# A Simple but Effective Approach to Improve Structured Language Model Output for Information Extraction

*Yinghao Li, Rampi Ramprasad, Chao Zhang;* Georgia Institute of Technology, Atlanta GA

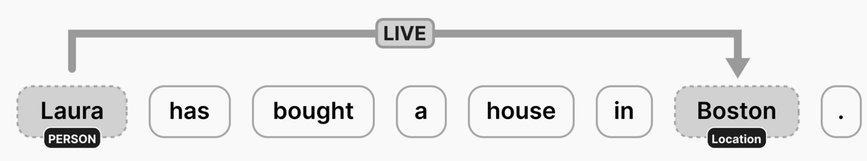
## Motivation

### ○ Information Extraction (IE)

*The process of automatically extracting **structured information** from unstructured or semi-structured documents*

Named Entity Recognition (NER)


Relation Extraction (RE)




### ○ IE with Language Models (LMs)

- ○ No need for training (which requires a lot of labels)
- ○ Zero-shot extraction of any kind of entities/relations
- ○ Great for smaller-scale project with large variance of label types

### ○ The Challenges

- ○ LMs are mostly trained with ····· unstructured data
- ○ IE demands ····· structured output

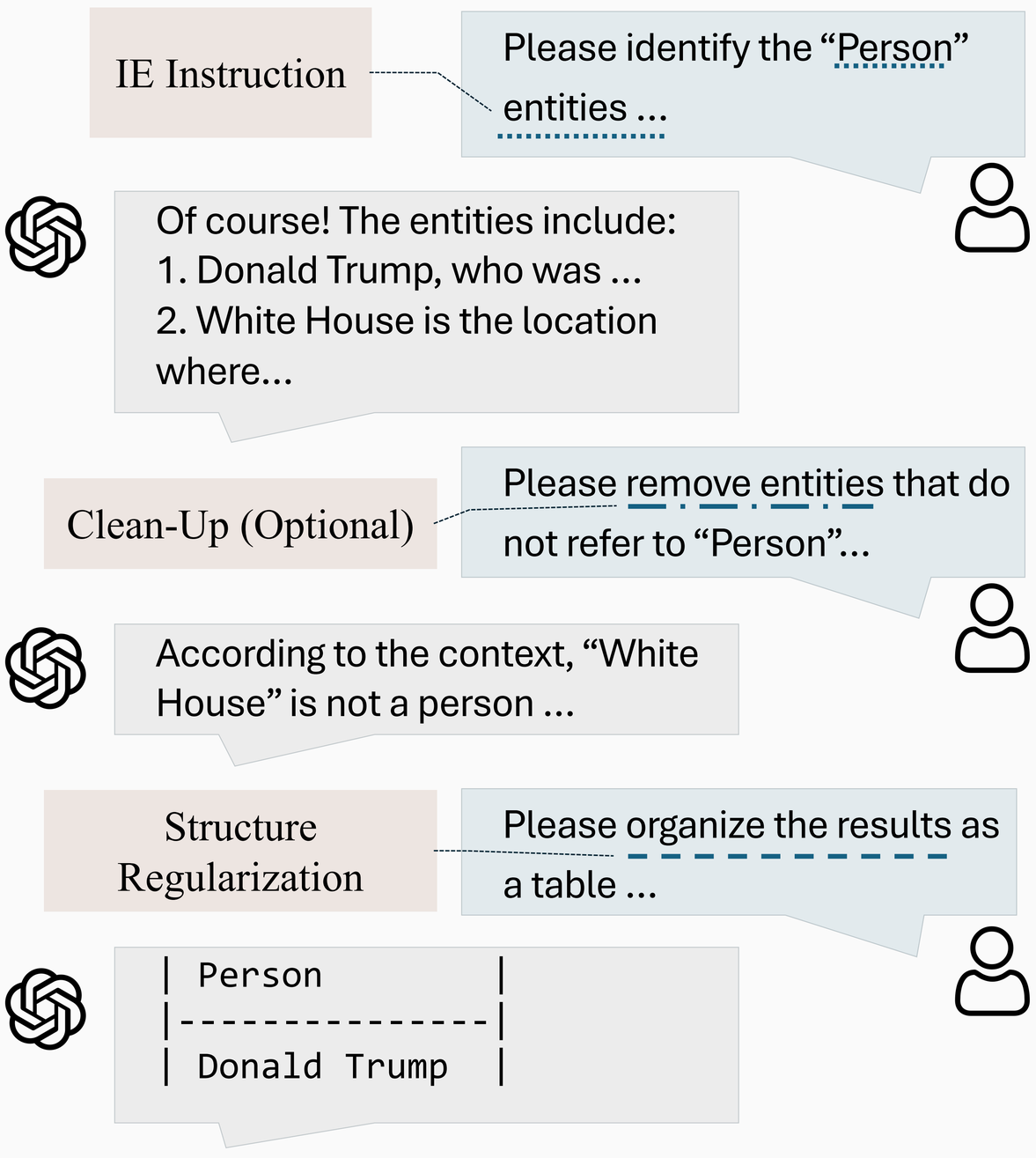
} Discrepancy!

*Mixing the IE instruction and the structure regularization instruction may confuse the model, increasing the difficulty of instruction following and resulting in sub-optimal performance.*

### ○ A Simple Solution

- ○ Formatting the output structure *after* generating answers

## Generate and Organize (G&O)




## Experiments

### ○ Datasets (Named Entity Recognition)

- ○ CoNLL 2003: general domain
- ○ NCBI Disease; BC5CDR: biomedical domain
- ○ PolyIE: materials science

### ○ Baselines (All built upon GPT-3.5-turbo)

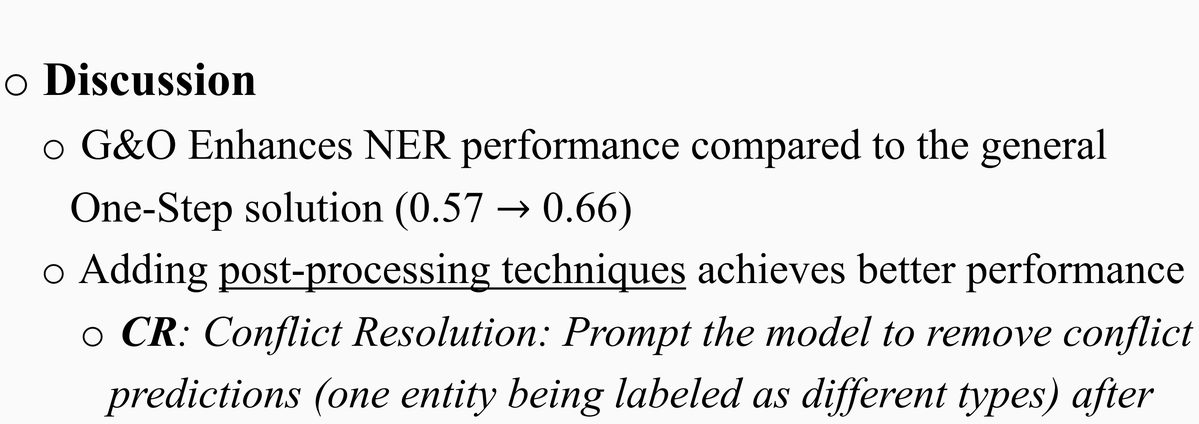
- ○ All-Entity-in-One (AEiO): identify entities of different types with one single question prompt; predicting **before** organizing.
- ○ One-Step: Let the model identify the entities and organize the results in the desired format **at the same time**.

| | Co03 | BC5CDR | NCBI | PolyIE | Average |
|---|---|---|---|---|---|
| AEiO | 0.54 | 0.62 | | 0.13 | 0.43 |
| One-Step | 0.47 | 0.70 | 0.65 | 0.47 | 0.57 |
| **G&O-NER** | 0.66 | 0.76 | 0.69 | 0.54 | **0.66** |
| - CoT | 0.66 | 0.66 | 0.57 | 0.46 | 0.59 |
| - Clean-Up | 0.70 | 0.74 | 0.65 | 0.51 | 0.65 |
| + CR | 0.68 | 0.77 | | 0.60 | **0.68** |
| + FT | 0.72 | 0.79 | 0.77 | 0.76 | **0.76** |

### ○ Discussion

- ○ G&O Enhances NER performance compared to the general One-Step solution (0.57 → 0.66)
- ○ Adding post-processing techniques achieves better performance
  - ○ ***CR**: Conflict Resolution: Prompt the model to remove conflict predictions (one entity being labeled as different types) after the label generation*
  - ○ ***FT**: BERT Fine-Tuning: Fine-tune a BERT model with GPT-predicted labels; can also be used to resolve conflicts*
- ○ Chain-of-Thought (CoT) plays an important role in G&O
  - ○ *Without CoT, G&O performs similar to One-Step*

Comparing G&O to baseline with different LLMs




Performance on Relation Extraction datasets