Index number	Technology used	Use in conversational agent	Scope of technology
Paper 1 (2017) **Related Papers: Paper 1, Paper 2**	1. Ontology based knowledge model 2. The knowledge integration module reasons over the knowledge base and picks possible reactions for the system. The dialogue manager then picks the best of the possible reactions. 3. Visual Scene Maker is used to determine agent's role in conversation (on which decisions are based) and to model non-verbal idle behaviour.	1. Information retrieval 2. Response generation 3. Non verbal communication	"KRISTINA is projected as an embodied companion for (elderly) migrants with language and cultural barriers in the host country and as a trusted information provision party and mediator in questions related to basic care and healthcare" Characteristics of agent 1. Retrieve multimedia 2. Understand concerns of user through multilingual verbal, facial and gestural input 3. Plan dialogue through ontology based reasoning 4. Communicate with verbal and non-verbal signals
Paper 2 (2019)	1. Models for question type and intent classification with slot filling 2. Semantic matching of user question to a set of results from a Question-Answer knowledge base 3. Sequential matching networks and neural multi-perspective sentence similarity networks for semantic matching 4. Dialogue framework to query missing information proactively	1. Customer support 2. Instructional assistance	1. Can provide instructional answers to a user question from a Question-Answer knowledge base
Paper 3 (2017)	1. Multi-task learning approach for training persona-based neural conversation models. 2. Learning uses conversation data across speakers and non-conversational text data pertaining to speaker and speaker roles to be modelled. 3. Model proposed does not rely on vast amounts of speaker-specific conversation data. 4. In the model proposed, the decoder of a SEQ2SEQ model and an autoencoder share parameters and are jointly trained. The SEQ2SEQ model generates a response given the previous context. The autoencoder generates the same response while incorporating the general conversation model of the persona.	1. Applicable to all extended conversations 2. May not be useful for single requests like asking for the weather 3. Can be used for target guided conversations where the systems adopts a persona to achieve a specific goal	1. Using the technique proposed in the paper we can add a speaker role (eg. support person), domain of expertise (eg. technical support for mobiles) and speaking style (eg. courteous) to a system. All three constitute the persona.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 4 (2020) **Related Papers: Paper 1**	1. Technique for target-guided open domain conversation. 2. Structured approach controls intended content of system response through keywords. Turn level supervised learning and knowledge relations between candidate keywords help attain smooth transitions in conversation. Approach can drive conversation towards a target. 3. Dynamic knowledge routing network considers semantic knowledge relations between keywords selected to predict the next keyword to be used (next topic of discourse). 4. Dual discourse-level target guided strategy to reach target smoothly. The strategy constrains the predicted keyword to be closer to the target at each turn. It also constrains the selected next response to contain the predicted keyword or a keyword closer to the target than the previous keyword.	1. Health checkups for the elderly can be more open ended while having the goal of checking their health status 2. Can be used along persona-based conversation models to help achieve the goal of the persona adopted by the system	1. When there is a specific goal to be reached, open ended conversations can be guided towards the goal while seeming like a natural conversation.
Paper 5 (2019)	1. Method for automatic conversational knowledge extraction from any preexisting chatbot (rule based chatbot is used in paper). 2. Method for building a new neural conversational agent using the knowledge extracted. 3. A question agent is used which has a large and noisy question database. The answer of the existing chatbot for each question is recorded. The dataset of questions and corresponding recorded answers are used as the training dataset when building the neural conversational agent. 4. Method is based on the hypothesis: "If the number of questions addressed to chatbot tends to infinity, the total number of unique answers rom the chatbot will tend to a stable number." Hypothesis is experimentally proven.	1. A chatbot performing a specific function (such as answering medical questions) can have its knowledge extracted and a new conversational agent can be made for the same function but with the ability to generate new answers. This allows it to better answer varied questions.	1. An existing chatbot (rule based or neural) can have its information extracted in order to create a new neural conversational agent. It enables machine to machine knowledge sharing.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 6 (2018) **Related Papers: Paper 1, Paper 2, Paper 3**	1. Framework for bootstrapping end-to-end dialogue agents for goal oriented dialogues in arbitrary domains 2. Framework is better than Wizard-of-Oz data collection both in terms of speed and diversity/coverage of dialogues. 3. Task schema (intent and slots to be filled) needs to be defined by developer for a new task but all task independent steps are automated. 4. First a simulated user bot and a domain agnostic system bot converse with each other to generate a sequence of natural language dialogues and their semantic parses. 5. Then humans rewrite the dialogues to make them more natural and provide context while preserving their meaning.	1. Chatbots for specific domains or domain specific parts of a conversational agent.	1. Technique proposed can be used to rapidly develop fully labelled dialogue datasets which can be used to develop goal oriented chatbots in any arbitrary domain.
Paper 7 (2021)	1. Framework to design nonverbal communication of a conversational agent. 2. Body movement and facial expressions are the two main nonverbal communications explored. 3. OCEAN personality model is used to generate facial and body expressions.	1. Conversational agent for a robot. 2. Embodied virtual conversational agent.	1. Framework proposed can be used to include nonverbal communication for any conversational agent. 2. Voice modulation and proper dialogue generation can be used with nonverbal communication to build a realistic conversational agent.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 8 (2020)	1. Intent discovery framework that mines conversation logs and generates labelled datasets for training intent models. 2. Density based clustering algorithm for unbalanced clustering is proposed. It is an extension of DBSCAN. The algorithm is used to cluster the representations of conversations into groups based on intent. 3. Domain experts then use utterances of the conversation and metadata/descriptions of the conversations to label each cluster. 4. A classifier is trained on this initial data. It is then used to label the remaining unlabelled conversations.	1. Chatbot for a particular domain or domain specific part of conversational agent of a domain in which vast amounts of conversational data exists.	1. Can be used to generate a dataset of conversations with labelled intent.

Index number	Technology used	Use in conversational agent	Scope of technology
<u>Paper 9 (2021)</u>	1. Includes new advances made in the Alexa Prize Socialbot Grand Challenge. 2. Semantic parsing is used to better represent and understand user utterances. In each turn of the conversation information is extracted from the user utterance and an inference is made to obtain knowledge from the information. The knowledge obtained is added to the knowledge graph of the system. 3. Common sense knowledge about topics is made available to the system by incorporating a conversational knowledge template (CKT), which holds the common sense attributes for each topic. Information from user utterances is extracted and answer set programming is used to make inferences based on the information and CKT. 4. Intent and topic classification in a conversation is done using a hierarchical RNN (HRNN). It consists of 2 RNNs, one to take a sequence of words from utterances and learn representations and the second to take a sequence of utterance representations and learn dialogue representations. Pre-trained models like BERT can also be used along with a neural network for classification. 5. Entity recognition and punctuation models are both BERT based models trained on their respective datasets. 6. Neural response generation model is trained using user utterances and the corresponding responses from another conversational agent. The model is not knowledge grounded since it is only trained using responses. Controlled response generator is also being experimented with. It has two controls 1. Empathy (to answer suitable when user has negative sentiment) 2. Topical control (response stays on same topic as user utterance).	1. Used in various aspect of the conversational agent	1. Paper has experimental techniques to improve various aspects of a conversational agent. 2. The reliability of the techniques is questionable as acknowledged by the paper.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 10 (2021)	1. Entity resolution features are used to rerank the natural language interpretation (domain, intent, slots) of a user utterance. 2. Reranking model weights are learnt based on a newly proposed loss term based on entity resolution signals (specifically when there is no match in the ER catalog). 3. For multi domain conversations, a score distribution matching method is proposed to calibrate reranking scores for different domains.So the reranking can be done independently for each domain and then the score distribution can be applied on it.	1. NLU part of conversational agent	1. The methods proposed can be used for accurate natural language interpretation of user utterances (identification of domain, intent and slots). 2. Can be used in multi-domain conversations.
Paper 11 (2020)	1. Policy driven neural response generation approach for knowledge grounded open domain dialog systems. 2. A dialog policy is to be used which plans the content and style of open domain responses in the form of an action plan. 3. The action plan includes new knowledge sentences related to the context of the conversation, the targeted purpose of each sentence in the response and order of sentences in the response. 4. A dialog policy model predicts the next action plan given the dialog context. 5. The neural response generators are jointly conditioned on all the action plan attributes and then target responses are produced. 6. Method for automatically annotating topical chat dataset is also given. Annotations are on attributes such as topic, dialog act.	1. Response generation part of conversational agent	1. The technique proposed generates responses which are controllable in terms of style and content. 2. The responses produced are more engaging and interesting as they respond to the user utterance and also introduce related topics smoothly.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 12 (2020) **Related Papers: Paper 1**	1. A recursive, hierarchical, frame-based representation is proposed to represent information extracted from user utterances. 2. A method to learn a representation template from data is given. It is formulated as a template-based tree decoding task. 3. The template consists of general forms of slot labels which are used as attributes. (eg. in plane ticket booking, origin and destination will both have the general slot label city name) 4. Duplication of nodes in the tree based template is minimised by adding a loss term corresponding to node duplication in the tree. 5. The template is filled using words directly from the user utterance.	1. Natural Language understanding 2. Intent classification	1. The representation proposed is to be used instead of the intent and slots representation as it can better represent more complex user requests with multiple intents in the same request. 2. It achieves significantly better results only on complex datasets.
Paper 13 (2021)	1. Slice based learning is used to improve skill routing system of a conversational agent for infrequently occurring (tail) queries. 2. Slice based learning is a technique in which performance of a model is improved for critical subsets of data (called slices). The model learns "expert representations" for the slices which are combined with an attention mechanism. 3. Intents to be monitored are defined and a pre trained model is extended to attend to these tail intents using SBL.	1. Some infrequent requests may be critical and failure to route to the correct skill is not acceptable (eg. user utterance indicating that they are not feeling well).	1. Slice based learning can be used to improve the skill routing of the conversational agent when met with a long or short tail request (infrequent requests with few or many words).

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 14 (2020)	1. Two designs for hypothesis rejection modules are given. Hypothesis rejection modules examine the hypothesis (domain, intent, named entities) generated by the NLU module and reject the hypothesis if it is incorrect. 2. The design in the paper uses the user utterance and the NLUs confidence score for the hypothesis in addition to the hypothesis itself. 3. First design is for a domain specific hypothesis rejection module. It examines the top hypothesis generated by the NLU for its domain and accepts/rejects it. 4. Second design is for a overall system hypothesis rejection module. It examines the top hypothesis across all domains and accepts/rejects it. 5. Both designs have the same model architecture. They differ in which domains can be accepted/ rejected. In the domain specific design, all other domains are always rejected. But in the overall system design any domain can be accepted/rejected.	1. It can be used to decrease user dissatisfaction with wrong responses. No response/ reconfirmation is better than a wrong response in some cases (eg. when the user asks a serious question and the response is light hearted).	1. Incorrect responses from the conversational agent are decreased as most false hypotheses are rejected by the technique proposed.
Paper 15 (2020) **Related Papers: Paper 1, Paper 2**	1. Architecture to parse simple and complex queries to extract information from them. The extracted information is stored as a hierarchical arrangement of intents and slots (parse tree). 2. Architecture can also be used to parse queries which don't conform to slot filling grammar or RNN grammar (No experimental results given for this). In this case extracted information cannot be stored in a parse tree. 3. Architecture based on Sequence to Sequence model (transformer used in paper) along with a pointer network. 4. The pointer network inputs the last hidden state of the encoder and at each step produces an intent/slot or a pointer to the user utterance. The pointers point to tokens in the user utterance which are used to fill particular slots.	1. NLU of conversational agent	1. The architecture proposed can be used to extract information from simple and complex queries. 2. It can also be used for queries of all three types 1. Slot filled 2. Parse tree 3. Queries not obeying either grammar.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 16 (2020)	1. Technique for fine tuning pre trained transformers to perform answer sentence selection in a particular domain. 2. A pre trained transformer is first fine tuned with a large dataset to transfer it into a model for general answer sentence selection. Dataset for this is provided by the authors. 3. The transformer is then fine tuned to adapt it to a domain.	1. Domain specific information retrieval (question answering)	1. Can be used to generate more robust and stable fine tuned models while reducing effort in selecting hyper-parameters. 2. Noisy data can also be used for fine tuning as the first step makes the model robust to noise. 3. A small dataset is enough to adapt the model to a domain. This is because it is already suited for general answer sentence selection.
Paper 17 (2020) **Related Papers: Paper 1**	1. Novel architecture for multi domain dialog state tracking (predicting current slot values based on conversation history). 2. Cross attention is used to model relationships between context and slots at different semantic levels (relationships between meanings at the word, phrase, sentence levels). 3. Self attention is used to resolve cross domain coreferences (tokens referring to the same entity).	1. Task oriented part of conversational agent	1. Accurate multi domain dialog state tracking without knowing domain ontologies beforehand. 2. Architecture can be used for new domains and unseen slot values as well.

Index number	Technology used	Use in conversational agent	Scope of technology
<u>Paper 18 (2020)</u>	1. The study collected 58 articles on conversational agents which can be used by businesses and answered the following research questions. 2. Question 1: How are machine learning methods used on knowledge extracted from interaction with user (self learning)? Only 6% of the articles use self learning methods. The methods used are: 1) Clustering of data which the system is not able to match to any known problems. New clusters which are formed represent new topics. 2) When the system cannot find matching answers, it creates a new AIML rule based on the user response. 3) User feedback on repetitive questions is considered when retraining the model. Model used is a BiLSTM with attention mechanism. 3. Question 2: How have conversational agents been personalised? Only 8% of the articles focused on personalisation. Some approaches are 1) In technical troubleshooting, personalised features of the user are used to identify the domain from the current query and past interactions. 4. Question 3: How have studies used generative techniques for response generation, LSTM, Seq2Seq or attention-based generation? 10% of the articles used generative techniques. Most articles used Seq2Seq models out of which some were implemented using LSTMs.	1. Performance improvement over time using self learning 2. Personalisation of agent 3. Generative techniques for response generation	1. Paper has various techniques of implementing conversational agents which are suitable to be used by a business.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 19 (2020) **Related Blog**	1. A multi-turn, open-domain chatbot called Meena is given. 2. Meena is trained end-to-end on data mined and filtered from social media conversations (dataset available). 3. Meena is a 2.6 billion parameter neural network. 4. Meena is trained to just reduce perplexity. 5. The authors propose a human evaluation metric called Sensibleness and Specificity Average (SSA) which captures how specific and informative the chatbots's responses are and also if they make sense. Ensuring specificity means that the bot cannot produce vague responses like "I don't know". 6. The experiments done show a strong correlation between SSA and perplexity. 7. Meena scores 72% on SSA and the full version of Meena scores 79% on SSA. Human score on SSA is 86%. Existing chatbots score around 56%. 8. Meena consists of a Seq2Seq model with the evolved transformer as the main architecture. 9. Meena was evaluated using 2 methods 1) Static - A dataset of 1477 multi-turn (1 to 3 turns) conversations was used for evaluation. 2) Interactive - Human evaluators could chat about anything they pleased. 10. Meena is compared with XiaoIce, Mitsuku, DialoGPT and Cleverbot.	1. Open domain conversation	1. Paper shows that a large, end to end deep learning model which is trained to reduce perplexity can achieve human like, open domain conversational skills. 2. Paper shows that perplexity is a good automatic evaluation metric that correlates well with human evaluation.

Index number	Technology used	Use in conversational agent	Scope of technology
Paper 20 (2019) **Related Papers: Paper 1**	1. Dialog self-play is used to create challenge datasets over multiple tasks and intents. 2. Challenge datasets can be used to evaluate an end-to-end CA for unseen dialog patterns, improving interpretability of the model by isolating specific reasoning capabilities. 3. Technique given is scalable, portable approach for generating task oriented dialog datasets from scratch which include complex, structured knowledge. 4. Banking domain chosen as it incorporates connected intents (transfer money, check balance). 5. User and system bots are used for dialog generation. For each conversation, a user profile is picked and conversation is carried out using actions (specification of speaker, intent and slots). 6. Actions are converted to uninstantiated pseudocode (intent specified but not slots). Human annotators convert this into natural language templates with entity placeholders instead of the specified entities. Final user utterance is produced by filling in templates from a KB. 7. Types of challenge sets 1) Out of Templates (OOT) 2) Out of Pattern (OOP) - Has test cases with unseen dialog flow. 8. Generating test cases 1) Turn compression - When slot requires reconfirmation, system normally notifies user and repeats slot filling question. In turn compression, user provides slot values before slot filling question. 2) New API - Number of slot values for an intent may not be provided or may be wrong during training. 3) Reordering - Order of slot filling questions is changed. 4) Another slot - User may provide slot value that is irrelevant for current intent but relevant for some other intent. 5) Audit More - When asked for slot reconfirmation, user also provides new values for other slots.	1. Task oriented conversational agents	1. Technique can be used to increase the interpretability of CA models. 2. Technique can also be used to create controlled datasets for task oriented CAs.

OTHER INTERESTING PAPERS:
1. https://link.springer.com/chapter/10.1007/978-3-030-29135-8_12
2. https://ieeexplore.ieee.org/abstract/document/8430301
3. https://dl.acm.org/doi/abs/10.1145/3267305.3274188
4. https://assets.amazon.science/27/7c/9f0b6c2b404296198d81c9a11cb2/fast-intent-classification-for-spoken-language-understanding-systems.pdf
5. https://assets.amazon.science/0f/f6/07dc357c4b6f96906a2aa1eb2c15/scipub-1347.pdf
6. https://assets.amazon.science/4c/ed/d4ea4e194b998bedba10870dce3e/multi-stage-and-multi-task-transfer-learning-for-dialogue-based-machine-reading-comprehension.pdf
7. https://assets.amazon.science/a6/cd/f76efcfb42b6ba4822fdcfc47a3b/just-ask-an-interactive-learning-framework-for-vision-and-language-navigation.pdf

CONVERSATIONAL AGENTS INVOLVING HUMAN INTERVENTION:
1. https://link.springer.com/chapter/10.1007/978-3-030-51310-8_25#Tab1
2. https://aclanthology.org/2020.emnlp-demos.20.pdf (https://aclanthology.org/2020.emnlp-demos.20/)

STUDIES ON HOW TO BUILD BETTER CONVERSATIONAL AGENTS:
1. https://arxiv.org/pdf/1901.06525.pdf (https://dl.acm.org/doi/abs/10.1145/3290605.3300705)
2. https://nottingham-repository.worktribe.com/index.php/preview/2293571/Large-small%20talk-CUI2019.pdf (https://dl.acm.org/doi/abs/10.1145/3342775.3342789)
3. https://www.researchgate.net/profile/Minjin-Rheu/publication/344093265_Systematic_Review_Trust-Building_Factors_and_Implications_for_Conversational_Agent_Design/links/5f8d9ab6458515b7cf8b79d2/Systematic-Review-Trust-Building-Factors-and-Implications-for-Conversational-Agent-Design.pdf (https://www.tandfonline.com/doi/abs/10.1080/10447318.2020.1807710)
4. https://www.sciencedirect.com/science/article/pii/S0167923618301404
5. https://www.mdpi.com/1660-4601/17/9/3189/pdf (https://www.mdpi.com/707644)
6. https://core.ac.uk/download/pdf/301379577.pdf
7. https://link.springer.com/chapter/10.1007/978-981-32-9515-5_51
8. https://dl.acm.org/doi/abs/10.1145/3411764.3445708

LINKS TO PAPERS GIVEN IN TABLE (LINKS TO RELATED PAPERS GIVEN UNDER LINK TO PAPER)
1. https://repositori.upf.edu/bitstream/handle/10230/34305/wanner_PAAMS2017_kris.pdf?sequence=1&isAllowed=y (https://link.springer.com/chapter/10.1007/978-3-319-59930-4_23)
    1. https://leichuan.github.io/files/sigmod20-mdx.pdf (https://dl.acm.org/doi/abs/10.1145/3318464.3386139)
    2. https://repositori.upf.edu/bitstream/handle/10230/32761/TenVentura_Inai_tow.pdf?sequence=1&isAllowed=y (https://link.springer.com/chapter/10.1007/978-3-319-67401-8_52)
2. https://www.mdpi.com/2078-2489/10/2/63/pdf (https://www.mdpi.com/412654)
3. https://arxiv.org/pdf/1710.07388.pdf (https://arxiv.org/abs/1710.07388)
4. https://ojs.aaai.org/index.php/AAAI/article/download/6390/6246 (https://ojs.aaai.org/index.php/AAAI/article/view/6390)
    1. https://arxiv.org/pdf/1905.11553.pdf (https://arxiv.org/abs/1905.11553)
5. https://www.sciencedirect.com/science/article/pii/S0957417419304932#fig0004
6. https://arxiv.org/pdf/1801.04871.pdf?ref=https://githubhelp.com (https://arxiv.org/abs/1801.04871)
    1. https://aclanthology.org/N18-3006.pdf (https://aclanthology.org/N18-3006/?source=post_page---------------------------)
    2. https://assets.amazon.science/6b/63/a5a118ee4d31baaa5ef3d21b9c30/alexa-conversations-an-extensible-data-driven-approach-for-building-task-oriented-dialogue-systems.pdf
7. http://repository.bilkent.edu.tr/bitstream/handle/11693/76892/A_conversational_agent_framework_with_multi_modal_personality_expression.pdf?sequence=1 (https://dl.acm.org/doi/abs/10.1145/3439795)
8. https://arxiv.org/pdf/2005.11014.pdf (https://arxiv.org/abs/2005.11014)
9. https://assets.amazon.science/e4/1a/798a51184d8a9bcf6f26501006f8/further-advances-in-open-domain-dialog-systems-in-the-third-alexa-prize-socialbot-grand-challenge.pdf
10. https://assets.amazon.science/81/56/2fdd6e7f4cb984b24ca35c98fcea/optimizing-nlr-reranking-using-entity-resoluton-signals-in-multi-domain-dialog-systems.pdf
11. https://assets.amazon.science/a5/75/82fc567f473c8b4759a503b53dac/policy-driven-neural-response-generation-for-knowledge-grounded-dialog-systems.pdf
12. https://assets.amazon.science/9d/03/07849f83482cb45634336f1813ff/scipub-1396.pdf

1. https://assets.amazon.science/11/1a/f47268a940bfbbe7cc29a8ba4700/scipub-1042.pdf
13. https://assets.amazon.science/5d/72/c8949e714972a6b973662cc8ce89/handling-long-tail-queries-with-slice-aware-conversational-systems.pdf
14. https://assets.amazon.science/b6/b8/2a116a1b4907b49483ce2226a4d7/design-considerations-for-hypothesis-rejection-modules-in-spoken-language-understanding-systems.pdf
15. https://assets.amazon.science/11/1a/f47268a940bfbbe7cc29a8ba4700/scipub-1042.pdf
    1. https://assets.amazon.science/9d/03/07849f83482cb45634336f1813ff/scipub-1396.pdf
    2. https://assets.amazon.science/df/a0/bbce7a9142068cc24bba4f319422/scalable-dialog-state-tracking-using-slot-semantics.pdf
16. https://assets.amazon.science/45/79/8aa8e0c8499a8adaaeb525c28206/scipub-1077.pdf
17. https://assets.amazon.science/df/a0/bbce7a9142068cc24bba4f319422/scalable-dialog-state-tracking-using-slot-semantics.pdf
    1. https://assets.amazon.science/11/1a/f47268a940bfbbe7cc29a8ba4700/scipub-1042.pdf
18. https://www.sciencedirect.com/science/article/abs/pii/S1574013719303193
19. https://arxiv.org/pdf/2001.09977.pdf
    1. https://ai.googleblog.com/2020/01/towards-conversational-agent-that-can.html
20. https://aclanthology.org/R19-1081.pdf (https://arxiv.org/abs/1910.07357)