

# 对话摘要最新进展简述

冯夏冲

2022.1

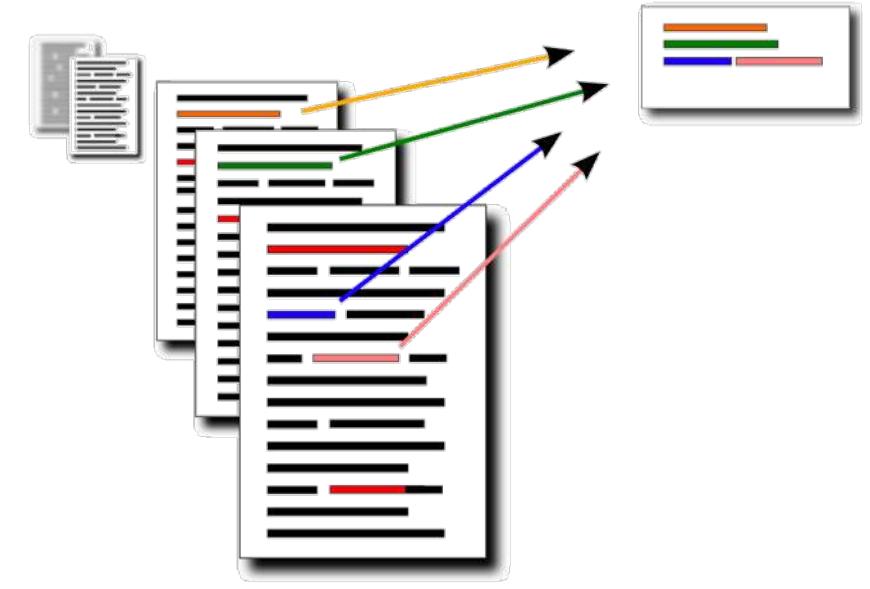
# 目录

- 任务介绍
- 相关工作
- 实验相关
- 未来趋势
- 总结

# 任务介绍

# 文本摘要

- 随着互联网产生的文本数据越来越多，  
文本信息过载问题日益严重。
- 因此，对各类文本进行一个“**简化**”处  
理显得非常必要，文本摘要便是其中一  
个重要的手段。
- 文本摘要：文本摘要旨在将文本或文本  
集合转换为包含**关键信息**的简短文本。



Source Text: Peter and Elizabeth took a taxi to attend the night party in the city.

While in the party, Elizabeth collapsed and was rushed to the hospital.

Summary: Peter

# 摘要任务分类



单文档



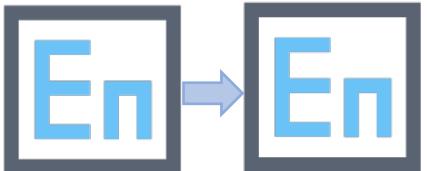
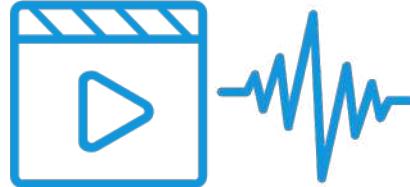
多文档



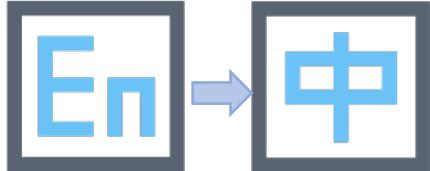
单模态



多模态



单语言



跨语言



新闻



专利



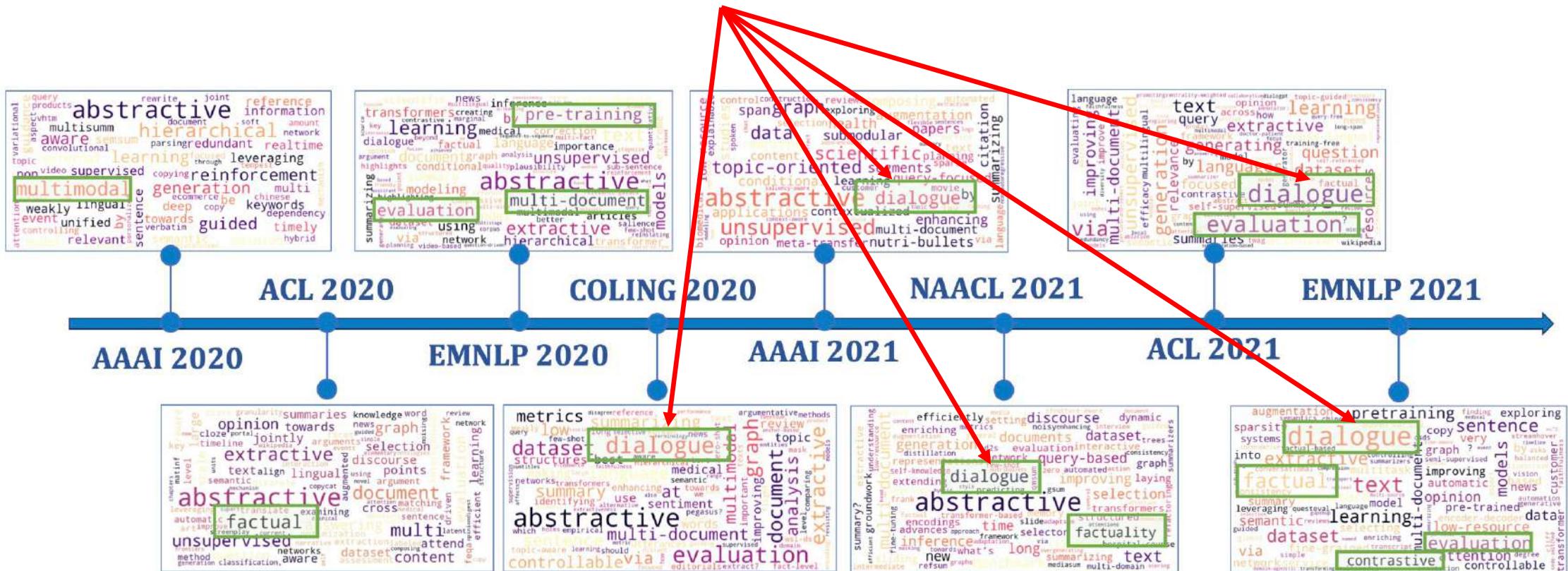
论文



对话

# 摘要任务进展

- 较为火热的方向：多模态摘要、多文档摘要、跨语言摘要、摘要评价、事实一致性研究、对话摘要



# 对话摘要

- 对话摘要关注对话类文本
  - 会议，闲聊，邮件，客服对话，医患对话，辩论等

部分会议
工业设计师：如果我们有电源支架呢？
界面设计师：你可以为支架和遥控器设计一些简洁的小设计。
项目经理：这会增加成本。
项目经理：我们需要改变最终的成本。
标准摘要
工业设计师建议在设备中加入一个电源支架，但最终被决定这不是一个有用的功能。

Meeting Minutes  
会议纪要

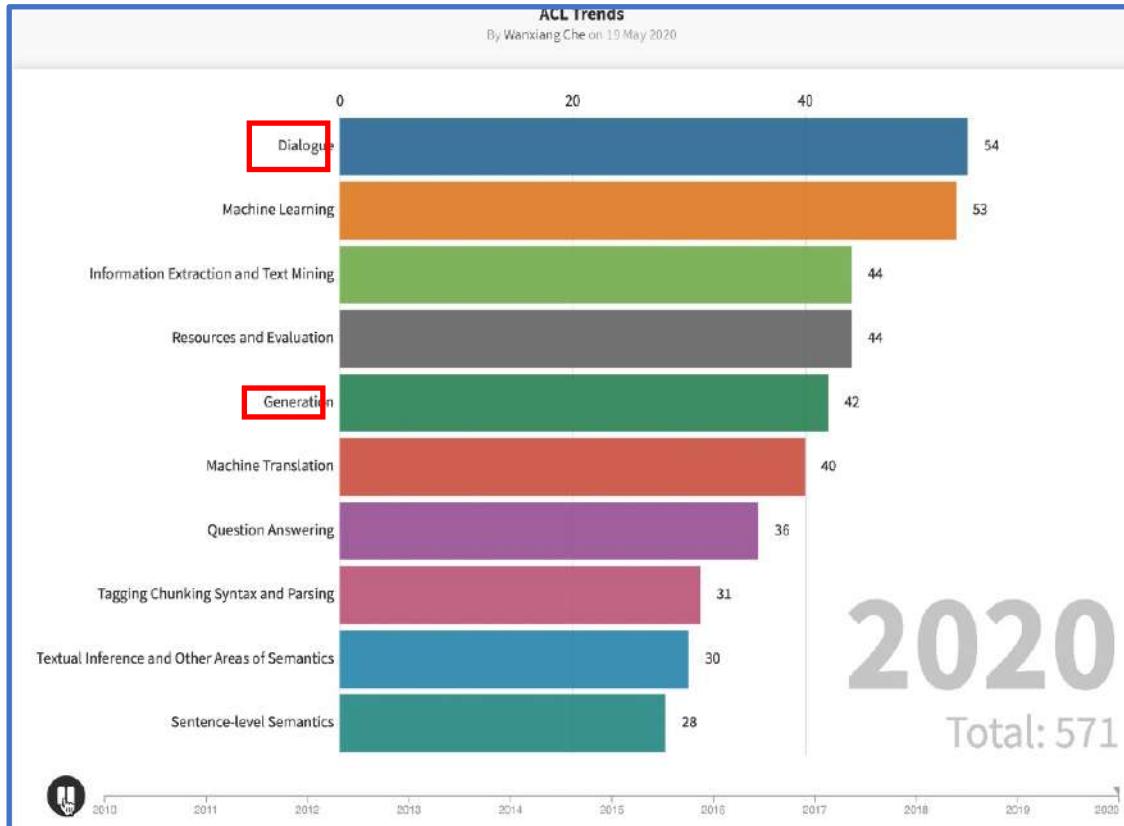
闲聊对话
鲍勃：老兄，你可以来接我一下吗？
汤姆：你在哪里？
鲍勃：在家，我的车坏了，我现在急需去上班，我需要你的帮助。
汤姆：我现在出发，10分钟之内到。
标准摘要
鲍勃的车坏了，汤姆会在10分钟内让他搭便车，送他去上班。

医患对话
医生：你最近有肿胀吗？
患者：时有时无。
医生：我知道了，什么时候开始的？
患者：大约在三周之前。
标准摘要
肿胀：大约三周之前开始，症状时有时无。

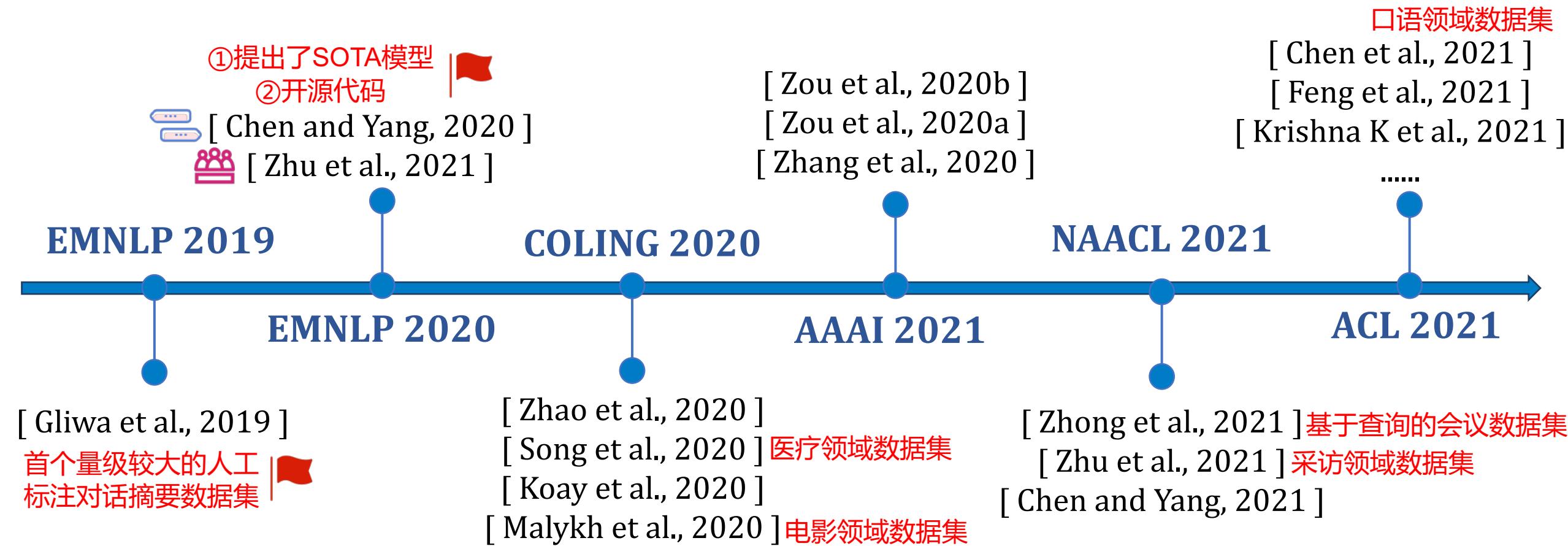
SOAP  
主观描述、客观观察、医生诊断、治疗计划

# 对话摘要的发展背景

- 对话摘要的发展得益于人机对话和文本生成技术的发展。



# 对话摘要的发展脉络



# 对话摘要的价值

对话类型

摘要示例

意义

## 会议摘要

### 部分会议

工业设计师：如果我们有电源支架呢?  
界面设计师：你可以为支架和遥控器  
设计一些简洁的小设计。  
项目经理：这会增加成本。  
项目经理：我们需要改变最终的成本。

### 标准摘要

工业设计师建议在设备中加入一个电源支架，但最终被决定这不是一个有用的功能。

帮助参会者捕捉冗长会议的核心内容，以便开展下一步工作。

## 闲聊对话摘要

### 闲聊对话

鲍勃：老兄，你可以来接我一下吗?  
汤姆：你在哪里?  
鲍勃：在家，我的车坏了，我现在急需去上班，我需要你的帮助。  
汤姆：我现在出发，10分钟之内到。

### 标准摘要

鲍勃的车坏了，汤姆会在10分钟内让他搭便车，送他去上班。

帮助说话人总结对话历史信息，快速开始新的对话。

## 医患对话摘要

### 医患对话

医生：你最近有肿胀吗?  
患者：时有时无。  
医生：我知道了，什么时候开始的?  
患者：大约在三周之前。

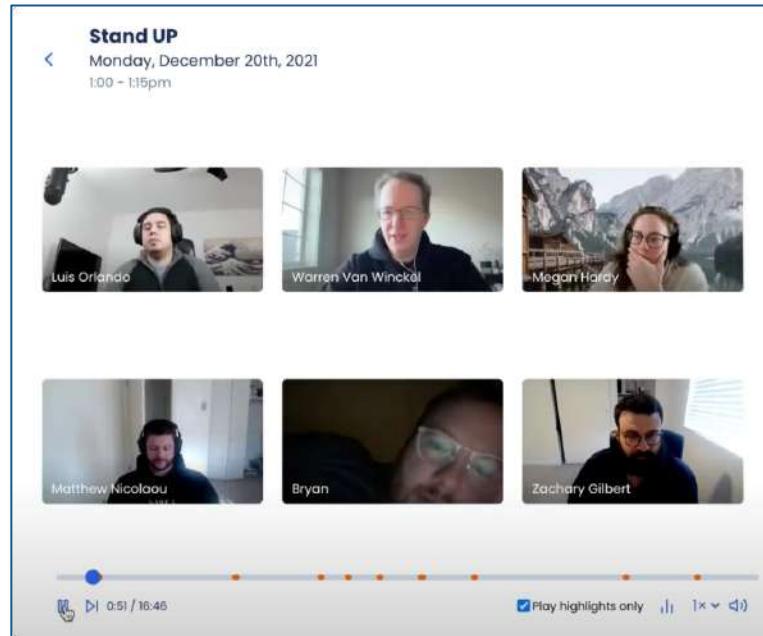
### 标准摘要

肿胀：大约三周之前开始，症状时有时无。

帮助医生集中于病人病情信息，剔除其他无用信息。

捕捉对话中的关键信息，帮助快速理解对话核心内容。

# 业界产品



Amazon

Headroom  
<https://www.youtube.com/watch?v=4qEi-eX46Cw>

A screenshot of a web browser window showing a POST request to the Symbl.ai API. The URL is https://api-labs.symbl.ai/v1/conversations/5776580840620288/summary?refresh=true. The request body contains a JSON object with a 'refresh' key set to 'true'. The response body shows a JSON summary of a conversation, including participant IDs, messages, and timestamps.

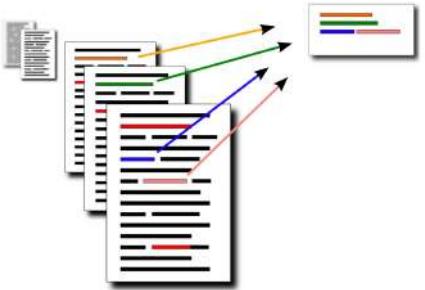
```
POST https://api-labs.symbl.ai/v1/conversations/5776580840620288/summary?refresh=true
Content-Type: application/json
{
  "refresh": true
}
{
  "summary": [
    {
      "id": "9749747991709224",
      "text": "Friend is a psychology major at millaps, and they like them. Speaker is the county coordinator for the re-election of the president."
    },
    {
      "id": "4868994204512768",
      "text": "I'm"
    },
    {
      "id": "5967999745303532"
    }
  ],
  "messages": [
    {
      "id": "9749747991709224",
      "text": "Friend is a psychology major at millaps, and they like them. Speaker is the county coordinator for the re-election of the president."
    },
    {
      "id": "4868994204512768",
      "text": "I'm"
    },
    {
      "id": "5967999745303532"
    }
  ]
}
```

Symbl.ai

# 相关工作

# 对话摘要的挑战

## 摘要生成



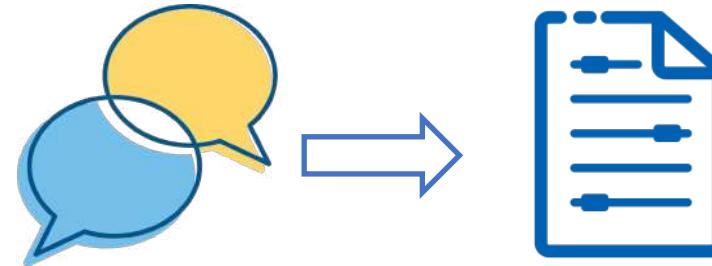
- 定位关键内容
- 摘要连贯性
- 摘要抽象性
- 事实一致性

## 对话理解



- 多人参与
- 结构丰富
- 主题漂移
- 指代频繁

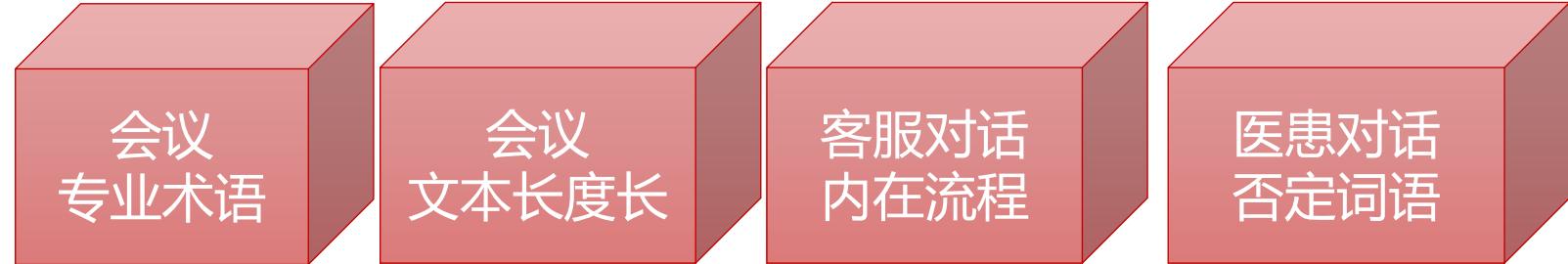
## 对话摘要



- 数据资源的挑战
- 对话建模的挑战
- 领域特定的挑战

# 对话摘要的挑战

## 领域特定的挑战



## 对话建模的挑战



## 数据资源的挑战



# | 对话摘要的挑战

**数据资源的挑战**

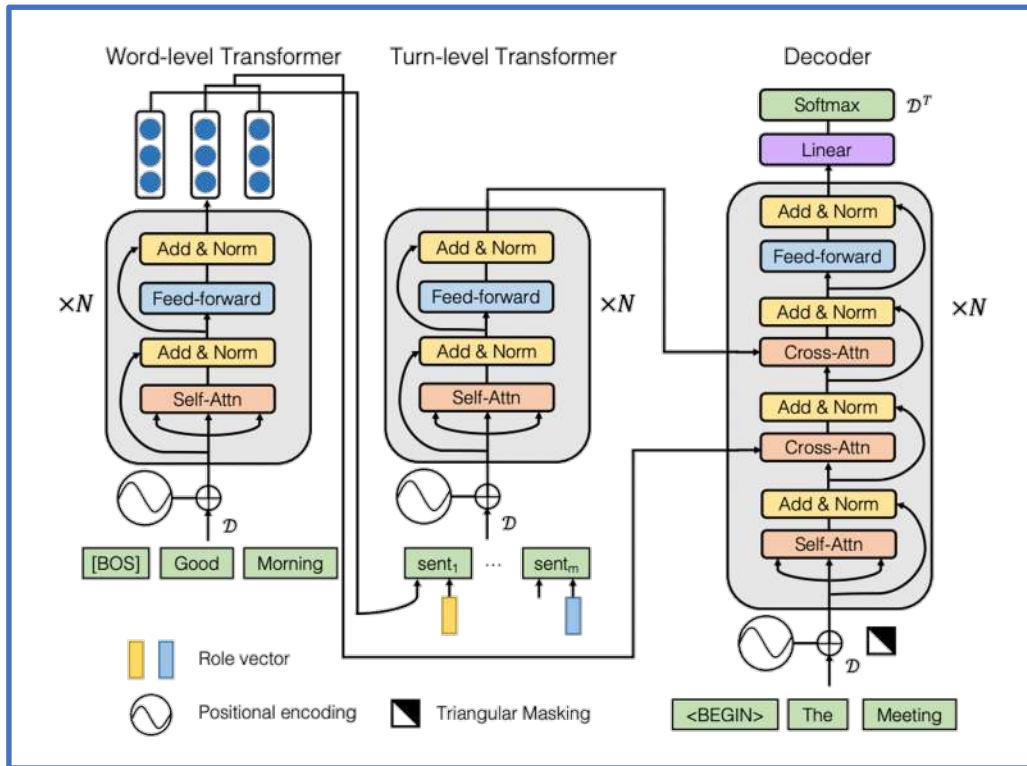
**数据稀缺**

# 数据集

编号	数据集	数据量	对话平均词语数	摘要平均词语数	说话人数量	生成式	抽取式	领域
1	AMI	137	4757.0	322.0	4.0	√	√	会议
2	ICSI	59	10189.0	534.0	6.2	√	√	会议
3	SAMSum	16.4k	83.9	20.3	2.2	√		闲聊
4	MediaSum	463.6k	1553.7	14.4	6.5	√		采访
5	QMSum	1.8k	9069.8	69.6	9.2	√		会议
6	SUMSCREEN	26.9k	6612.5	337.4	28.3	√		电视节目
7	SumTitles	21.4k	423.06	55.03	4.88	√		电影
8	DialogSum	13.4k	131	13.8	-	√		口语
9	LCSPIRT	38500	684.3	75	2	√		警方审讯（中）
10	EMAILSUM	2.5k	233.2	68.5	-	√		邮件
11	CSDS	10k	390.0	83.21	2	√		客服（中）
12	TODSum	9.9K	186.9	45.4	2	√		客服（英）

# 借助预训练：领域外数据

- 使用新闻摘要数据预训练模型

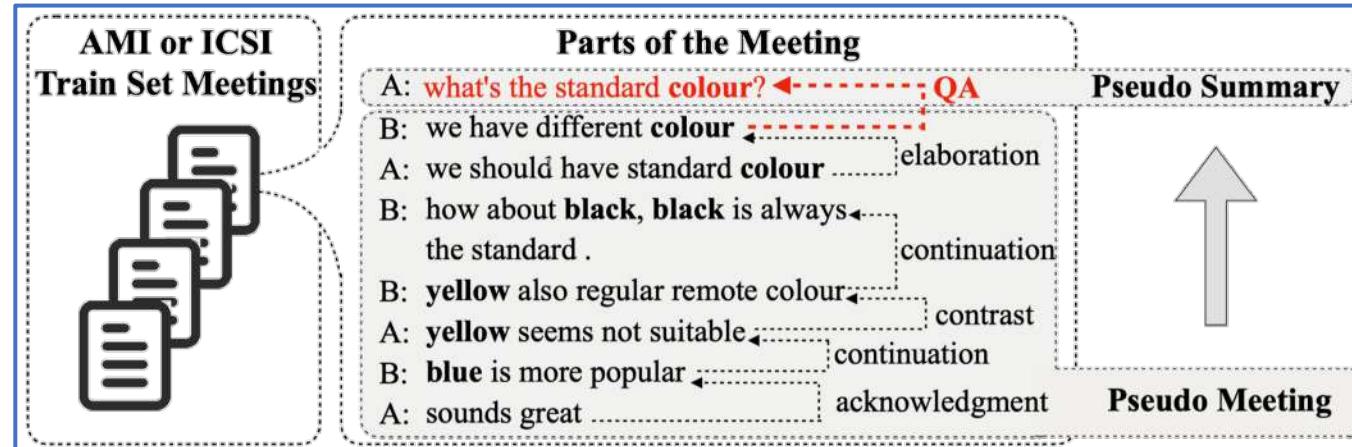


Model	ROUGE-1	R-2	R-SU4
<b>AMI</b>			
HMNet	<b>53.0</b>	<b>18.6</b>	<b>24.9</b>
–pretrain	48.7	18.4	23.5
–role vector	47.8	17.2	21.7
–hierarchy	45.1	15.9	20.5
<b>ICSI</b>			
HMNet	<b>46.3</b>	<b>10.6</b>	<b>19.1</b>
–pretrain	42.3	<b>10.6</b>	17.8
–role vector	44.0	9.6	18.2
–hierarchy	41.0	9.3	16.8

Table 3: Ablation study of HMNet.

# | 借助预训练：领域内数据

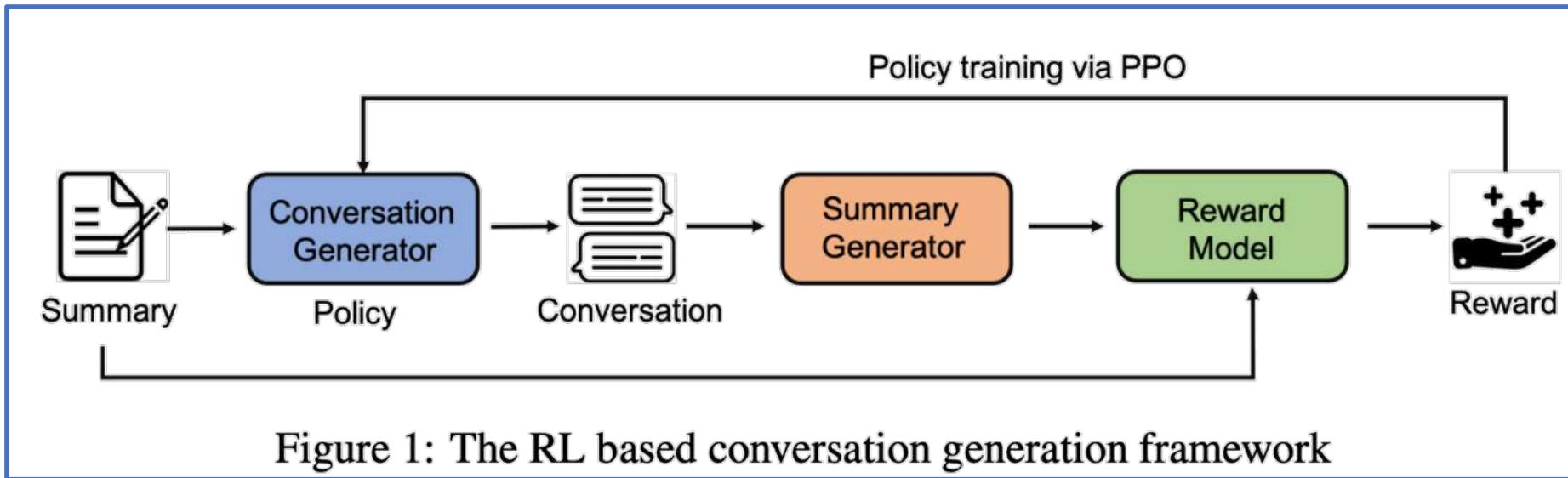
- 构造伪造会议摘要数据集用于预训练
  - “问题”会引起“讨论”，“问题”包含了“讨论”的核心内容。



AMI Pseudo Corpus	ICSI Pseudo Corpus
# of Original Data	97
# of Pseudo Data	1539
Avg.Tokens	124.44
Avg.Sum	13.18
	107.44
	11.97

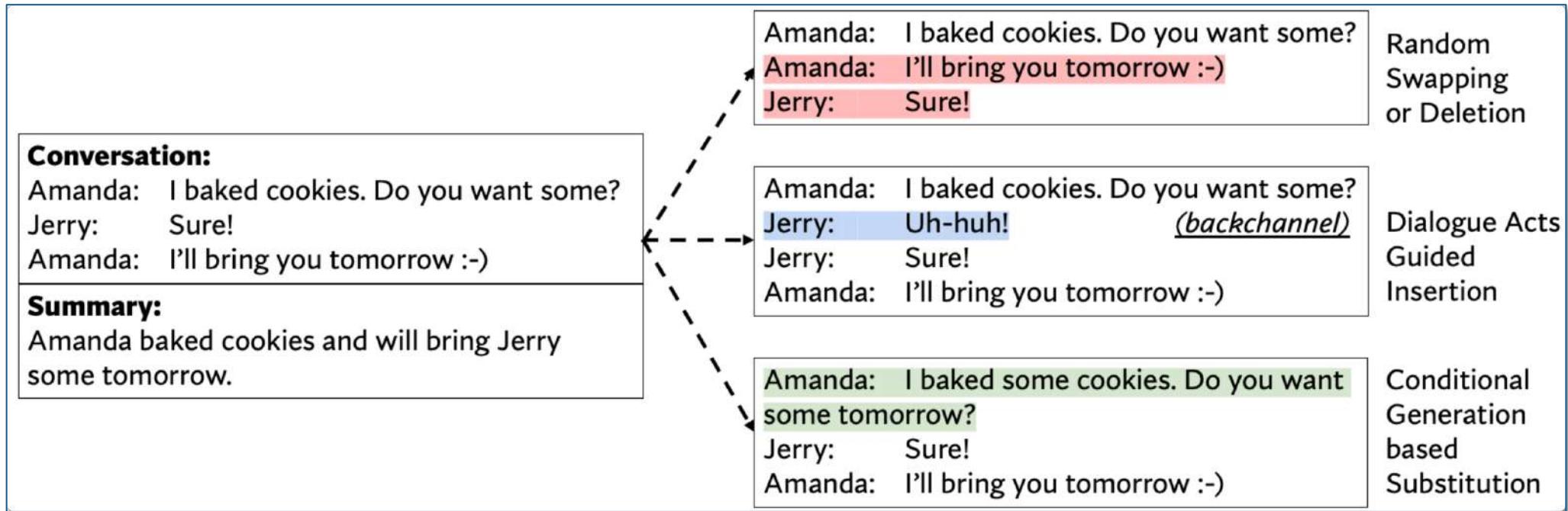
# 数据增强①

- 从摘要生成对话，进行数据增强



# 数据增强②

- 操作对话内容，进行数据增强。



# 无监督方法

- 基于相似度选择主题句+降噪自编码器

①  
训练句子  
相似度计  
算模型

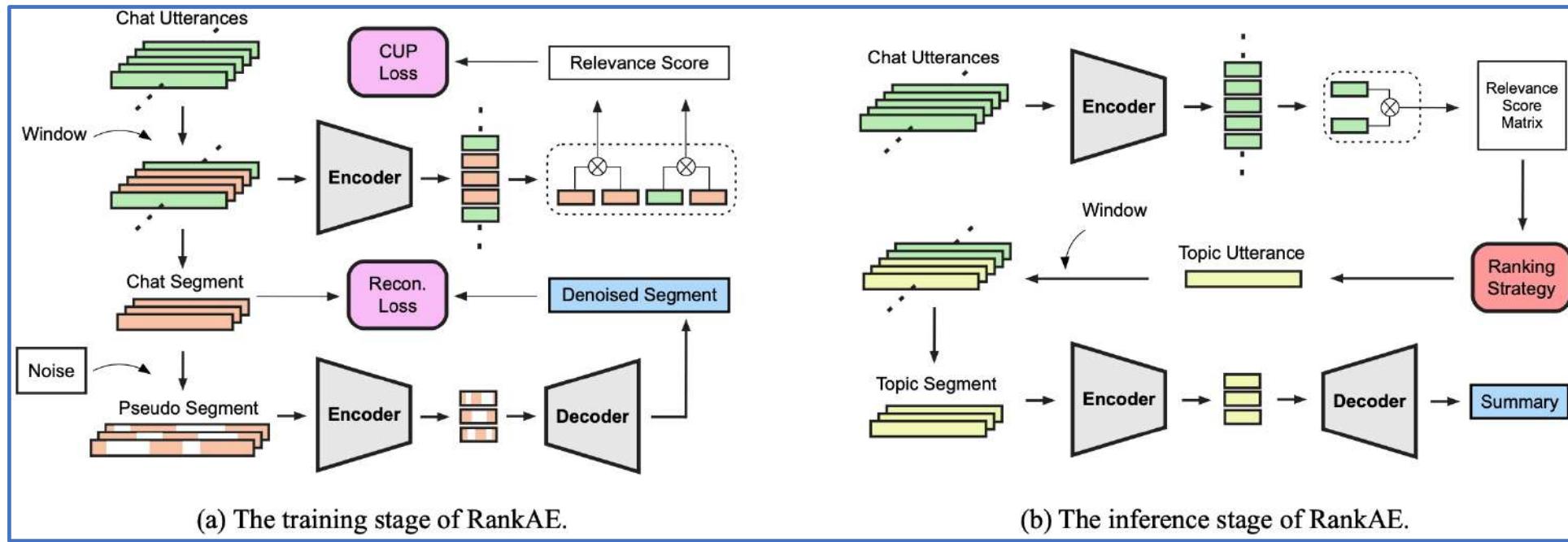
②  
训练降噪  
自编码器

③  
根据句子相似  
度，使用  
MMR算法选  
择主题句

④  
生成摘要

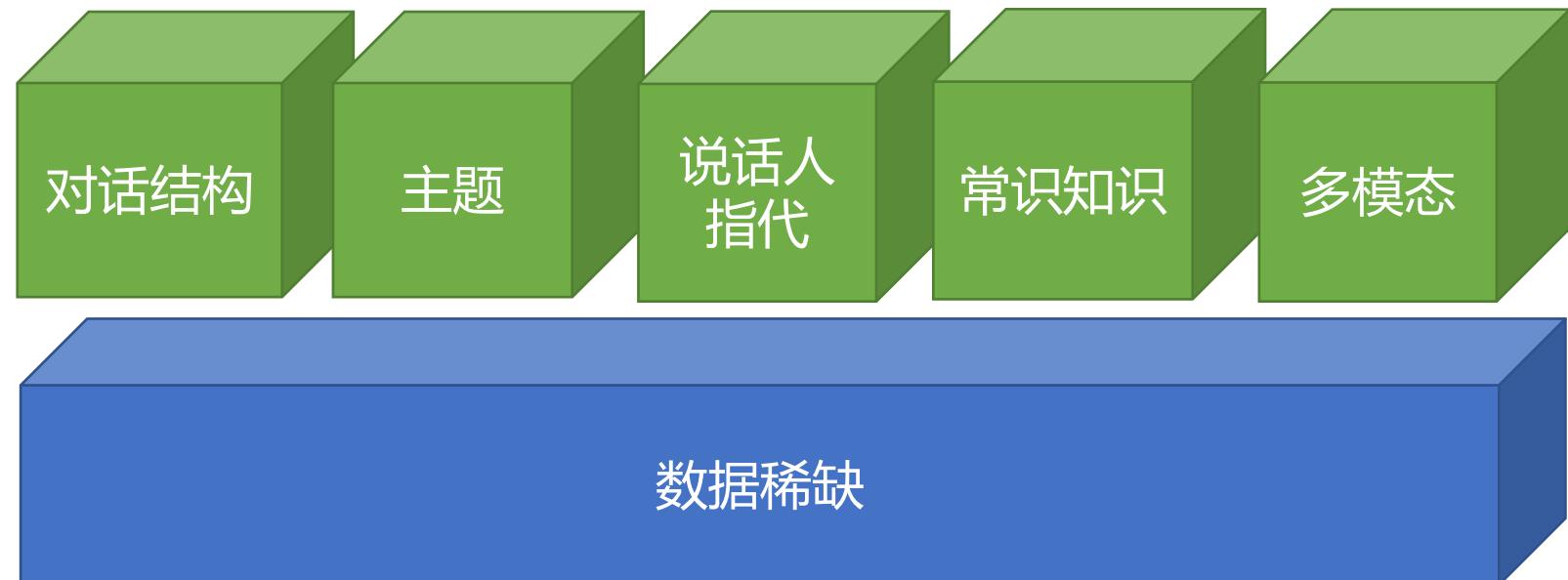
训练

测试



# 对话摘要的挑战

对话建模的挑战



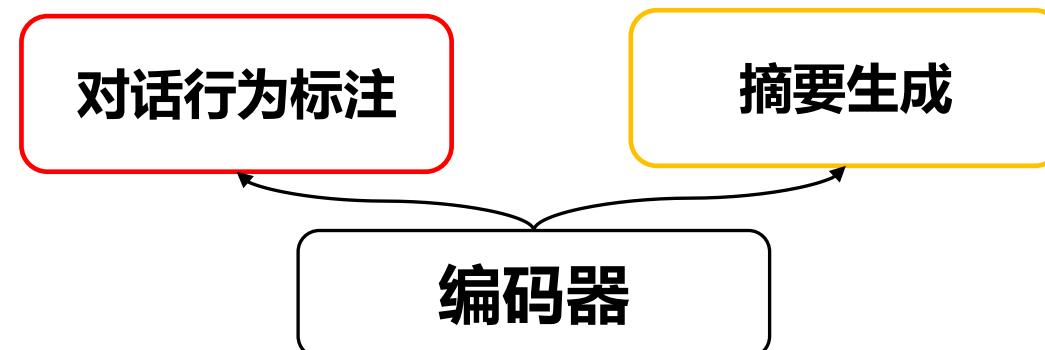
数据资源的挑战

# 对话行为信息

- 对话行为 (Dialogue Act) 指示了句子在对话中的作用与影响

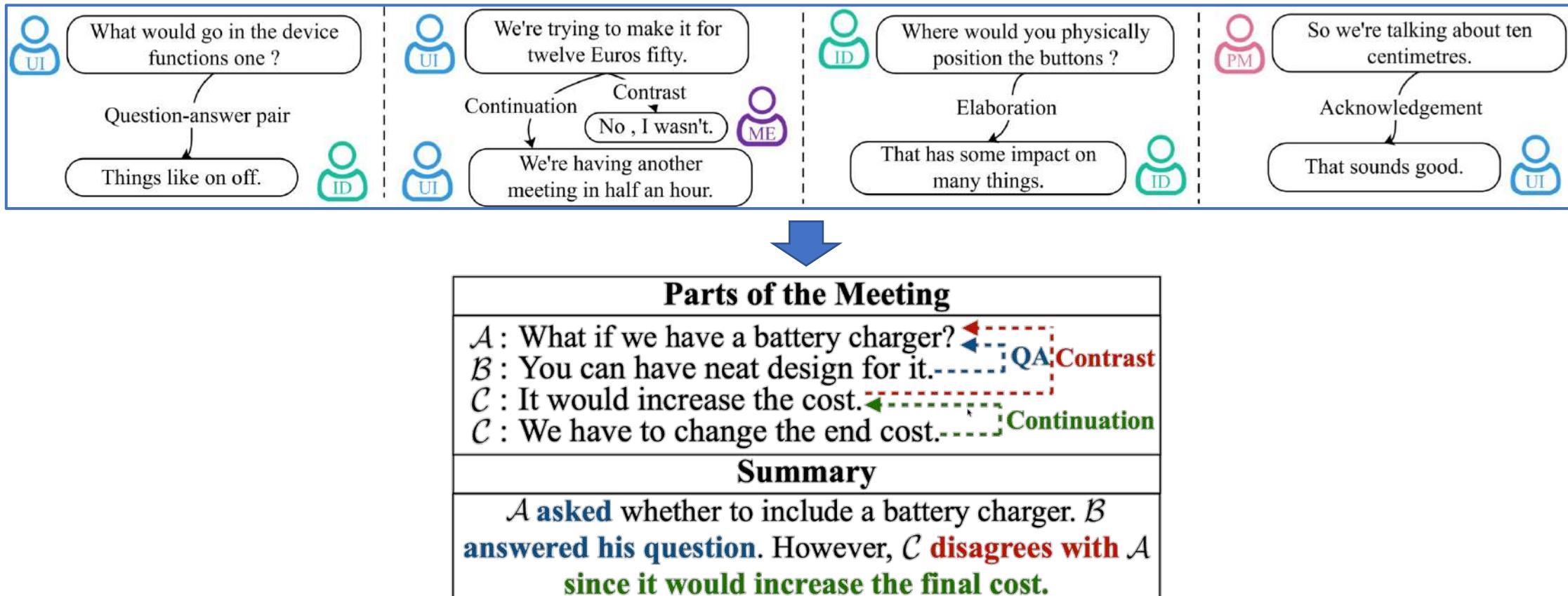
Multi-Party Dialogue	Dialogue Act
A: mm-hmm .	Backchannel
B: mm-hmm .	Backchannel
C: then , these are some of the remotes which are different in shape and colour , but they have many buttons .	Inform
C: so uh sometimes the user finds it very difficult to recognise which button is for what function and all that .	Inform
D: so you can design an interface which is very simple , and which is user-friendly .	Inform
D: even a kid can use that .	Inform
A: so can you got on t t uh to the next slide .	Suggest
Summary: alternative interface options	

- 模型：多任务学习



# 对话结构信息

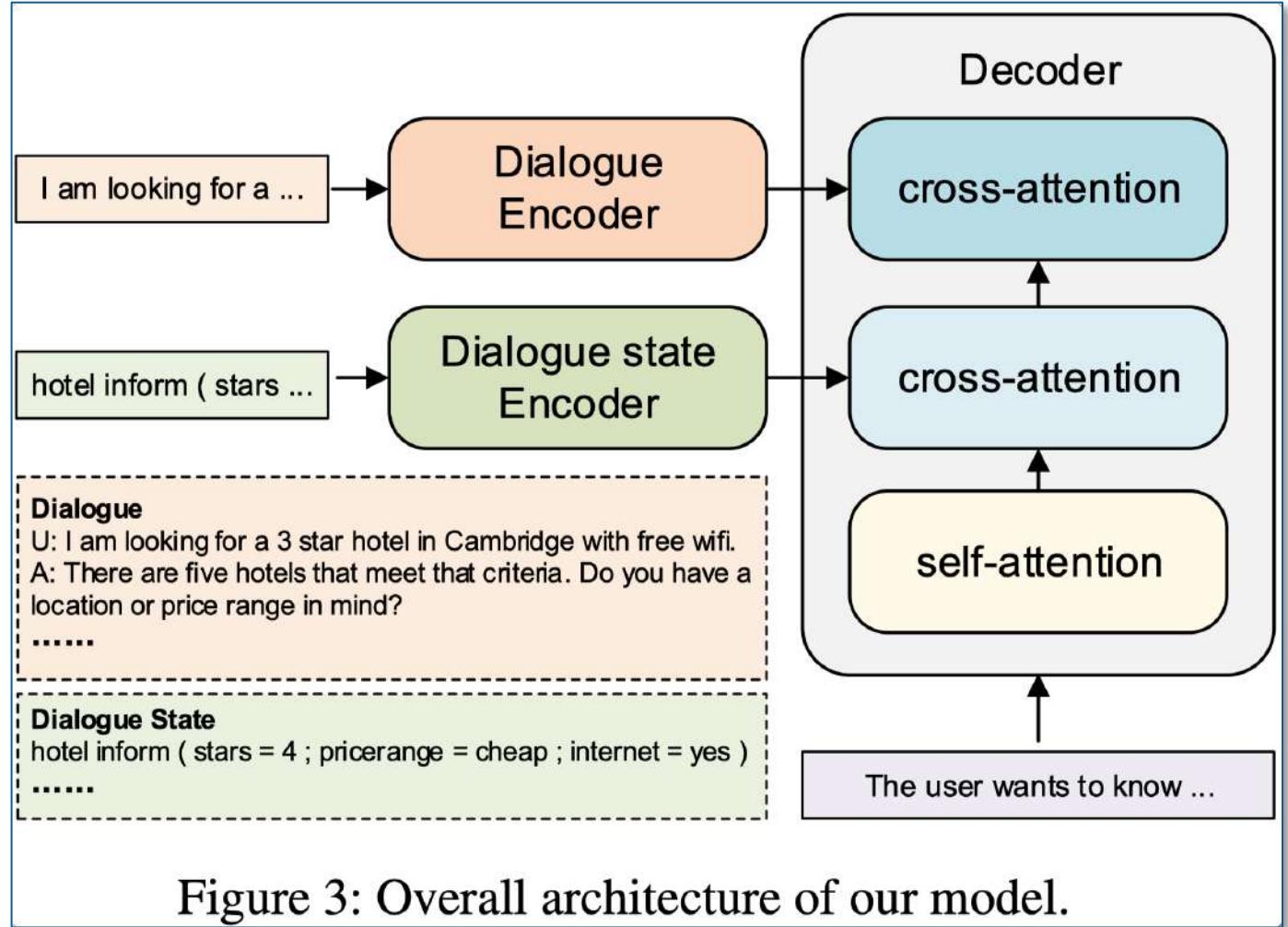
- 对话篇章结构指示了句子之间的交互关系



# 对话状态信息

Dialogue	
[domain1: hotel]	[dialogue state]
U: I am looking for a <u>3_star</u> hotel in Cambridge with free wifi.	hotel: (stars, 3), (area, Cambridge), (internet, yes)
A: There are five hotels that meet that criteria. Do you have a location or price range in mind?	
U: I would like something cheap.	
A: Unfortunately, none of the cheap hotels have free wifi. Would you prefer to try a different price range or go with a hotel without free wifi?	hotel: (stars, 3), (area, Cambridge), (internet, yes), (price range, cheap)
U: How about a hotel with a <u>4_star rating</u> ?	hotel : (stars, 4), (area, Cambridge), (internet, yes), (price range, cheap)
A: There are many what area would you like to stay in?	.....
[domain2: restaurant]	[dialogue state]
U: I also need a place to eat that serves italian food and should be in the same price range as the hotel.	restaurant: (food, italian), (price range, cheap)
.....	.....
[domain3: taxi]	[dialogue state]
U: I want to book a taxi from the hotel to the restaurant, that arrives at 11:30.	taxi: (arrives at, 11:30)
.....	.....
<b>Ground Truth</b>	
The user wants to know the <u>address</u> of the hotel, and the hotel is <u>4-star</u> and has internet. ...	
<b>BART Prediction</b>	
The user wants to know the <u>price range</u> of the hotel, and the <u>3-star</u> hotel is in the west and has internet. ...	
<b>BART+DS Prediction</b>	
the user wonders what the <u>address</u> of the hotel, and the <u>4-star</u> hotel has internet and cheap pricing. ...	

Figure 1: An example of TODSum with dialogue states. There exists three challenges: **Factual Inconsistency** (factual errors in generated summaries like 3-star), **Repetition** and **Negotiation** (see *hotel*), and **Multiple Domains**.



# 主题信息

- 主题漂移 (Topic Drift) 是对话中的一种常见现象

对话级别

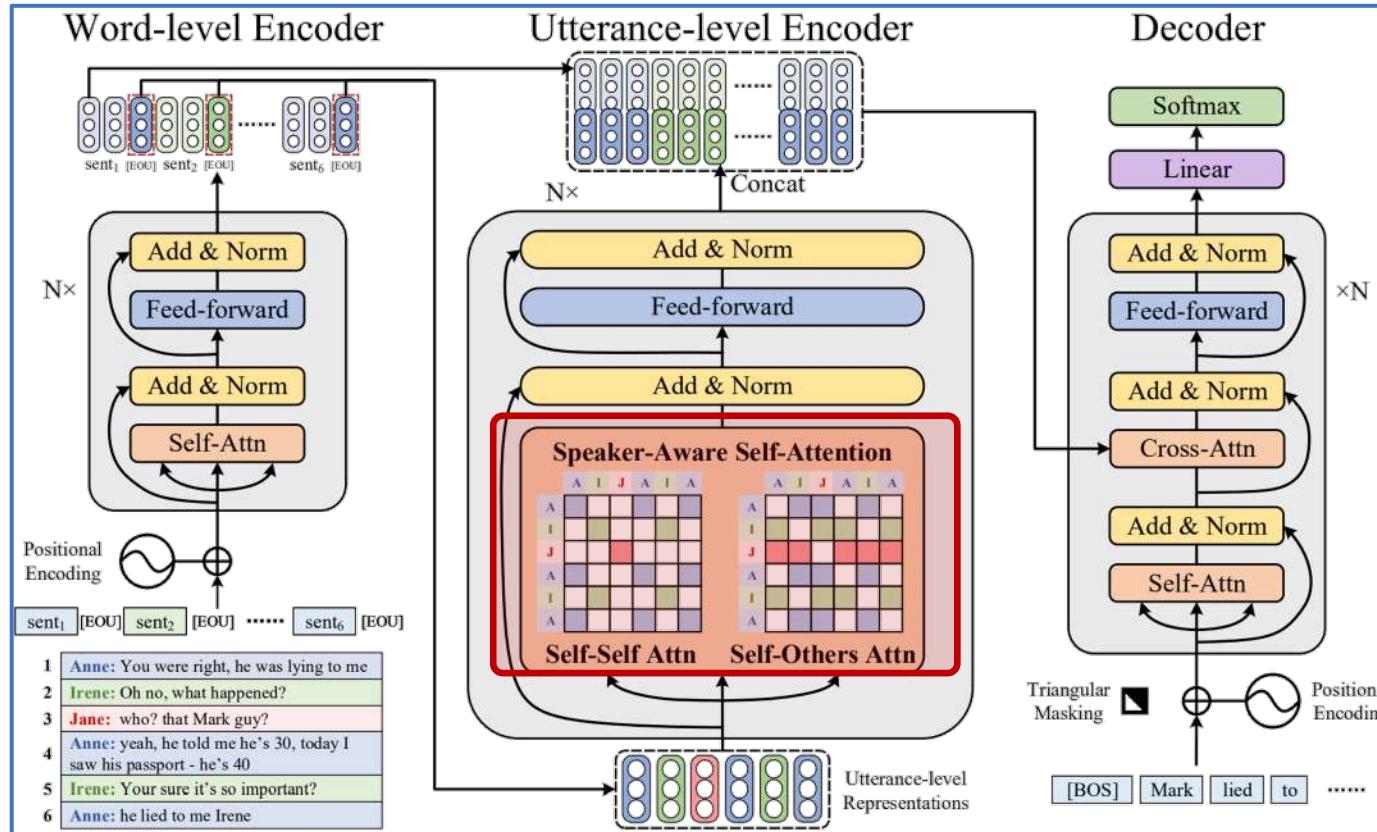
句子级别

对话级别	主题级别	阶段级别
<b>Conversation</b> James: Hey! I have been thinking about you : ) Hannah: Oh, that's nice ; ) James: What are you up to? Hannah: I'm about to sleep James: I miss u. I was hoping to see you Hannah: Have to get up early for work tomorrow James: What about tomorrow? Hannah: To be honest I have plans for tomorrow evening James: Oh ok. What about Sat then? Hannah: Yeah. Sure I am available on Sat James: I'll pick you up at 8? Hannah: Sounds good. See you then.	<b>Topic View</b> Greetings Today's plan Plan for tomorrow Plan for Saturday Pick up time	<b>Stage View</b> Openings Intention Discussion Conclusion
<b>Summary</b> James misses Hannah. They agree for James to pick Hannah up on Saturday at 8.		

Table 1: Example conversation from SAMSum (Gliwa et al., 2019) with its topic view and stage view (extracted by our methods), and the human annotated summary.

# 参与者信息

- 同一说话人之间的注意力机制 (Self-Self Attn)
- 不同说话人之间的注意力机制 (Self-Others Attn)



# 参与者信息

## Dialogue Content:

John: I missed our 5-year college reunion. I was down with a terrible flu.

Mary: Let me fill you in on the gossips!

John: Oh, please

Mary: Tony and Bell split up.

John: What?! They have been together for 8 years!

Mary: Yeah, Bell met a new guy. He is really handsome, by the way. He came with her to the reunion.

John: Was Tony there? Must have been awkward....

Mary: Yeah, Tony still wants to be friends for the sake of the children, but I think Bell prefers a clean cut.

**From John's Perspective:** Mary sent John some gossip from her college reunion. John missed as he was down with the flu.

**From Tony's Perspective:** Tony and Bell split up after 8 years of marriage. Tony still wants to be friends for the sake of the children.

**Comprehensive Planning:** {John, Mary, Tony, Bell}

**Output:** Mary sent John some gossip from her college reunion. John missed the reunion. Tony and Bell split up. Bell met a new guy. He came with her to the reunion.

**Baseline Model w/o Conditional Generation:** Mary sent John some gossip from her college reunion. John missed the reunion.

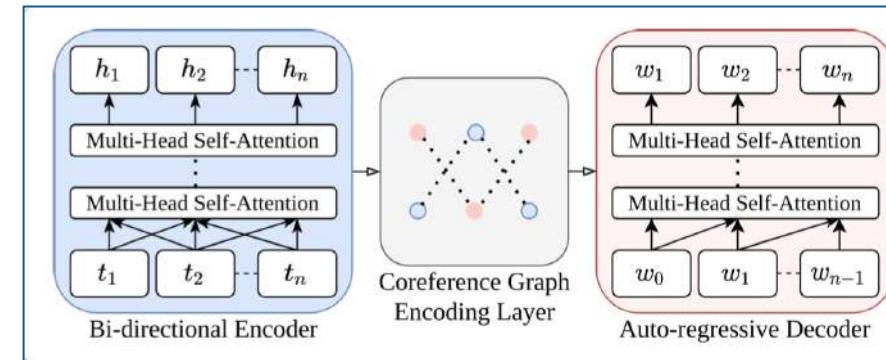
# 共指信息

**Max:** Know any good sites to buy clothes from?  
**Payton:** Sure :) <file\_other> <file\_other> <file\_other>  
**Max:** That's a lot of them!  
**Payton:** Yeah, but they have different things so I usually buy things from 2 or 3 of them.  
**Max:** I'll check them out. Thanks.  
...  
**Max:** Do you like shopping?  
**Payton:** Yes and no.  
**Max:** How come?  
**Payton:** I like browsing, trying on, looking in the mirror and seeing how I look, but not always buying.  
...  
**Max:** So what do you usually buy?  
**Payton:** Well, I have 2 things I must struggle to resist!  
**Max:** Which are?  
**Payton:** Clothes, ofc ;)  
**Max:** Right. And the second one?  
**Payton:** Books. I absolutely love reading!  
...

**Base Model:** Payton is looking for good places to buy clothes. He usually buys things from 2 or 3 of them. He likes browsing and trying on clothes. Max likes reading books.

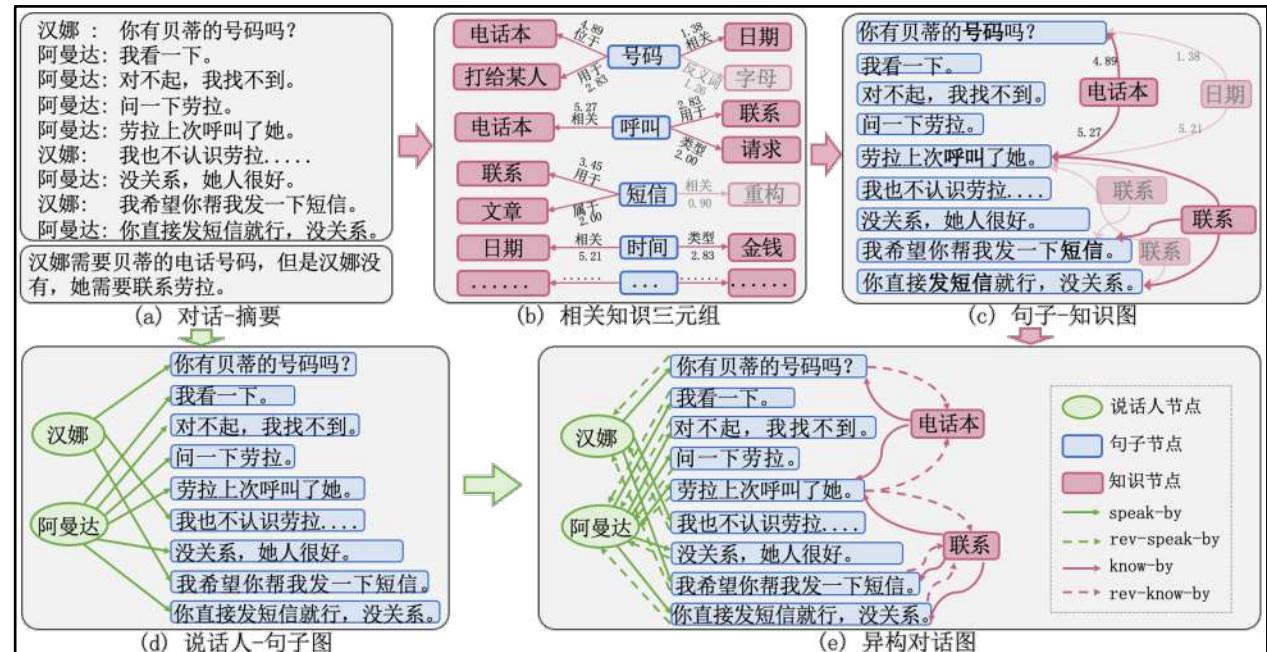
**Coreference-Aware Model:** Max will check out some good places to buy clothes. Payton likes browsing, trying on, looking in the mirror and seeing how she looks. Payton loves reading.

0 Riley : 1 Chloe is on tv!!  
2 James : on which channel?  
2 James : never mind, 2 i've found it  
2 James : What is 1 she doing? 2 i don't get it  
0 Riley : this is a programme in which women undergo a complete metamorphosis.  
0 Riley : OMG 1 she looks pretty gorgeous!



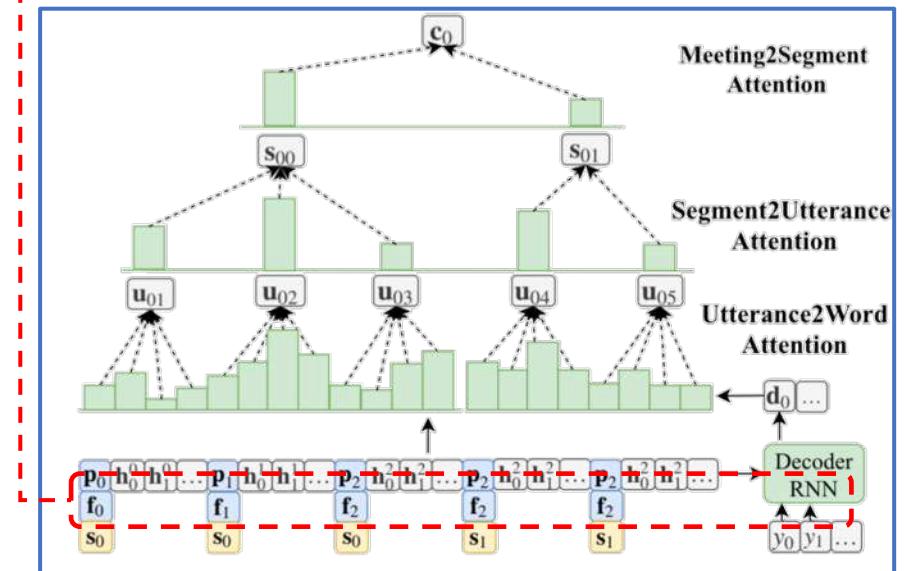
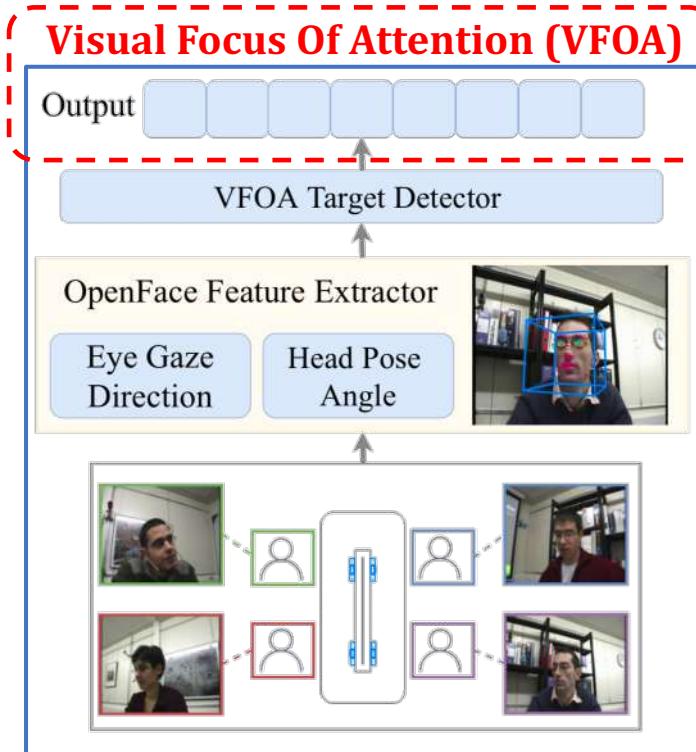
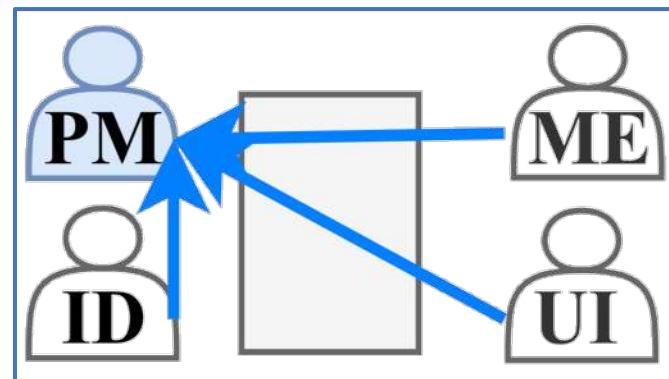
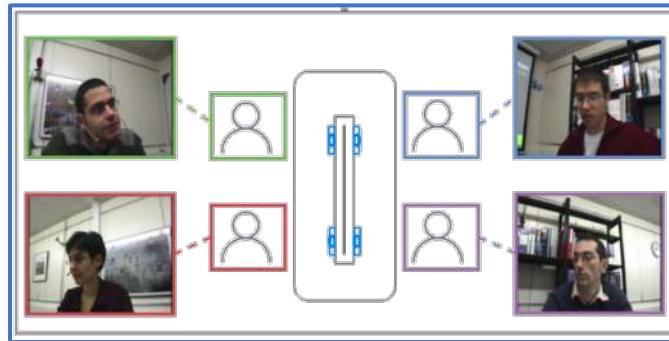
# 常识知识

- 对话参与者通过自己的常识知识理解对话内容，做出回复



# 多模态信息

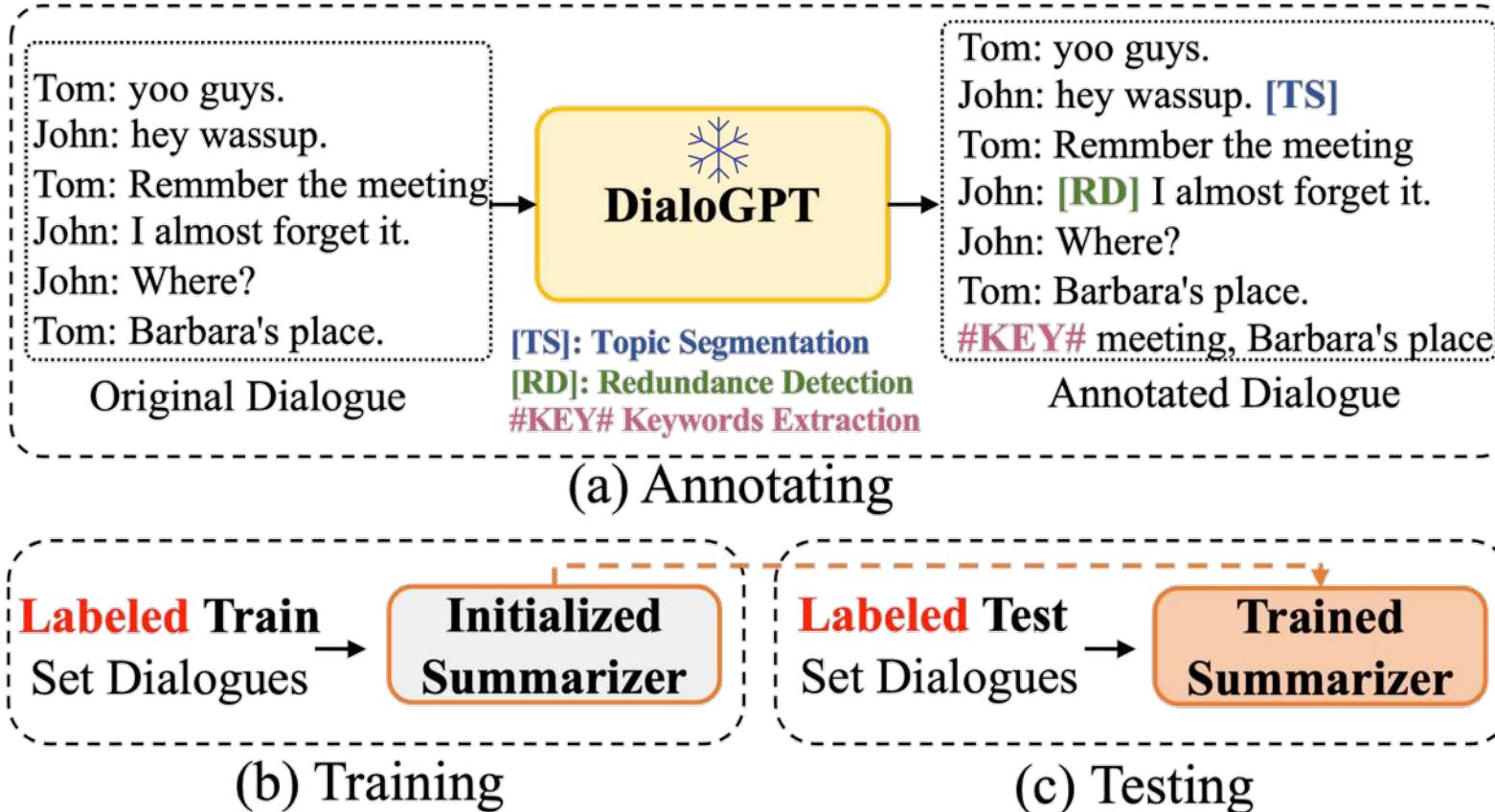
- 融入多模态信息定位关键内容



说话人被其他参与者注视的时间越长，该说话者的信息越重要。

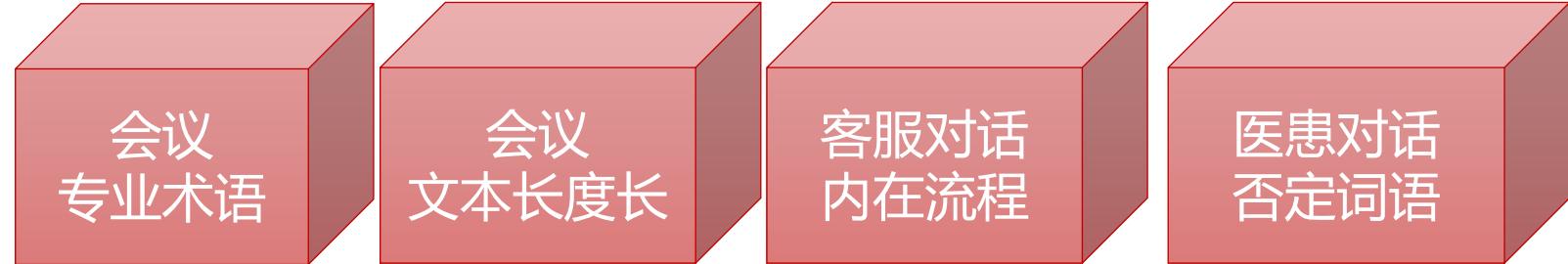
# 将预训练语言模型作为无监督标注器

- 关键词抽取、冗余句检测、主题分割



# 对话摘要的挑战

## 领域特定的挑战



## 对话建模的挑战



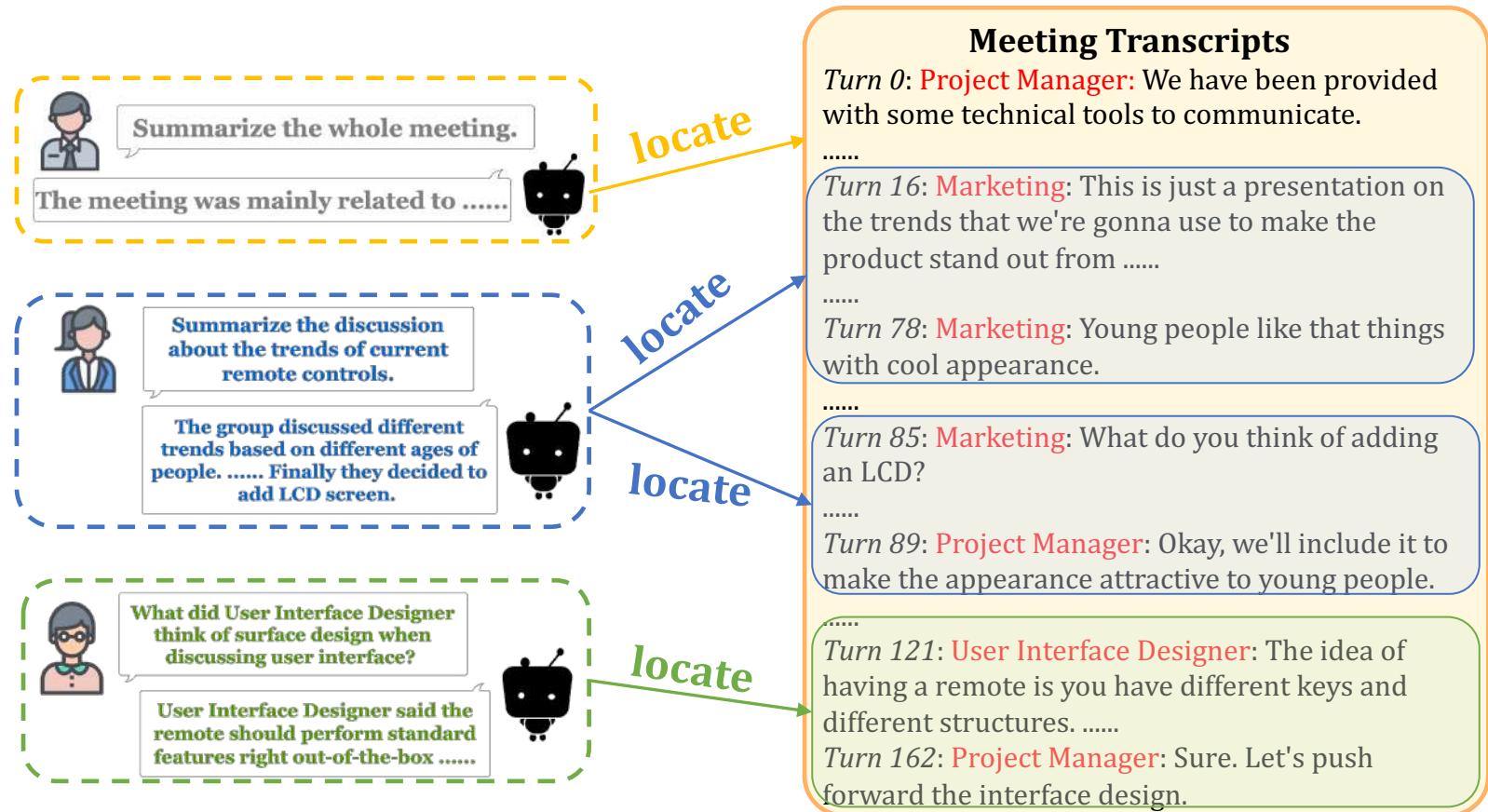
## 数据资源的挑战





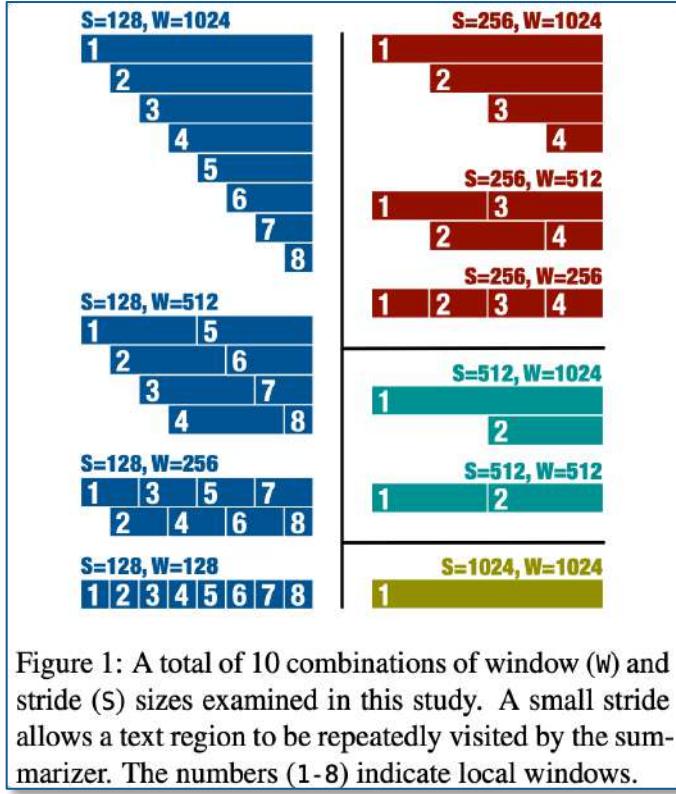
# 基于查询的会议摘要

- 各取所需，灵活度更高。

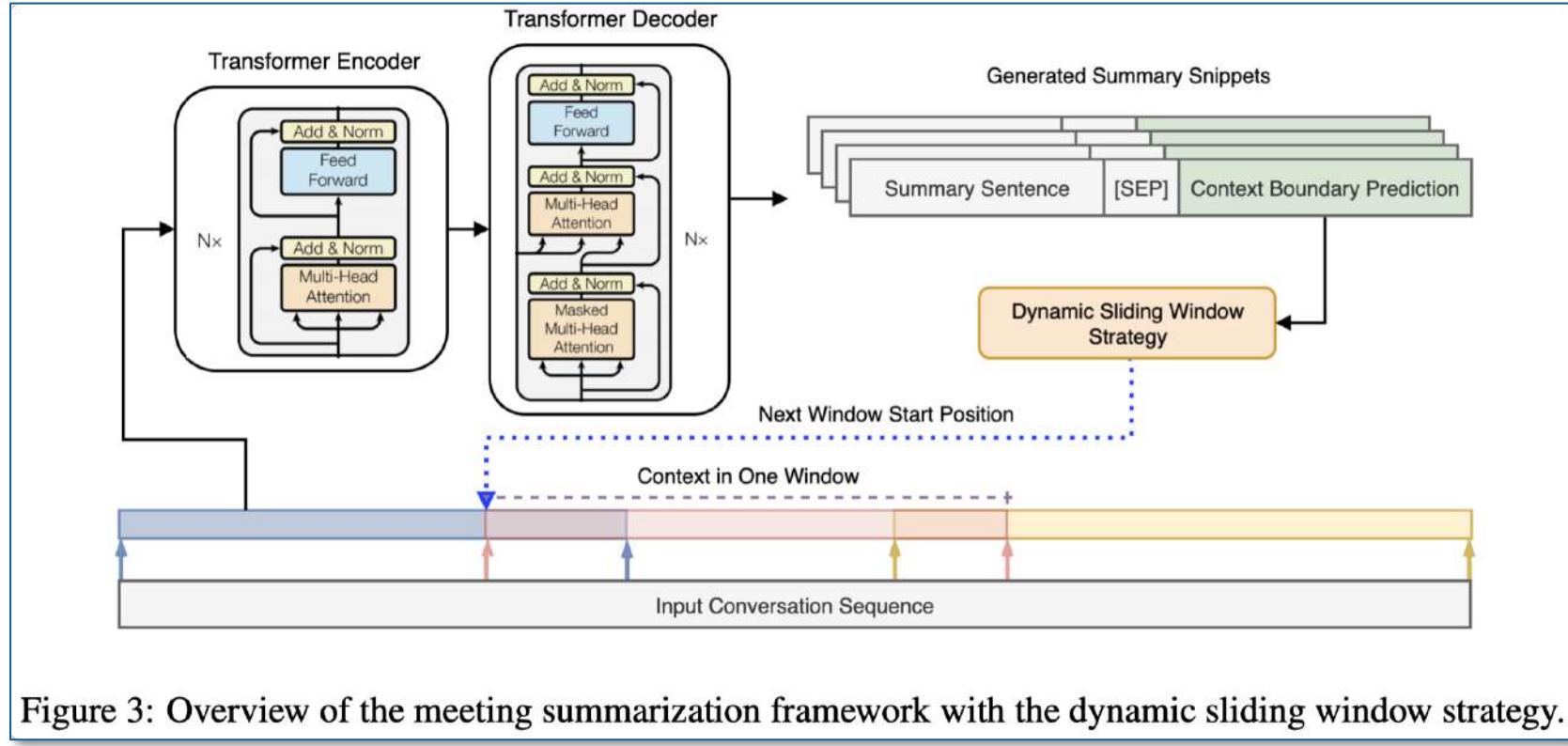




# 滑动窗口机制



静态窗口

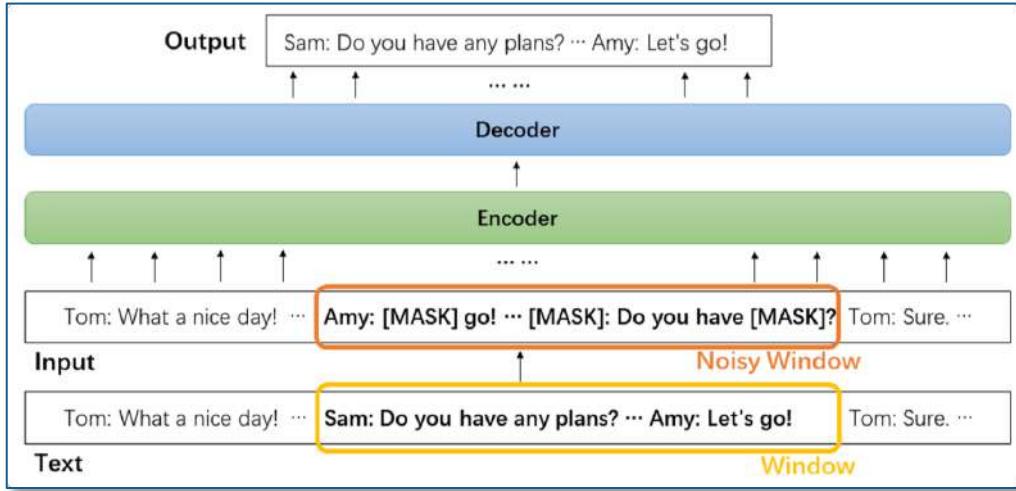


动态窗口



# 长文本预训练语言模型

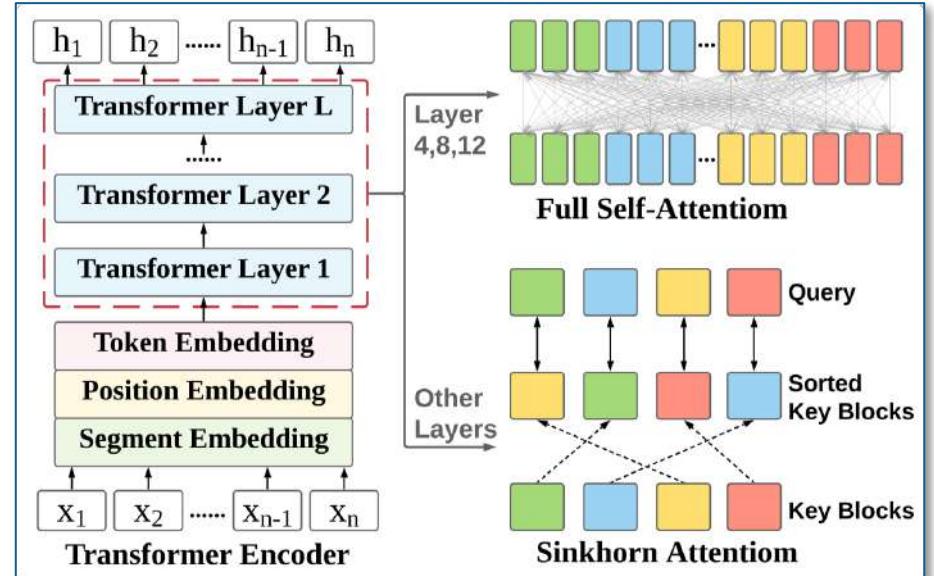
## 预训练 模型



## 预训练 任务

Noise Type	Original Dialogue	Noisy Dialogue
Speaker Mask	Tom: The weather is good today!	[MASK]: The weather is good today!
Turn Splitting	Tom: The weather is good today! Do you have any plans? How about we go to play basketball?	Tom: The weather is good today! [MASK]: Do you have any plans? [MASK]: How about we go to play basketball?
Turn Merging	Tom: The weather is good today! Do you have any plans? Bob: I still have homework to do today. I'm afraid I can't go out to play.	Tom: The weather is good today! Do you have any plans? I still have homework to do today. I'm afraid I can't go out to play.
Text Infilling	Tom: The weather is good today! Do you have any plans? How about we go to play basketball?	Tom: The weather is [MASK] Do you have [MASK] any plans? [MASK] we go to play basketball?
Turn Permutation	Tom: Do you have any plans? Bob: How about we go to play basketball? Sam: I still have homework to do today. I'm afraid I can't go out to play.	Sam: I still have homework to do today. I'm afraid I can't go out to play. Tom: Do you have any plans? Bob: How about we go to play basketball?

## 稀疏注意力机制





# 长文本编码

- 探索了三种编码长文本的方法
  1. Extended transformer models such as [Longformer](#),
  2. [Retrieve-then-summarize](#) pipeline models with several dialogue utterance retrieval methods, and
  3. [Hierarchical](#) dialogue encoding models such as HMNet.



# 医患对话中的主题信息

- 医生针对不同的症状进行询问

肿胀

头痛

胸痛

头晕

a)

[Nurse] Hi Mr.#name#, you were discharged on #date#. There are some questions I'd like to check with you.  
 [Patient] Ok, Ok.  
 [Nurse] Well, have you been experiencing swelling recently?  
 [Patient] Swelling? It comes and go, comes and go.  
 [Nurse] Comes and go ... I see .. #repetition#  
 [Nurse] ... #pause#... When did it start?  
 [Patient] Let me see, started from three weeks ago.

...  
 [Nurse] Are you experiencing any headache right now as we speak?  
 [Patient] Umm ... #back-channel#  
 [Nurse] Let me check, the last time you told me is sometimes at night.  
 [Patient] Oh, right, only a bit.

...  
 [Nurse] Still feel some chest pain or chest discomfort?  
 [Patient] Yes, my head is... #false-start# no, the pain is much better.  
 Still feel headache though ... #topic-drift#

...  
 [Nurse] Any giddiness or palpitation?  
 [Patient] Palpitation? Do not have-- #interruption#  
 [Nurse] Well ... Do you-- #interruption#  
 [Patient] and no giddiness, no, nothing.

...  
 [Nurse] Ok, you need to check your heartrate everyday.  
 [Nurse] Do you know how to use the device?  
 [Patient] Yes, yes, no problem.

...

b)

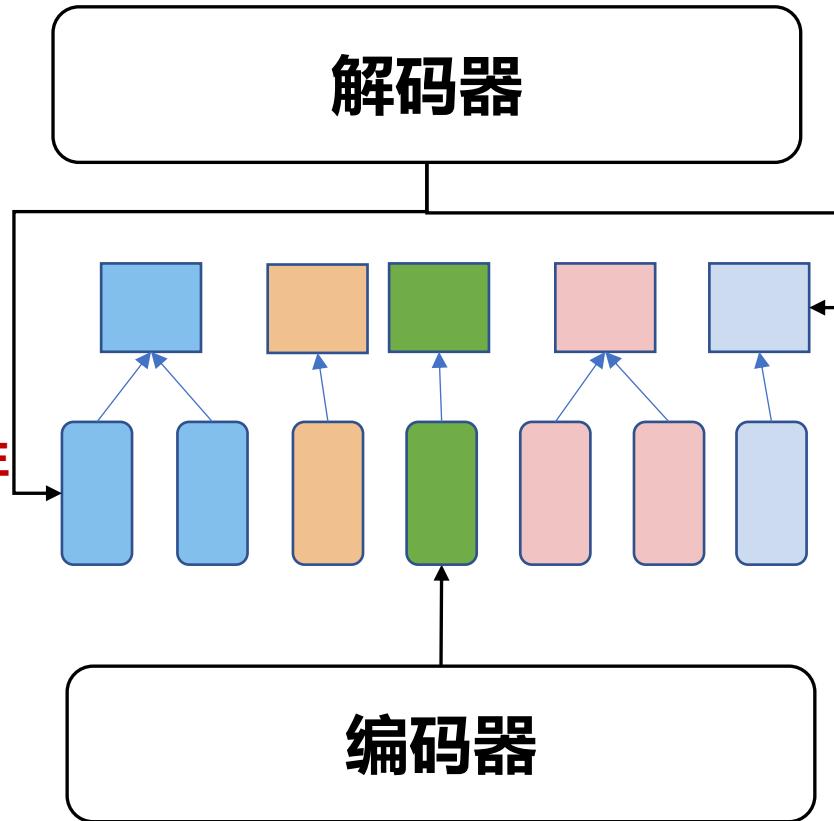
Swelling: started from three weeks ago, comes and go.  
 Headache: sometimes, at night, only a bit.  
 Chest pain: much better.  
 Dizziness: none.

词语级别注  
意力机制

解码器

主题级别注  
意力机制

编码器

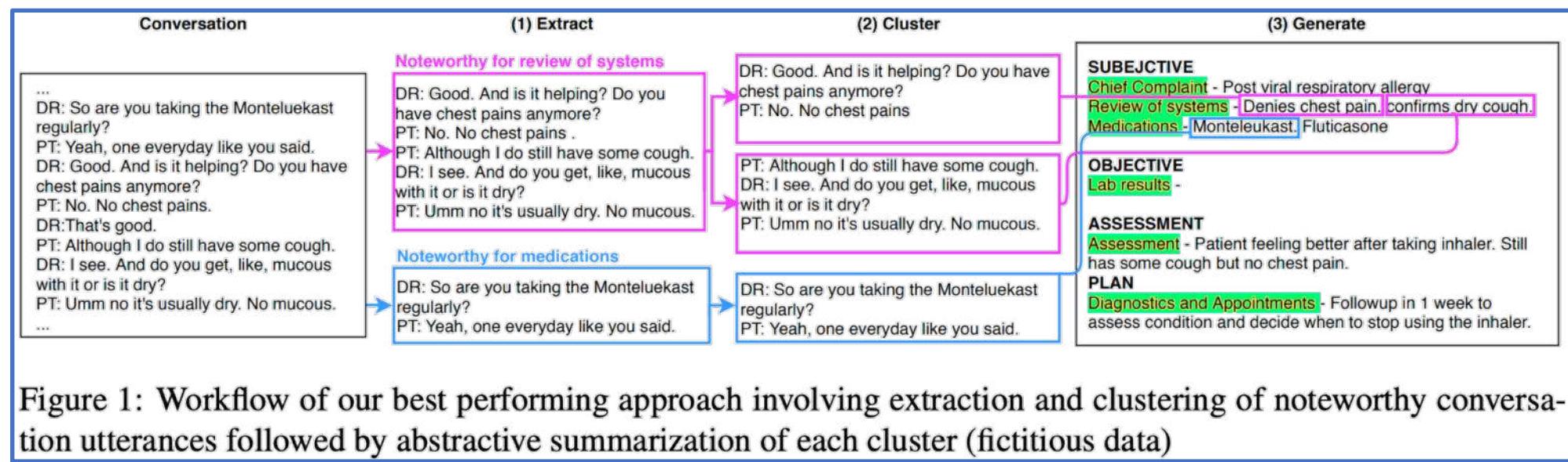




# 医患对话摘要

## • 医疗对话摘要：SOAP

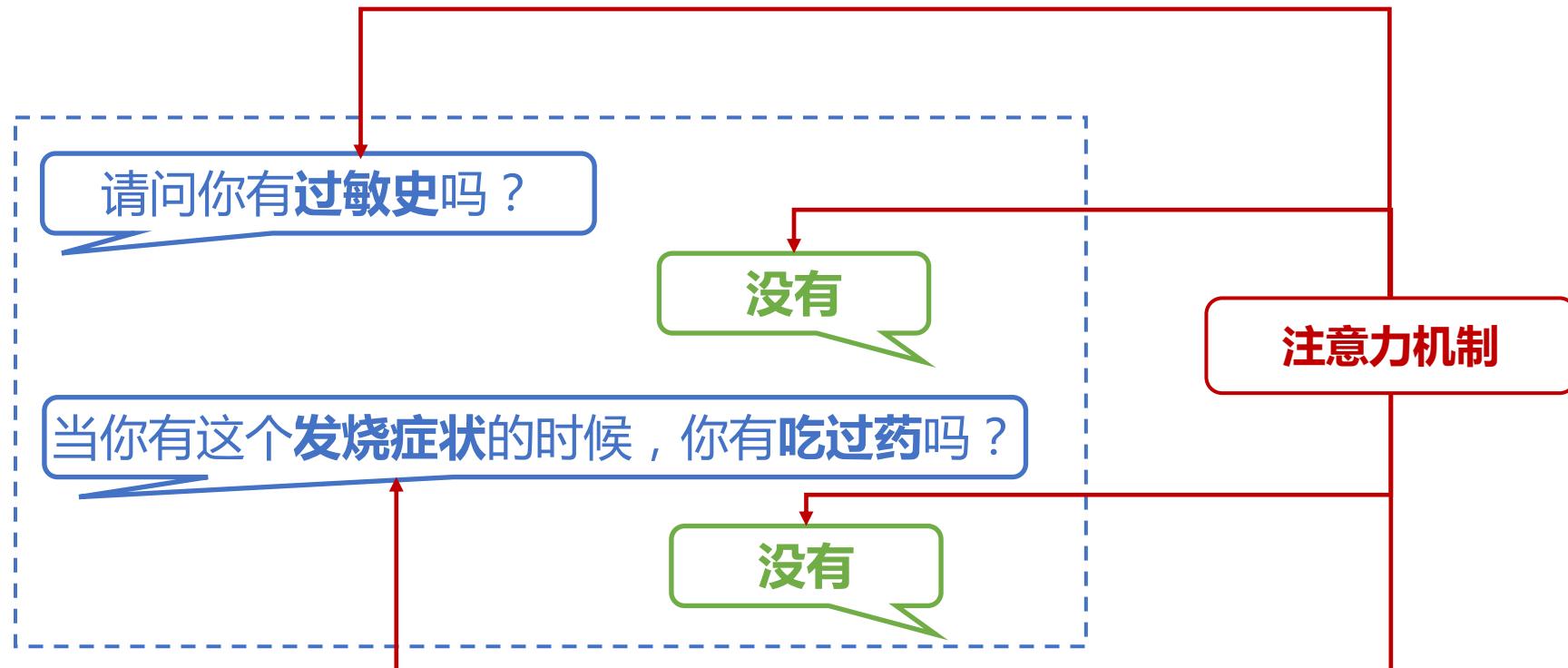
- (S)ubjective information reported by the patient
- (O)bjective observations, e.g., lab results
- (A)ssements made by the doctor (typically, the diagnosis)
- (P)lan for future care





# 医患对话中的否定回答

- 医患对话中的否定回答需要额外注意

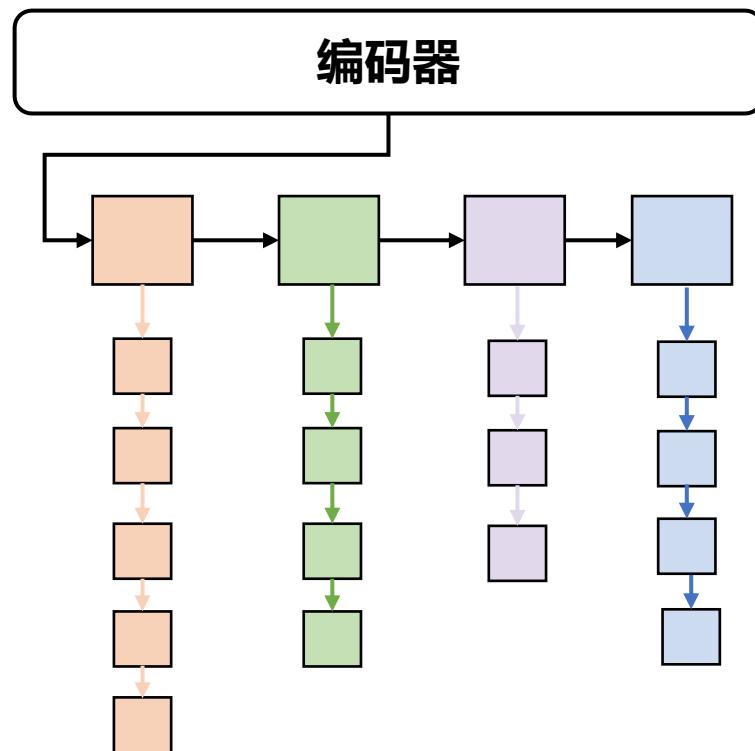




# 客服对话的内在结构

- 客服对话存在隐式的流程结构

Dialogue
AGENT: Hello, what can I do for you?
USER: What's the standard of electric vehicles for the Express.
AGENT: Do you have a car?
AGENT: Or are you going to buy a car?
USER: I am hesitating which car to buy. One is Jianghuai EV Seven, the other is BYD YUAN.
AGENT: OK, you can fulfill the table in this link (link info) with the type of vehicle you wish to check. We will give you feedback in seven days.
USER: I have not bought yet.
USER: Can you check it now?
AGENT: I am quite sorry for that. A specialist on this issue will check it and call you back.
AGENT: They will give a precise answer for your question.
USER: OK.
AGENT: Thanks for your understanding. What else can I do for you?
USER: Nothing, thanks. Bye.
AGENT: Thank you. Have a nice day.
Summary
The user's question was about the standard of EV car for the Express. He asked the standard to decide which car to buy. I told the user to fill in the type of the cars in our system and we would give feedback in seven days. The user approved the result. The user hung up.
Key point sequence
Question description → Solution → User approval → End





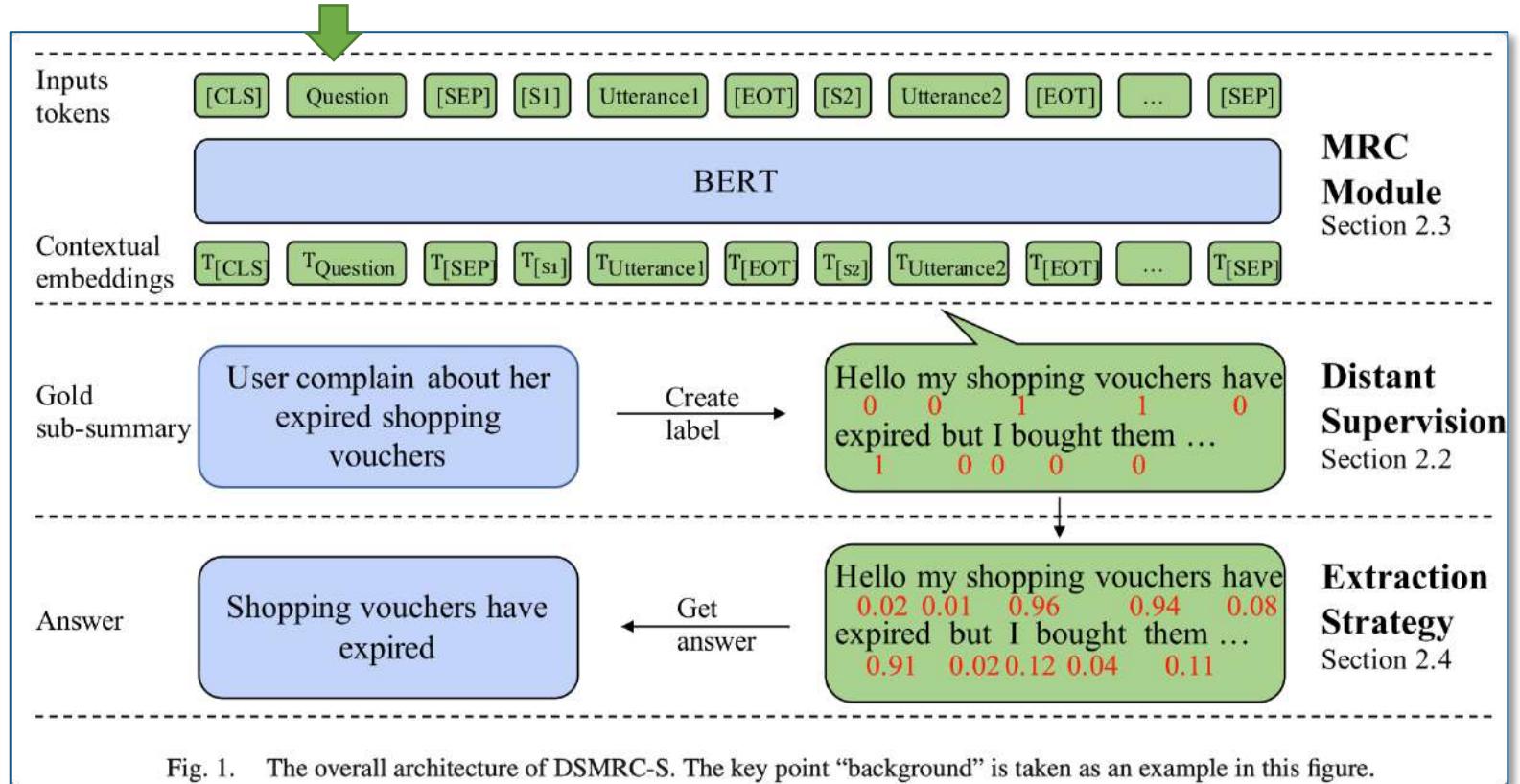
# 客服对话的内在结构

What is the **background** of the call?

What is the **purpose** of the user's call?

What is the **key** of the question?

What is your **solution**?





# 客服对话数据集

Dialogue	
0 Q: 我的物流怎么没有进展呢? (Why is my shipping information not updating?)	
1 A: 这边帮您查一下。 (I'll check it for you.)	
2 Q: 我购买的雨伞, 怎么没有更新进展呀? (Why hasn't the purchased umbrella been updated with the shipping information?)	
3 A: 正在运输去上海松江分拨中心。 (It is being transported to Shanghai Songjiang Distribution Center.)	
4 A: 运输途中是不显示的。 (The shipping information is not shown in transit.)	
5 A: 等下午到上海[地址], 就能更新新信息。 (New information will be updated when the goods arrive in Shanghai [address] in the afternoon.)	
6 Q: 承诺今天到, 不知可否到呢? 我明天早上出差用。 (You promised to arrive today. Can you make it? I'm on a business trip tomorrow morning.)	
7 A: 今天到的。 (It will arrive today.)	
8 Q: 谢谢。 (Thanks.)	
9 A: 请问还有其他问题要查询吗? (Do you have any other questions to inquire?)	
10 A: 好的, 不客气。 (Okay, you're welcome.)	
11 Q: 没有了。 (There are no questions.)	

(a)

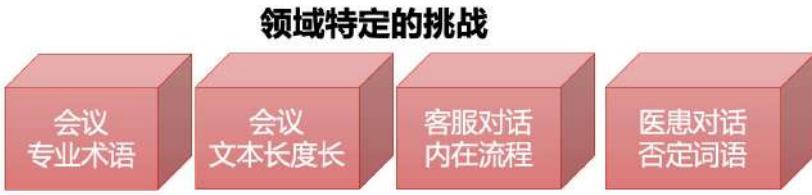
Fine-grained Annotation			
User identity: 顾客 / Customer			
Topics	Questions	Answers	Well-formed answers
物流全程跟踪 (delivery tracking)	用户询问为何物流信息没有更新。 (0, 2) (The customer asks why shipping information is not updating.)	客服回应由于货物在运输中, 因此物流信息不会更新。 (3, 4) (The agent replies because the goods are in transit, the shipping information is not shown.)	(Same as the answer)
物流周期 (delivery time)	用户询问货物能否今天到达。 (6) (The customer asks if the goods will arrive today.)	客服表示会的。 (7) (The agent says it will.)	客服表示货物今天会送达的。 (6, 7) (The agent says the goods will be delivered today.)

(b)

<b>Overall summary:</b> 用户询问为何物流信息没有更新。客服回应由于货物在运输中, 因此物流信息不会更新。 (The customer asks why shipping information is not updating. The agent replies because the goods are in transit, the shipping information is not shown.) 用户询问货物能否今天到达。客服表示会的。 (The customer asks if the goods will arrive today . The agent says it will. )
<b>User summary:</b> 用户询问为何物流信息没有更新。 (The customer asks why shipping information is not updating.) 用户询问货物能否今天到达。 (The customer asks if the goods will arrive today.)
<b>Agent summary:</b> 客服回应由于货物在运输中, 因此物流信息不会更新。 (The agent replies because the goods are in transit, the shipping information is not shown. ) 客服表示货物今天会送达的。 (The agent says the goods will be delivered today. )

(c)

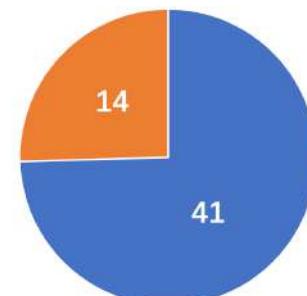
# EMNLP 2021相关情况



对话建模的挑战



数据资源的挑战



■ 非对话摘要 ■ 对话摘要

An Exploratory Study on  
**Long** Dialogue  
Summarization: What  
Works and What's Next

Leveraging Pretrained Models  
for Automatic Summarization  
of **Doctor-Patient**  
Conversations

**Topic-Aware** Contrastive  
Learning for Abstractive  
Dialogue Summarization

Give the Truth: Incorporate  
**Semantic Slot** into Abstractive  
Dialogue Summarization

CSDS: A Fine-grained Chinese  
**Dataset** for Customer Service  
Dialogue Summarization

GupShup : Summarizing Open-  
Domain **Code-Switched**  
Conversations

Simple Conversational **Data  
Augmentation** for Semi-  
supervised Abstractive Dialogue  
Summarization

**Low-Resource** Dialogue  
Summarization with  
Domain-Agnostic Multi-  
Source Pretraining

# 相关机构

复旦

**AAAI 2021**  
Topic-Oriented Spoken Dialogue Summarization for Customer Service with Saliency-Aware Topic Modeling

**AAAI 2021**  
Unsupervised Summarization for Chat Logs with Topic-Oriented Ranking and Context-Aware Auto-Encoders

**EMNLP 2021**  
Low-Resource Dialogue Summarization with Domain-Agnostic Multi-Source Pretraining

微软

**EMNLP Findings 2021**  
A Hierarchical Network for Abstractive Meeting Summarization with Cross-Domain Pretraining

**NAACL Short 2021**  
MediaSum: A Large-scale Media Interview Dataset for Dialogue Summarization

**EMNLP Short 2021**  
An Exploratory Study on Long Dialogue Summarization: What Works and What's Next

**AAAI 2022**  
DialogLM: Pre-trained Model for Long Dialogue Understanding and Summarization.

佐治亚理工

**EMNLP 2020**  
Multi-View Sequence-to-Sequence Models with Conversational Structure for Abstractive Dialogue Summarization

**NAACL 2021**  
Structure-Aware Abstractive Conversation Summarization via Discourse and Action Graphs

**EMNLP 2021**  
Simple Conversational Data Augmentation for Semi-supervised Abstractive Dialogue Summarization

A\*STAR

**ASRU 2019**  
Topic-aware Pointer-Generator Networks for Summarizing Spoken Conversations

**SIGIR 2021**  
Coreference-Aware Dialogue Summarization

**SummDial@SIGDial 2021**  
Dynamic Sliding Window for Meeting Summarization

**EMNLP 2021**  
Controllable Neural Dialogue Summarization with Personal Named Entity Planning

北邮

**COLING 2020**  
Improving Abstractive Dialogue Summarization with Graph Structures and Topic Words

**EMNLP 2021 Findings**  
Give the Truth: Incorporate Semantic Slot into Abstractive Dialogue Summarization

**TODSum**: Task-Oriented Dialogue Summarization with State Tracking

# 实验相关

# 对话摘要实验技巧

- Trick1 : name substitution
  - 将人名替换常用名（保持性别不变）
  - 有效，ROUGE-L增加0.8
- Trick2 : negation scope highlighting
  - 外部工具指明：“I don’t know what to do” --> “I don’t <NEG> know what to do <\NEG>”
  - 无效，效果低了
- Trick3 : multi-task learning on common sense tasks
  - Short Story Ending Prediction , Commonsense Generation , Commonsense Knowledge Base Construction
  - 有效，ROUGE-L增加一个点左右
- Trick4 : pretraining on an in-domain dialogue corpus
  - 继续预训练BART在PersonaChat和Reddit comments
  - 有效，ROUGE-L增加0.9个点左右。

# 排行榜

## AMI and ICSI

Model	AMI			ICSI		
	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-1	ROUGE-2	ROUGE-L
<i>Extractive Methods</i>						
TextRank [Mihalcea and Tarau, 2004]	35.19	6.13	15.70	30.72	4.69	12.97
SummaRunner [Nallapati <i>et al.</i> , 2017]	30.98	5.54	13.91	27.60	3.70	12.52
<i>Abstractive Methods</i>						
UNS [Shang <i>et al.</i> , 2018]	37.86	7.84	13.72	31.73	5.14	14.50
PGN [See <i>et al.</i> , 2017]	42.60	14.01	22.62	35.89	6.92	15.67
Sentence-Gated [Goo and Chen, 2018]	49.29	19.31	24.82	39.37	9.57	17.17
TopicSeg [Li <i>et al.</i> , 2019a]	51.53	12.23	25.47	-	-	-
TopicSeg+VFOA [Li <i>et al.</i> , 2019a]	53.29	13.51	26.90	-	-	-
HMNet [Zhu <i>et al.</i> , 2020]	52.36	18.63	24.00	45.97	10.14	18.54
PGN( $\mathcal{D}_{\text{ALL}}$ ) [Feng <i>et al.</i> , 2021]	50.91	17.75	24.59	-	-	-
DDAMS [Feng <i>et al.</i> , 2020a]	51.42	20.99	24.89	39.66	10.09	17.53
DDAMS+DDADA [Feng <i>et al.</i> , 2020a]	53.15	22.32	25.67	40.41	11.02	19.18
<i>Pre-trained Language Model-based Methods</i>						
Longformer-BART [Fabbri <i>et al.</i> , 2021]	54.81	20.83	25.98	43.40	12.19	19.29
Longformer-BART-arg [Fabbri <i>et al.</i> , 2021]	55.27	20.89	24.94	44.51	11.80	19.19

Table 2: Leaderboard of meeting summarization task on AMI [Carletta *et al.*, 2005] and ICSI [Janin *et al.*, 2003] datasets. We adopt reported results from published literatures [Feng *et al.*, 2020a] and corresponding publications. The results of Longformer [Fabbri *et al.*, 2021] are obtained by evaluating the output files provided by the author.

## SAMSum

Model	R-1	R-2	R-L
<i>Extractive Methods</i>			
LONGEST-3	32.46	10.27	29.92
TextRank [Mihalcea and Tarau, 2004]	29.27	8.02	28.78
<i>Abstractive Methods</i>			
DynamicConv [Wu <i>et al.</i> , 2019]	33.69	10.88	30.93
Transformer [Vaswani <i>et al.</i> , 2017]	36.62	11.18	33.06
PGN [See <i>et al.</i> , 2017]	40.08	15.28	36.63
Fast Abs RL [Chen and Bansal, 2018]	41.95	18.06	39.23
D-HGN [Feng <i>et al.</i> , 2020b]	42.03	18.07	39.56
TGDGA [Zhao <i>et al.</i> , 2020]	43.11	19.15	40.49
<i>Pre-trained Language Model-based Methods</i>			
DialoGPT [Zhang <i>et al.</i> , 2020d]	39.77	16.58	38.42
UniLM [Dong <i>et al.</i> , 2019]	47.85	24.23	46.67
PEGASUS [Zhang <i>et al.</i> , 2020a]	50.50	27.23	49.32
BART [Lewis <i>et al.</i> , 2020]	52.98	27.67	49.06
S-BART [Chen and Yang, 2021]	50.70	25.50	48.08
FROST [Narayan <i>et al.</i> , 2021]	51.86	27.67	47.52
CODS [Wu <i>et al.</i> , 2021]	52.65	27.84	50.79
MV-BART [Chen and Yang, 2020]	53.42	27.98	49.97
BART( $\mathcal{D}_{\text{ALL}}$ ) [Feng <i>et al.</i> , 2021]	53.70	28.79	50.81

Table 3: Leaderboard of chat summarization task on SAMSum [Gliwa *et al.*, 2019] dataset, where “R” is short for “ROUGE”. We adopt reported results from published literatures [Gliwa *et al.*, 2019; Wu *et al.*, 2021] and corresponding publications. The results of MV-BART [Chen and Yang, 2020] are obtained via running the open-source code. The results of S-BART [Chen and Yang, 2021] are obtained by evaluating the output file provided by the author.

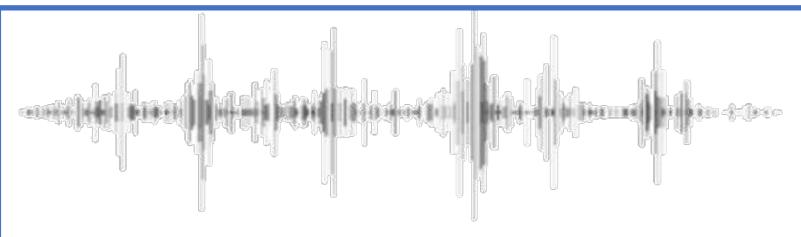
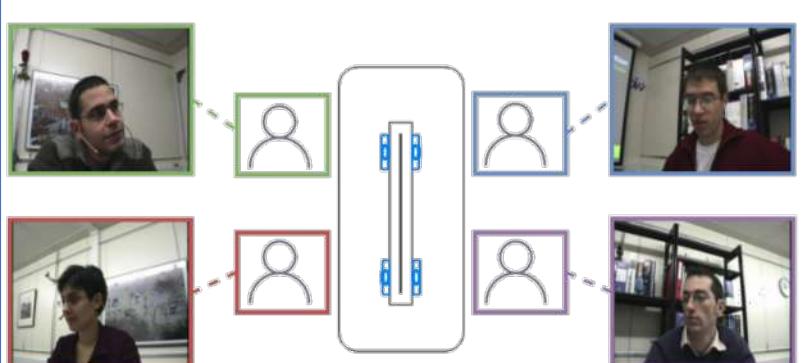
# 可用资源

- 所有对话摘要论文、代码整合
  - <https://github.com/xcfcode/Summarization-Papers#dialogue>
- 微软长文本预训练对话摘要模型
  - <https://github.com/microsoft/DialogLM>
- 微软层次化会议摘要模型
  - <https://github.com/microsoft/HMNet>

# 未来趋势

# 未来趋势①：多模态对话摘要

- 同步的多模态

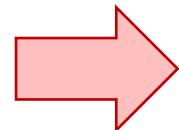


音频

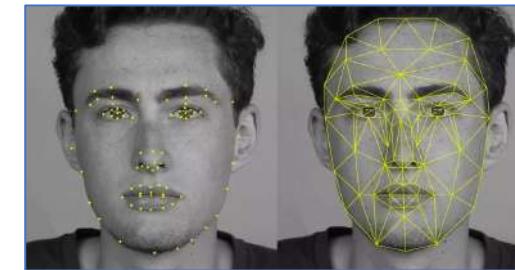
We're developing a remote control which you probably already know.

文本

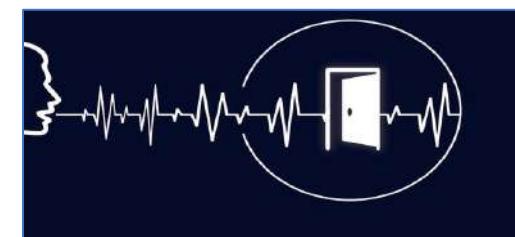
潜在问题



数据隐私性

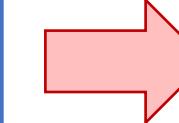


面部特征

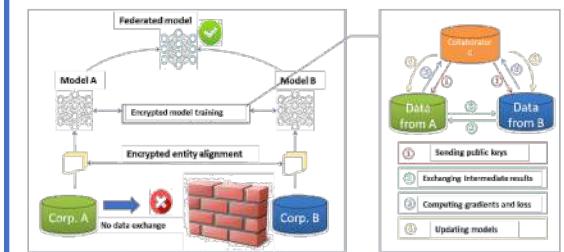
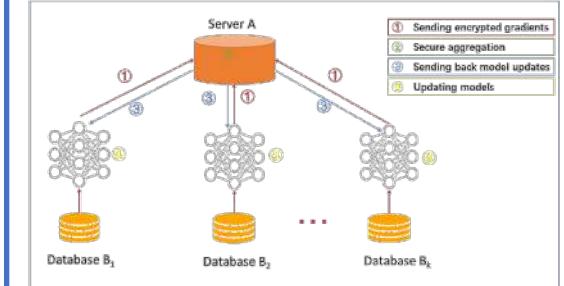


声纹特征

可行方案

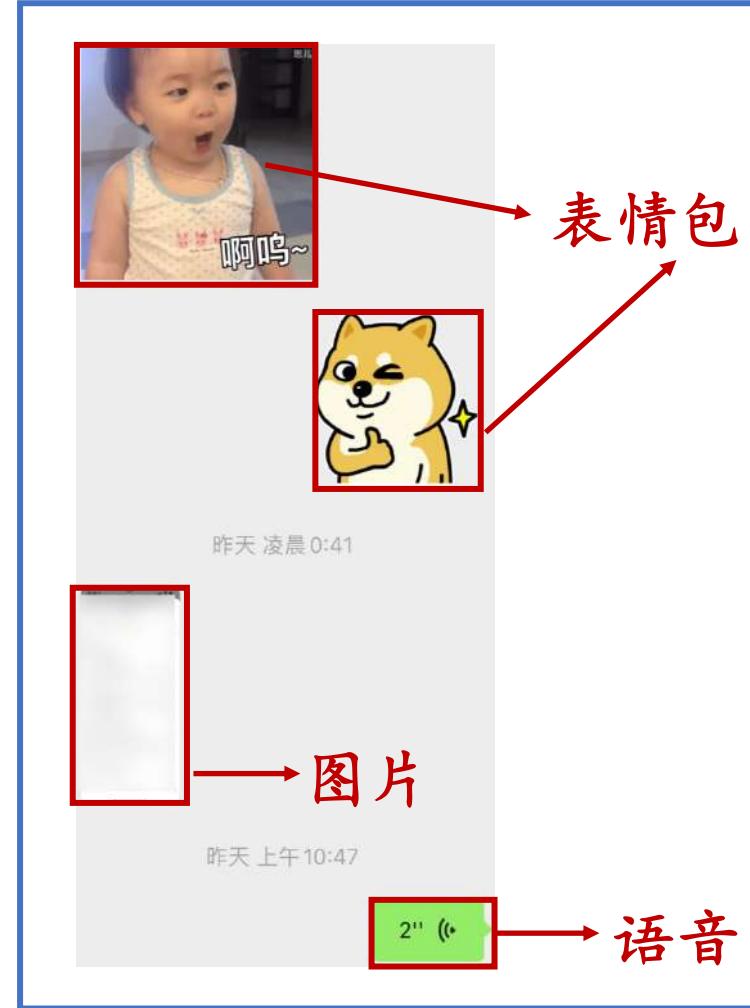


联邦学习



# | 未来趋势①：多模态对话摘要

- 异步的多模态
  - 文本
  - 图片（静态）
  - 表情包（静态+动态）
  - 视频（动态）
  - 语音
- 相关方向延伸
  - 情感分析



# 未来趋势②：多领域对话摘要

- 对话数据形式不一，各有特点。



电影对话



采访



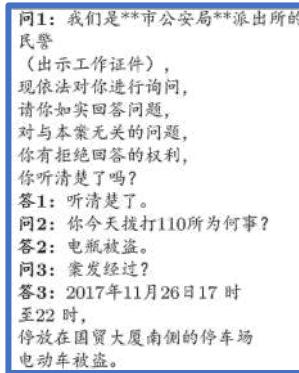
闲聊



邮件

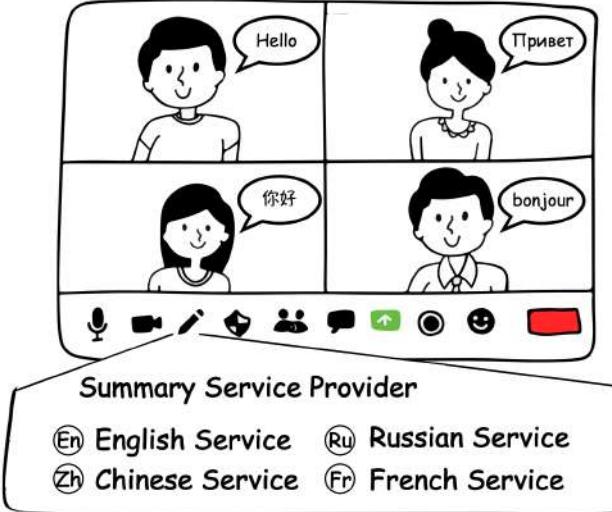


会议



派出所报警

# 未来趋势③：多语言对话摘要



English Dialogue	
<b>Elliot</b>	: I can't talk rn , I'm rly busy.
<b>Elliot</b>	: Can I call u back in about 2 hours?
<b>Jordan</b>	: Not really , I'm going to a funeral.
<b>Jordan</b>	: I'll call you tonight , ok?
<b>Elliot</b>	: Sure
<b>Elliot</b>	: Whose funeral is it?
<b>Jordan</b>	: My colleague's , <b>Brad</b> .
<b>Jordan</b>	: I told you about him , he had a <b>liver cancer</b> .
<b>Elliot</b>	: I'm so sorry man , I hope u're ok.
<b>Elliot</b>	: I'll call u at <b>8 pm</b> .
Generated Summaries ( <i>One-to-many</i> )	
English	<b>Elliot</b> can't talk because he's busy. <b>Jordan</b> is going to a funeral for his colleague, <b>Brad</b> , who had a <b>liver cancer</b> . <b>Elliot</b> will call him at <b>8 pm</b> . [71.19-42.11-50.85]
Chinese	乔丹要去参加他的同事布拉德的葬礼。他得了肝癌。 埃利奥特将在晚上8点给乔丹打电话 [66.67-40.00-35.09]
Russian	Джордан собирается на похороны своего коллеги Брэда, у него рак печени. Элиот позвонит Джордана в 20: 00. [58.38-30.00-38.10]
French	<b>Elliot</b> ne peut pas parler parce qu'il est occupé. <b>Jordan</b> va au funeral de son collègue, <b>Brad</b> , qui a un <b>cancer du foie</b> . Il appellera <b>Elliot</b> à <b>20 h</b> . [68.97-42.86-55.17]
Arabic	جوردن هو الذي يذهب إلى جنازة زميلها براد لديه سرطان الكبد. إليوت سوف يدعوه له في الساعة الثامنة مساءً. [57.78-27.91-31.11]
Spanish	<b>Elliot</b> no puede hablar porque está ocupado. <b>Jordan</b> va a un funeral de su colega, <b>Brad</b> , que tuvo un <b>cáncer de hepática</b> . <b>Eliot</b> llamará a <b>Jordan</b> a las <b>8 p.m.</b> [60.71-29.63-39.29]

Leon:	kya tujeh abhi tak naukari nahi mili?
Arthur:	nahi bro, abhi bhi unemployed :D
Leon:	hahaha, LIVING LIFE
Arthur:	mujeh yeh bahot acha lagta hai, dopahar ko jagata hoon, sports dekhta hoon - ek Aadmi ko aur kya chahiye?
Leon:	a paycheck? ;)
Arthur:	mean mat bano ...
Leon:	but seriously, mere dosth ke company mein ek junior project manager offer hai, tujeh interest hai?
Arthur:	sure thing, tere pass details hai?
Leon:	<file_photo>
English Summary:	<b>Arthur</b> is still unemployed. <b>Leon</b> sends him a job offer for junior project manager position. <b>Arthur</b> is interested.

Table 1: Example of a code-switched Hi-En conversation and the corresponding En summary. ■: En words, ■: transliterated Hi words, ■: language-agnostic words such as named entities and punctuation marks

# 未来趋势④：事实一致性

**Category 1 - Missing Information:** The content of the generated summary is incomplete compared to the reference.

**Example:**

[Reference Summary] *Williams invites Ms. Blair for a coffee. They will go to her favourite coffee place near the square in a side alley at 2 p.m.*

[Model-Generated Summary] *Ms. Blair is going to a coffee place near the square in a side alley.*

**Category 2 - Redundant Information:** There is redundant content in the generated summary compared to the reference.

**Example:**

[Reference Summary] *Paula helped Charlotte with correct pronunciation of "Natal Lily."*

[Model-Generated Summary] *Charlotte asks Paula how to pronounce the name of the plant "Natal Lily." Paula confirms that the stress on the second syllable is 2nd.*

**Category 3 - Circumstantial Error:** Circumstantial information (e.g., date, time, location) about the predicate doesn't match the reference.

**Example:**

[Reference Summary] *The USA was founded in 1776.*

[Model-Generated Summary] *The USA was founded in 1767.*

**Category 4 - Wrong Reference Error:** A pronoun is with an incorrect or nonexistent antecedent, or a personal named entity in the generated summary is in the place of a different personal entity in the reference.

**Example:**

[Reference Summary] *Mohit asked Darlene about the test.*

[Model-Generated Summary] *Darlene asked Mohit about the test.*

**Category 5 - Negation Error:** This encompasses factual errors resulting from missing or erroneous negation in the generated summary compared to the reference.

**Example:**

[Reference Summary] *Justin likes books.*

[Model-Generated Summary] *Justin does not like books.*

**Category 6 - Object Error:** This covers factual errors resulting from incorrect direct or indirect objects (for non-personal entities only; errors of this nature involving personal entities are designated as Wrong Reference Errors).

**Example:**

[Reference Summary] *Tara raised her glass.*

[Model-Generated Summary] *Tara raised her spoon.*

**Category 7 - Tense Error:** This encompasses factual errors resulting from discrepancies in grammatical tense between the generated summary and the reference.

**Example:**

[Reference Summary] *The children will go to the library.*

[Model-Generated Summary] *The children went to the library.*

**Category 8 - Modality Error:** This includes factual errors resulting from modal discrepancies, such getting words like "may", "should", "could" wrong, between the generated summary and the reference.

**Example:**

[Reference Summary] *School may be cancelled today.*

[Model-Generated Summary] *School is cancelled today.*

# 总结

# 总结

- 到目前为止，耗时两年时间，对话摘要相关工作完成了多种特征的尝试，初步完成对话系统+文本生成+文本摘要三个方向的成果整合，使得对话摘要前进一步。
- 现有工作呈现出领域化的趋势，研究者倾向于在某一领域扎根研究，解决领域内的独特问题。
- 已经摆脱研究初期纯学术道路，研究者更努力使相关研究得以落地。
- 对话摘要的研究依旧会被划分在大研究范围之下，例如多领域、多语言、多模态等。

**谢谢 !**

[xiachongfeng1996@gmail.com](mailto:xiachongfeng1996@gmail.com)