

CUA-SPEEDRUN

Representative tasks

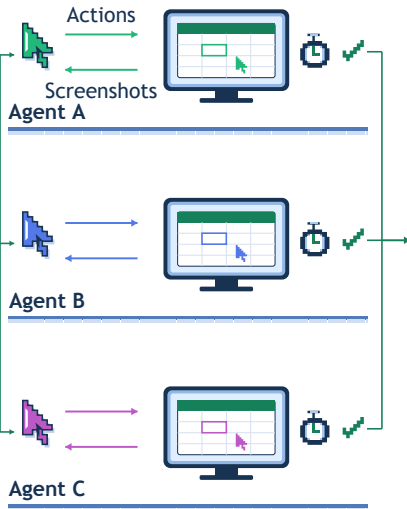
OSWorld

OSWorld2

CUA-World

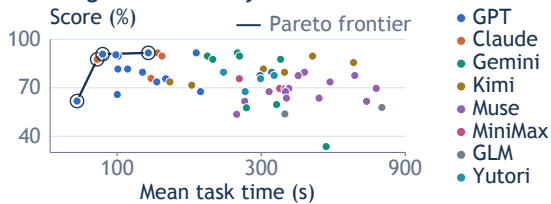
MyPCBench

Same tasks and runtime



Frontier CUA Analysis

No single model family dominates the Pareto frontier



More reasoning sometimes lowers cost

Gemini 3 Flash Preview

Low	\$1.40
Medium	\$0.59 (-58%)

Faster I/O sometimes slows agents

GPT-6 Astra at low effort

Normal	89.5s
Fast	99.0s (+11%)