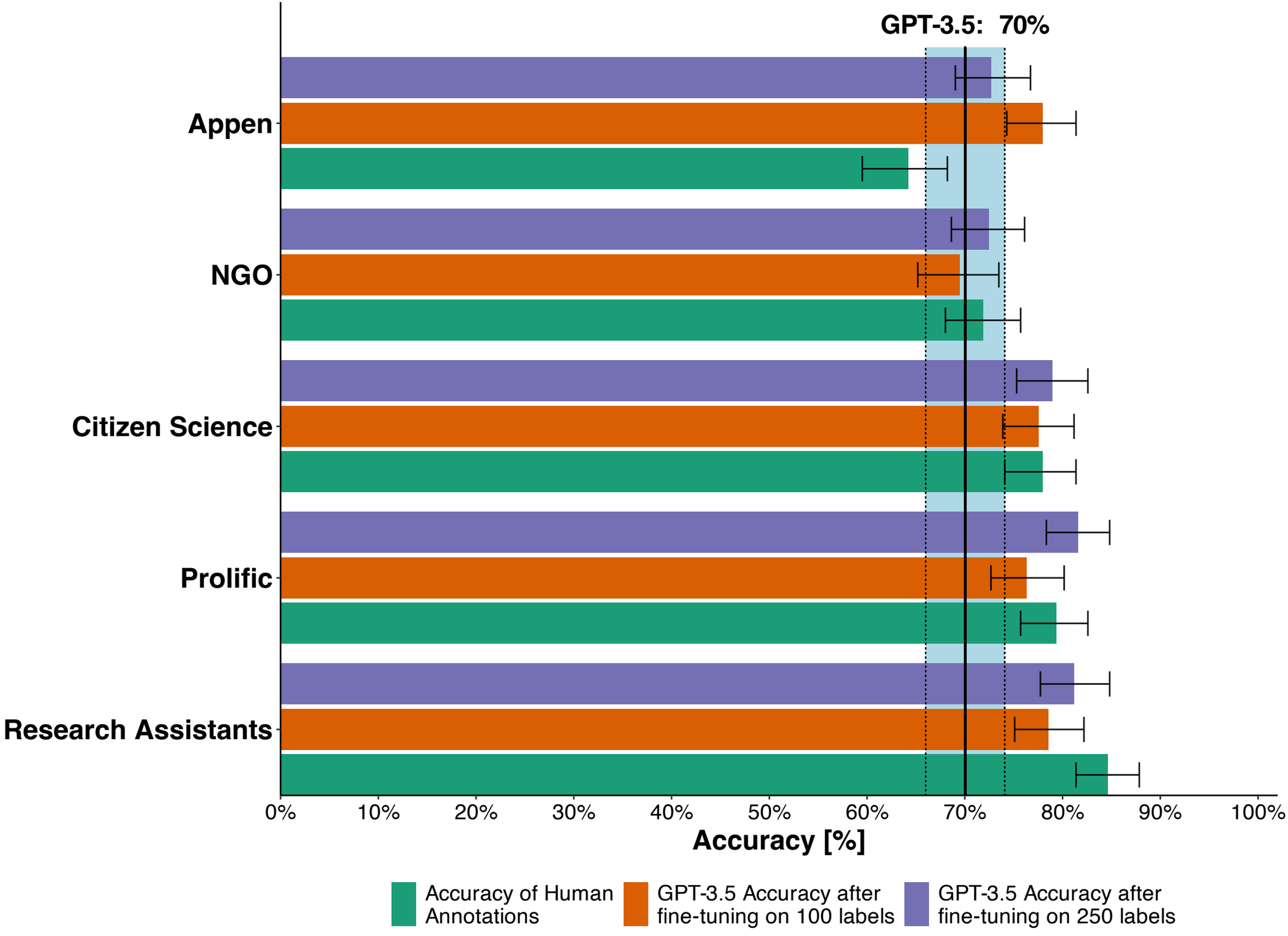
GPT-3.5: 70%
Appen
NGO
Citizen Science
Prolific
Research Assistants
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Accuracy [%]
Accuracy of Human Annotations
GPT-3.5 Accuracy after fine-tuning on 100 labels
GPT-3.5 Accuracy after fine-tuning on 250 labels