## A Conversation Pairs and Response Distribution

We examined the distribution for query-response pairs and the statistic result is shown in the Figure 3. Questions that exceed 100 words and responses that are longer than 40 words are excluded. This led to 88% of the original pairs.

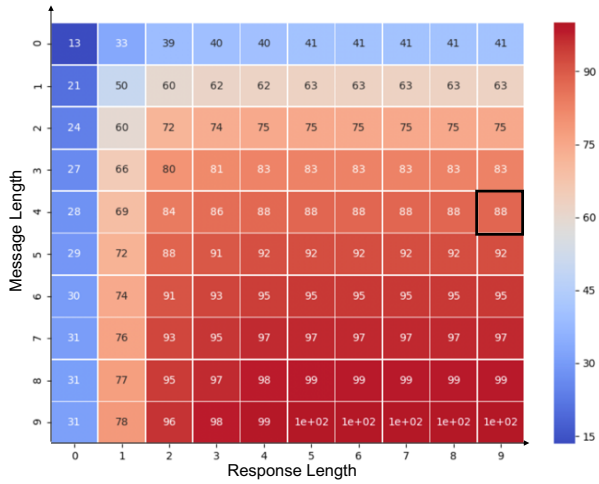


Figure 3: Message and response length distribution.

Figure 4 illustrates the distribution of the number of responses under same question in our dataset. We found that questions with 2 to 3 responses are the majority.

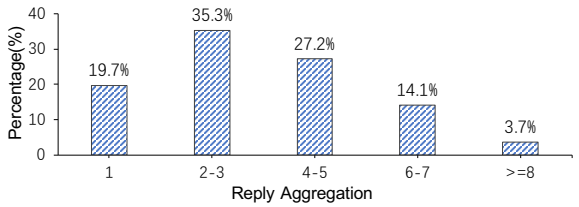


Figure 4: The average number of replies per question.

## B Reliability of API-based User Attribute Information

To examine the reliability of the extracted information, we conducted human annotation validations. Specifically, we randomly selected 50 users from our population set together with their attributes. And we shared the user attributes and source comments with annotators and asked them to judge whether the user attributes corresponded to the comments. If the source comment truly reflects user's corresponding attributes, they should give label "Right", otherwise "Not Right". For attribute types with more than one value, such as possessions, "Right" means all the values are truly related and "Partly right" means some are related and some are not and "Not Right" means all the values are not related.

| **Attribute** | Pets | Family | Residence | Favorites |
|---|---|---|---|---|
| **Percentage** | 28 | 70.9 | 46 | 56 |
| **Attribute** | Partner | Possession | Gender | Self-description |
| **Percentage** | 38 | 98 | 100 | 86 |

Table 8: Coverage rate for each attribute in those sampled 50 users.

| **Attribute** | Pets | Family | Residence | Favorites |
|---|---|---|---|---|
| **Right /** % | 85.7 | 82.4 | 82.6 | 96.4 |
| **Partly Right /** % | 14.3 | 8.8 | 0 | 0 |
| **Not Right /** % | 0 | 8.8 | 17.4 | 3.6 |
| **Attribute** | Partner | Possession | Gender | Self-description |
| **Right /** % | 100 | 83.7 | 86.0 | 65.1 |
| **Partly Right /** % | 0 | 12.2 | 0 | 18.6 |
| **Not Right /** % | 0 | 4.1 | 14.0 | 16.3 |

Table 9: Attribute reliability annotation results.

Table 8 shows the coverage rate for each attribute in selected users. The distribution aligns well with the overall coverage rate and shows that the sampled users are representative. Table 9 shows the reliability annotation result. The percentage for each attribute in annotation result is calculated among users with value in that attribute. As shown in Table 9 all of the attributes show a high reliability rate by considering "Right" and "Partly Right". Although all the user attributes are inferred from what user has said about himself/herself, there still exists information that does not represent his/her personal attributes. Table 10 shows examples of positive and negative label results for some attributes.

| **Source Comments** | **Consistency?** |
|---|---|
| *Gender:* | |
| I am a thin girl that has trouble... | Right |
| If I was a girl who was... | Not Right |
| *Favorites:* | |
| I LOVE Halloween and delight in... | Right |
| I like the force. | Not Right |
| *Residence:* | |
| I live in San Diego county. | Right |
| I live just outside of Boston. | Not Right |
| *Possession:* | |
| I have it on my pandora playlist. | Right |
| I wanted to start up my own prison. | Not Right |
| *Self-description:* | |
| I'm a rapper who... | Right |
| I was basically a zombie. | Not Right |
| *Family:* | |
| I asked my mother if she loved me... | Right |
| I met our mother-a documentary... | Not Right |

Table 10: Annotation examples for different attribute types with attributes in red.

## C Diverse Responses for Models

Table 12 and Table 14 show additional responses generated by different models on our dataset. With profile and related comments, our model can generate not only user-attribute related entity but also is capable of capturing different users' attitudes towards other people or things. For example, the responses generated by GSMN properly reflected the user's positive or negative opinions when Trump was mentioned.

## D Consistency Experiments

In Table 11 and 15, we show all 20 questions we designed and profiles of four sampled users in the manual evaluation experiments. From question no.1 to question no.10, the questions are related to basic personal information and from question no.11 to question no.20, the questions are related to personal attitudes. The answer of every two adjacent questions should be considered consistent. Table 16 shows examples of top related comments for given questions and given user.

| **Questions for basic information:** |
|---|
| 1. Where are you from ? |
| 2. Where do you come from ? |
| 3. Do you have a girlfriend or boyfriend ? |
| 4. Are you single ? |
| 5. What do you do for living ? |
| 6. What is your job ? |
| 7. Are you male or female ? |
| 8. What is your gender ? |
| 9. What pet do you have ? |
| 10. Do you have any pets ? |
| **Questions for personal attitudes:** |
| 11. Who is your most reliable one? |
| 12. Who do you rely on most? |
| 13. What is your most upset experience? |
| 14. What upsets you most? |
| 15. What is your attitude toward Trump? |
| 16. Will you vote for Trump? |
| 17. Who is your favourite person? |
| 18. Who do you love most? |
| 19. What do you like to do most? |
| 20. What do you enjoy most? |

Table 11: Questions in human evaluation experiments.

| |
|---|
| **Question:** Who is the one person you could rely on if you were in a real bad situation? |
| **Seq2Seq:** My best friend. |
| **User A**: *Gender: male*; *Family: wife*<br>**Gold:** Wife or guy I used to do things with a long time ago.<br>**GSMN:** My wife. She's my wife. |
| **User B**: *Gender: female*; *Family: husband*<br>**Gold:** My dad. Guy would melt the ice caps if I asked him to.<br>**GSMN:** My husband. He's a good guy. |
| **User C**: *Gender: male*; *Partner: girlfriend*<br>**Gold:** My girlfriend.<br>**GSMN:** My ex. She' s my best friend. |

Table 12: Example responses generated by our GSMN.

| |
|---|
| **User A:** *Residence:Canada* |
| **Question:** Where are you from ? |
| **Seq2Seq:** I' m in the UK.<br>**Speaker:** I' m in California.<br>**GMN:** New York city.<br>**GSMN:** I' m from Canada. |
| **Question:** Where do you come from? |
| **Seq2Seq:** I' m from US<br>**Speaker:** I' m in Florida.<br>**GMN:** I' m in the US.<br>**GSMN:** I' m from Canada. It's pretty <unk>. |
| **User B:** *Comments:I love dogs, I really do.* |
| **Question:** What do you like to do most ? |
| **Seq2Seq:** I like to watch movies.<br>**Speaker:** I like to play video games. It' s so much fun.<br>**GMN:** I' ve always wanted to be able to play video games.<br>**GSMN:** I like to play with my dogs. I love them. |
| **Question:** What do you enjoy most? |
| **Seq2Seq:** I like to play video games.<br>**Speaker:** I like to go to the gym.<br>**GMN:** Being in a good mood.<br>**GSMN:** I like to play with my dogs. |

Table 13: Examples from consistency evaluation.

| |
|---|
| **Question:** Writing a christmas card to Mr. Trump! What should it say? |
| **Seq2Seq:** I'm going to buy a new one. # |
| User A:<br>**Gold:** You' re a piece of garbage and we all hate you so much.<br>**GSMN:** I'm not a Trump supporter. |
| User B:<br>**Gold:** Thank you for doing a great job.<br>**GSMN:** It's a wonderful life. |
| **Question:** Gamers of Reddit what, is the most overrated game that critics have shilled for? |
| **Seq2Seq:** I' m not sure if it' s a game or not, but I' d say <unk>. # |
| User A:<br>**Gold:** Most AAA games.<br>**GSMN:** The last of us. It' s so boring. |
| User B:<br>**Gold:** Fortnite, and other battle royales.<br>**GSMN:** Fortnite, it' s a bad game. |
| **Question:** What was the last thing you laughed at? |
| **Seq2Seq:** I' ve seen a lot of funny jokes. |
| User A:<br>**Gold:** A comment on reddit about a guy.<br>**GSMN:** I laughed at a meme. |
| User B:<br>**Gold:** My best friend' s toddler dipping his cookie in ketchup.<br>**GSMN:** My friend' s birthday. |
| User C:<br>**Gold:** My cat headbutting me in the face.<br>**GSMN:** My cat' s laugh. |
| **Question:** You have one movie to watch for the rest of your life , what is it ? |
| **Seq2Seq:** I' d like to see a movie called "<unk>". # |
| User A:<br>**Gold:** Star Wars The Force Awakens.<br>**GSMN:** I' m a fan of Star Wars. |
| User B:<br>**Gold:** Prestige [digit] watches.<br>**GSMN:** Requiem for a Dream, <unk>. |

Table 14: Example responses generated by GSMN and baseline. # indicates poor-quality response.

| **User1:** |
|---|
| Gender: male<br>Residence: Canada<br>Pets: dog<br>Family: sister; father; mother<br>Partner girlfriend<br>Favorites: ice cream<br>Self-description: good artist; newbie<br>Possessions: team rocket hoodie; video games; junk food addiction |
| **User2:** |
| Gender: female<br>Residence: Germany<br>Pets: cat<br>Family: sister; father; mother<br>Partner: husband<br>Favorites: honey whine; buckwheat; beef<br>Self-description: christian bit; smartest person<br>Possessions: university subject chemistry |
| **User3:** |
| Gender: male<br>Residence: Illinois<br>Pets: cat; dog<br>Family: father; mother<br>Partner: wife<br>Favorites: pumpkin cheesecake; chili dogs;<br>Self-description: native American man<br>Possessions: American accent; hearing loss; church family; food aversion; stomach problem |
| **User4:** |
| Gender: male<br>Pets: dog<br>Family: father<br>Partner: girlfriend<br>Favorites: adventure |

Table 15: Sampled user profiles in human evaluation experiments.

| **Question: What upsets you most?** |
|---|
| User1: People who pronounce my name wrong. |
| User2: When people around me show self restraint it reminds me of my failures and this feels bad. |
| User3: It really hurts when someone says something. |
| User4: Anything unhealthy. Also, anything that makes someone feel sad. |
| **Question: Who is your favorite person?** |
| User1: My sister and my dog. |
| User2: My husband who is my best friend and my biggest supporter. |
| User3: Most popular girls in school. # |
| User4: My guitar. And my lamp. And my girlfriend. # |
| **Question: What do you like to do most?** |
| User1: I love food more than people . Though if my dog wanted some of my food, I' d love him enough to share it. |
| User2: The most magical thing for me is and has always been winter solstice. The rebirth of the sun. I always bake a sweet sun-bread. |
| User3: I love dogs, I really do. |
| User4: I like adventure time. Feels like old school cn. |

Table 16: Top retrieved comments for given questions in human evaluation experiments. # indicates that the retrieved top comments are not well related to our designed questions.