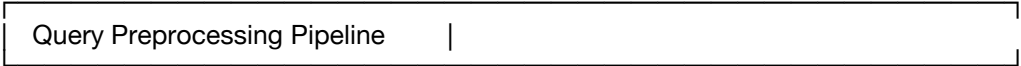


Raw Query



1. Normalization
2. Entity Extraction
3. Intent Classification
4. Query Expansion
5. Context Injection



Enhanced Query Package → Router

Stage 1A: Normalization

****What it does:****

- Spelling correction ("Vizanagram" → "Vizianagaram")
- Acronym expansion ("NEP" → "National Education Policy 2020")
- Synonym mapping ("PTR" = "Pupil-Teacher Ratio" = "Student-Teacher Ratio")
- Date parsing ("last year" → 2024, "2016-17" → academic year normalization)

****Example:****

Input: "dropout rate in vizanagram last yr"

Output: {
 normalized: "dropout rate in Vizianagaram 2024",
 corrections: ["vizanagram→Vizianagaram", "yr→year"],
 expansions: ["dropout rate = students leaving school"]
}

****Implementation approach:****

- Build domain dictionary: 500 education terms + synonyms
- AP-specific gazetteer: 26 districts, 100+ schemes, 50+ acts
- Regex patterns for dates, GO numbers, section references

Stage 1B: Entity Extraction (NER)

****Extract & Tag:****

Entity Type	Examples	Metadata
	----- ----- -----	
District	Vizianagaram, Anantapur	`admin_level: district`
Scheme	Nadu-Nedu, Amma Vodi, Jagananna Gorumudda	`scheme_category: infrastructure/finance`
Metric	PTR, GER, dropout rate, NAS scores	`metric_type: input/output/outcome`
Legal Reference	Section 12, RTE Act, Article 21A	`doc_type: Act/Rule, section_id`
GO Number	GO MS 34/2023	`go_pattern: detected`
Year/Period	2016-17, FY 2023-24	`temporal_type: academic/fiscal`
School Category	government, private, aided	`school_type`
Social Category	SC, ST, OBC, EWS	`student_category`

****Example:****

...

Query: "SC student dropout in Vizianagaram govt schools 2016-17"

Entities: {
 district: "Vizianagaram",
 student_category: "SC",
 metric: "dropout_rate",
 school_type: "government",
 year: "2016-17",
 admin_level: "district"
}

...

****Stage 1C: Intent Classification (Multi-Label)****

****Primary Intents:****

- `factual_lookup`: "What is Section 12?"
- `data_query`: "Show me dropout rates"
- `compliance_check`: "Are schools RTE-compliant?"
- `comparison`: "Compare AP vs national PTR"
- `historical_evolution`: "How has policy changed?"
- `gap_analysis`: "Policy vs implementation vs outcomes"
- `drafting_request`: "Create a briefing note on..."
- `inspiration_seeking`: "What do Nordic countries do for teacher training?"

****Secondary Intents (can have multiple):****

- `needs_legal_basis`: True/False
- `needs_implementation_details`: True/False
- `needs_data_evidence`: True/False
- `needs_judicial_precedent`: True/False
- `needs_global_context`: True/False
- `needs_temporal_analysis`: True/False

****Example:****

...

Query: "Are Vizianagaram schools RTE-compliant on PTR? Show legal basis and data."

Primary intent: compliance_check

Secondary intents: {
 needs_legal_basis: True,
 needs_data_evidence: True,
 needs_implementation_details: True,
 needs_global_context: False
}

...

****Stage 1D: Query Expansion (Critical for Recall)****

****Three expansion strategies:****

****1. Synonym Expansion****

...

Query: "teacher shortage"

Expanded: "teacher shortage" OR "teacher vacancy" OR "unfilled posts"
OR "teacher gap" OR "staff shortage"

...

****2. Related Concept Expansion****

...

Query: "midday meal"

Expanded: "midday meal" + related:[
"Jagananna Gorumudda",
"PM POSHAN",
"nutritional support",
"MDM scheme"

]
...

****3. Legal Context Expansion****

...

Query: "SMC powers"

Expanded: "School Management Committee" + legal_context:[
"AP School Education Management Act 1998",
"RTE Section 21",
"SMC composition",
"SMC functions"

]
...

****When to expand:****

- Always for Normal QA Mode
- Aggressively for Brainstorming Mode (3x expansion)
- Selectively for Pro Mode (only high-confidence expansions)

**Stage 1E: Context Injection (Memory Layer)**

****Add conversation context:****

...

Current query: "What about primary schools?"

Conversation history:

- User asked about "PTR in Vizianagaram" 2 messages ago
- User asked about "dropout rates" 1 message ago

Context injection:

"What about primary schools?"

→ "What about PTR in primary schools in Vizianagaram?"

...

****Add user profile context (if available):****

...

User role: District Education Officer, Vizianagaram

Context injection:

- Auto-filter to Vizianagaram district
- Prioritize implementation-focused results over theory

...

Part 2: The Three Modes Architecture

This is where your system becomes **truly intelligent**.

MODE 1: Normal QA Mode (Fast & Precise)

Use case: Quick factual queries, policy lookups, data retrieval

Characteristics:

- Single-pass retrieval
- Top-5 results only
- No query expansion beyond synonyms
- No internet search
- Answer in <3 seconds
- Citation: minimal (1-2 sources)

Query Processing:

```
Query → Normalize → Entity Extract → Intent Classify → Route to Agent(s)
↓
Agent Retrieval (Top-5, no reranking)
↓
Quick Merge
↓
Answer (2-3 sentences + citations)
```

Example:

Q: "What is Section 12 of RTE?"

A: Section 12(1)(c) of the RTE Act mandates that private unaided schools reserve 25% seats for economically weaker sections and disadvantaged groups. [RTE Act 2009, Section 12, Clause 1(c)]

Response time: 1.8s

Optimization:

- Cache frequent queries
- Pre-compute embeddings
- No cross-agent orchestration unless absolutely needed

MODE 2: Brainstorming Mode (Broad & Exploratory)

Use case: Policy research, comparative analysis, exploring options

Characteristics:

- Multi-pass retrieval with expansion
- Top-20 results, diverse ranking
- Aggressive query expansion (3x)
- **Internet search enabled** (global policies, research papers)
- Answer in <15 seconds
- Citation: comprehensive (5-10 sources, mix of internal + external)
- Shows multiple perspectives, alternatives, trade-offs

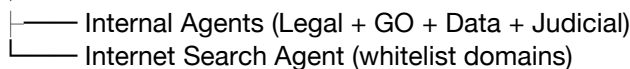
****Query Processing:****

...

Query → Full Enhancement Pipeline → Expand Aggressively



Route to Multiple Agents (parallel)



Retrieve Top-20 per agent



Rerank + Diversify (avoid redundancy, maximize coverage)



Synthesize Multi-Perspective Answer



Show: AP Policy + Other States + Global Examples + Research Evidence

...

****Example:****

...

Q: "How can we improve teacher training in AP?"

Answer Structure:

1. Current AP Policy:

- SCERT AP guidelines on teacher training [cite]
- Budget allocation: ₹X cr (FY 2024) [cite]

2. Legal Framework:

- NCTE norms for teacher education [cite]
- AP RTE Rules on in-service training [cite]

3. Implementation Gap:

- Only 60% teachers received training (UDISE 2023) [cite]
- District-wise variation: [data]

4. Other Indian States:

- Kerala's Samagra Shiksha approach [cite external]
- Karnataka's digital teacher training platform [cite external]

5. Global Inspiration:

- Finland's teacher development model [cite: OECD report]
- Singapore's teacher growth framework [cite: research paper]

6. Research Evidence:

- Meta-analysis on effective teacher training [cite: academic paper]

7. Potential Options:

[Synthesized recommendations based on all above]

Sources: 12 internal (AP docs) + 8 external (research + global)

Response time: 12s

...

****Internet Search Strategy for Brainstorming:****

****Trigger conditions:****

- Query contains: "how to", "best practices", "global", "international", "compare", "other countries"
- Intent: `inspiration_seeking` or `comparison`
- Mode: Brainstorming

****Search domains (whitelist):****

Domain Category	Examples	Use Case
International Orgs	unesco.org, worldbank.org, oecd.org	Global benchmarks, frameworks
Academic	.edu, scholar.google.com, researchgate.net	Evidence-based practices
Gov (other countries)	gov.uk, education.gov.au	Comparative policies
Think Tanks	brookings.edu, rand.org	Policy analysis
Indian States	*.gov.in (other states)	Interstate comparison

****Search process:****

...

1. Rewrite query for web search:

"improve teacher training AP"

→ "effective teacher professional development policies education"

2. Add constraints:

site:unesco.org OR site:oecd.org OR site:worldbank.org

filetype:pdf (prefer authoritative reports)

3. Fetch top-10 results

4. For each result:

- Download and snapshot
- Extract text
- Create chunks
- Embed
- Tag as: source_type="external", domain, date_accessed

5. Merge with internal results:

- Weight: 70% internal, 30% external
- Label clearly in answer: [External: OECD Report 2023]

...

**MODE 3: Pro Mode / DeepThink Mode (Comprehensive & Rigorous)**

****Use case:**** Policy drafting, legal analysis, compliance audits, reports for senior leadership

****Characteristics:****

- Multi-pass iterative retrieval
- Top-50 results, exhaustive coverage
- Query expansion + relation traversal + temporal analysis
- Internet search for validation (not inspiration)
- Answer in <60 seconds (worth the wait)
- Citation: exhaustive (15-30 sources, full provenance)
- Shows reasoning chain, confidence levels, gaps, conflicts

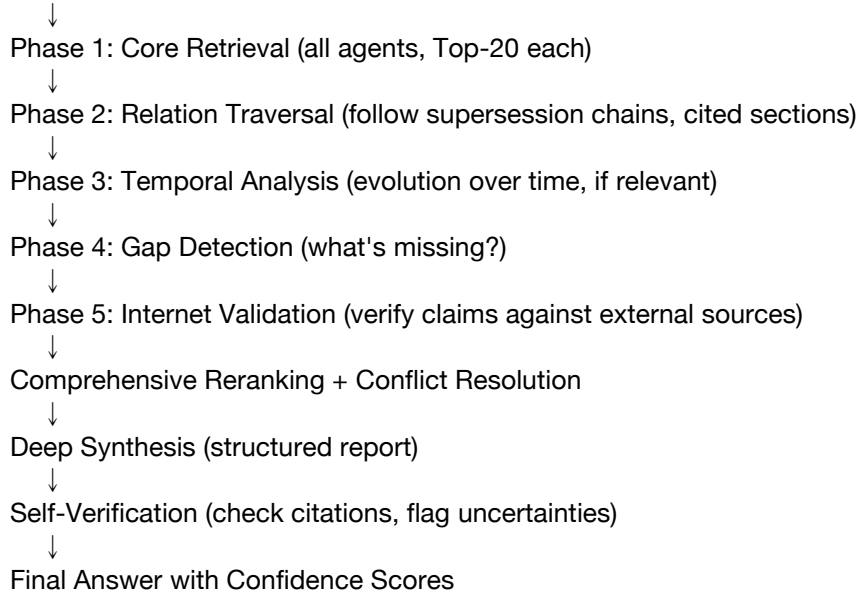
****Query Processing:****

...

Query → Full Enhancement → Intent Deep Analysis



Multi-Phase Retrieval:



Example:

...

Q: "Create a comprehensive briefing note on RTE Section 12(1)(c) implementation in AP: legal basis, GOs, judicial precedents, budget, outcomes, gaps, and recommendations."

Answer Structure:

EXECUTIVE SUMMARY

[2-3 paragraph TL;DR with key findings and confidence: 92%]

1. LEGAL FRAMEWORK (Confidence: 98%)

- 1.1 Constitutional Basis
 - Article 21A [cite]
- 1.2 RTE Act Provisions
 - Section 12(1)(c) text and analysis [cite]
- 1.3 AP RTE Rules
 - Rule X on admission process [cite]
- 1.4 Amendments & Clarifications
 - Timeline of changes [5 citations]

2. ADMINISTRATIVE IMPLEMENTATION (Confidence: 85%)

- 2.1 Current GO
 - GO MS 67/2018 (latest) [cite]
- 2.2 Supersession Chain
 - GO 45/2010 → GO 23/2015 → GO 67/2018 [relation graph]
- 2.3 Departmental Circulars
 - [4 citations]

3. JUDICIAL INTERPRETATION (Confidence: 90%)

- 3.1 Supreme Court
 - Society for Unaided Schools vs UOI (2012) [cite + summary]
- 3.2 AP High Court
 - [3 key cases with holdings]
- 3.3 Pending Litigations
 - [2 cases, status]

4. FINANCIAL PROVISIONS (Confidence: 75%)

4.1 Budget Allocation

- FY 2023-24: ₹X cr [cite: Budget docs]

4.2 Reimbursement Status

- ₹Y cr disbursed, Z% pending [cite: CAG report]

4.3 Per-Student Cost

- [calculation + cite]

5. IMPLEMENTATION OUTCOMES (Confidence: 80%)

5.1 Enrollment Data

- X students enrolled under 12(1)(c) in 2022-23 [cite: UDISE+]
- District-wise breakdown [table]

5.2 School Compliance

- Y% private schools compliant [cite]

5.3 Grievances

- Z complaints on Meekosam [cite: if available]

6. GAP ANALYSIS (Confidence: 70%)

[Identified gaps between law → implementation → outcomes]

- Legal gap: [if any]
- Implementation gap: Reimbursement delays [evidence]
- Data gap: No school-level tracking

7. COMPARATIVE PERSPECTIVE (External sources)

7.1 Other States

- Kerala's approach [cite: state doc]
- Tamil Nadu's model [cite]

7.2 International

- Chile's voucher system [cite: research paper]
- [Clearly marked as inspiration, not prescription]

8. RECOMMENDATIONS

[Evidence-based, cited recommendations]

ANNEXURES

- Full text of Section 12(1)(c)
- GO 67/2018 excerpt
- Budget allocation table
- District-wise enrollment data

CITATIONS (28 total)

- 18 internal (Acts, GOs, UDISE, CAG)
- 6 judicial
- 4 external (research, comparative)

CONFIDENCE ASSESSMENT

- Overall: 85%
- Legal framework: 98% (complete)
- Implementation: 85% (good coverage)
- Outcomes: 80% (data limitations)
- Financial: 75% (partial budget data)

KNOWN LIMITATIONS

- No data on private school-level compliance
- Reimbursement data only until FY 2022-23
- Pending litigations outcome not known

Response time: 45s
Chunks retrieved: 127
Relations traversed: 15
External sources: 4
...

****Internet Search Strategy for Pro Mode:****

****Purpose:**** Validation and triangulation, not inspiration

****Process:****

1. ****After internal retrieval****, identify key claims
2. ****Verify claims**** against external authoritative sources
3. ****Flag discrepancies****: "Internal doc says X, but OECD report says Y"
4. ****Add context****: "AP policy aligns with WHO guidelines on..."
5. ****Citation standards****: Prefer original sources over secondary

****Search for:****

- Academic papers validating policy approaches
- International standards (UNESCO, UNICEF guidelines)
- Comparative government reports
- Research on similar policies in other contexts

****Do NOT search for:****

- Basic facts already in corpus
- Opinion pieces or blogs
- Unverified sources

****Part 3: The Bridge Table Architecture (Critical for Pro Mode)****

The bridge table is what makes ****intelligent linking**** possible across verticals.

****Bridge Table Purpose:****

Connect policy topics → legal basis → implementation → evidence → outcomes

****Schema:****

...

policy_bridge:

- topic_id (PK)
 - topic_name (e.g., "RTE Section 12(1)(c) Implementation")
 - legal_doc_ids[] (array: doc IDs of Acts, Rules)
 - go_doc_ids[] (array: implementing GOs)
 - judicial_doc_ids[] (array: relevant cases)
 - metric_codes[] (array: related metrics from metrics table)
 - scheme_names[] (array: related schemes)
 - keywords[] (array: for retrieval)
 - confidence_score (how complete is this mapping?)
 - last_updated
 - curator (human/system)
- ...

****How Bridge Table Enhances Each Mode:****

****Normal QA Mode:****

- If query matches a bridge topic exactly → fast path

- Return pre-linked sources without orchestration
- Example: "SMC powers" → fetch SMC bridge → instant answer

****Brainstorming Mode:****

- Use bridge as starting point
- Expand beyond bridge to find related topics
- Example: Query about "SMC" → bridge to SMC → also fetch community participation topics

****Pro Mode:****

- Bridge is the skeleton, flesh it out with exhaustive retrieval
- Validate bridge completeness (are all links still valid?)
- Detect gaps (is a new GO missing from bridge?)

**Bridge Table Population Strategy:**

****Phase 1: Manual (Week 1-2)****

- Curator (you + policy expert) manually creates 20-30 core topic bridges
- Topics: RTE 12(1)(c), SMC, PTR norms, MDM, Nadu-Nedu, Amma Vodi, teacher recruitment, etc.

****Phase 2: Semi-Automated (Week 3-4)****

- System suggests bridges based on relation extraction
- Human reviews and approves
- Example: System detects GO 45/2023 cites Section 12 → suggests adding to RTE 12(1)(c) bridge

****Phase 3: Automated Maintenance (Month 2+)****

- Daily: scan new GOs for mentions of bridged topics
- Weekly: re-validate all bridge links (are GOs superseded? are metrics updated?)
- Monthly: curator review of auto-suggested updates

**Bridge Table as Metadata for Retrieval:**

When routing query to agents, ****check bridge table first:****

Query: "RTE 12(1)(c) implementation status"

1. Check bridge table: topic="RTE Section 12(1)(c)"

→ Found bridge with:

- legal_doc_ids: [doc_123, doc_456]
- go_doc_ids: [doc_789, doc_890]
- metric_codes: ['ews_enrollment', 'reimbursement_rate']

2. Route with metadata:

Legal Agent: focus on doc_123, doc_456

GO Agent: focus on doc_789, doc_890

Data Agent: query for 'ews_enrollment', 'reimbursement_rate'

3. Result: Faster, more accurate retrieval

**Part 4: Internet Search Layer (The Inspiration Engine)**

For Brainstorming and Pro modes, the internet layer is ****critical but controlled****.

**Two-Tier Internet Strategy:**

****Tier 1: Authoritative Validation (Always)****

- Purpose: Verify claims, add credibility
- Domains: UN agencies, World Bank, OECD, peer-reviewed journals
- Usage: "According to UNESCO guidelines..."
- Labeling: [External: UNESCO 2023]

****Tier 2: Inspiration Seeking (Brainstorming only)****

- Purpose: Find innovative approaches from other contexts
- Domains: Other countries' gov sites, think tanks, case studies
- Usage: "Finland's approach to teacher training..."
- Labeling: [Inspiration: Finland Ministry of Education]

**Internet Search Workflow:**

...

1. Query Analysis



2. Determine if internet search needed:

- Intent: inspiration_seeking? → Yes
- Intent: compliance_check? → No
- Mode: Brainstorming/Pro? → Consider
- Mode: Normal QA? → No

3. If Yes → Formulate web search query:

Original: "How to improve teacher training in AP?"

Web query: "effective teacher professional development education policy"

4. Search with constraints:

site:unesco.org OR site:worldbank.org OR site:oe.cd.org OR site:.edu
filetype:pdf (prefer reports over web pages)

5. Fetch & Process:

- Download top-10 results
- Store snapshot (GCS bucket)
- Extract text
- Quality check (min length, coherence)
- Create chunks
- Embed
- Tag: source_type='external', domain, confidence_score

6. Integrate with internal results:

- Show separately: "Internal Policy" vs "Global Context"
- Weight: 70-80% internal, 20-30% external
- Never let external override internal legal requirements

...

**Internet Source Ranking (for education policy):**

****Tier S (Highest Authority):****

- UNESCO official publications
- OECD Education Reports
- World Bank Education Global Practice
- UNICEF policy briefs
- National academies (NAS, Royal Society)

****Tier A (High Quality):****

- Peer-reviewed education journals
- Government education ministries (OECD countries)
- Brookings, RAND, other reputable think tanks

- University research centers (.edu)

****Tier B (Contextual):****

- NGO reports (Pratham, Azim Premji Foundation)
- Indian state government reports
- NITI Aayog publications

****Tier C (Use with Caution):****

- News articles (for recent developments only)
- Conference papers
- Working papers (not peer-reviewed)

****Blocked:****

- Blogs, opinion sites
- Unverified sources
- Commercial ed-tech promotional content

**Citation Format for External Sources:**

Internal source:

[RTE Act 2009, Section 12, MoE India]

External source (authoritative):

[External: UNESCO Global Education Monitoring Report 2023, p. 45]

External source (inspiration):

[Inspiration: Finland Ministry of Education - Teacher Training Framework 2022]

External source (research):

[Research: Darling-Hammond et al. (2017), "Effective Teacher Professional Development", Learning Policy Institute]

**Part 5: Mode Selection Logic**

How does the system decide which mode to use?

**Option A: User Explicitly Selects**

UI: [Normal] [Brainstorming] [Pro]

User clicks → mode set

**Option B: System Auto-Suggests**

Query: "What is Section 12?"

System suggests: Normal mode (simple factual)

Query: "How to improve teacher retention?"

System suggests: Brainstorming mode (exploratory)

Query: "Create briefing note on RTE implementation"

System suggests: Pro mode (comprehensive)

**Option C: Hybrid (Recommended)**

...

Default: Normal mode

System detects complex queries → suggests upgrade:

"This query may benefit from Brainstorming mode for diverse perspectives. Switch?"

User can override anytime

...

Mode Detection Heuristics:

Query Pattern	Suggested Mode	Reasoning
"What is...", "Show me...", "Which GO..."	Normal	Factual lookup
"How to...", "Best practices...", "Improve..."	Brainstorming	Needs ideas
"Create report...", "Comprehensive analysis...", "Draft note..."	Pro	Needs depth
Contains "compare", "global", "international"	Brainstorming	Needs external context
Contains "legal opinion", "compliance audit"	Pro	Needs rigor
Short query (<5 words)	Normal	Quick answer needed
Long query (>20 words)	Pro	Complex, multi-faceted

Part 6: Query Enhancement Examples (All Modes)

Let me show how the same query gets enhanced differently:

Original Query: "teacher training"

****Normal QA Mode Enhancement:****

...

1. Normalize: "teacher training" (no change)
2. Expand: "teacher training" OR "teacher professional development" OR "in-service training"
3. Entities: None detected
4. Intent: factual_lookup
5. Route to: Legal Agent (for norms) + GO Agent (for schemes)
6. Metadata filter: doc_type IN ('Act', 'Rule', 'GO'), topic='teacher_training'
7. Retrieval: Top-5
8. Internet search: No

...

****Brainstorming Mode Enhancement:****

...

1. Normalize: "teacher training"
2. Expand Aggressively:
 - "teacher training" OR "teacher professional development" OR "continuing education"
 - PLUS related concepts: "teacher capacity building", "pedagogical training", "subject knowledge enhancement", "digital pedagogy training"
3. Entities: None specific (broaden scope)
4. Intent: inspiration_seeking
5. Route to: Legal + GO + Data (outcomes) + Internet
6. Metadata filter: Broader - include frameworks, guidelines, reports
7. Retrieval: Top-20 per agent
8. Internet search: YES
 - "effective teacher professional development policies"
 - "teacher training international best practices"
 - "teacher continuous professional development frameworks"
9. Diversification: Ensure results cover different aspects (pre-service, in-service, digital, subject-specific)

...

****Pro Mode Enhancement:****

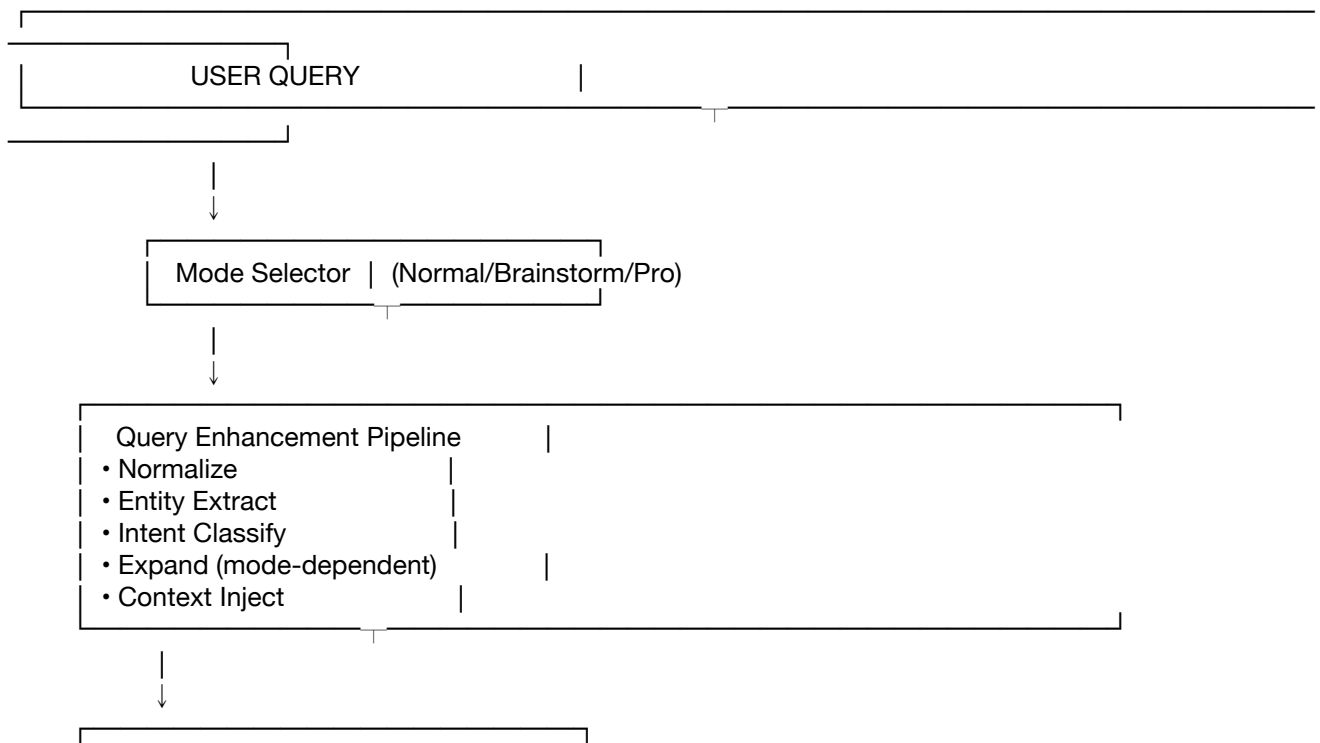
...

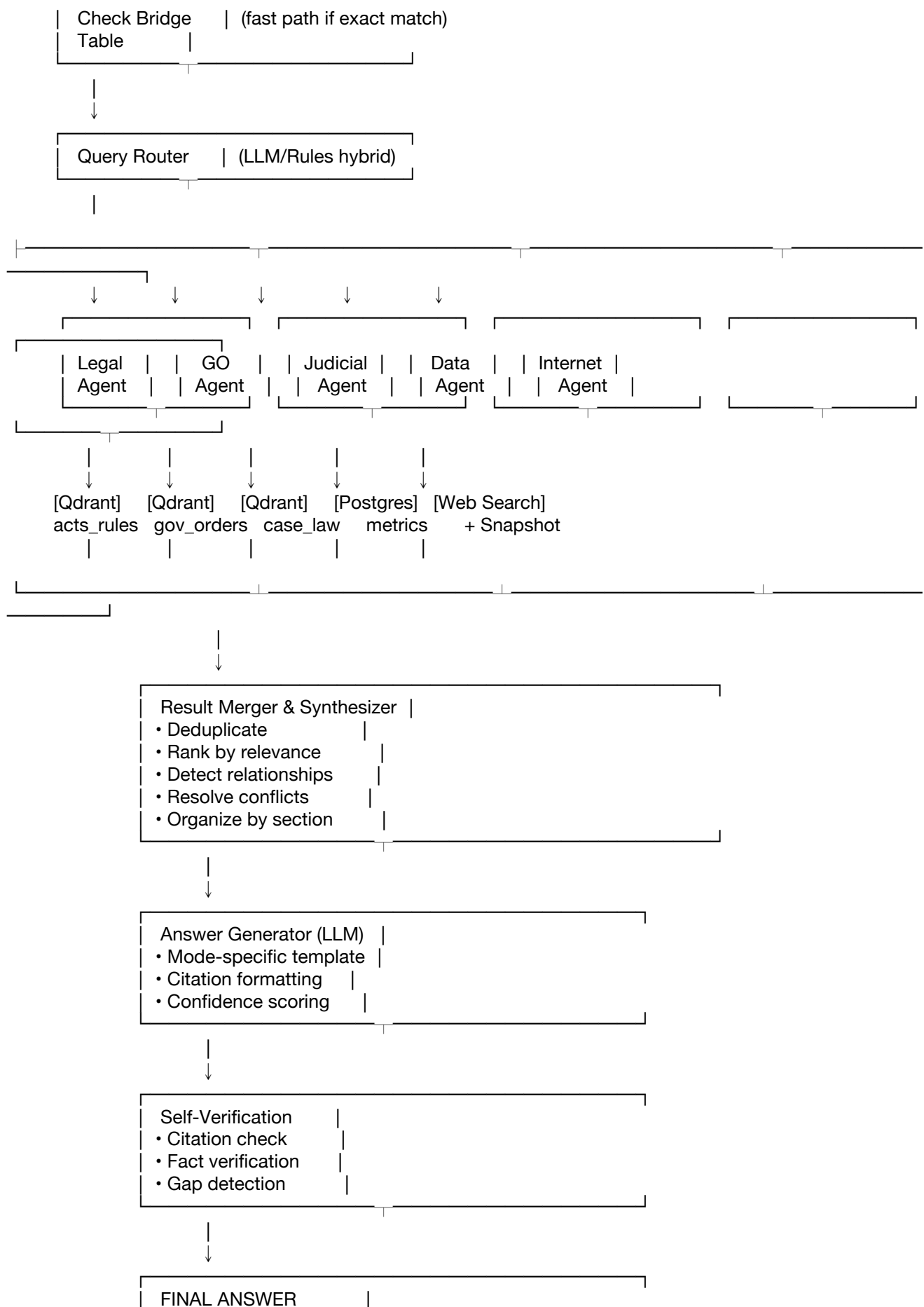
1. Normalize: "teacher training"
2. Expand Systematically:
 - Legal terms: "teacher education", "NCTE norms", "teacher qualification"
 - Implementation terms: "SCERT training", "DIETs", "in-service programs"
 - Outcome terms: "teacher competency", "training effectiveness"
3. Entities: None specific but prepare for district-level disaggregation
4. Intent: comprehensive_analysis
5. Route to: ALL agents (Legal + GO + Data + Judicial + Internet)
6. Metadata: No filters initially (exhaustive search), then organize by:
 - Legal framework
 - Implementation (GOs + schemes)
 - Budget/finance
 - Outcomes (UDISE teacher qualification data, NAS correlation)
 - Judicial (any cases on teacher training mandates)
7. Retrieval: Top-50, multi-phase:
 - Phase 1: Core docs
 - Phase 2: Related docs via bridge table
 - Phase 3: Temporal analysis (evolution of policy)
 - Phase 4: External validation
8. Internet search: YES (validation focus)
 - "teacher professional development research meta-analysis"
 - "OECD teacher training effectiveness"
 - Search for specific models: "Finland teacher training", "Singapore teacher development"
9. Relation traversal: Follow all "implements", "cites", "amends" chains
10. Gap analysis: What's mandated vs what's implemented vs outcomes

...

**Part 7: The Complete Architecture (Final Vision)**

...





+ Citations	
+ Confidence	
+ Sources Used	
+ Known Gaps	