

Post-Processed DialogSum Data

- The original speaker is masked and replaced with #PERSON1, #PERSON2.
- But some speaker names are mentioned in the dialogue. This may cause mismatch for models to process and unnatural.
- We recovered them with an automatic way and released the speaker recovered version.

Old Version

- **#Person1#:**
 - Hi, Mr. Smith. I'm Doctor Hawkins. Why are you here today?
- **#Person2#:**
 - I found it would be a good idea to get a check-up.
- **#Person1#:**
 - Yes, well, you haven't had one for 5 years. You should have one every year.
- **#Person2#:**
 - I know. I figure as long as there is nothing wrong, why go see the doctor?
- **#Person1#:**
 - Well, the best way to avoid serious illnesses is to find out about them early. So try to come at least once a year for your own good.
- **#Person2#:**
 - Ok.
- **#Person1#:**
 - Let me see here. Your eyes and ears look fine. Take a deep breath, please. Do you smoke, Mr. Smith?

...

New Version

- **Doctor Hawkins:**
 - Hi, Mr. Smith. I'm Doctor Hawkins. Why are you here today?
- **Mr. Smith:**
 - I found it would be a good idea to get a check-up.
- **Doctor Hawkins:**
 - Yes, well, you haven't had one for 5 years. You should have one every year.
- **Mr. Smith:**
 - I know. I figure as long as there is nothing wrong, why go see the doctor?
- **Doctor Hawkins:**
 - Well, the best way to avoid serious illnesses is to find out about them early. So try to come at least once a year for your own good.
- **Mr. Smith:**
 - Ok.
- **Doctor Hawkins:**
 - Let me see here. Your eyes and ears look fine. Take a deep breath, please. Do you smoke, Mr. Smith?

...

InsructDS - EMNLP 2023

A demo page for <https://arxiv.org/abs/2310.10981>

The image shows a web interface for InsructDS. It consists of two main panels. The left panel contains input fields for 'Question', 'Dialogue', 'Dataset Index', 'Sample Index', and 'Preferred Output Length'. Each field has a placeholder text: 'Input your own question. Leave it blank to extract directly from the dataset.' for Question, and 'Input your own dialogue. Leave it blank to extract directly from the dataset.' for Dialogue. The 'Dataset Index', 'Sample Index', and 'Preferred Output Length' fields are accompanied by sliders and numeric input boxes showing values 1, 1, and 0 respectively. At the bottom of the left panel are 'Clear' and 'Submit' buttons. The right panel contains a 'Question' field, a 'Dialogue' field, and a 'Summary' field, all with placeholder text.

Question

Input your own question. Leave it blank to extract directly from the dataset.

Dialogue

Input your own dialogue. Leave it blank to extract directly from the dataset.

Dataset Index

1

Sample Index

1

Preferred Output Length

0

Clear

Submit

Question

Dialogue

Summary

- After hosting the demo program, you can access this page from
0.0.0.0:8080 (local machine)
xxx:xxx:xxx:xxx:8080 (remote machine)

InsructDS - EMNLP 2023

A demo page for <https://arxiv.org/abs/2310.10981>

Question

Input your own question. Leave it blank to extract directly from the dataset.

Dialogue

Input your own dialogue. Leave it blank to extract directly from the dataset.

Dataset Index

1

Sample Index

1

Perferred Output Length

0

Clear

Submit

Question

No quesiton provided. Below is a general summarization of the dialogue.

Dialogue

Hannah: Hey, do you have Betty's number?
Amanda: Lemme check
Hannah: <file_gif>
Amanda: Sorry, can't find it.
Amanda: Ask Larry
Amanda: He called her last time we were at the park together
Hannah: I don't know him well
Hannah: <file_gif>
Amanda: Don't be shy, he's very nice
Hannah: If you say so..
Hannah: I'd rather you texted him
Amanda: Just text him 😊
Hannah: Urgh.. Alright
Hannah: Bye
Amanda: Bye bye

Summary

Amanda doesn't have Betty's number. Larry called her last time they were at the park. Hannah doesn't know Larry well.

- Click, Submit.
- It will extract the 1st sample (sample index button) from dataset 1 (dataset index), and summarize.

InsructDS - EMNLP 2023

A demo page for <https://arxiv.org/abs/2310.10981>

Question

Input your own question. Leave it blank to extract directly from the dataset.

Dialogue

Input your own dialogue. Leave it blank to extract directly from the dataset.

Dataset Index

2

Sample Index

1

Perferred Output Length

0

Clear

Submit

Question

Who does Hannah need to contact?

Dialogue

Hannah: Hey, do you have Betty's number?
Amanda: Lemme check
Hannah: <file_gif>
Amanda: Sorry, can't find it.
Amanda: Ask Larry
Amanda: He called her last time we were at the park together
Hannah: I don't know him well
Hannah: <file_gif>
Amanda: Don't be shy, he's very nice
Hannah: If you say so..
Hannah: I'd rather you texted him
Amanda: Just text him 😊
Hannah: Urgh.. Alright
Hannah: Bye
Amanda: Bye bye

Summary


Betty's number

- Change to dataset 2 (SAMSum_QDS)
- It will have one question, and the summary (answer) is based on the question.

A demo page for <https://arxiv.org/abs/2310.10981>

Question

Who does Hannah need to contact to get Betty's number?



Dialogue

Hannah: Hey, do you have Betty's number?

Amanda: Lemme check

Hannah: <file_gif>

Amanda: Sorry, can't find it.

Amanda: Ask Larry

Amanda: He called her last time we were at the park together

Hannah: I don't know him well

Hannah: <file_gif>

Amanda: Don't be shy, he's very nice

Hannah: If you say so..


Hannah: I'd rather you texted him

Amanda: Just text him 😊

Hannah: Urgh.. Alright

Hannah: Bye

Amanda: Bye bye



Dataset Index

2

Sample Index

1

Perferred Output Length

0

Clear

Submit

Question

Who does Hannah need to contact to get Betty's number?

Dialogue

Hannah: Hey, do you have Betty's number?

Amanda: Lemme check

Hannah: <file_gif>

Amanda: Sorry, can't find it.

Amanda: Ask Larry

Amanda: He called her last time we were at the park together

Hannah: I don't know him well

Hannah: <file_gif>

Amanda: Don't be shy, he's very nice

Hannah: If you say so..

Hannah: I'd rather you texted him

Amanda: Just text him 😊

Hannah: Urgh.. Alright

Hannah: Bye

Amanda: Bye bye

Summary

Larry, who called her last time they were at the park.

- You can then copy-paste to ask your own questions.