Human Evaluation Questions

You are being invited to participate in a research study titled - ToD4IR: A Humanising Task-Oriented Dialogue System for Industrial Robots. This study is being done by Chen Li (Department of Materials and Production, Aalborg University), Xiaochun Zhang (The School of Management Science and Computer, Anhui University of Finance & Economic), Dimitrios Chrysostomou (Department of Materials and Production, Aalborg University) and Hongji Yang (The School of Computing and Mathematica Sciences, University of Leicester).

Your participation is entirely voluntary, and you can withdraw at any time. You are free to omit any question.

1.   Engaging - How much would you prefer to talk to the ToD4IR?

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not at all	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	⬭	Absolutely

2.   Engaging - Which version of the ToD4IR is more likely to hold your attention and mke you want to hear more ?

Mark only one oval.

- ⬭ ToD4IR-gpt2
- ⬭ ToD4IR-gpt2-large
- ⬭ ToD4IR-gpt2-xl
- ⬭ ToD4IR-gpt-neo(1.3B)
- ⬭ ToD4IR-gpt-neo(2.7B)
- ⬭ None of them

3.   Interesting - Do you think it is interesting to talk to the ToD4IR?

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not interesting	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Super interesting

4.   Interesting - Which version arouses your curiosity or replies you something new or useful ?

Mark only one oval.

◯ ToD4IR-gpt2

◯ ToD4IR-gpt2-large

◯ ToD4IR-gpt2-xl

◯ ToD4IR-gpt-neo(1.3B)

◯ ToD4IR-gpt-neo(2.7B)

◯ None of them

5.   Humanlike - Do you think ToD4IR sounds more human?

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not at all	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Absolutely

6.   Humanlike - Which version of ToD4IR is more natural and personable ?

Mark only one oval.

○ ToD4IR-gpt2

○ ToD4IR-gpt2-large

○ ToD4IR-gpt2-xl

○ ToD4IR-gpt-neo(1.3B)

○ ToD4IR-gpt-neo(2.7B)

○ None of them

7.   Knowledgeable - Do you think ToD4IR is more knowledgeable?

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Not at all	○	○	○	○	○	○	○	○	○	○	Absolutely

8.   Knowledgeable - Which version of ToD4IR seems more well informed and confident in the information?

Mark only one oval.

○ ToD4IR-gpt2

○ ToD4IR-gpt2-large

○ ToD4IR-gpt2-xl

○ ToD4IR-gpt-neo(1.3B)

○ ToD4IR-gpt-neo(2.7B)

○ None of them

Google Forms