

Typefaces and the Perception of Humanness in Natural Language Chatbots

Heloisa Candello

Claudio Pinhanez

Flavio Figueiredo

CHI'17 Proceedings of the 2017 CHI Conference on Human Factors in
Computing Systems

Pages 3476-3487

目次

- 背景
- 目的
- 使用する書体
- RQ
- 実験内容
- 実験結果
- まとめ・所感

背景

- 現在チャットボットは,世界中で使われている
- 近年ではオープンソースライブラリやAPIを利用することで,簡単にチャットボットを作成できるようになった.
- 過去には,ユーザがボットに与える影響に関する研究がなされてきた.

目的

- チャットボットの人間性を認識する上で、書体は認識に影響を与えるか.

使用する書体

- OCR

A B C D E F G H I J K L M

- Bradley

A B C D E a b c d e

- Georgia

A B C D E F G H I J K L M

- Helvetica

A B C D E F G H I J K L M

RQ

- RQ1:機械的な書体である”OCR”は,チャット内で機械と認識されるのか.
- RQ2:人間が書いた書体に近い”Bradley”は,チャット内で人間と認識されるのか.

実験内容

1. 参加者はチャットで行われた**10**個の対話を見てそれが人間の会話か機械の会話かを判断する.
2. 1 の操作を書体を変えてもう一度行う.

参加者は**40**歳未満の成人のポルトガル人で,**AI**を使用する職業やデザイナーなどを対象とした.

アンケート内容

1. あなたの年齢は.
2. あなたの職業は.
3. あなたの性別は.
4. 人工知能を使用したシステムと会話をしたことがありますか.
5. 話し手(金融アドバイザー)の書体が違っていただけに気がつきましたか.

実験結果

OCR

Percentage of cases perceived as MACHINE in the OCR Experiment						
95% confidence intervals for an exact binomial test on the OCR Experiment.						
Statistically significant results indicating that average fractions are above random chance (50%) are marked with * ($p < 0.1$); ** ($p < 0.05$); *** ($p < 0.01$)						
		count	OCR	Georgia	Helvetica	All Fonts
Overall		90	63 ± 11 %*	51 ± 11 %	51 ± 11 %	55 ± 6 %
Designers	Yes	15	60 ± 28 %	53 ± 27 %	67 ± 28 %	60 ± 16 %
	No	75	64 ± 12 %*	51 ± 12 %	48 ± 12 %	54 ± 7 %
Work in an AI company	Yes	46	74 ± 15 %**	59 ± 15 %	54 ± 15 %	62 ± 9 %**
	No	44	52 ± 16 %	43 ± 15 %	48 ± 15 %	48 ± 9 %
Gender	Male	49	57 ± 15 %	55 ± 15 %	57 ± 15 %	56 ± 8 %
	Female	39	72 ± 17 %*	46 ± 16 %	46 ± 16 %	55 ± 9 %
	No Ans.	2	50 ± 49 %	50 ± 49 %	0 ± 0 %	33 ± 29 %
Used AI before	Yes	65	66 ± 13 %*	51 ± 13 %	58 ± 13 %	58 ± 7 %*
	No	25	56 ± 21 %	52 ± 21 %	32 ± 17 %	47 ± 12 %
Perceived different fonts	Yes	71	59 ± 12 %	48 ± 12 %	55 ± 12 %	54 ± 7 %
	No	11	91 ± 32 %*	73 ± 34 %	27 ± 21 %	64 ± 19 %
	Now	8	62 ± 38 %	50 ± 34 %	50 ± 34 %	54 ± 21 %
All p-values are computed after a benjamini hochberg correction with FDR=0.1						

Bradley

Percentage of cases perceived as MACHINE in the Bradley Experiment						
95% confidence intervals for an exact binomial test on the OCR Experiment.						
Statistically significant results indicating that average fractions are above random chance (50%) are marked with * ($p < 0.1$); ** ($p < 0.05$); *** ($p < 0.01$)						
		count	Bradley	Georgia	Helvetica	All Fonts
Overall		109	48 ± 10 %	57 ± 10 %	61 ± 10 %*	55 ± 6 %
Designers	Yes	28	54 ± 20 %	68 ± 20 %	43 ± 18 %	55 ± 11 %
	No	81	46 ± 11 %	53 ± 11 %	68 ± 11 %**	56 ± 6 %
Work in an AI company	Yes	45	42 ± 15 %	58 ± 16 %	78 ± 15 %***	59 ± 9 %
	No	64	52 ± 13 %	56 ± 13 %	50 ± 13 %	53 ± 7 %
Gender	Male	57	47 ± 13 %	61 ± 14 %	63 ± 14 %	57 ± 8 %
	Female	49	47 ± 14 %	53 ± 15 %	61 ± 15 %	54 ± 8 %
	No Ans.	3	67 ± 57 %	33 ± 32 %	33 ± 32 %	44 ± 31 %
Used AI before	Yes	61	49 ± 13 %	51 ± 13 %	62 ± 13 %	54 ± 8 %
	No	48	46 ± 14 %	65 ± 15 %	60 ± 15 %	57 ± 9 %
Perceived different fonts	Yes	98	46 ± 10 %	59 ± 10 %	61 ± 10 %	55 ± 6 %
	No	6	67 ± 44 %	33 ± 29 %	67 ± 44 %	56 ± 25 %
	Now	5	60 ± 45 %	40 ± 35 %	60 ± 45 %	53 ± 27 %
All p-values are computed after a benjamini hochberg correction with FDR=0.1						

Tables 2 and 3: Percentage of cases perceived as MACHINE in OCR (left) and Bradley (right) experiments.

実験結果

- RQ1:機械的な書体である”OCR”は,チャット内で機械と認識されるのか.

Table2の結果より,OCRは機械であると認識されたと言える.

実験結果

- RQ2:人間が書いた書体に近い”Bradley”は、チャット内で人間と認識されるのか.

Table3の結果より,人間と認識されることはなかった.

まとめ

- 機械らしいフォントを用いたOCRは機械と認識されたが,人が手書きで書いたようなフォントは人間であると認識されなかった.
- 今回は書体に注目して実験を行ったが,書体に加え,色や絵文字を使えばまた違った結果が得られる可能性がある.

所感

- フォントを用いて,**Bot**らしい印象を与えることは可能であるが,人間らしい印象を与えることは難しいのかもしれない.
- **Bot**の話し方や話す内容などで**Bot**と感じさせない研究は多そうだが,書体に注目する例は珍しいと思った.