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

Heydar Soudani[1,2], Elizabeth Lingg[2], Faegheh Hasibi[1], Navid Rekabsaz[2]

[1]Radboud University [2]Thomson Reuters Labs

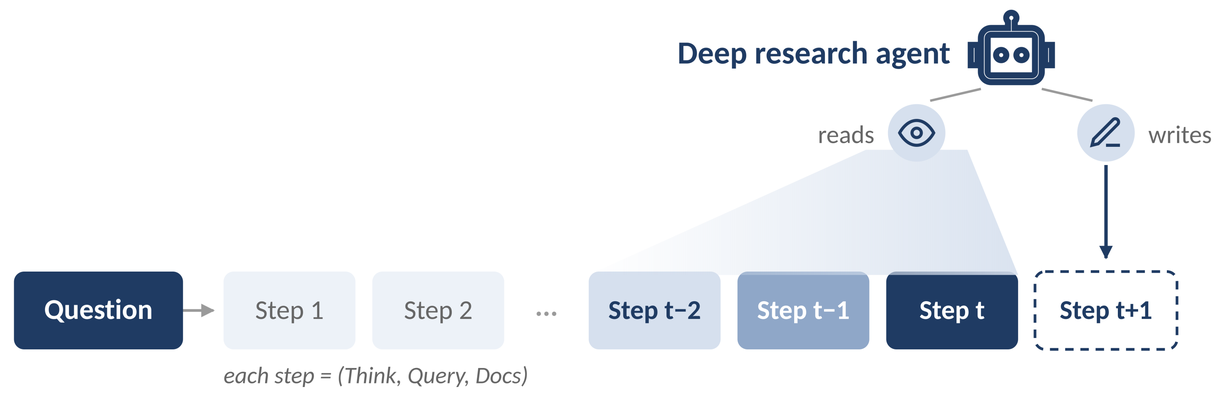
1 Introduction

Deep Research Agent (DRA)

- 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.

Challenge

- Prior work improves DRA components in isolation (retrieval, reranking, query reformulation).
- **How effectively do these components work together over a long reasoning trajectory to achieve the final objective?**



2 Intermediate Steps Analysis: Reasoning Stagnation

AIM

Measure whether each step makes progress toward the answer, by computing recall, accuracy, and the type of retrieved documents at every intermediate step of the trajectory.

FINDING

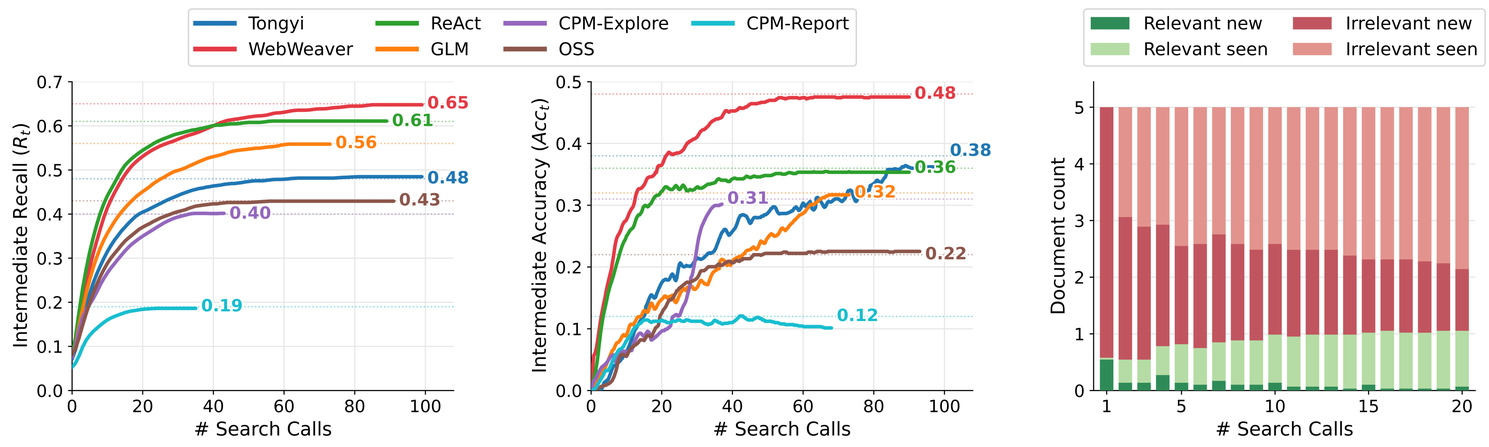
Stagnant reasoning: the majority of iterations contribute little or no improvement to final performance.

- Agents are not fully aware of what they have already gathered.
- Agents struggle to retrieve alternative evidence needed for the final answer.

SOLUTION

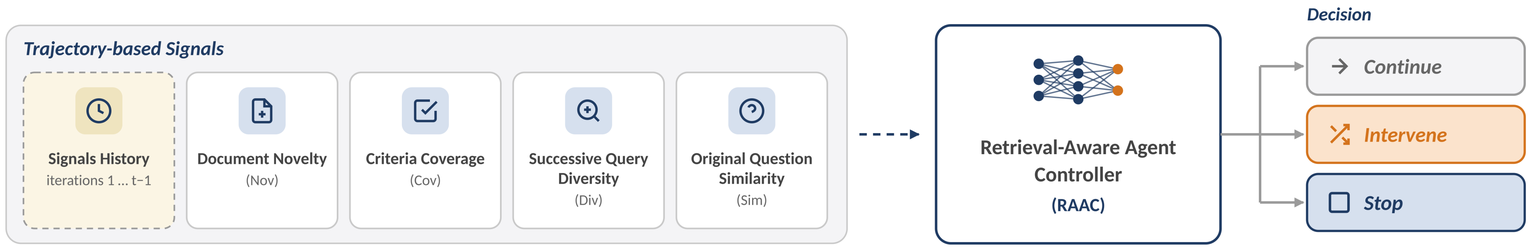
Agents need trajectory-level, retrieval-aware signals to guide their reasoning process.

EXP. SETUP 7 agents (4 architecture families) · 3 datasets: BrowseComp-Plus (BCP), NeuCLIR, LegalSearch · top-5 retrieval



3 Retrieval-Aware Agent Controller (RAAC)

Agent Controller



Document Novelty (Nov)
Fraction of retrieved documents not seen in any earlier iteration. Primary signal: low novelty = recycling old evidence.

Criteria Coverage (Cov)
Share of the question's information needs already addressed (uncovered / partial / covered). Confirms stop vs. intervene.

Successive Query Diversity (Div)
1 – cos(current query, previous query). Low diversity means the agent is looping on near-identical queries.

Original Question Similarity (Sim)
cos(current query, user question). A guardrail: when it drops, exploration is drifting off-topic.

- **Continue** — trajectory still productive; no intervention.
- **Intervene** — stale but recoverable: a critical re-thinker explains why the search is stuck and issues a substantially different query.
- **Stop** — exhausted: generate the final answer now.

Example

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

Gold answer: *The Light* **Gold documents (8):** *19916, 1518, 82164, 72044, 47600, 40969, 63688, 60658*

RAAC

Original trajectory — 22 steps

1	Search: *music group named after French …*	no new relevant docs · recall 0	no cand.
7	Search: *Saint Etienne last music video …*	+1 new relevant · recall 0.125	no cand.
8	Search: *The Ark Korean band disbanded …*	+5 new relevant · recall 0.625	The Light
12	Search: *The Ark 'On My Own' music …*	+1 new relevant · recall 0.125	The Light
15	Search: *The Ark kpop officially …*	+1 new relevant · recall 0.125	The Light
22	Search: *The Ark kpop second song …*	no new relevant docs · recall 0	The Light

Answer: **The Light** ✓

Trajectory with RAAC — 8 steps

1	Search: *music group named after French …*	no new relevant docs · recall 0	no cand.	Continue
3	Search: *Saint Vincent music group …*	no new relevant docs · recall 0	no cand.	Continue
CS	Search: *band named after French saint …*	+1 new relevant · recall 0.125	no cand.	Intervene → Critical rethinking
4	Search: *Ladies Code K-pop group …*	+2 new relevant · recall 0.25	The Light	Continue
5	Search: *The Ark K-pop group French …*	+4 new relevant · recall 0.5	The Light	Continue
8	Search: *The Ark K-pop group officially …*	no new relevant docs · recall 0	The Light	Stop

Answer: **The Light** ✓

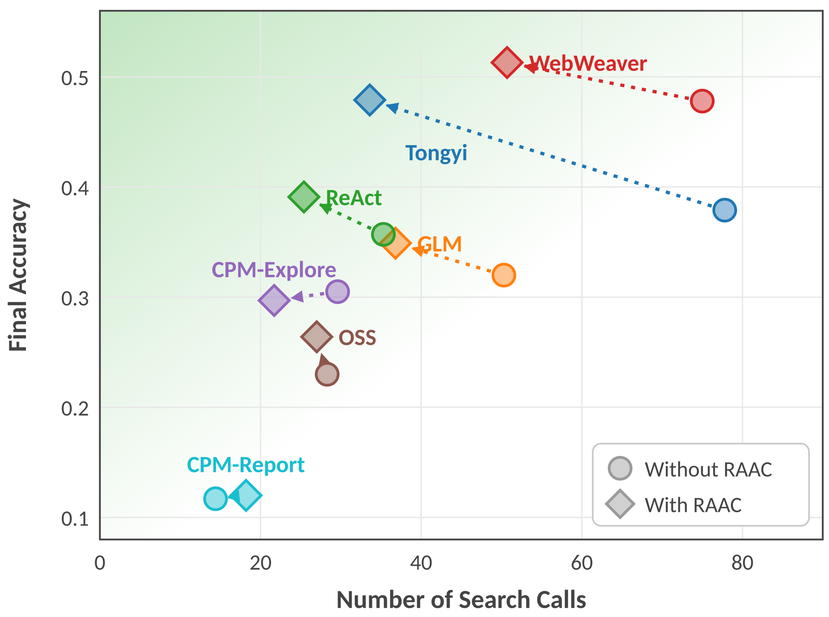
Relevant new · Relevant seen · Irrelevant (new or seen)

Results

+6.5
recall (BCP, avg.)
NeuCLIR +1.4 · Legal +4.5

+3.2 pts
accuracy (BCP, avg.)
up to +10 pts (Tongyi-DR)

−14
search calls (BCP, avg.)
NeuCLIR −4.4 · Legal −5.6



- **Efficiency:** fewer search calls on all three datasets while recall and accuracy are maintained or improved; controller overhead ≈ 10% of latency.
- **Effectiveness:** RAAC on the SOTA agents gives the best overall results: WebWeaver+RAAC on BCP (recall 67.2, acc. 51.3, 33% fewer calls); Tongyi-DR+RAAC on NeuCLIR and LegalSearch (45–58% fewer calls).
- **Adaptive:** where the base agent stops too early (CPM-Report), RAAC adds iterations and performance rises.
- **Signals:** Document Novelty is the most informative signal (AUROC up to 0.81 vs. new-item recall) and the best primary signal; Coverage helps only when human-written criteria exist (BCP).
- **Actions:** both actions help. Queries where RAAC stopped: 35% win vs. 18% loss in recall; intervene-only: 33% vs. 18% (BCP). Intervention redirects to higher-recall paths; stopping does not cut productive runs.

4 Takeaways

- **Reasoning stagnation is pervasive.** Across 7 DRAs and 3 datasets most iterations add no progress; agents cannot tell what they have found, or when to redirect or stop.
- **Trajectory signals expose it, and novelty matters most.** Four unsupervised retrieval-aware signals; document novelty is the most informative, coverage a useful complement.
- **A lightweight controller turns the signals into action.** RAAC (continue / intervene / stop) is agent-agnostic and training-free, sets new SOTA, and cuts search calls by a third.
- **Next: make agents natively trajectory-aware.** Train agents on these signals, or learn the controller, instead of bolting one on.

EMNLP2026 Budapest, Hungary

heydar.soudani@ru.nl