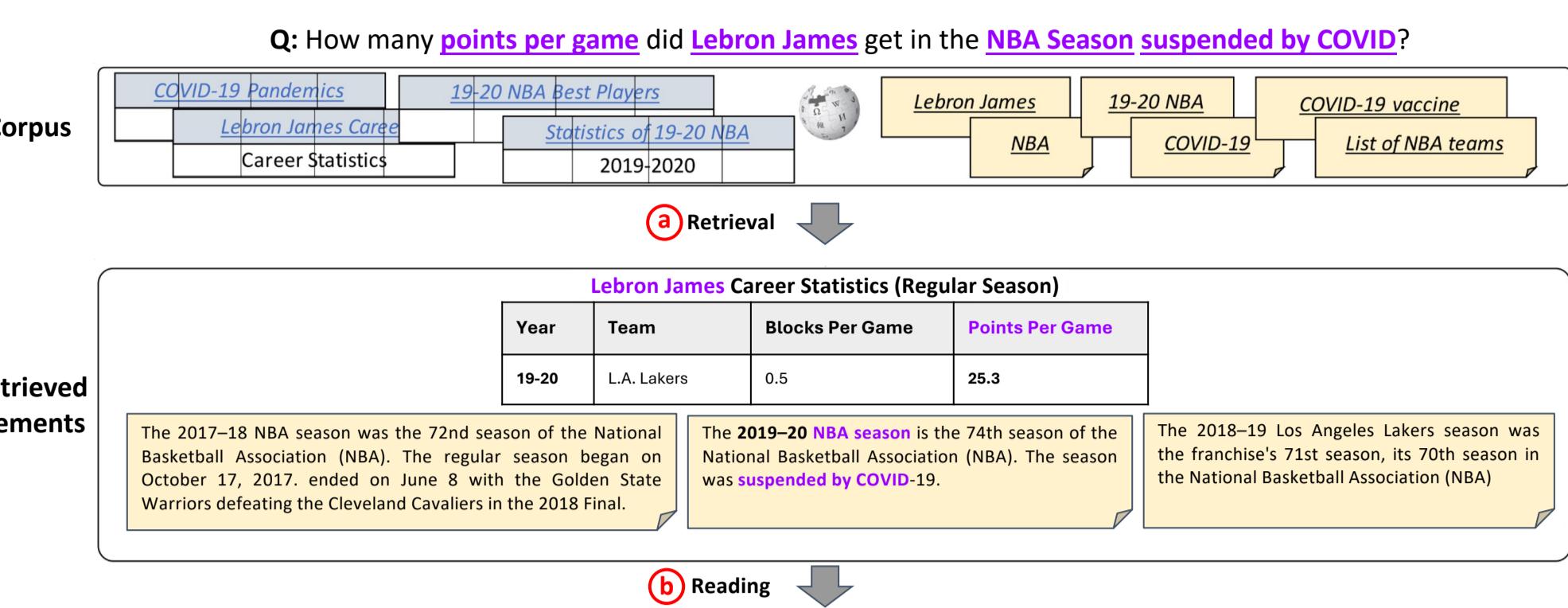


Open Table and Text Question Answering

Goal: Generate answer to a question by extracting answer strings from retrieved elements from a fixed corpus, a set of passages and tables.



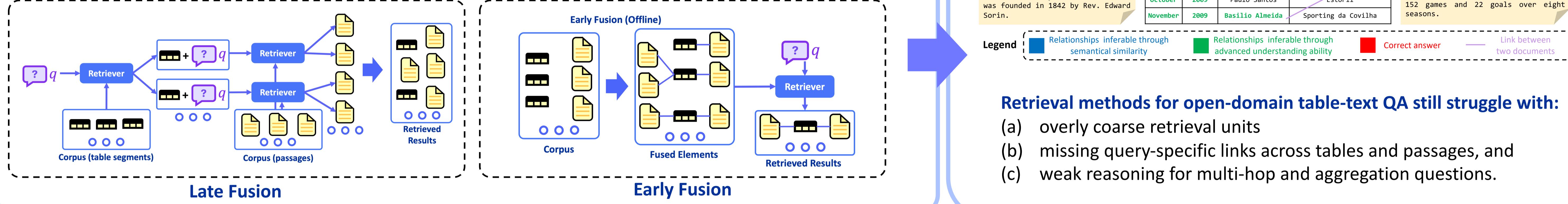
Standard OTT-QA System:

- (a) Retrieval : Retrieve elements from a fixed corpus of passages and tables with a retriever.
- (b) Reading : Analyze retrieved elements to provide an answer to the given question with a reader.

Previous Methods

Existing studies are categorized into two main approaches based on when this relationship is considered:

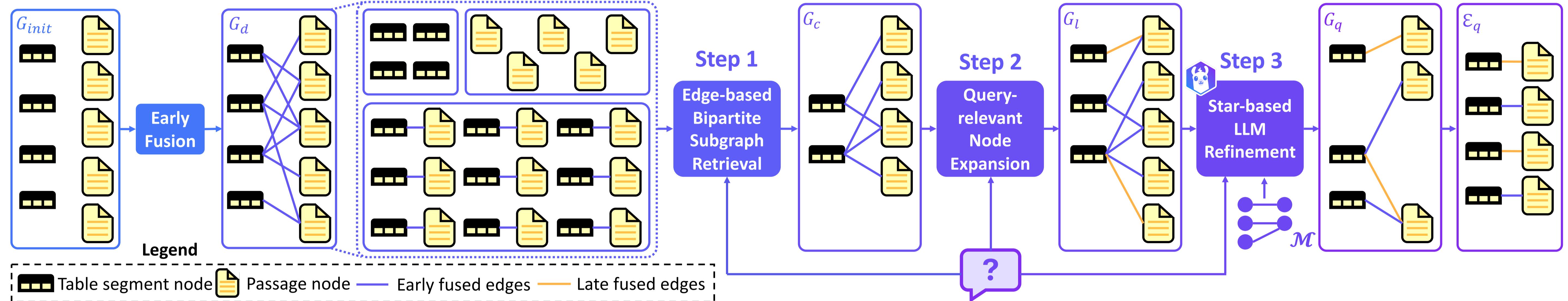
- **Late fusion:** Relationship is considered after the query is provided.
- **Early fusion:** Relationship between tables and passages is considered before the query is provided.



How does HELIOS push the frontier?

HELIOS reframes retrieval as locating a query-relevant subgraph from a bipartite data graph constructed via early fusion between table segments and passages. It introduces a three-stage, granularity-aware pipeline that harmonizes the strengths of both **early** and **late fusion** techniques while incorporating **LLM reasoning**.

System Overview

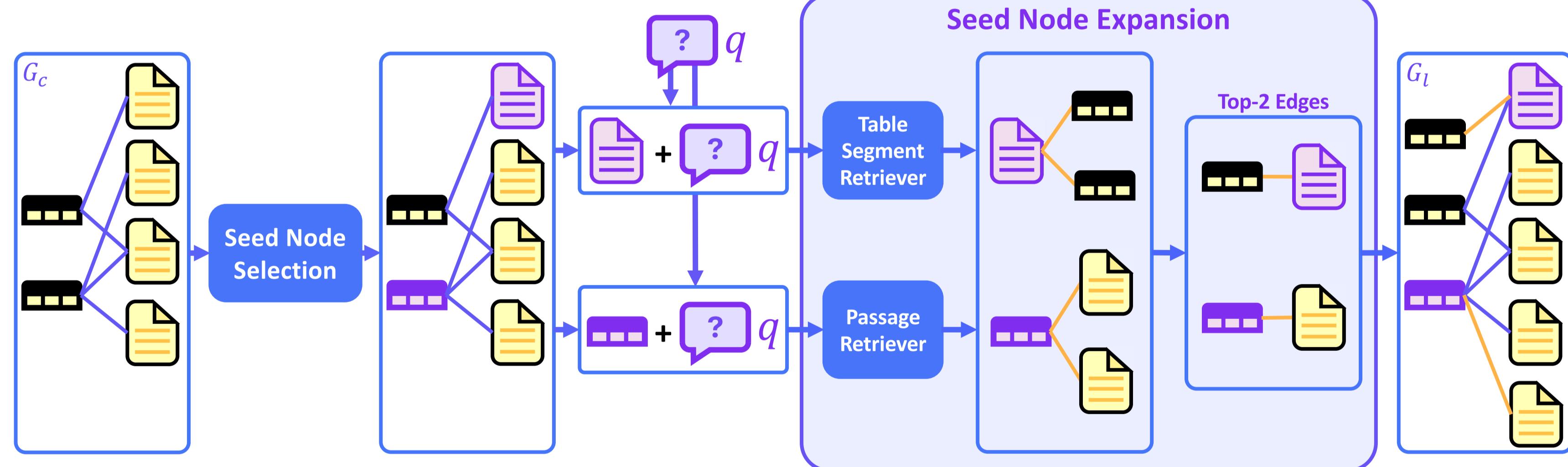


1. **Edge-based Bipartite Subgraph Retrieval** (Early Fusion): Offline entity linking builds the graph. An edge-level multi-vector retriever extracts a compact, high-precision candidate subgraph.

2. **Query-relevant Node Expansion** (Late Fusion): Identifies nodes most aligned with the query and selectively expands them, restoring essential query-dependent links.

3. **Star-based LLM Refinement** (LLM Reasoning): Decomposes the expanded graph into star-shaped subgraphs. An LLM performs aggregation and multi-hop reasoning, retaining only verified evidence.

Query-relevant Node Expansion



The overall procedure of query-relevant node expansion

- The beam width b is set as 2 in this example.
- The purple-colored nodes indicate the selected seed nodes.

Evaluation

* Please read our paper for full experiment results!

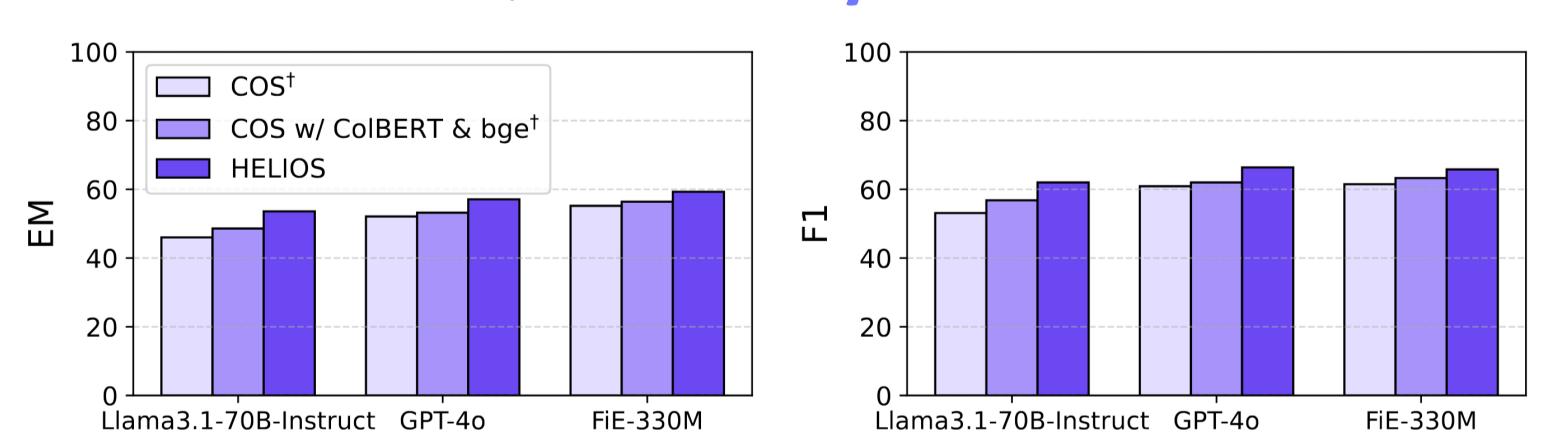
Retrieval Accuracy Comparison

Model	AR@2	AR@5	AR@10	AR@20	AR@50	nDCG@50	HITS@4k
Iterative Retriever	—	—	—	—	—	—	27.2
Fusion Retriever	—	—	—	—	—	—	52.4
OTTER ¹	31.4	49.7	62.0	71.8	82.0	25.9	70.1
DOTTER ¹	31.5	51.0	61.5	71.9	80.8	26.7	70.3
CORE ¹	35.3	50.7	63.1	74.5	83.1	25.4	77.2
COS ¹	44.4	61.6	70.8	79.5	87.8	33.6	81.8
COS w/ ColBERT & bge ¹	49.6	68.2	78.7	85.0	91.7	36.5	85.9
DOTTER + COS & LLM ¹	50.0	62.4	70.0	76.2	84.7	34.7	—
HELIOS	63.3	76.7	85.0	90.4	94.2	47.0	91.8

HELIOS consistently outperforms all competitors on AR@k across different k values.

- upto **39.9%** higher nDCG@50 compared to SOTA
- upto **12.2%** higher HITS@4k compared to SOTA

End-to-End QA Accuracy



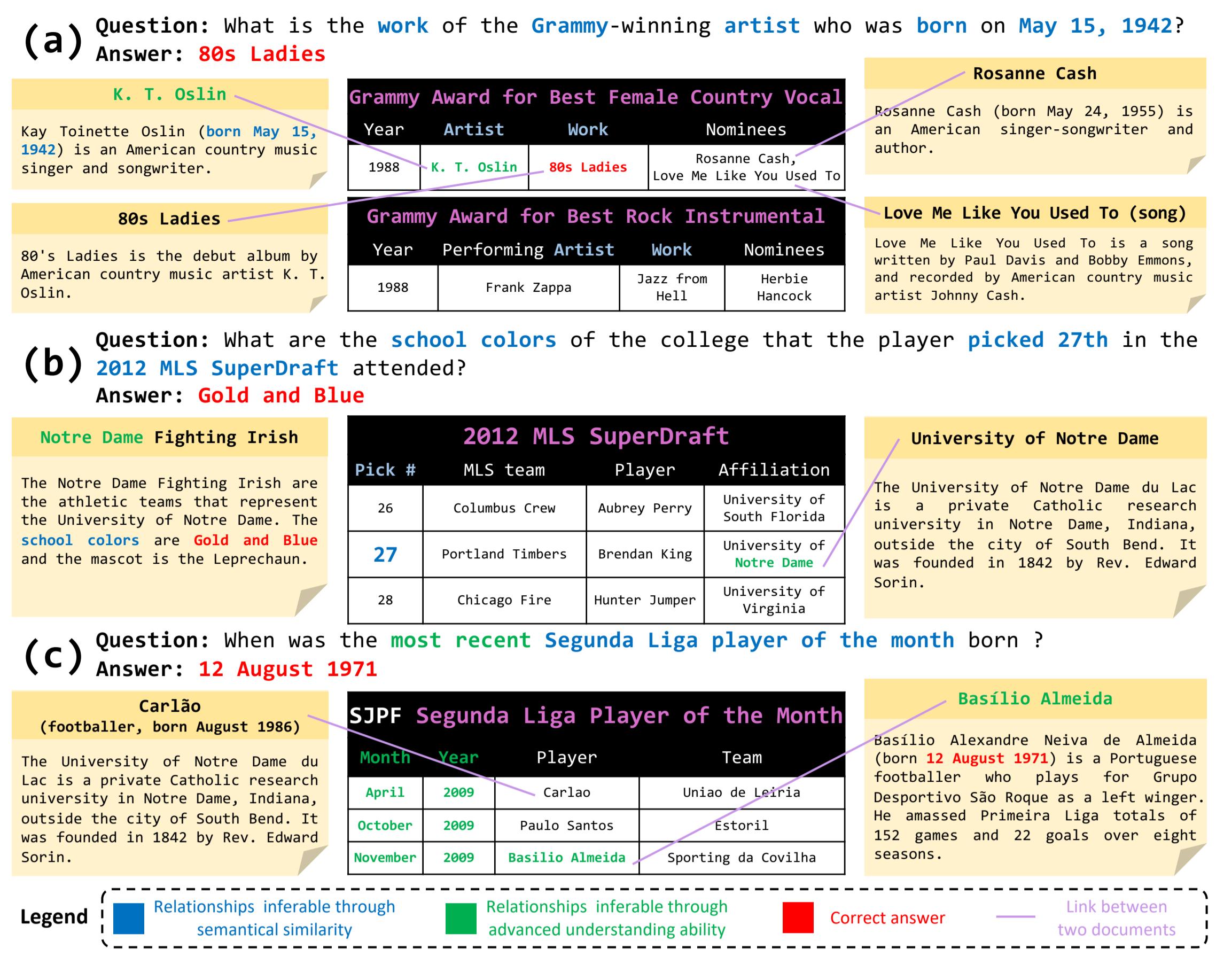
HELIOS consistently outperformed all reader models, achieving an average **EM gain of 7.5%**, and an average **F1 gain of 6.6%** over COS with ColBERT and BGE.

Ablation Study & Algorithm Execution Time

Algorithm	Execution Time (s)	nDCG@50
DotTER	0.08	26.7
CORE	4.13	25.4
COS	3.75	33.6
COS w/ ColBERT & bge	5.46	36.5
HELIOS	5.14	47.0
HELIOS w/ Finetuned SLR	4.76	47.6
w/o QNE	—	45.1
w/o SLR	2.16	46.5
w/o (SLR & Edge Reranker)	1.11	42.1

HELIOS achieved the best retrieval accuracy • reaching **47.6 nDCG@50** with finetuned SLR, • while maintaining **comparable execution time**, and • ablation confirms that both **SLR** and **QNE** provide meaningful gains.

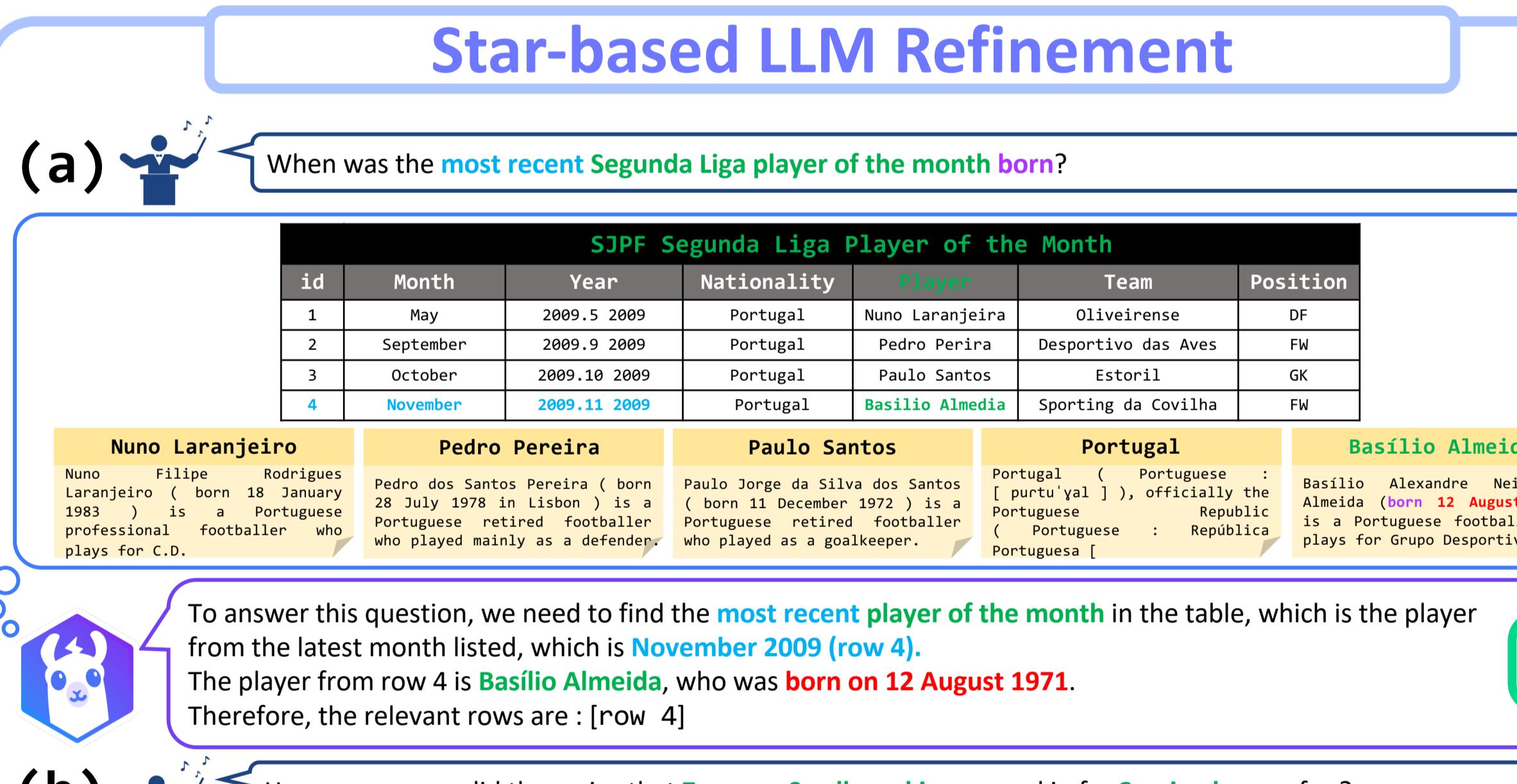
What problem do we solve?



Retrieval methods for open-domain table-text QA still struggle with:

- (a) overly coarse retrieval units
- (b) missing query-specific links across tables and passages, and
- (c) weak reasoning for multi-hop and aggregation questions.

System Overview



1. **Edge-based Bipartite Subgraph Retrieval** (Early Fusion): Offline entity linking builds the graph. An edge-level multi-vector retriever extracts a compact, high-precision candidate subgraph.

2. **Query-relevant Node Expansion** (Late Fusion): Identifies nodes most aligned with the query and selectively expands them, restoring essential query-dependent links.

3. **Star-based LLM Refinement** (LLM Reasoning): Decomposes the expanded graph into star-shaped subgraphs. An LLM performs aggregation and multi-hop reasoning, retaining only verified evidence.

Star-based LLM Refinement

(a) When was the most recent Segunda Liga player of the month born?

id	Month	Year	Nationality	Player	Team	Position
1	May	2009.5	Portugal	Nuno Laranjeira	Oliveirense	DF
2	September	2009.9	Portugal	Pedro Pereira	Desportivo das Aves	FW
3	October	2009.10	Portugal	Paulo Santos	Estoril	GK
4	November	2009.11	Portugal	Basílio Almeida	Sporting da Covilhã	FW

Nuno Laranjeira (born 1981) is a Portuguese footballer who plays for C.D. Odivelas. Pedro Pereira (born 1981) is a Portuguese retired footballer who played mainly as a defender. Paulo Santos (born 1981) is a Portuguese retired footballer who played mainly as a goalkeeper. Basílio Almeida (born 1981) is a Portuguese footballer who plays for Sporting da Covilhã.

To answer this question, we need to find the **most recent player of the month** in the table, which is the player from the latest month listed, which is November 2009 (row 4).

The player from row 4 is Basílio Almeida, who was born on 12 August 1971.

Therefore, the relevant rows are : [row 4]

(b) How many years did the series that Zuzanna Szadkowska appeared in for 3 episodes run for?

Guiding Light	Zuzanna Szadkowska
Guiding Light (1978) is an American television soap opera. It is listed in Guinness World Records as the longest-running drama in television in American history, broadcast on CBS for 57 years...	Year Title Role Notes

First, Zuzanna Szadkowska appeared in the series "Guiding Light".

Second, the linked passage titled "Guiding Light" mentions that it ran for 57 years on CBS and had a 19-year broadcast on radio.

Therefore, relevant passages are: ["Guiding Light"]

The star-based LLM refinement consists of two phases: column-wise aggregation and passage verification.

- (a) shows a successful case of the column-wise aggregation module in resolving a complex query.
- (b) shows a successful case of the passage verification module in addressing the query.

Conclusion

- We presented HELIOS, a novel table-text retrieval method that **harmonizes** the strengths of both **early** and **late fusion** techniques while incorporating **LLM reasoning**.
- It addresses the limitations of competitors by introducing a **multi-granular retrieval system** that optimally balances granularity across retrieval stages.
- Experiments on OTT-QA show that it **surpasses SOTA models**, achieving a 42.6% AR@2 improvement and a 39.9% nDCG@50 gain.