system can capture additional user information as the user interacts more with the system and therefore provide more tailored results. Petition, 68-69. Thus, Kellner discloses a functional positive feedback loop just as Barbara does. In other words, as I explained above, the continued success of subsequent interpretations and the resulting function of Barbara's system is reliant on the continual correct recognition of the utterance text and intent and building the database with that history. Similarly, Kellner's positive feedback loop and continued function of its system is based on increased interactions with the system based on lowered inhibition thresholds when the user gets pleasant responses. Just like I described with respect to Barbara, this positive feedback loop from Kellner is entirely inconsistent with Ross's periodic purging of the conversational record. Kellner's continued use of a user model to adapt the style of speech and/or the behavior or the user to increase user interaction with the system would be completely

undermined if the conversational history used to collect information about the style of speech to better adapt future responses was purged. Thus, a POSITA would not be motivated to combine Kellner with Barbara and Ross.

135. Accordingly, in my opinion, Ground 2 based on the combination of Barbara, Ross, and Kellner fail to render obvious the challenged claims.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on the information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed on September 9, 2025 in Davis, CA.

Houman Homayoun, Ph.D.