arXiv:2603.07241v1 [cs.LG] 7 Mar 2026

从网页内容分布匹配的视角重新思考深度研究

Zixuan Yu [*1] **Zhenheng Tang** [*2] **Tongliang Liu** [3] **Chengqi Zhang** [4] **Xiaowen Chu** [5] **Bo Han** [6]

摘要

尽管已集成搜索工具，深度搜索智能体仍常常面临推理驱动的查询与底层网页索引结构之间的错位。现有框架将搜索引擎视为静态工具，导致查询要么过于粗略，要么过于细碎，难以检索到精确证据。我们提出WeDAS，一种网页内容分布感知框架，将搜索空间的结构特征纳入智能体的观测空间。我们方法的核心是查询-结果对齐分数，这一指标用于量化智能体意图与检索结果之间的兼容性。为克服对动态网络建立索引的不可行性，我们引入少样本探测机制，通过有限的查询访问迭代估计该分数，使智能体能够根据局部内容格局动态重新校准子目标。作为即插即用模块，WeDAS 在四个基准测试上持续提升子目标完成度与准确率，有效弥合高层推理与底层检索之间的差距。

1. 引言

大语言模型（LLM）的持续扩展与架构改进，显著增强了其认知与推理能力（Achiam 等人，2023；Kaplan 等人，2020年；Guo 等人，2025）。凭借处理复杂逻辑和进行长程规划的能力，这些模型已从简单的文本处理器转变为自主能够进行深度搜索（Zhu 等人，2025）和深度研究（Huang 等人，2025）的智能体。这类系统利用大语言模型的潜在推理能力，通过迭代检索和高阶分析思维，在广阔的信息空间中探索，标志着开放领域中多维问题求解范式的转变（Huang 等人，2025）。

尽管智能体推理取得了显著进展（Ke 等人，2025；Kaplan 等人，2020），大语言模型（LLM）的内部认知深度与其信息获取的实际精确度之间仍存在巨大差距（Luo 等人，2025）。这些智能体的核心智能日趋成熟，但其在开放网络中导航的能力仍受持续存在的信息噪声瓶颈所阻碍（Xi 等人，2025）。我们认为，深度研究难以有效开展，主要障碍并非智能体的核心推理能力，而是其从网络环境中检索最相关网页的感知能力（Qian 与 Liu，2025；Liu 等人，2025）。这一瓶颈源于智能体的语言意图与搜索引擎索引的潜在信息分布之间存在结构性错位（Li 等人，2025b；Zhu 等人，2025）。智能体无法感知网页内容分布，使这一信息获取差距进一步加剧，并导致查询粒度校准普遍失效。粗略查询会带来大量无关噪声，而过于具体的查询则会造成检索稀疏性（Li 等，2025d；Huang 等人，2025）。

为量化所生成搜索查询与检索结果之间的语义连贯性，我们采用TF-IDF（Luhn，1957年）、杰卡德相似度（Jaccard，1901年）以及归一化莱文斯坦相似度（莱文斯坦，1966年）来衡量其相关性。如图1所示，成功轨迹与失败轨迹之间的相似度指标存在明显的分布偏移，证实了上文所讨论的结构性错位。这些轨迹收集自MiroThinker-v1.0-30B（Team等人，2025）求解GAIA（Mialon等人，2023）任务时的记录；若最终答案正确，则将该轨迹标记为成功，否则标记为失败。受这些挑战的启发，我们探究以下核心研究问题：

自主智能体如何感知

[*]共同贡献 [1]中山大学计算机科学与工程学院 [2]香港科技大学计算机科学与工程系 [3]悉尼大学悉尼人工智能中心 [4]香港理工大学数据科学与人工智能系 [5]香港科技大学（广州）数据科学与分析专修 [6]香港浸会大学计算机科学系 TMLR 团队。通讯作者：Bo Han<bhanml@comp.hkbu.edu.hk>。

预印本。*2026*年*3*月*10*日。

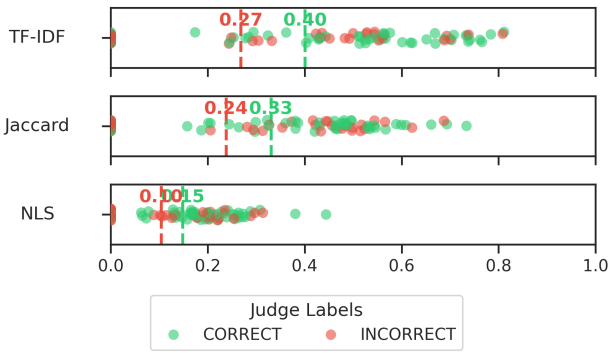

图 *1.* 成功与失败轨迹的查询—观测对齐指标（TF-IDF、Jaccard 和归一化莱文斯坦相似度）分布，凸显了智能体生成的查询与检索到的网页内容之间的结构性错位。

网页内容的分布，从而自适应地校准其搜索方向和查询粒度？

为解决这一问题，我们提出查询-结果对齐分数（QRAS），这是一种旨在量化搜索查询与当前信息格局之间契合度的指标。QRAS 将查询与搜索输出之间的接口分解为三个不同维度，从而对检索效能进行细致评估。该指标使智能体能够评估其目标的可搜索性，并依据经验反馈自适应地优化其轨迹。

基于这一指标，我们提出网页内容分布感知搜索（WeDAS），这是一个赋予智能体环境感知能力的框架。通过整合 QRAS，WeDAS 使智能体能够在最终确定搜索路径之前，预判网页内容的密度和相关性。这种前瞻性可确保检索结果与研究目标紧密契合，从而显著提高复杂调查中的信噪比。

这种方法面临的主要技术挑战在于，互联网规模庞大且缺乏结构，因此无法事先了解内容分布。我们通过引入一种内容分布探测机制来解决这一问题，该机制针对特定研究子目标对潜在查询空间进行采样。此机制构建可用信息的代表性映射，提供必要的反馈，帮助在查询广度与具体性之间权衡。通过揭示在线数据的内在密度，我们的框架引导智能体校准搜索轨迹，确保后续工作聚焦于效用高的信息区域，同时避开冗余或稀疏的子空间。我们的贡献如下：

•我们将**查询-结果对齐分数**（QRAS）形式化为一种易于处理的指标，用于细粒度量化搜索效果（第4节）。

• 我们提出了**网页内容分布感知搜索**（WeDAS），这是一种新颖的框架，利用迭代式少样本探测来描绘网络的潜在信息地形（第5.1节）。

• 通过在开放式研究基准上进行实验，我们证明 WeDAS 提升了搜索轨迹的信息增益（第 5.2 节）。

2. 相关工作

2.1. 搜索引擎

传统搜索引擎历来依赖基于统计框架的词汇匹配，例如 TF-IDF（Luhn，1957）和 BM25（Jones 等人，2000年），这些方法优先考虑精确的关键词重叠和词频（Zhu 等人，2025）。这些方法利用倒排索引结构，基于启发式的评分高效地对文档进行排序（Li 等人，2025e）。尽管这些方法计算稳健且高度可扩展，但从根本上受限于词汇不匹配问题（Li 等人，2025b）。由于依赖表层词项匹配，当查询与文档未使用完全相同的术语时，这些方法往往无法捕捉二者之间的潜在语义关系；这一局限促使检索范式转向神经网络和基于语义的检索范式（Thakur 等人，2021）。

2.2. 工具增强型大语言模型

工具增强型大语言模型代表着从静态生成器向主动问题解决者的重大转变，使其能够与外部工具交互，从而克服自身固有的知识截止问题和幻觉（Qu 等人，2025）。这一框架的核心能力之一是集成网络搜索引擎，使大语言模型能够充当动态接口，获取现实世界的信息（Zhu 等人，2025）。与通常基于预定义本地语料库运行的传统检索增强生成（RAG）不同（Huang 和 Huang，2024年），搜索增强系统利用开放网络检索最新信息和长尾知识。

标准集成遵循重写-检索-阅读流程。具体而言，模型首先通过改写查询，使用户意图更符合搜索引擎的要求；随后执行检索，最后综合生成有依据的回答（Ma 等人，2023）。然而，这些传统方法往往受限于静态的单轮交互模式，因此难以有效处理复杂的多步查询。近期研究因此转向智能体式搜索范式和深度研究架构（Xi 等人，2025）。Search-R1 和 WebThinker 等系统利用强化学习和自主规划，开展迭代式

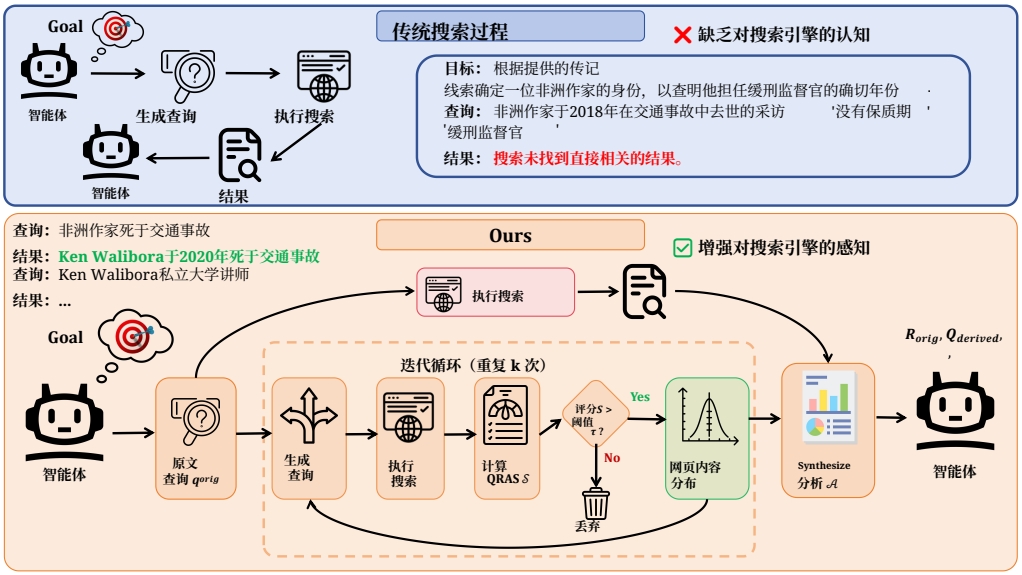

图2。网　页内容分布感知搜索框架　（WeDAS）

浏览与推理，从而能够更深入地挖掘信息（Jin 等人，2025；Li 等人，2025d）。

2.3. 深度搜索智能体

与通常依赖单轮检索的传统工具增强大语言模型不同，深度搜索智能体能够自主理解复杂的用户意图、规划多阶段搜索策略，并在不同来源中执行动态的多轮检索（Huang 等人，2025）。当前研究将这类智能体的架构范式划分为三种不同的结构（Xi 等人，2025）：

并行结构，将复杂意图分解为相互独立的子查询，或通过丰富查询内容，同时涵盖多个视角（Zhao等人，2025）。**顺序结构**，利用反思驱动循环或主动规划，根据中间反馈和累积推理自适应地优化搜索动作（Chen等人，2025b）。**混合结构**，例如基于树或基于图的方法，将并行探索与顺序推理结合起来，以优化复杂的信息寻求轨迹（Chen等人，2024年）。

尽管取得了这些进展，现有方法主要侧重于使模型适应固定工作流，却忽视了工具自身的演进。因此，搜索引擎往往仍是静态组件，而不是智能体迭代发现过程中的动态协同推理伙伴。

3. 预备知识

3.1. 大语言模型

我们将智能体工作流中的大语言模型（LLM）形式化为参数化函数 π_θ。令 $\mathcal{X}$ 表示自然语言字符串的空间。给定高层研究目标 $x \in \mathcal{X}$ 以及智能体累积的交互历史 $S_t = \{(a_i, o_i)\}_{i=1}^{t-1}$，其中 a_i 是第 i 个工具增强动作（例如发出查询或打开页面），o_i 是相应的观测，大语言模型充当随机策略，将当前上下文映射为关于下一动作 a_t 的分布：

$$a_t \sim P_{\pi_\theta}(\cdot \mid x, S_t) \tag{1}$$

具体而言，大语言模型将拼接后的序列$[x;\ S_t]$作为由词元组成的序列 $\mathbf{t} = (t_1,\ t_2,\ \ldots\ldots,\ t_n)$进行处理。输出生成通常是一个自回归过程，其中生成动作 a_t 的概率可分解为：

$$P(a_t \mid x, S_t) = \prod_{j=1}^{L} P(t_j \mid x, S_t, t_{<j}; \theta) \tag{2}$$

其中，L 是生成的动作字符串的长度。

在我们的框架中，我们使用不同的提示词复用同一个骨干模型：π_θ 表示行动者，负责生成动作（当 a_t 对应搜索动作时，该动作由查询 q_t参数化），而另设一个元评估器$\mathcal{M}_\theta$（第5.1节）用于计算对齐分数$\varphi(q, \tilde{o})$，并根据检索证据综合得出可行性分析 $\mathcal{A}$。

3.2. 搜索引擎

搜索引擎充当智能体与开放域环境 $\mathcal{W}$ 之间的接口。我们将环境定义为一个非结构化、高维度的语料库 $\mathcal{D} = \{d_1, d_2, \ldots, d_N\}$，其中 N表示已索引网页的总数。从数学角度看，搜索引擎被建模为一个检索函数 f_{SE}：$\mathcal{Q} \to \mathcal{O}$，其中$\mathcal{Q} \subset \mathcal{X}$表示可能的搜索查询空间。

给定智能体在时间 t 生成的查询 $q_t \in \mathcal{Q}$，搜索引擎会返回结构化观测 o_t，即由排名前 k 的文档组成的列表：

$$o_t = f_{SE}(q_t; \mathcal{W}) = \{d^{(1)}, d^{(2)}, \ldots, d^{(k)}\}. \quad (3)$$

其中，$d^{(j)} \in \mathcal{D}$ 表示检索到的第 j 个文档。每个文档 d 均由一个元组（标题、摘要、URL）构成。在我们的实验中，我们使用 Serper Google 搜索 API 实例化 f_{SE}；每次调用都会返回 $k = 10$ 条结果。为了进行评分和相似度计算，我们进一步定义文本化观察 $\tilde{o}_t$，将返回的摘要（以及可选的标题）串联成一个字符串；除非另有说明，我们使用仅串联摘要的方式作为 o_t。

在本研究中，我们将 f_{SE} 视为**不可微的黑箱函数**，这意味着智能体无法访问底层的排序梯度或完整的索引 $\mathcal{D}$。因此，核心挑战在于通过迭代探测来近似响应分布 $P(o_t \mid q_t, \mathcal{W})$，以克服由网页索引的离散性和易变性所导致的相关性瓶颈。

3.3. 深度搜索智能体工作流

我们将深度搜索智能体形式化为一个分层决策系统，该系统通过与动态网络环境的迭代交互来解决开放域任务$\mathcal{T} \in \mathcal{X}$。该过程始于规划阶段，其中任务分解算子 Φ_{plan} 将输入映射为 n原子子问题：$\mathcal{S} = \{s_1, s_2, \ldots, s_n\} = \Phi_{\text{plan}}(\mathcal{T})$。对于每个子问题 s_i，智能体进入推理–动作循环。在每一步 t，智能体维护交互历史$S_{t,i}$（用于子问题 s_i），采样中间推理轨迹$c_{t,i} \sim \pi_{\text{reason}}(S_{t,i}, s_i)$，并选择工具增强动作$a_{t,i} \sim \pi_{\text{act}}(c_{t,i})$（例如，发出查询或打开页面）。当$a_{t,i}$ 对应于搜索动作时，它由查询$q_{t,i}$参数化。

随后，环境返回一个观测 $o_{t,i}$，交互历史更新为 $S_{t+1,i} = \{S_{t,i}, a_{t,i}, o_{t,i}\}$。当满足停止条件时，此循环终止，并生成一个子结果 r_i。最后，在总结阶段，聚合算子 Ψ 将所有子结果合成为最终答案：$A = \Psi(\mathcal{T}, \{r_1, \ldots, r_n\})$。整体流程总结见算法 1。在本节中，我们

算法 1深度研究智能体框架

输入： 任务 $\mathcal{T}$，推理策略 π_θ，限制 T_{max}
输出： Synthesized Response A
$\{s_i\}_{i=1}^n \leftarrow \Phi_{\text{plan}}(\mathcal{T})$ {任务分解}
$\mathcal{K} \leftarrow \emptyset$ {知识收集}
对于 $i = 1$ **至** n，**执行**
  $\xi \leftarrow (s_i)$ {初始化轨迹}
  当 $|\xi| < T_{max}$ **时执行**
    $c \sim \pi_\theta(\xi)$ {潜在推理状态}
    $a \sim \pi_{\text{act}}(c)$
    如果 $a =$ 终止 **那么**
      跳出
    结束如果
    $o \leftarrow \text{Env.step}(a)$ {搜索或浏览交互}
    $\xi \leftarrow \xi \circ (a, o)$ {更新轨迹}
  结束 while
  $\mathcal{K} \leftarrow \mathcal{K} \cup \{\text{总结}(\xi)\}$
结束 for
$A \leftarrow \Psi(\mathcal{T}, \mathcal{K})$ {最终信息综合}
返回 A

用 q 表示搜索查询，用 o 表示相应的检索到的观测（例如，文档的前 k 项列表），并用 S_t 表示交互历史。

4. 查询-结果对齐分数

在本节中，我们用 q 表示搜索查询，o 表示相应的结构化检索观测（例如，文档的前k 个列表），并用 $\tilde{o}$ 表示其文本化形式（定义见第 3.2 节），该形式由返回的摘要构建而成。我们用 S_t 表示交互历史。

4.1. 信息增益的定义

我们将搜索轨迹视为一个证据获取过程，它能够降低关于潜在真实答案随机变量 $A^\star$（即任务的正确解）在给定智能体交互历史 S_t 时的不确定性。令

$$H(A^\star \mid S_t) = -\mathbb{E}_{A^\star \sim P(\cdot \mid S_t)}\big[\log P(A^\star \mid S_t)\big] \quad (4)$$

表示第 t 步的条件香农熵。在发出查询 q_t 并从网络环境 $\mathcal{W}$获得观测 o_t后，历史更新为 $S_{t+1} = \{S_t, q_t, o_t\}$。在实践中，每当需要计算对齐相关分数时，我们都会将结构化检索结果 o_t 转换为其文本化形式 $\tilde{o}_t$（第 4 节），并应用$\varphi(q_t, \tilde{o}_t)$。

逐点信息增益和期望信息增益。给定一个已实现的观测—

值 o_t，我们将**逐点信息增益**定义为

$$IG(o_t; A^\star \mid S_t, q_t) \triangleq H(A^\star \mid S_t) - H(A^\star \mid S_t, q_t, o_t). \tag{5}$$

由于智能体在看到 o_t 之前就选择了 q_t，因此决策目标是**期望信息增益**。这里，q_t 是用于索引观测分布$P(o_t \mid q_t, \mathcal{W})$的决策变量，而不是关于 $A^\star$ 的额外信息来源。

$$\mathrm{EIG}(q_t \mid S_t) \triangleq \mathbb{E}_{o_t \sim P(o_t \mid q_t, \mathcal{W})}\big[IG(o_t; A^\star \mid S_t, q_t)\big]. \tag{6}$$

因此，最优策略 π_θ 旨在选择能够最大化 $\mathrm{EIG}(q_t \mid S_t)$ 的查询。然而，在开放域设置中，后验分布 $P(A^\star \mid S_t, q_t, o_t)$ 无法解析且维度很高，因此无法直接估计 EIG；此外，带噪检索还会造成持续存在的相关性瓶颈。

KL 形式。逐点信息增益衡量熵减，对于某个特定实现可能为负。然而，众所周知，期望信息增益等价于后验分布与（以查询为条件的）先验分布之间的期望 KL 散度（贝叶斯惊异）：

$$\begin{aligned}\mathrm{EIG}(q_t \mid S_t) &= \mathbb{E}_{o_t \sim P(\cdot \mid q_t, \mathcal{W})}\Big[D_{\mathrm{KL}}\big(P(A^\star \mid S_t, q_t, o_t) \,\|\, P(A^\star \mid S_t)\big)\Big] \\ &= I(A^\star; o_t \mid S_t, q_t).\end{aligned} \tag{7}$$

由于 q_t 是在观测 o_t 之前选定的，且其本身不会揭示关于 $A^\star$ 的新证据，因此有 $P(A^\star \mid S_t, q_t) = P(A^\star \mid S_t)$。因此，最大化 EIG 就是寻找预期会引发较大后验偏移的查询。

潜在相关性与代理目标。我们引入一个二元潜变量 $z \in \{0, 1\}$，用于指示检索到的观测是否包含新的任务相关证据($z = 1$)，还是主要包含噪声/冗余($z = 0$)。

假设4.1（先选择查询，再进行观测）。在观测 o_t 之前选定查询 q_t，且查询本身不会揭示关于 $A^\star$ 的新证据。因此 $P(A^\star \mid S_t, q_t) = P(A^\star \mid S_t)$。

假设4.2（相关性条件后验）。存在一个随机变量 z，使得

$$P(A^\star \mid S_t, q_t, o_t) = \sum_{z \in \{0,1\}} P(A^\star \mid S_t, q_t, o_t, z)\, P(z \mid S_t, q_t, o_t). \tag{8}$$

此外，当 $z = 0$（没有新证据）时，该观测产生的后验偏移可以忽略不计：$P(A^\star \mid S_t, q_t, o_t, z = 0) \approx P(A^\star \mid S_t, q_t)$。

假设4.3（有界条件后验偏移）。令$\Delta(o_t, z) \triangleq D_{\mathrm{KL}}\big(P(A^\star \mid S_t, q_t, o_t, z) \,\|\, P(A^\star \mid S_t)\big)$ 表示以潜在相关性z为条件的后验偏移。我们假设：（1）如果 $z = 0$, $\Delta(o_t, z = 0) \approx 0$。（2）如果 $z = 1$，则存在常数 $\Delta_{\min}, \Delta_{\max} > 0$，使得$\Delta(o_t, z = 1) \in [\Delta_{\min}, \Delta_{\max}]$。

命题4.4（基于相关性的EIG上界）。在假设*4.1–4.3*下，期望信息增益的上界为期望相关性：

$$\mathrm{EIG}(q_t \mid S_t) \;\le\; \Delta_{\max}\, \mathbb{E}_{o_t \sim P(o_t \mid q_t, \mathcal{W})}\big[P(z = 1 \mid S_t, q_t, o_t)\big]. \tag{9}$$

证明概要。回顾一下，期望信息增益定义为 $\mathrm{EIG}(q_t \mid S_t) = \mathbb{E}_{o_t}[D_{\mathrm{KL}}(P(A^\star \mid o_t) \,\|\, P(A^\star))]$。

根据假设4.2，后验分布 $P(A^\star \mid o_t)$ 是一个混合分布 $\sum_z P(z \mid o_t) P(A^\star \mid o_t, z)$。由于 KL 散度 $D_{\mathrm{KL}}(\cdot \,\|\, Q)$ 是凸函数，我们应用詹森不等式，得到：

$$D_{\mathrm{KL}}(P(A^\star \mid o_t) \,\|\, P(A^\star)) \le \sum_{z \in \{0,1\}} P(z \mid o_t)\, \Delta(o_t, z).$$

应用假设4.3后，与 $z = 0$对应的项消失（因为 $\Delta \approx 0$），而 $z = 1$ 项的上界为$\Delta_{\max}$。因此：

$$D_{\mathrm{KL}}(\dots) \le P(z = 1 \mid o_t) \Delta_{\max}.$$

最后，对两边的 o_t 取期望，证明即告完成。

□

从 o_t 到 $\tilde{o}_t$，再到 φ。在实践中，我们只能观察到文本化摘要 $\tilde{o}_t$（第 3.2 节），而非完整的结构化 o_t。因此，我们用 $P(z = 1 \mid S_t, q_t, \tilde{o}_t)$ 近似 $P(z = 1 \mid S_t, q_t, o_t)$。最后，我们采用有界对齐分数 $\varphi(q_t, \tilde{o}_t)$ 作为这一相关性概率的排序一致代理（即，较大的 φ 应对应较大的$P(z = 1 \mid S_t, q_t, \tilde{o}_t)$），并将其用于筛选和轨迹控制。

4.2. 衡量搜索过程的信息增益

我们采用三种互补指标来衡量智能体生成的查询 q 与检索到的观测之间的对齐程度。具体而言，我们计算 q 与文本化观察 $\tilde{o}$（拼接前-k 个返回摘要所得）之间的相似度：**(i) TF-IDF 相似度**，定义为加权词向量之间的余弦相似度，$\mathrm{Sim}_{\text{tf-idf}}(q, \tilde{o}) = \frac{\mathbf{v}_q \cdot \mathbf{v}_{\tilde{o}}}{\|\mathbf{v}_q\| \, \|\mathbf{v}_{\tilde{o}}\|}$；**(ii) Jaccard 相似度**，用于量化词元重叠，$J(q, \tilde{o}) = \frac{|T_q \cap T_{\tilde{o}}|}{|T_q \cup T_{\tilde{o}}|}$；以及 **(iii) 归一化莱文斯坦相似度（NLS）**，通过归一化编辑距离捕捉形态接近度，$\mathrm{Sim}_{\text{nls}}(q, \tilde{o}) = 1 - \frac{\mathrm{Lev}(q, \tilde{o})}{\max(|q|, |\tilde{o}|)}$。从经验来看，单个搜索步骤的效用通常与这些对齐分数呈正相关。

尽管上述指标为评估查询与内容之间的对应关系提供了稳健的描述性统计框架，但将其集成到深度搜索智能体框架中时，仍存在固有局限。具体而言，这些相似度属于事后描述性统计，而非事前预测性启发式方法。虽然它们能够量化信息稀疏性或噪声的程度，却缺乏内在机制，无法为智能体提供用于查询优化的方向性信息。因此，若仅依赖这些传统指标，智能体就无法根据网络内容密度的动态变化自主校准其粒度。

为了解决这一局限并形式化反馈信号，我们提出了**查询-结果对齐分数（QRAS）**，将其作为一个有界效用指标 $\varphi(q,\ \tilde{o}) \in [\![0,\ 10]\!]$，其中 $\tilde{o}$表示根据排名前 k 的检索片段构建的文本化观察。在我们的实现中，元评估器以查询和排名前 k 的检索片段作为输入，并输出总体分数及简短分析。φ 并不直接近似全局熵减，而是通过以下三个潜在维度对搜索质量进行局部、多方面的评估：

- **主题相关性（s_{rel}）：** 查询 q 与文本化观察 $\tilde{o}$之间的语义一致性。
- **信息密度（s_{den}）：** $\tilde{o}$中非冗余、任务显著命题的集中程度。
- **噪声鲁棒性（s_{noi}）：** $\tilde{o}$中无关或具有干扰性内容占比的倒数。

我们将总分定义为三个子分数的非加权平均值：

$$\varphi(q,\tilde{o}) = \frac{1}{3}\big(s_{\text{rel}}(q,\tilde{o}) + s_{\text{den}}(q,\tilde{o}) + s_{\text{noi}}(q,\tilde{o})\big). \quad (10)$$

这种分解为逐步搜索效用提供了一个可解释的代理指标，并可在无需完整展开搜索过程的情况下，系统地评估探索效率。

5. 方法论

5.1. 网页内容分布感知搜索

基于所提出的对齐分数，我们引入**网络内容分布感知搜索（WeDAS）**，使智能体能够动态校准其探索策略。WeDAS 集成了一个反馈循环，其中大语言模型（LLM）充当元评估器 $\mathcal{M}_\theta$。

如算法2所述，对于每个查询–观测对$(q,\ \tilde{o})$，$\mathcal{M}_\theta$ 计算对齐分数 $\varphi(q,\ \tilde{o})$ 并生成定性可行性分析 $\mathcal{A}$，针对当前轨迹。通过用这些增强原始观测元评估，智能体获得对局部网页内容分布的经验估计。这种认知能够自适应地控制查询粒度，在宽泛的探索性探测与定向查询之间动态切换，从而在固定的交互预算下提高搜索效率。

元评估器接口。在我们的实现中，元评估器 $\mathcal{M}_\theta$ 执行**整体评分**：它以查询 q 及其对应的文本化观察结果 $\tilde{o}$（由排名前k 的搜索摘要拼接而成）作为输入，并输出三个子分数 $\big(s_{\text{rel}}\ q\ \tilde{o}\ \ s_{\text{den}}\ q\ \tilde{o}\ \ s_{\text{noi}}\ q\ \tilde{o}\big) \in (,\ )$ $[0,10]^3$ $\varphi\ q\ \tilde{o}\ \in [0,10]$ （，）（，）以及它们的平均值（，）。

动态阈值。我们不使用固定的分数阈值，而是在固定的探测预算下维护一小组效用较高的探测查询，同时始终将初始查询 q_0 保留为主要执行查询。

令 $\mathcal{P}_t$ 表示截至第 t次迭代收集到的派生探测元组集合，其中每个元组的形式为$(q,\ s,\ \mathcal{A})$，其中 $s = \varphi(q,\ \tilde{o})$，而 $\tilde{o}$ 是搜索查询 q后返回的文本化观察。我们设定

$$\tau_t = \min_{(q,s,\mathcal{A})\in\mathcal{P}_t} s, \quad (11)$$

并移除得分最低的探测查询。这个过程可视为一种在线 Top 集合维护规则：每次迭代都丢弃当前表现最差的探测查询，并将其余得分较高的探测查询保留为引导候选项。

提供给智能体的输出。在探测过程中，每个派生查询都会返回一个元组$(q,\ s,\ \mathcal{A})$，其中包含查询、其对齐分数以及元评估器分析；这些内容将作为后续查询生成的引导信号。相比之下，智能体仅使用初始查询 q_0 检索到的搜索结果来构建依据并收集证据。

候选生成算子。在算法 2中，$\Gamma_{\text{gen}}(\cdot)$表示一个候选生成算子（由大语言模型实现），它将当前查询集扩展为语义多样化变体，用于探测。

5.2. 通过少样本查询探测进行分布估计

为解决查询优化中的不确定性，我们提出一种**迭代式少样本探测** 策略，以近似高效用信息的局部分布。由于网络的规模和异质性使得对 $P(o \mid q,\ \mathcal{W})$ 进行全局建模难以实现，因此我们通过以下步骤估计局部 信息地形。

候选生成。给定当前查询集 $\mathcal{Q}_{\text{cur}}$，智能体会生成语义多样化变体 $\mathcal{Q}_{\text{cand}}$，以探测查询空间的不同区域。

对齐评估与剪枝。对于每个候选项$q \in \mathcal{Q}_{\text{cand}}$，智能体执行搜索以获得 o 并

算法 2 WeDAS

输入： 初始查询 q_0，最大探测迭代次数 T
$\mathcal{P} \leftarrow \emptyset$ {已收集的探测元组}
对于 $t = 1$ **到** T **执行**
  $q_t \leftarrow \Gamma_{\text{gen}}(q_0, \mathcal{P})$ {生成一个派生探针查询}
  $\tilde{o}_t \leftarrow$ 搜索(q_t) {返回文本化片段}
  $s_t \leftarrow \varphi(q_t, \tilde{o}_t)$
  $\mathcal{A}_t \leftarrow \mathcal{M}_\theta(q_t, \tilde{o}_t)$ {简短分析，来自元评估器}
  $\mathcal{P} \leftarrow \mathcal{P} \cup \{(q_t, s_t, \mathcal{A}_t)\}$
  $(q^-, s^-, \mathcal{A}^-) \leftarrow \arg\min_{(q,s,\mathcal{A}) \in \mathcal{P}} s$
  $\mathcal{P} \leftarrow \mathcal{P} \setminus \{(q^-, s^-, \mathcal{A}^-)\}$ {去掉一个得分最低的探针}
结束循环
返回 Synthesize$(\mathcal{P})$ {仅返回指导元组（非证据）}

计算 $s = \varphi(q, o)$。然后，我们应用基于阈值的剪枝：

$$\mathcal{Q}_{\text{next}} = \{q \in \mathcal{Q}_{\text{cand}} \mid \varphi(q, o) > \tau\}. \tag{12}$$

此筛选确保在第 $t+1$ 次迭代中，仅扩展估计效用较高的候选项。

综合与元分析。 达到迭代上限 T 或收敛后，智能体调用 Synthesize(·)，聚合探测元组（q、o、$\varphi(q, o)$），并生成定性分析 $\mathcal{A}$。由此产生的分布感知上下文有助于减少检索噪声，并改善后续决策。

6. 实验

6.1. 设置

基准测试。 为通过实证验证所提出框架的有效性，我们在四个具有代表性的基准测试上评估了我们的方法：**(1)Browsecomp**（Wei 等人，2025）旨在评估智能体通过复杂的网页浏览和推理策略检索隐蔽、多方面信息的能力。**(2)Browsecomp-zh**(Zhou 等人，2025）是 Browsecomp 的扩展，专注于中文网页内容。**(3) GAIA**(Mialon 等人，2023）侧重于多模态推理和工具使用能力。遵循先前的工作（Li 等人，2025d；Wu 等人，2025；Li 等人，2025年），我们仅使用纯文本验证划分中的103个样例子集。**(4)xbench-ds**（Chen等人，2025年）是一个复杂、面向任务的基准测试，旨在评估人工智能智能体自主调用工具的能力，尤其关注穷尽信息检索和多步搜索推理。

基线智能体框架。 我们采用 MIROFLOW（Team，2025年），这是一个复杂的智能体框架，专为结构化任务分解和多工具编排而设计。MIROFLOW 提供了一个模块化环境，使智能体能够在长程任务中保持较高的逻辑一致性，因此特别适用于满足目标基准测试所需的复杂检索和推理要求。在实验表格中，我们将方法与直接推理基线（例如 Qwen-2.5-32B/72B（Team，2025b）、GPT-4o/GPT-4.1（Achiam 等人，2023）、QwQ-32B、o4-mini）以及具有代表性的智能体系统/开源搜索智能体（例如 WebSailor（Li 等人，2025年）、Search-o1（Li 等人，2025c）、WebThinker（Li 等人，2025d）、ASearcher-Web（Gao 等人，2025）、WebDancer（Wu 等人，2025））进行比较，并在条件允许时纳入闭源智能体。

模型。 为评估我们的框架在不同模型架构和对齐范式下的通用性与稳健性，我们选择两个最先进的大语言模型作为主代理：（1）**MiroThinker-v1.0-30B**（Team 等，2025）是一个开源模型，基于通义千问3-30B-A3B（Team，2025b）基础架构初始化而成。通过针对性微调和专门的后训练，该模型旨在智能体推理和迭代验证方面表现出色。MiroThinker-v1.0-30B 针对与 Miroflow 智能体框架的无缝集成和架构对齐进行了战略性优化。（2）**GPT-5-mini**（Singh 等，2025）是来自 Ope-nAI 的最先进闭源模型，作为通用推理引擎运行。在我们的实验中，GPT-5-mini 直接集成到该框架中，未进行领域特定适配。详细的实验设置及流水线中使用的提示词见附录 A 和附录 B。

6.2. 主要结果

表 1 总结了统一智能体框架下的主要结果：我们仅替换探测模块，同时保持工作流程的其余部分不变。在两种骨干模型上，集成 WeDAS 均能持续提升搜索密集型基准测试上的 pass@3，这表明所提出的分布感知探测在很大程度上与模型无关。

表 1 表明，WeDAS 通常能够提升性能，但提升幅度因配置而异。MiroThinker-v1.0-30B 在噪声较多的多语言环境（如 BrowseComp-zh）中表现出色，尽管其在 XBench-DS 上有所退步。总体而言，表 2表明，WeDAS 显著增强了开源智能体的能力，为闭源系统提供了具有竞争力的替代方案。

6.3. 分析

为进一步理解为何WeDAS 能缓解信息与噪声之间的瓶颈，我们进行了细粒度的

表1. 多个框架和基础模型的性能比较。结果采用 **pass@3** 指标进行评估。我们的方法（WeDAS）集成到不同架构中，以展示其即插即用的有效性。粗体数值表示各模型类别中的最佳性能。缩写： BC 表示 BrowseComp，BC-zh 表示 BrowseComp-zh，XBD 表示 XBench-DS，Miro-30B 表示 MiroThinker-v1.0-30B。

框架 模型		BrowseComp	BrowseComp-zh	GAIA	XBench-DS
WEBSAILOR	WEBSAILOR-32B	16.05	39.28	64.72	68.00
	WEBSAILOR-72B	18.96	42.14	66.85	69.00
MIROFLOW	MIRO-30B	33.00	53.00	74.76	**89.00**
	GPT-5-mini	29.00	44.00	73.79	61.00
WeDAS	MIRO-30B	**35.00**	**58.00**	**75.73**	86.00
	GPT-5-mini	30.00	43.00	74.76	70.00

分析，通过测量生成的查询与检索结果之间的查询–观测对齐程度，对 GAIA 进行了研究。

如表3所示，WeDAS 在所有相似度维度（TF-IDF、Jac- card 和归一化莱文斯坦相似度）上始终优于基线。值得注意的是，这种改进具有普遍性，在成功和失败的试验中均有所体现。在成功案例中，我们的方法在所有指标上都取得了显著提升，表明迭代式探测机制能够有效引导智能体生成与实际网页内容分布在语义上更加一致的查询。更为关键的是，失败案例中的表现进一步有力地证明了我们方法的稳健性。虽然基线的对齐分数在失败试验中有所下降，但 WeDAS 在三个指标上仍保持相对较高的分数。这种差异表明，即使未能检索到最终答案，WeDAS 仍能使智能体保持对环境的连贯认知，避免其偏离到无关的噪声中。这些结果从实证角度支持了我们的假设：WeDAS 能够根据来自网页响应的实时反馈，动态校准查询的粒度，从而有效弥合这一差距。

6.4. 消融实验

本节探讨 WeDAS 对最大探测迭代次数 T 的敏感性，该参数决定了分布感知探测过程的深度。除非另有说明，我们使用 GPT-5-mini 作为智能体骨干模型，并在 GAIA 基准测试上进行消融实验。我们改变 $T \in \{0, 1, 2\}$，同时保持分数阈值 τ 不变。

总体而言，与不进行探测（$T = 0$）相比，启用探测（$T \geq 1$）能够提升性能，这支持了我们的观点：分布感知探测有助于智能体校准查询粒度并减少检索噪声。将T 从 1 增至 2 后，pass@3 表现相近，但 pass@1 略有下降，这表明在固定交互条件下，收益会逐渐递减——

表2. 四个基准上的 Pass@1 准确率（%）。我们将提出的框架与直接推理基线方法以及具有代表性的开源/闭源智能体系统进行比较。粗体表示每列中的最佳表现。缩写： BC 代表 BrowseComp，BC-zh 代表 BrowseComp-zh，XBD 代表 XBench-DS，Miro-30B 代表 MiroThinker-v1.0-30B。

类别 模型 / 框架		BrowseComp	BrowseComp-zh	GAIA	XBench-DS
直接推理	QWEN-2.5-32B	0.6	3.9	13.6	8.7
	QWEN-2.5-72B	0.6	7.0	14.6	12.7
	GPT-4O	0.6	6.2	17.5	18.0
	GPT-4.1	1.5	14.4	22.3	17.0
	QWQ-32B	0.5	10.0	22.3	10.7
	O4-MINI	**6.1**	**15.2**	**33.3**	**22.3**
闭源-来源	GROK-DEEPRESEARCH	-	12.9	-	52.0
	豆包深度思考	-	26.0	-	50.0
	OpenAI 深度研究	**51.5**	**42.9**	**67.4**	-
开源-来源	SEARCH-O1-32B	2.8	17.9	39.8	25.0
	WEBTHINKER-32B-RL	2.8	7.3	48.5	24.0
	ASEARCHER-WEB-32B	5.2	15.6	52.8	42.1
	WEBDANCER-QWQ-32B	3.8	18.0	51.5	38.3
	WEBSAILOR-32B	10.50	25.50	53.20	53.30
	WEBSAILOR-72B	12.00	30.10	55.40	55.00
	GPT-5-mini +MIROFLOW	15.00	28.00	51.46	38.00
	GPT-5-MINI +WEDAS	17.00	25.00	57.28	44.00
	MIRO-30B + MIROFLOW	17.00	34.00	63.11	**73.00**
	MIRO-30B + WeDAS	**26.00**	**41.00**	**66.99**	72.00

表3. 基线与 WeDAS 之间查询—结果对齐的不同指标比较。

方法	结果	TF-IDF ↑	Jaccard ↑	NLS ↑
基线	成功	0.401	0.331	0.148
	失败	0.269	0.238	0.104
WeDAS	**成功**	**0.431**	**0.364**	**0.165**
	失败	**0.292**	**0.241**	**0.108**

动作预算。

7. 结论

我们从推理驱动的查询生成与由搜索引擎索引隐式塑造的网页内容分布之间的结构失配这一视角重新审视深度研究，并提出 WeDAS 作为一种即插即用框架，通过迭代式少样本探测和元评估信号赋予智能体分布感知能力。该框架的核心是查询-结果对齐分数（QRAS），它提供可操作的多维评估，衡量检索结果与智能体意图的对齐程度；实验表明，这种分布感知设计能够通过缓解信息噪声瓶颈，在许多场景中提升性能，不过收益可能因模型和任务而异。

表 4. 使用不同最大探测迭代次数在 GAIA 上进行的消融 T（以 GPT-5-mini 作为智能体模型）。

T	pass@1	pass@3
0	49.51	72.82
1	**57.28**	**74.76**
2	54.37	**74.76**

影响声明

本文所呈现工作的目标是通过提升自主深度研究智能体的效率与准确性，推动机器学习领域的发展。通过将智能体推理与网页内容的结构分布对齐，我们的框架降低了大规模信息检索所带来的计算开销与噪声。尽管这些进展提升了人工智能在学术与专业研究中的效用，但也要求我们持续警惕自动化系统传播或汇聚有偏网页内容的可能性。除此之外，我们的工作还可能带来许多潜在的社会影响，但我们认为其中没有任何一项必须在此特别强调。

参考文献

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., 等。Gpt-4 技术报告。*arXiv* 预印本 *arXiv:2303.08774*，2023。

Chen, K., Ren, Y., Liu, Y., Hu, X., Tian, H., Xie, T., Liu, F., Zhang, H., Liu, H., Gong, Y., 等。xbench：通过职业对齐的真实世界评测追踪智能体生产力扩展。*arXiv* 预印本 *arXiv:2506.13651*，2025年。

Chen, M., Sun, L., Li, T., Sun, H., Zhou, Y., Zhu, C., Wang, H., Pan, J. Z., Zhang, W., Chen, H., 等。通过强化学习让大语言模型借助搜索进行推理学习。*arXiv* 预印本 *arXiv:2503.19470*，2025b。

Chen, Z., Liu, K., Wang, Q., Liu, J., Zhang, W., Chen, K. 和 Zhao, F. MindSearch：模仿人类思维，激发深度人工智能搜索器。*arXiv* 预印本 *arXiv:2407.20183*，2024年。

Gao, J., Fu, W., Xie, M., Xu, S., He, C., Mei, Z., Zhu, B. 和 Wu, Y. 超越十轮：借助大规模异步强化学习，解锁长程智能体搜索。*arXiv* 预印本 *arXiv:2508.07976*，2025。

Guo, D., Yang, D., Zhang, H., Song, J., Zhang, R., Xu, R., Zhu, Q., Ma, S., Wang, P., Bi, X., 等。DeepSeek-R1：通过强化学习激励大语言模型的推理能力。*arXiv* 预印本 *arXiv:2501.12948*，2025年。

Huang, Y. 和 Huang, J. 大语言模型检索增强文本生成综述。*arXiv* 预印本 *arXiv:2404.10981*，2024年。

Huang, Y., Chen, Y., Zhang, H., Li, K., Zhou, H., Fang, M., Yang, L., Li, X., Shang, L., Xu, S., 等。深度研究智能体：系统性考察与路线图。*arXiv* 预印本 *arXiv:2506.18096*，2025。´Jaccard, P. 阿尔卑斯山和汝拉山脉一部分地区植物分布的比较研究。沃州自然科学学会公报，37:547–579，1901年。

Jin, B., Zeng, H., Yue, Z., Yoon, J., Arik, S., Wang, D., Zamani, H. 和 Han, J. Search-R1：通过强化学习训练大语言模型进行推理并利用搜索引擎。*arXiv* 预印本 *arXiv:2503.09516*，2025。

Jones, K. S., Walker, S. 和 Robertson, S. E. 信息检索的概率模型：开发与比较实验：第 2 部分。信息处理与管理，36(6):809–840，2000年。

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J. 和 Amodei, D. 神经语言模型的扩展定律。*arXiv* 预印本 *arXiv:2001.08361*，2020年。

Ke, Z., Jiao, F., Ming, Y., Nguyen, X.-P., Xu, A., Long, D. X., Li, M., Qin, C., Wang, P., Savarese, S. 等。大语言模型推理前沿综述：推理扩展、推理学习与智能体系统。*arXiv* 预印本 *arXiv:2504.09037*，2025。

Langley, P. 撰写机器学习论文。载于 Langley, P.（编），第 *17* 届国际机器学习会议论文集（*ICML 2000*），第 1207–1216 页，斯坦福，加利福尼亚州，2000年。Morgan Kaufmann。

Lcvenshtcin, V. 能够纠正删除、插入和反转的二进制码。在《苏联物理学文献》，第 10 卷，1966年。

Li, K., Zhang, Z., Yin, H., Zhang, L., Ou, L., Wu, J., Yin, W., Li, B., Tao, Z., Wang, X., 等。WebSailor：探索超越人类推理能力的网络智能体。*arXiv* 预印本 *arXiv:2507.02592*，2025年。

Li, M., Lv, X., Zou, J., Chen, T., Zhang, C., An, S., Nie, E. 和 Zhou, G. 预训练模型和大语言模型时代的查询扩展：全面综述。*arXiv* 预印本 *arXiv:2509.07794*，2025b。

Li, X.、Dong, G.、Jin, J.、Zhang, Y.、Zhou, Y.、Zhu, Y.、Zhang, P.和Dou, Z. Search-o1：智能体搜索增强的大型推理模型。*arXiv*预印本 *arXiv:2501.05366*，2025c。

Li, X.、Jin, J.、Dong, G.、Qian, H.、Wu, Y.、Wen, J.-R.、Zhu, Y.和Dou, Z. WebThinker：赋予大型推理模型深度研究能力。*arXiv*预印本 *arXiv:2504.21776*，2025d。

Li, X., Jin, J., Zhou, Y., Zhang, Y., Zhang, P., Zhu, Y., 和 Dou, Z. 从匹配到生成：生成式信息检索综述。*ACM* 信息系统汇刊, 43(3):1–62, 2025e.

Liu, B., Li, X., Zhang, J., Wang, J., He, T., Hong, S., Liu, H., Zhang, S., Song, K., Zhu, K., 等. 基础智能体的进展与挑战：从受脑启发的智能到进化、协作与安全系统。*arXiv* 预印本 *arXiv:2504.01990*, 2025.

Luhn, H. P. 文献信息机械化编码与搜索的统计方法。*IBM* 研究与开发杂志, 1(4):309–317, 1957.

Luo, J., Zhang, W., Yuan, Y., Zhao, Y., Yang, J., Gu, Y., Wu, B., Chen, B., Qiao, Z., Long, Q., 等. 大语言模型智能体：方法论、应用与挑战综述。*arXiv* 预印本 *arXiv:2503.21460*, 2025.

Ma, X., Gong, Y., He, P., Zhao, H., and Duan, N. 检索增强大语言模型中的查询改写。载于*2023*年自然语言处理实证方法会议论文集，第5303–5315页，2023年。

Mialon, G., Fourrier, C., Wolf, T., LeCun, Y., and Scialom, T. GAIA：通用人工智能助手的基准测试。载于第十二届国际学习表征会议，2023年。

Qian, H. and Liu, Z. 知识的气息：运用信息觅食优化搜索增强推理。*arXiv* 预印本 *arXiv:2505.09316*，2025年。

Qu, C., Dai, S., Wei, X., Cai, H., Wang, S., Yin, D., Xu, J., and Wen, J.-R. 大语言模型的工具学习：一项综述。计算机科学前沿，19(8):198343，2025年。

Singh, A., Fry, A., Perelman, A., Tart, A., Ganesh, A., El-Kishky, A., McLaughlin, A., Low, A., Ostrow, A., Ananthram, A. 等人。OpenAI GPT-5 系统说明卡。*arXiv* 预印本 *arXiv:2601.03267*，2025。

Team, M., Bai, S., Bing, L., Chen, C., Chen, G., Chen, Y., Chen, Z., Chen, Z., Dong, X. 等人。Mirothinker：通过模型、上下文与交互式规模化推动开源研究智能体的性能边界。*arXiv* 预印本 *arXiv:2511.11793*，2025。

团队，M. A. Miroflow：一个高性能的开源研究智能体框架。https://github.com/MiroMindAI/MiroFlow，2025年。

团队，Q. 通义千问3技术报告，2025b。URL https://arxiv.org/abs/2505.09388。

Thakur, N., Reimers, N., R¨uckl´e, A., Srivastava, A. 与 Gurevych, I.。BEIR：用于信息检索模型零样本评估的异构基准。*arXiv* 预印本 *arXiv:2104.08663*，2021。

Wei, J., Sun, Z., Papay, S., McKinney, S., Han, J., Fulford, I., Chung, H. W., Passos, A. T., Fedus, W., 和 Glaese, A. Browsecomp：面向浏览智能体的简单但具有挑战性的基准测试。*arXiv* 预印本 *arXiv:2504.12516*，2025。

Wu, J., Li, B., Fang, R., Yin, W., Zhang, L., Tao, Z., Zhang, D., Xi, Z., Fu, G., Jiang, Y., 等。WebDancer：迈向自主信息搜寻智能体。*arXiv* 预印本 *arXiv:2505.22648*，2025。

Xi, Y., Lin, J., Xiao, Y., Zhou, Z., Shan, R., Gao, T., Zhu, J., Liu, W., Yu, Y., 和 Zhang, W. 大语言模型深度搜索智能体综述：范式、优化、评估与挑战。*arXiv* 预印本 *arXiv:2508.05668*，2025。

Zhao, S., Yu, T., Xu, A., Singh, J., Shukla, A., and Akkiraju, R. ParallelSearch：通过强化学习训练大语言模型拆解查询，并并行搜索子查询。*arXiv* 预印本 *arXiv:2508.09303*，2025。

Zhou, P., Leon, B., Ying, X., Zhang, C., Shao, Y., Ye, Q., Chong, D., Jin, Z., Xie, C., Cao, M., 等。BrowseComp-zh：面向中文大语言模型网页浏览能力的基准测试。*arXiv* 预印本 *arXiv:2504.19314*，2025。

Zhu, Y.、Yuan, H.、Wang, S.、Liu, J.、Liu, W.、Deng, C.、Chen, H.、Liu, Z.、Dou, Z. 和 Wen, J.-R. 面向信息检索的大语言模型：综述。*ACM* 信息系统汇刊，44(1):1–54，2025。

附录

A. 实验设置

由于计算成本，我们并未在完整的 BrowseComp 和 BrowseComp-zh 测试集上进行评估；而是从每个基准测试中均匀采样 100 个实例，并在这些子集上报告结果。对于所有实验，我们采用统一的解码配置，温度为 $= 1.0$, top-$p = 0.95$，min-$p = 0.0$，top-$k = -1$，最大词元数为 $= 16384$。主代理执行的最大交互轮次设为 600，每轮最多进行 10 次工具调用。对于我们的方法，元评估器（QRAS）和探测查询生成模块均使用 GPT-5-mini 进行实例化。所有实验均在单块 NVIDIA RTX PRO 6000 GPU 上运行。

B. WeDAS 中使用的提示词

为促进可复现性，我们提供了框架中使用的确切系统指令（提示词） .

B.1. 探测查询生成

候选生成算子 Γ_{gen} 使用以下提示词，指示大语言模型扮演“搜索关键词专家”，将对当前搜索结果的分析转化为一组策略性、基于关键词的后续查询。

你是一名搜索关键词专家。你的任务是分析一组“查询—结果”对，并确定下一步的策略性搜索关键词。

任务
评估结果。如果结果揭示了新信息，则生成 2-5 个后续查询。如果结果无果或问题已得到完整解答，则返回空列表。

严格查询约束

- **必须简洁**：每个查询必须是简短短语或由 1-3 个关键词组成的列表。
- **不得使用句子**：请勿生成自然语言问题（例如“我该如何.....”）。
- **关键词重点**：使用结果中出现的具体实体、技术术语、错误代码或唯一标识符。
- **字数限制**：每条查询最多 6 个词。

决策逻辑

1. **修剪（返回 []）**：
   - 结果无关，质量低下，或属于“404/未找到”页面。
   - 该结果已完全满足信息需求。
2. **扩展（返回 [简短关键词]）**：
   - 该结果提到了一个需要进一步调查的特定库、应用程序接口或概念
   - 该结果使用专业术语。

输出格式
返回一个 JSON 对象：

```
{
    "分析": "选择这些关键词的简要说明。",
    "派生查询": [
        "关键词1 关键词2",
        "specific term site:github.com"
    ]
}
```

图 *3.* 候选生成算子的系统指令（Γ_{gen}）。

B.2. 对齐评分与元评估

提示词 b 以下内容由元评估器 $\mathcal{M}_\theta$ 用于计算查询-结果对齐分数（QRAS).

角色
你是一名内容相关性评估器。请评估多个搜索结果集与各自搜索查询的匹配程度。

输入
你将收到一个项目列表。每个项目都包含一个“查询”和“结果”。

评分维度（每项 0-10 分）

1.****主题相关性****：结果是否确实讨论了相关主题？ 2.****信息密度****：摘要是否包含潜在答案（数字、日期、分数），还是只有笼统的信息？ 3.****噪声水平****：结果集中有多少内容无关（广告、SEO 垃圾内容）？（0 表示全是噪声，10 表示完全没有噪声）

输出格式（仅限 JSON）
返回一个 JSON 对象，包含一个 "evaluations" 的键，其中包含一个对象列表：

```
{
    "evaluations": [
        {
            "查询": <查询>,
            "整体相关性评分": <0-10>,
            "维度评分": {
                "主题相关性": <0-10>,
                "信息密度": <0-10>,
                "噪声水平": <0-10>
            },
            "分析": "<简短摘要>"
        },
        ...
    ]
}
```

图4. 元评估器的系统指令（$\mathcal{M}_\theta$）。