

# DESIGNING BOTS TO ACT NICE

Words can be found in any direction (including diagonals) and can overlap each other. Use the word bank below.

T A K I V I D E O N O L Y S M A Z O Q E  
D Z B M N R B S L A I P T W D G F U V T  
I F J R C N M Q G J N R E K N N W P C A  
D L J O C V A K K O A D F I I I I K T B  
O V V B O A T N R T C U A T M T K R C E  
T P I O M L H K E Z U I S C P O N G U D  
Y B F T P U O G T V L H L H E V O D F S  
S Q L S U E Y M A I T O A E E P I X Z D  
G G W Y T S L H L R U F K N D U S K U S  
X T R K E L Z U U T R I I D R A W E R T  
E V A T R E H M G U E T R A I T S O M Z  
T M T L S A P A E A L D X M I S D V I J  
O R Z A L R Y N R L E U X J B L U E N S  
B E S S U N Z I H D T X E T N O C R E L  
T T X A F I B T J H A R M L E S S L C E  
A N A B P N S Y A Z G Q L L E Q N O R D  
H U C G L G X O O U T P U T Z F X R A O  
C O C M E K M D B W I A F Y N O E D F M  
F C N E H N J E N I H C A M R S R S T I  
H O N E S T T G N I N I A R T E R P J T

## Word Bank

- |              |              |                 |             |
|--------------|--------------|-----------------|-------------|
| 1. harmless  | 2. deepmind  | 3. regulate     | 4. agent    |
| 5. video     | 6. strategy  | 7. minecraft    | 8. output   |
| 9. helpful   | 10. counter  | 11. math        | 12. traits  |
| 13. chatbot  | 14. culture  | 15. computers   | 16. safety  |
| 17. honest   | 18. machine  | 19. robots      | 20. debate  |
| 21. learning | 22. upvoting | 23. humanity    | 24. virtual |
| 25. kitchen  | 26. basalt   | 27. overlords   | 28. reward  |
| 29. context  | 30. models   | 31. pretraining | 32. values  |