



MindSearch: Mimicking Human Minds Elicits Deep AI Searcher

Zehui Chen^{1*}, Kuikun Liu^{2*}, Qiuchen Wang¹, Jiangning Liu², Wenwei Zhang², Kai Chen², Feng Zhao¹

¹University of Science and Technology of China ²Shanghai AI Laboratory



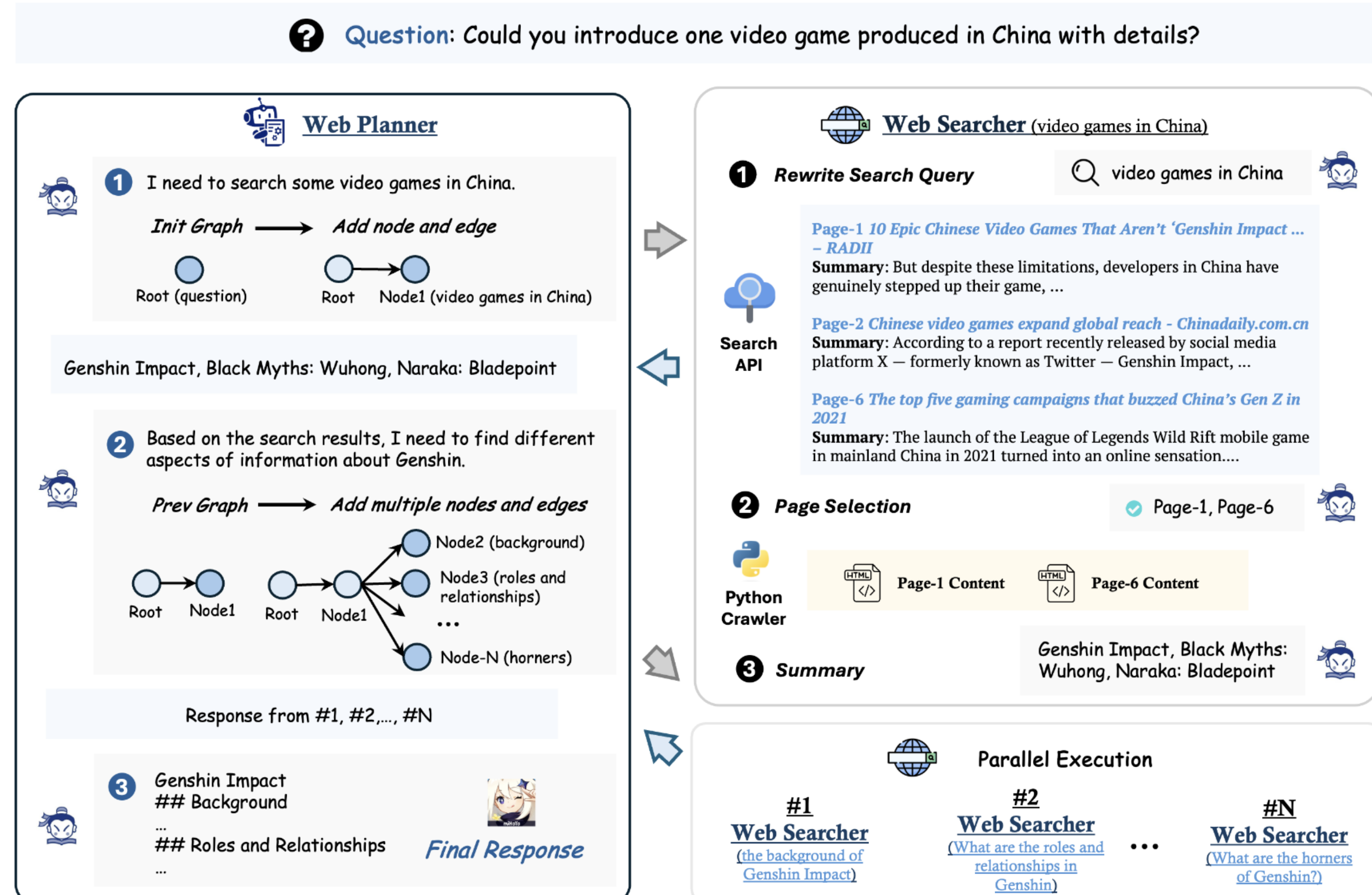
Motivation

opaque and low-quality reasoning process



Framework

Multi-Agent with WebPlanner & WebSearcher

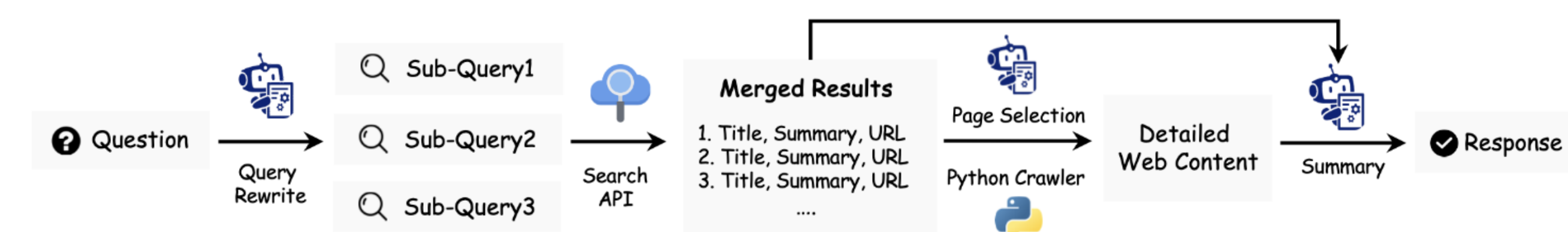


Methods

WebPlanner: Planning as Graph Construction



WebSearcher: Web Browsing with Hierarchical Retrieval

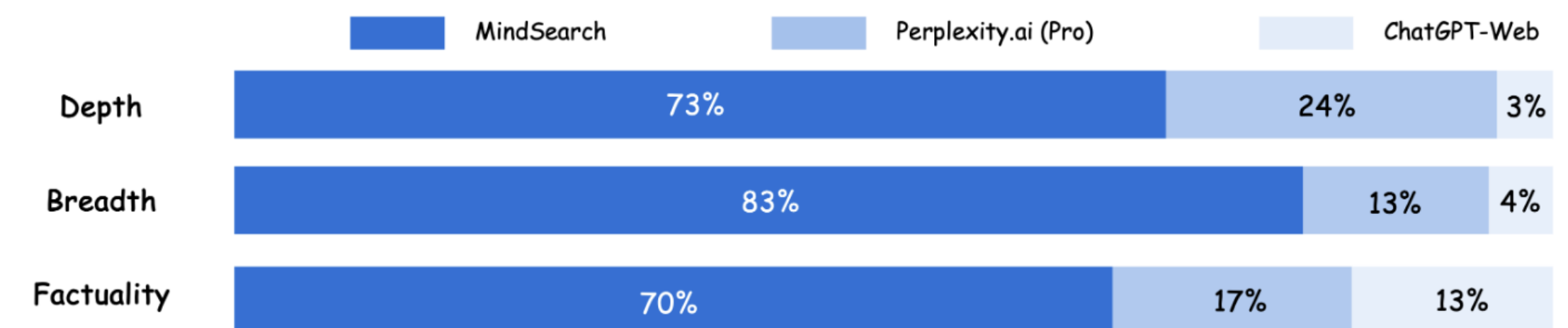


Experiments

Close-Set QA Tasks

Model	Bamboogle	Musique			HotpotQA			AVG
		2-hop	3-hop	4-hop	Easy	Medium	Hard	
Closed-Sourced LLM (GPT-4o)								
w/o Search Engine	70.4	54.0	22.0	20.0	73.0	69.0	66.0	53.5
ReAct Search	75.2	48.0	25.0	13.3	81.0	73.0	70.0	55.1
MindSearch	76.8	60.0	35.0	14.6	80.0	74.0	78.0	59.8
Open-Sourced LLM (InternLM2.5-7b-chat)								
w/o Search Engine	34.0	28.0	10.0	17.3	47.0	26.0	40.0	28.9
ReAct Search	55.2	38.0	17.0	16.0	69.0	56.0	49.0	42.9
MindSearch	67.8	46.0	20.0	18.6	69.0	66.0	57.0	49.2

Open-Set QA Tasks



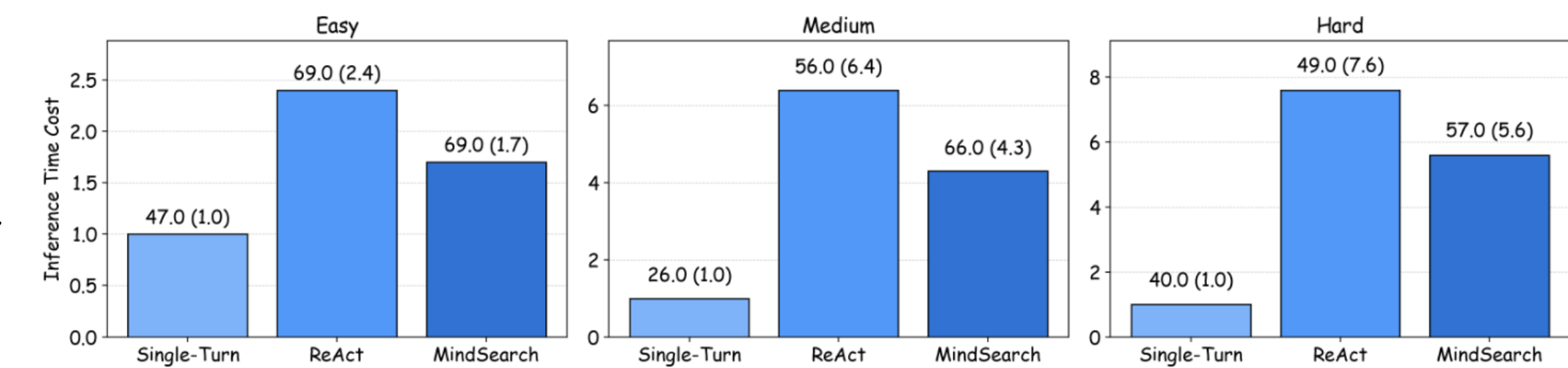
Analysis

Ablations on WebPlanner & WebSearcher

Strategy	Performance
-	37.6
ReAct	58.0
CodeAct	61.3
Ours	64.0

Method	Performance
Ours	64.0
w/o MQG	60.6
w/o PPC	63.3
w/o PS	58.0

Inference cost vs. search performance



Resource

