|  |  |  |
| --- | --- | --- |
| Model | Prompt | Krippendorff's α |
| GPT-4 | Zhou | 0.84 (0.77, 0.9) |
| GPT-4 | Kojima | 0.8 (0.73, 0.87) |
| GPT-4 | Zhou-instruction | 0.8 (0.73, 0.87) |
| GPT-4 | Articulate | 0.8 (0.72, 0.87) |
| GPT-4 | Plan | 0.78 (0.71, 0.85) |
| GPT-4 | Elaborate | 0.78 (0.71, 0.85) |
| GPT-4 | Rephrase | 0.78 (0.7, 0.85) |
| GPT-4 | Self-critique | 0.76 (0.68, 0.84) |
| GPT-4 | Converse | 0.74 (0.66, 0.82) |
| GPT-4 | Direct | 0.72 (0.64, 0.8) |