Java Migration Success Rate: Our Framework vs Baselines
(Java 8 → 21, Maximal Migration)

Model	Success Rate
Sonnet 4.5 (Ours)	88.0%
Sonnet 4 (Ours)	80.0%
Sonnet 3.5 (Ours)	67.5%
Gemini 2.5 Flash (FreshBrew JDK17)	52.3%
Gemini 2.5 Flash (FreshBrew JDK21)	49.8%
GPT-4.1 (FreshBrew JDK17)	47.1%
GPT-4o (FreshBrew JDK17)	30.9%
SD-Feedback (MigrationBench)	27.3%
OpenRewrite (FreshBrew JDK17)	7.0%
o3-mini (FreshBrew JDK21)	4.5%

50%

X-axis: Success Rate (%) — 0, 20, 40, 60, 80, 100

Legend:
- Our Multi-Agent Framework (with Web Search)
- MigrationBench SD-Feedback Baseline
- FreshBrew JDK 21 Baseline
- FreshBrew JDK 17 Baseline