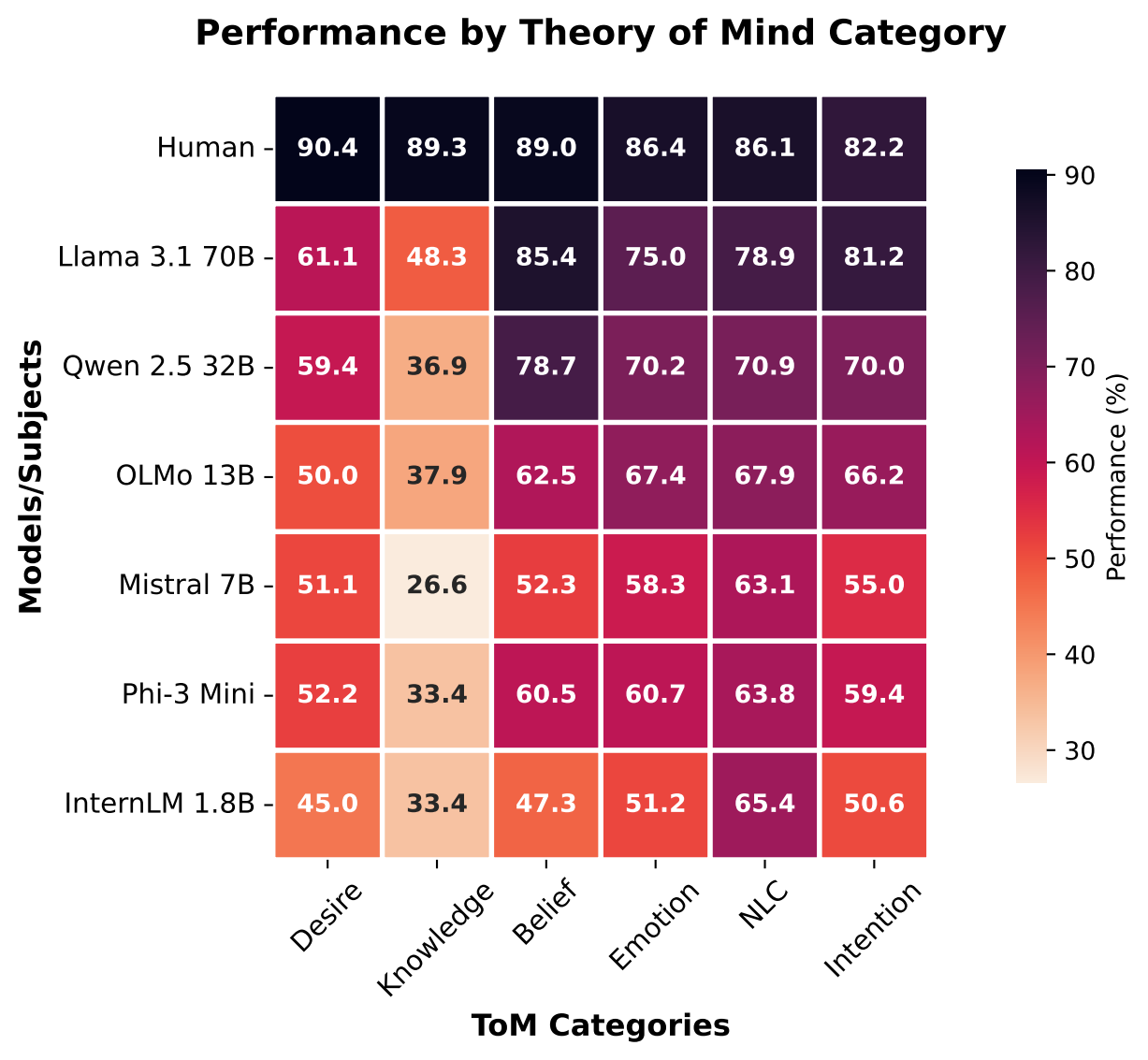
Performance by Theory of Mind Category
Models/Subjects
Human
Llama 3.1 70B
Qwen 2.5 32B
OLMo 13B
Mistral 7B
Phi-3 Mini
InternLM 1.8B
90.4 89.3 89.0 86.4 86.1 82.2
61.1 48.3 85.4 75.0 78.9 81.2
59.4 36.9 78.7 70.2 70.9 70.0
50.0 37.9 62.5 67.4 67.9 66.2
51.1 26.6 52.3 58.3 63.1 55.0
52.2 33.4 60.5 60.7 63.8 59.4
45.0 33.4 47.3 51.2 65.4 50.6
Desire
Knowledge
Belief
Emotion
NLC
Intention
ToM Categories
Performance (%)
90
80
70
60
50
40
30