

## Background

- Identifying relevant knowledge to be used in document-grounded conversational systems is critical to effective response generation.
- Knowledge Identification (KI)** is the task of locating knowledge in a long document that is relevant to the current user query given the conversation context.

### Dialogue Context

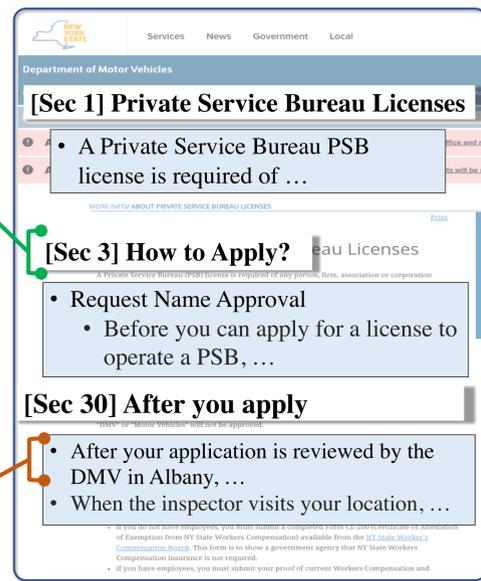
**[User]:** Hi, can you tell me something about the private service bureau licenses?

**[Agent]:** Do you want to apply for a PSB?

**[User]:** No, I was being curious. Just in case, what should I do if I apply for PSB?

**[Agent]:** Your application will be reviewed in Albany's DMV. After that, it will be sent to your local DMV office and you'll be scheduled for an inspection.

### Grounding document



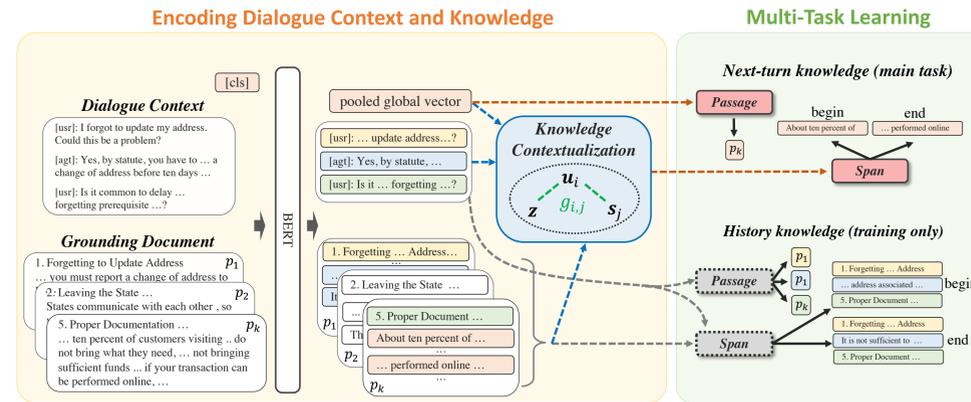
## Challenges & Solutions

- Challenge 1** The grounding document can be long.
- Solution** A multi-passage knowledge reader that selects the most relevant passage from which the final answer string is selected.
- Challenge 2** Relevance of information depends on: (1) What has been asked. (2) What has been already communicated. (3) Who said what.
- Solution** Dialogue-contextualized passage representations and a multitask learning framework with objectives to identify knowledge for the next turn, as well as used knowledge for previous turns.

## Datasets

- Doc2Dial [3]
  - 4.8k goal-oriented dialogues in 4 social-welfare domains.
  - The blind held-out test set has an additional Covid-19 domain.
  - Each turn is grounded in a sequence of knowledge spans in a given document.
- Wizard of Wikipedia (WoW) [2]
  - Over 20k social chat conversations.
  - Dev/Test set has two subsets with conversation topics seen or unseen in training.
  - Each turn is grounded in one or no sentence from 7 Wikipedia passages.

## Method Overview



### Knowledge Contextualization

The contextualized span representation  $\hat{s}_j = [s_j, \hat{s}_j, \tilde{s}_j]$  combines the original span vector  $s_j$  with

- Gated pooling with user turns  $\hat{s}_j$ .
- Gated pooling with agent turns  $\tilde{s}_j$ .

### Next-Turn Knowledge Identification ( $\mathcal{L}_{next}$ )

Applies linear layers on  $\mathbf{z}$  and  $\hat{s}_j$  to predict the gold passage, begin and end spans.

### History Knowledge Identification ( $\mathcal{L}_{hist}$ )

Leverages  $\mathbf{u}_i$  and  $s_j$  to predict the gold passage, begin and end spans for each history turn.

### Posterior Regularization ( $\mathcal{L}_{adv}$ )

Incorporates a  $f$ -divergence based posterior regularization mechanism [1] during training.

### Joint objective with tunable parameters

$$\mathcal{L} = \mathcal{L}_{next} + \alpha\mathcal{L}_{hist} + \beta\mathcal{L}_{adv}$$

## Evaluation Results on test sets

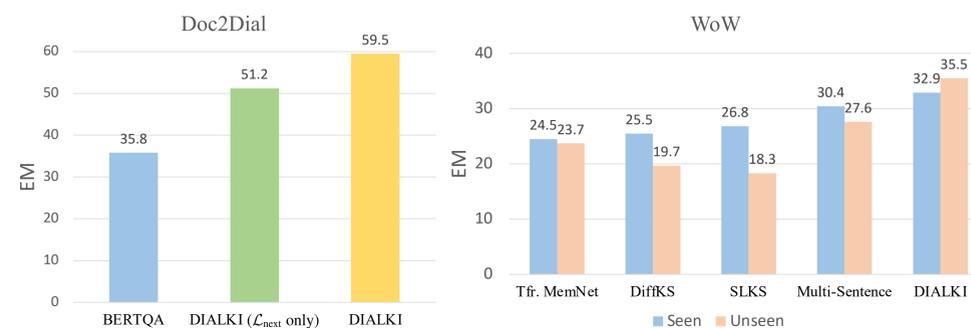


Figure 1. Exact Match (EM) scores on test sets. We observe similar trends in F1.

## Ablations on Dev Sets

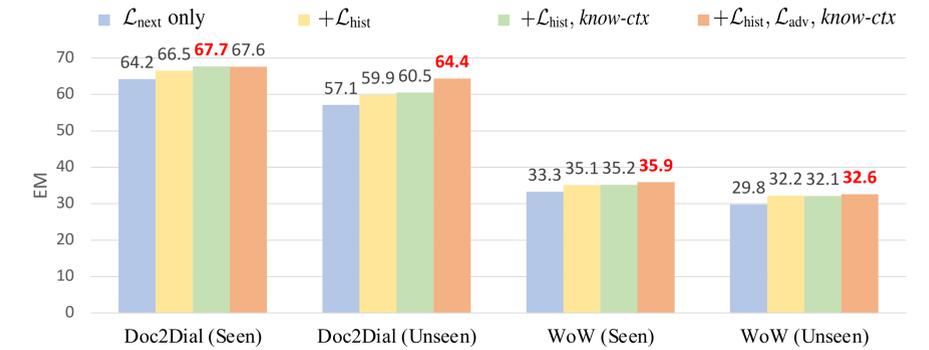


Figure 2. Exact Match (EM) scores on test sets. We observe similar trends in F1.

- Seen / unseen data during training:
  - Doc2Dial: seen / unseen grounding documents.
  - WoW: seen / unseen conversations with topics.
- Overall, adding each component of our model proves to be effective.

## Impact of KI on Response Generation

KI Model	Knowledge Input	sacrebleu
-	full doc	22.84
BERTQA-Token	pred span	21.42
DIALKI	pred span	25.16
DIALKI	pred span & passage	25.84

Table 1. Response generation results on Doc2Dial dev set. KI stands for Knowledge Identification.

- Using knowledge predicted by DIALKI leads to improvements in the sacrebleu score.

## Takeaways

- Addresses knowledge identification in conversational systems with long grounding documents using:
  - passage representations contextualized with the dialogue-document history.
  - multi-task learning and posterior regularization.
- Achieves state-of-the-art results on two dialogue datasets.

## References

- H. Cheng, X. Liu, L. Pereira, Y. Yu, and J. Gao. Posterior differential regularization with  $f$ -divergence for improving model robustness. In *Proceedings of the 2021 Conference of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, 2021.
- Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. Wizard of wikipedia: Knowledge-powered conversational agents. In *International Conference on Learning Representations*, 2019.
- Song Feng, Hui Wan, Chulaka Gunasekara, Siva Patel, Sachindra Joshi, and Luis Lastras. doc2dial: A goal-oriented document-grounded dialogue dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8118–8128. Online, November 2020. Association for Computational Linguistics.