



Thomson
Reuters

When Deep Research Agents Stagnate: Enhancing Reasoning with Retrieval-Aware Agent Control

Heydar Soudani^{1,2}, Elizabeth Lingg², Faegheh Hasibi¹, Navid Rekabsaz²

¹Radboud University

²Thomson Reuters Labs

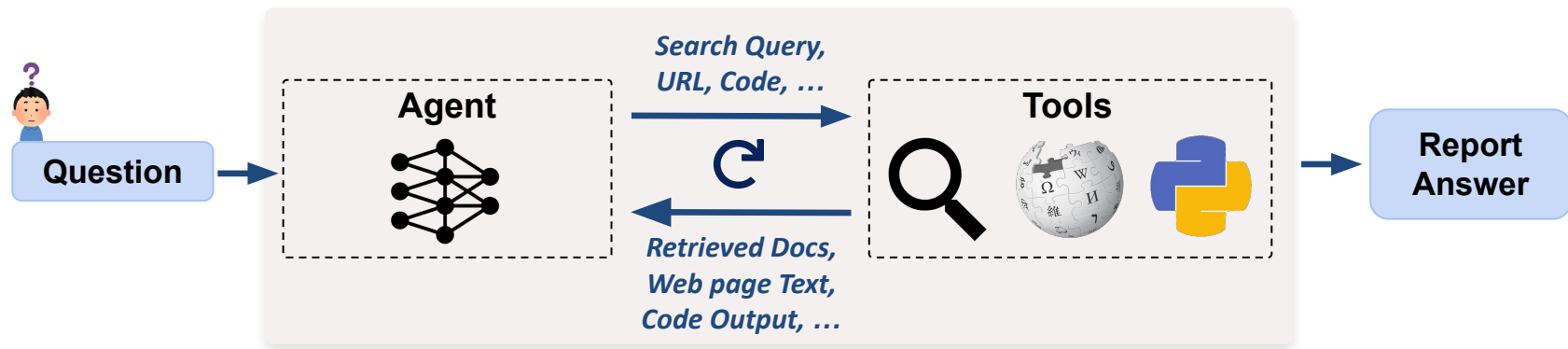
August 2026



Deep Research Agent (DRA)

Definition: the model is trained or prompted to use tools, like a search engine, while it thinks

Designed to perform *long-horizon*, multi-step information seeking and reasoning



DRAs: Example

Example: Search-R1, GLM, GPT-oss, ...



DRAs Components

Reasoner

Searcher

WorkFlow



Task: plan for the user task, reason over collected information, generate answer based on the gathered information

Ways to improve: high quality training data, increase the size of LLM, design reward function and training recipe

DRAs Components

Reasoner

Searcher

WorkFlow



The most common tool

Task: gather information

Ways to improve: training reasoning-aware retriever, incorporating IR tasks like query reformulation and reranking

DRAs Components

Reasoner

Searcher

WorkFlow



Task: Communication between the reasoner and the searcher, Context management, Represent gathered information for the reasoner

Ways to represent: Raw context (ReAct-style), Outline, Report

Critical Question

How effectively do these components work together over long reasoning trajectories to achieve the final objective?





Intermediate Steps Analysis

Intermediate Recall

- **Definition:** proportion of gold documents that have been retrieved up to that step
- Measure how effectively the retriever discovers relevant information throughout the reasoning trajectory

$$R_t = |D_g \cap D_{1:t}| / |D_g|, \quad D_{1:t} = \bigcup_{i=1}^t D_i$$

Intermediate Accuracy

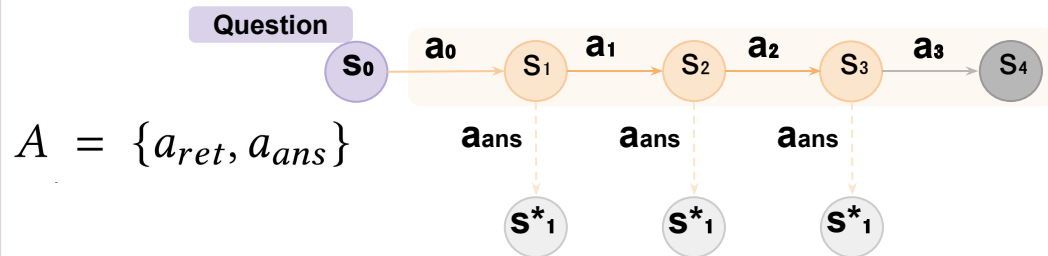
- Spawn a conceptual branch from the main reasoning trajectory by prompting the agent to generate a secondary reasoning step
- Allow the agent to output no candidate if sufficient information has not yet been gathered

Answer Candidate Generation - Format

Based on the evidence gathered so far, provide a concise answer.
IMPORTANT: The `<answer>` tag must contain **ONLY** the short factual answer, no explanations, no reasoning, no bullet points. Put all reasoning inside `<think>`.

`<think>` your reasoning over the collected evidence `</think>`

`<answer>` your short answer `</answer>`



Document Types

Relevant New:

The document is labeled as a gold document and has not been seen before in the trajectory

Relevant Seen:

The document is a gold document and has already been observed in the trajectory

Irrelevant New:

The document is not a gold document and has not been seen before in the trajectory

Irrelevant Seen:

The document is not a gold document but has already been observed in the trajectory

Experimental Setups

Agent	Backbone / Model	Size	Workflow
AgentCPM-Report	Qwen2.5	7B	Report-guided
GPT-oss-20B	gpt-oss-20b	20B	ReAct-style loop
AgentCPM-Explore	Qwen2.5	7B	ReAct-style loop
GLM-4.7	GLM-4.7-flash	30B	ReAct-style loop
ReAct	Claude Sonnet 4.5	-	ReAct-style loop
WebWeaver	Claude Sonnet 4.5	-	Outline-guided
Tongyi-DR	Qwen-2.5	30B	ReAct-style loop

Datasets: Deep Research

BrowseComp-Plus: contains Relevance Judgement and Final Answer

Q: What is the name of this celebrity:

1. She grew up in various cities due to her parents' professions
2. Her last theatre performance was between 1988-1994
3. As of December 2023, she was married twice
4. She and her second husband produced a play between 1970 and 1975 for which they were prosecuted, spending many nights in prison
5. As of December 2023, she only had one child in her second marriage

Ground truth answer: Tzeni Karezi

Gold doc IDs: {'69237': 2, '19085': 1, '73129': 2, '73639': 2, '40600': 2, '74532': 2, '6392': 2}

Datasets: Information Retrieval

NeuCLIR24 and **LegalSearch (UK)** contain human-verified supporting documents

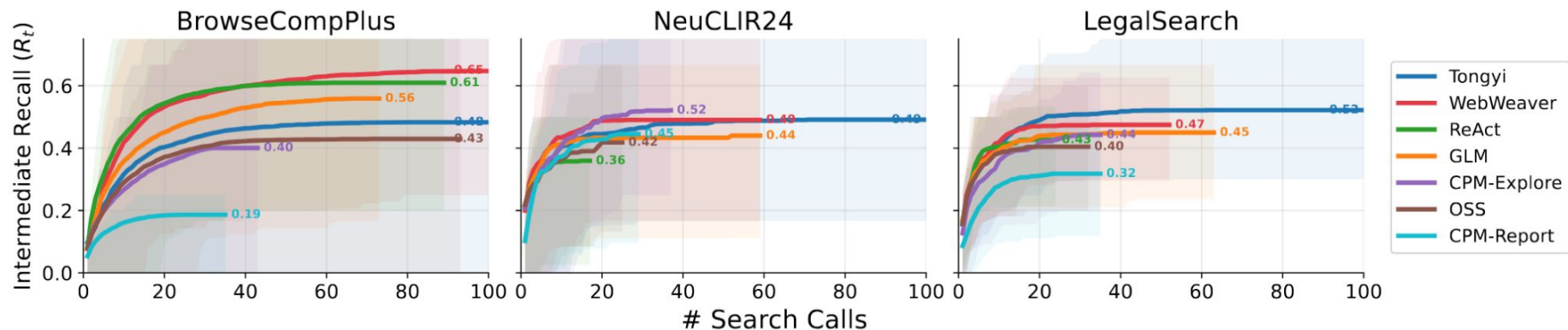
Q: What is the impact of the COVID-19 pandemic on suicide rates in Japan?

Gold doc IDs: {'69237': 2, '19085': 1, '73129': 2, '73639': 2, '40600': 2, '74532': 2, '6392': 2}

Retriever: Qwen3-Embedding-4B

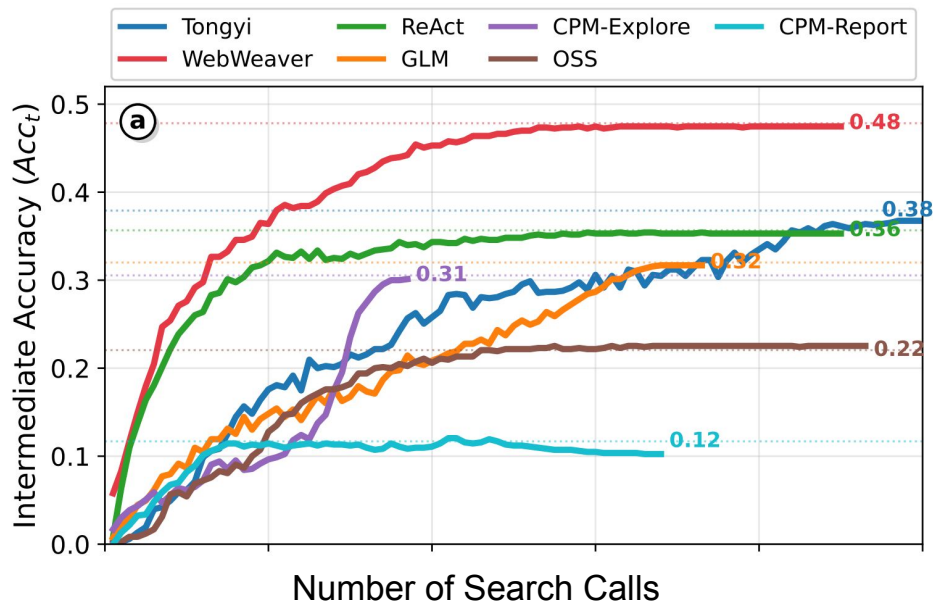
Result: Intermediate Recall

RQ1: How do DRAs' search and reasoning performances change throughout the reasoning trajectory?



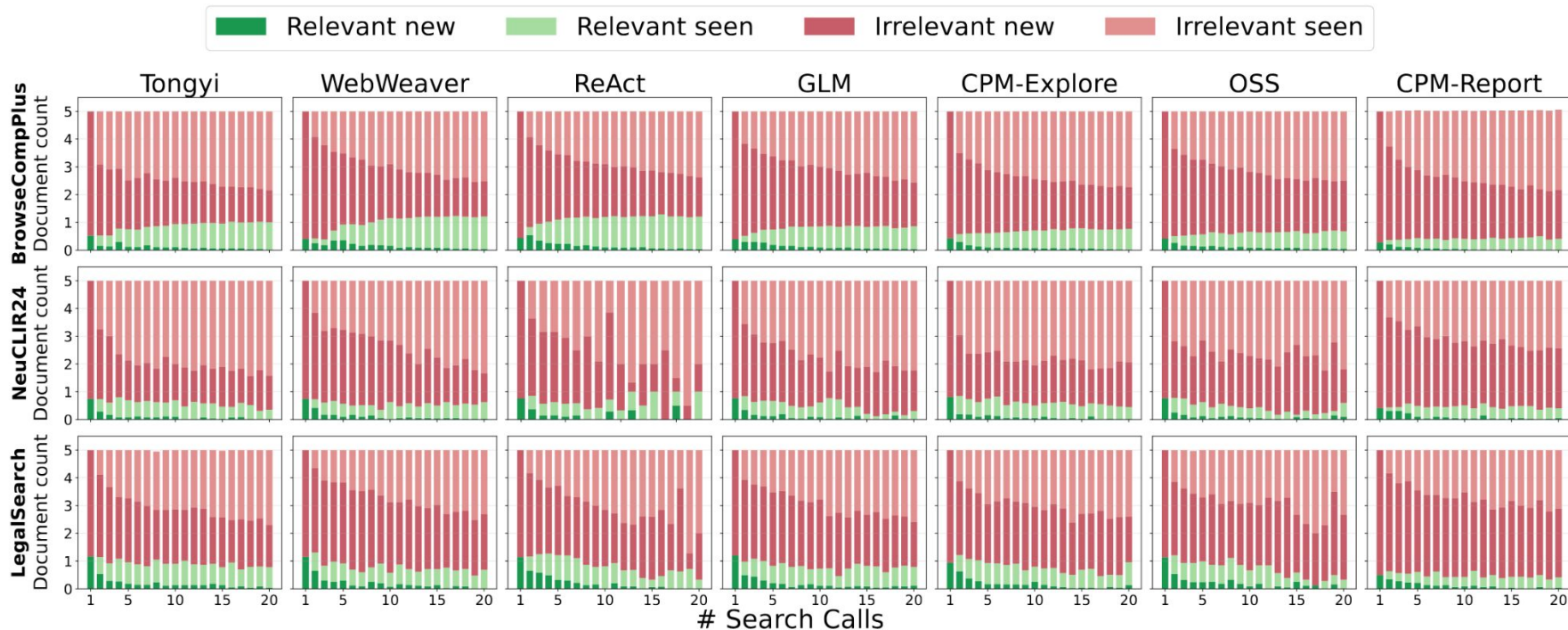
Result: Intermediate Accuracy

RQ1: How do DRAs' search and reasoning performances change throughout the reasoning trajectory?



Result: Document Types

RQ2: How does the distribution of document types observed by the agent evolve across iterations?



Findings

- Agents often suffer from ***Reasoning Stagnation***

The majority of iterations contribute *little or no improvement* to final performance, the agent does not know what is happening in its own trajectory, so it cannot change its search strategy, and it cannot decide when to stop.

Retrieval-Aware Agent Controller (RAAC)

DRAs Stagnate Causes

- (1) Limited awareness of information coverage
- (2) Difficulty shifting criteria or exploring new aspects

Solution: trajectory-level retrieval-aware signals to guide the reasoning process

Retrieval-aware Signals

Document Novelty (Nov)

Fraction of newly
retrieved documents
in an iteration

Criteria Coverage (Cov)

What information is
needed to fully
accomplish the task

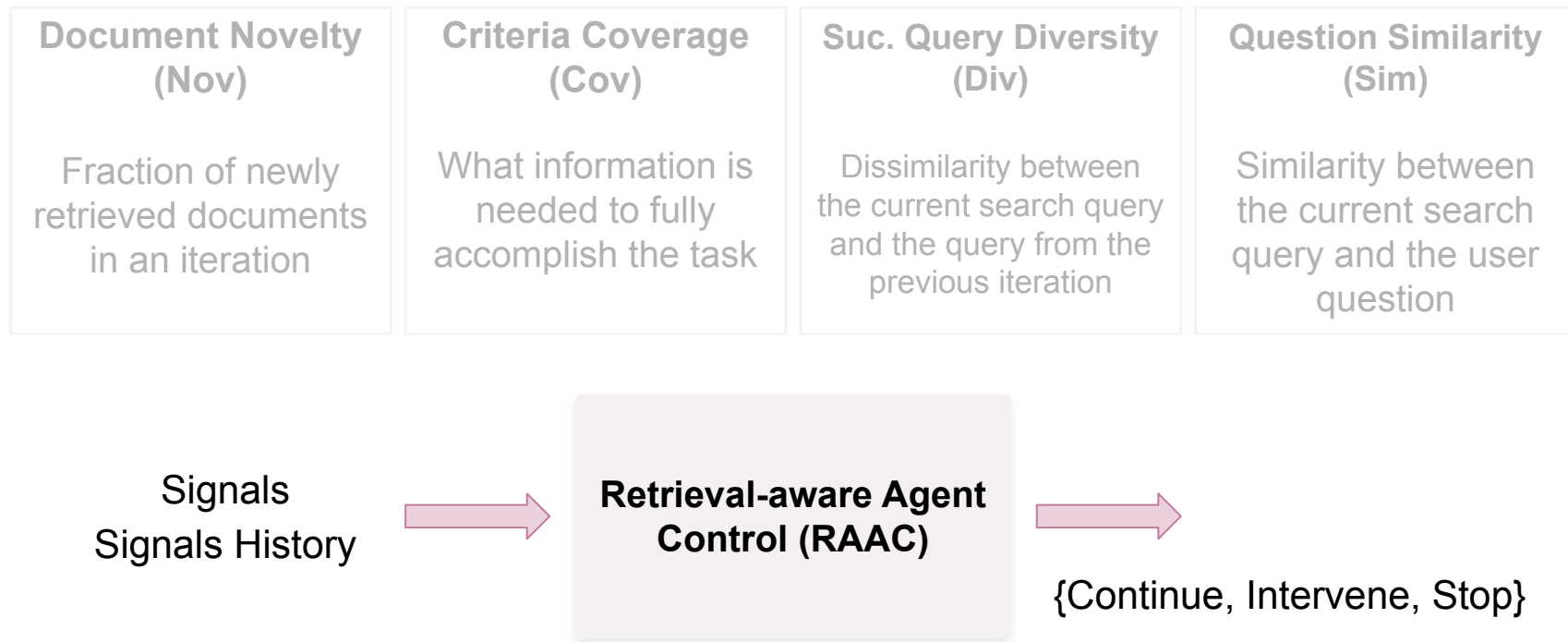
Suc. Query Diversity (Div)

Dissimilarity between
the current search query
and the query from the
previous iteration

Question Similarity (Sim)

Similarity between
the current search
query and the user
question

Agent Controller



RAAC: Example

Relevant Documents Ids: 19916, 1518, 82164, 72044, 47600, 40969, 63688, 60658

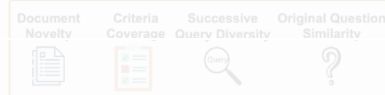
Gold Answer: *The Light*

? **Task:** A French saint inspired the name of this music group, which disbanded between 2010 and 2020 (exclusive). Later, two ...

New-item Recall	Interm. Answer	Agent	Original Trajectory
0.0	No Candidate		1 Search : music group named after French ... Doc ids: [59705, 22467, 47854, 56192, 26675]
			⋮
0.125	No Candidate		7 Search : Saint Etienne last music video ... Doc ids: [60658, 5804, 99078, 76827, 43575]
			⋮
0.625	The Light		8 Search : The Ark Korean band disbanded ... Doc ids: [19916, 1518, 82164, 72044, 47600]
			⋮
0.125	The Light		12 Search : The Ark 'On My Own' music ... Doc ids: [60658, 47600, 72044, 40969, 19916]
			⋮
0.125	The Light		15 Search : The Ark kpop officially ... Doc ids: [1518, 19916, 72044, 47600, 63688]
			⋮
0.0	The Light		22 Search : The Ark kpop second song ... Doc ids: [72044, 47600, 19916, 82164, 60658]
			Last Answer : The Light ✓

New-item Recall	Interm. Answer	Agent	Trajectory with RAAC
0.0	No Candidate		1 Search : music group named after French ... Doc ids: [59705, 22467, 47854, 56192, 26675]
			⋮
0.0	No Candidate		3 Search : Saint Vincent music group ... Doc ids: [26675, 59705, 37517, 31657, 36947]
			⋮
0.125	No Candidate		CS Search : band named after French saint ... Doc ids: [60658, 99729, 5804, 79021, 38904]
			⋮
0.25	The Light		4 Search : Ladies Code K-pop group ... Doc ids: [60658, 47600, 40969, 78345, 43033]
			⋮
0.5	The Light		5 Search : The Ark K-pop group French ... Doc ids: [19916, 1518, 82164, 72044, 47600]
			⋮
0.0	The Light		8 Search : The Ark K-pop group officially ... Doc ids: [72044, 1518, 19916, 47600, 82164]
			Last Answer : The Light ✓

Trajectory-based Signals



Retrieval-Aware Agent Controller (RAAC)

Continue



Intervene

Continue

Continue

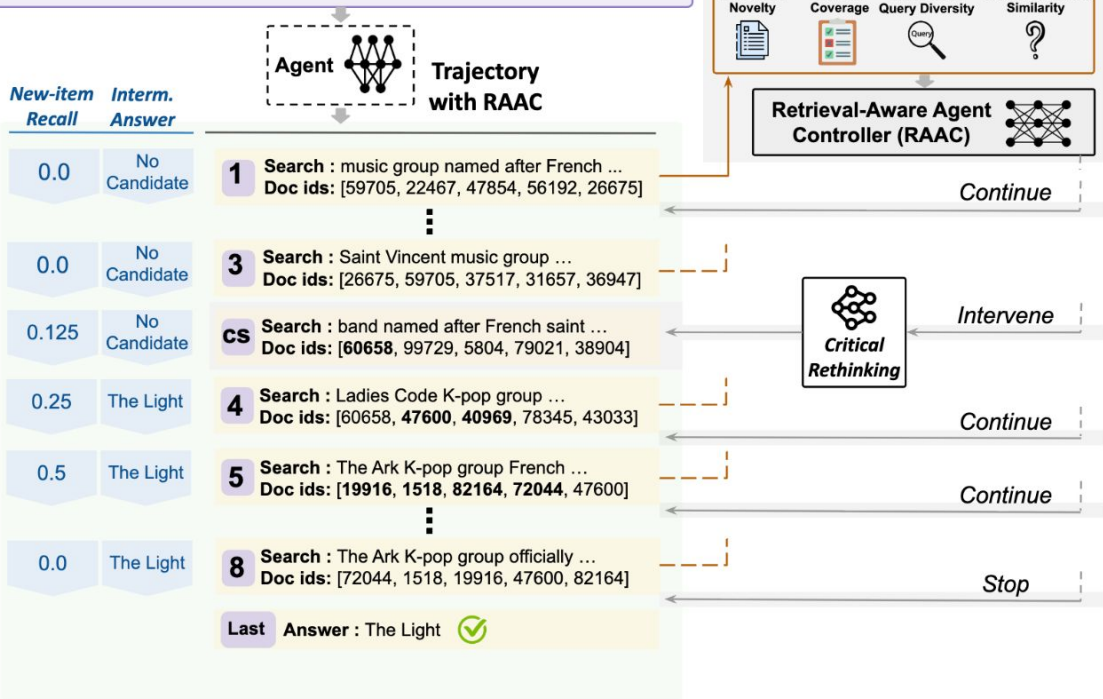
Stop

RAAC: Example

Relevant Documents Ids: 19916, 1518, 82164, 72044, 47600, 40969, 63688, 60658

Gold Answer: *The Light*

Task: A French saint inspired the name of this music group, which disbanded between 2010 and 2020 (exclusive). Later, two ...





Evaluation

RAAC Performance

RQ3) What is the effect of RAAC in mitigating reasoning stagnation?

Agent	BrowseCompP.			NeuCLIR		LegalSearch	
	R	Acc	#S↓	R	#S↓	R	#S ↓
CPM-Report	18.6	11.7	14.4	44.5	20.6	31.8	19.4
+ RAAC	30.5[†]	12.0	18.2	48.7	22.2	40.8[†]	21.6
GPT-oss-20b	43.0	23.0	28.3	41.8	12.2	40.5	10.2
+ RAAC	50.7[†]	26.4[†]	27.0	43.7	14.0	43.8	10.8
CPM-Explore	40.1	30.5	29.6	52.1	25.5	44.2	23.8
+ RAAC	48.0[†]	29.7	21.7	42.5 [†]	12.5	45.2	12.2
GLM-4.7	55.9	32.0	50.3	43.9	12.1	45.0	19.8
+ RAAC	59.2[†]	34.9[†]	36.8	46.8	17.1	50.0	18.6
ReAct	60.9	35.7	35.3	35.9	5.6	42.6	8.7
+ RAAC	62.6	39.1	25.4	39.1	5.7	47.5[†]	8.4
WebWeaver	64.9	47.8	75.0	49.0	19.2	47.5	19.9
+ RAAC	67.2[†]	51.3[†]	50.7	50.0	20.3	51.6	22.1
Tongyi-DR	48.4	37.9	77.8	49.1	62.1	52.2	53.8
+ RAAC	59.0[†]	47.9[†]	33.6	55.6[†]	34.4	56.5	22.8

Signals Effect

RQ4) What is the effect of each signal on controlling DRAs behavior?

Ag.	Sig.	BrowseCompP.			NeuCLIR		LegalSearch	
		R	Acc	#S↓	R	#S↓	R	#S ↓
OSS	<i>N-C-D</i>	50.7	26.4	27.0	39.3	11.6	40.2	10.0
	<i>N-D</i>	48.5 [†]	24.5 [†]	25.5	43.7	14.0	43.8	10.8
	<i>C-D</i>	41.4 [†]	17.2 [†]	12.3	39.3	10.4	40.7	8.9
GLM	<i>C-N-D</i>	59.2	34.9	36.8	44.1	12.0	47.2	13.9
	<i>N-D</i>	56.3 [†]	33.9	35.3	46.8	17.1	50.0	18.6
	<i>C-D</i>	45.1 [†]	26.9 [†]	14.7	41.1 [†]	10.0	44.4 [†]	11.2
WebWeaver	<i>C-N-D</i>	67.2	51.3	50.7	47.9	19.8	50.1	19.1
	<i>N-D</i>	65.7	51.8	42.1	50.0	20.3	51.6	22.1
	<i>C-D</i>	56.7 [†]	43.2 [†]	25.6	46.5	16.9	46.2	16.4
Tongyi	<i>C-N-D</i>	59.0	47.9	33.6	47.9 [†]	21.1	49.7 [†]	14.4
	<i>N-D</i>	56.9	49.1	31.3	55.6	34.4	56.5	22.8
	<i>C-D</i>	48.1 [†]	45.6 [†]	18.1	47.8 [†]	16.9	43.6 [†]	12.1

Signals Effect

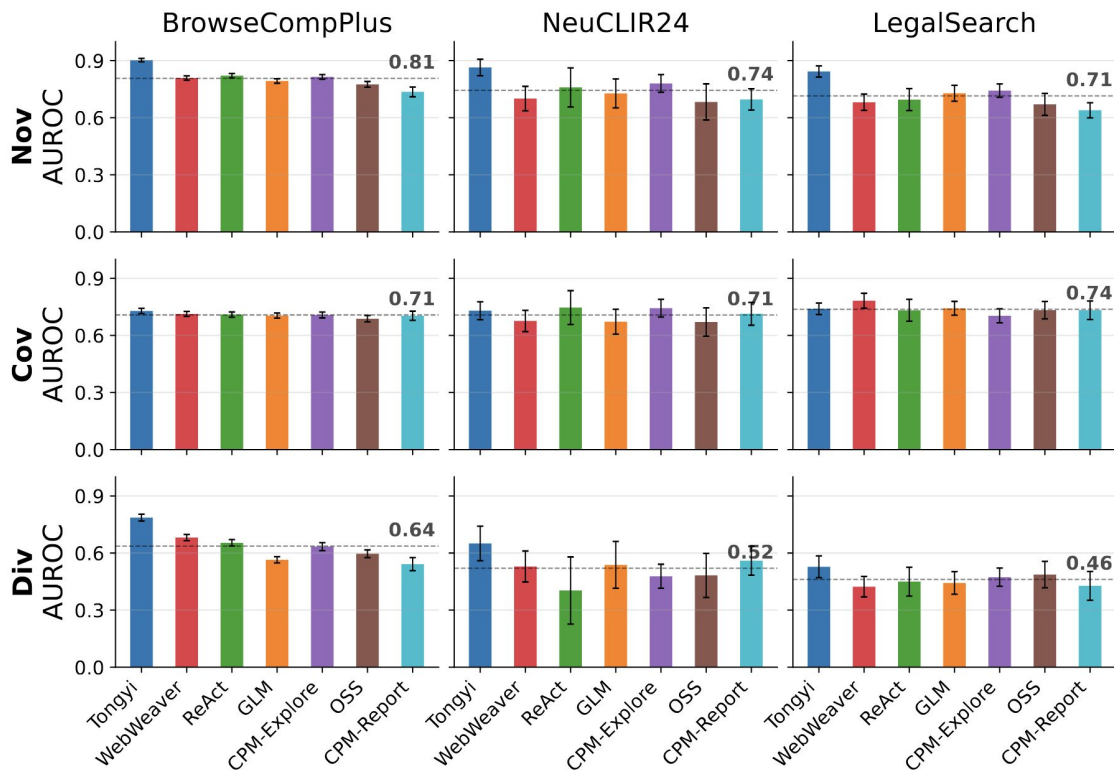
RQ4) What is the effect of each signal on controlling DRAs behavior?

New-item Recall: fraction of new relevant documents retrieved at iteration t that have not been retrieved in previous iterations

$$\text{NiR}_t = \frac{|(D_t \cap D_g) \setminus D_{1:t-1}|}{|D_g|}.$$

Signals Effect

RQ4) What is the effect of each signal on controlling DRAs behavior?



Takeaways

- **Reasoning stagnation.** First systematic analysis of DRAs' intermediate behavior. Most iterations add no progress, agents are unaware of what they have already found
- **Agents need a trajectory-level view.** Four unsupervised retrieval-aware signals give the agent a view of where its trajectory is going
- **That view must drive the agent's actions.** RAAC, a lightweight and agent-agnostic controller: {continue, intervene, stop}.

Across 7 agents and 3 datasets: higher recall and accuracy with fewer search calls