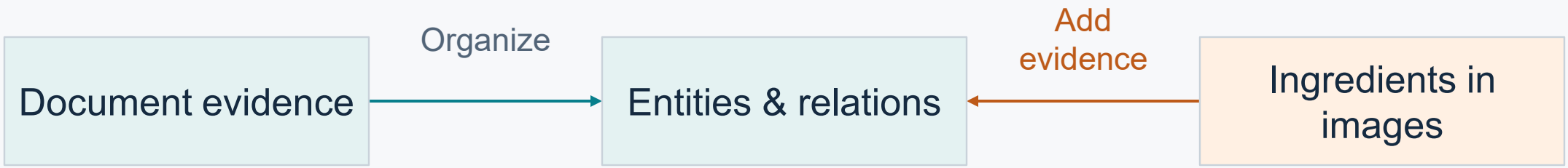
GraphRAG & Multimodal Knowledge Graphs

How connected knowledge grounds an answer



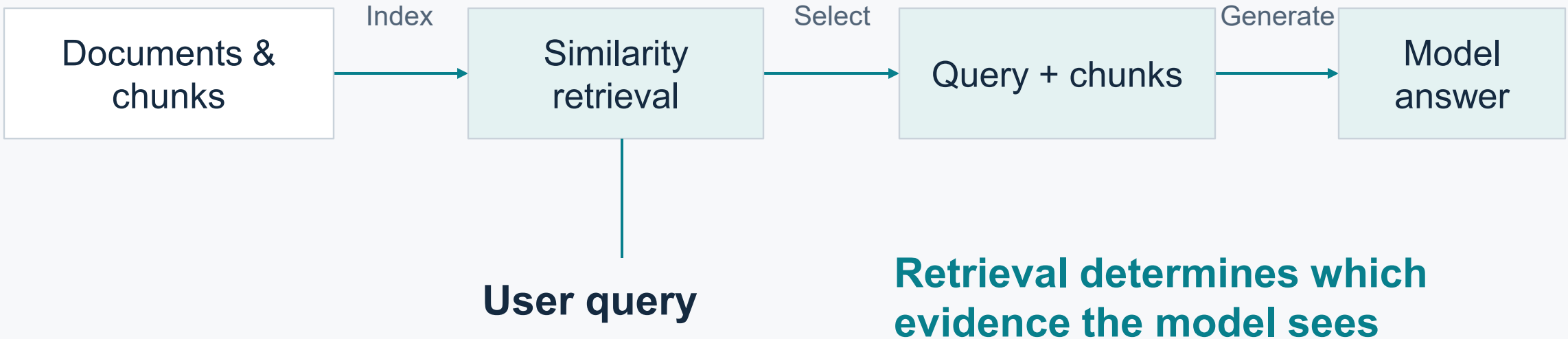
University introduction · 90 minutes · Two technical articles

Learning goals and class plan

Compare retrieval	Explain chunk retrieval, relational lookup and global summaries
Trace evidence	Find sources; handle missing or conflicting records
Build multimodal links	Connect dishes, ingredients and image regions with named relations
Transfer with limits	Design a new case: input, process, output and limits

Intro 10m · GraphRAG 30m · Build 12m · Multimodal 26m · Review 12m

How retrieved chunks enter an answer



Retrieval determines which evidence the model sees

Retrieved text enters this context; this does not retrain the model.

Three questions, three evidence needs

Question type	Campus dining question	Evidence needed
Specific lookup	What is in Kung Pao Chicken?	One explicit record
Relational lookup	Which stalls serve chicken dishes?	Stall–dish–ingredient relations
Global summary	What topics occur in these records?	Representative evidence across the corpus

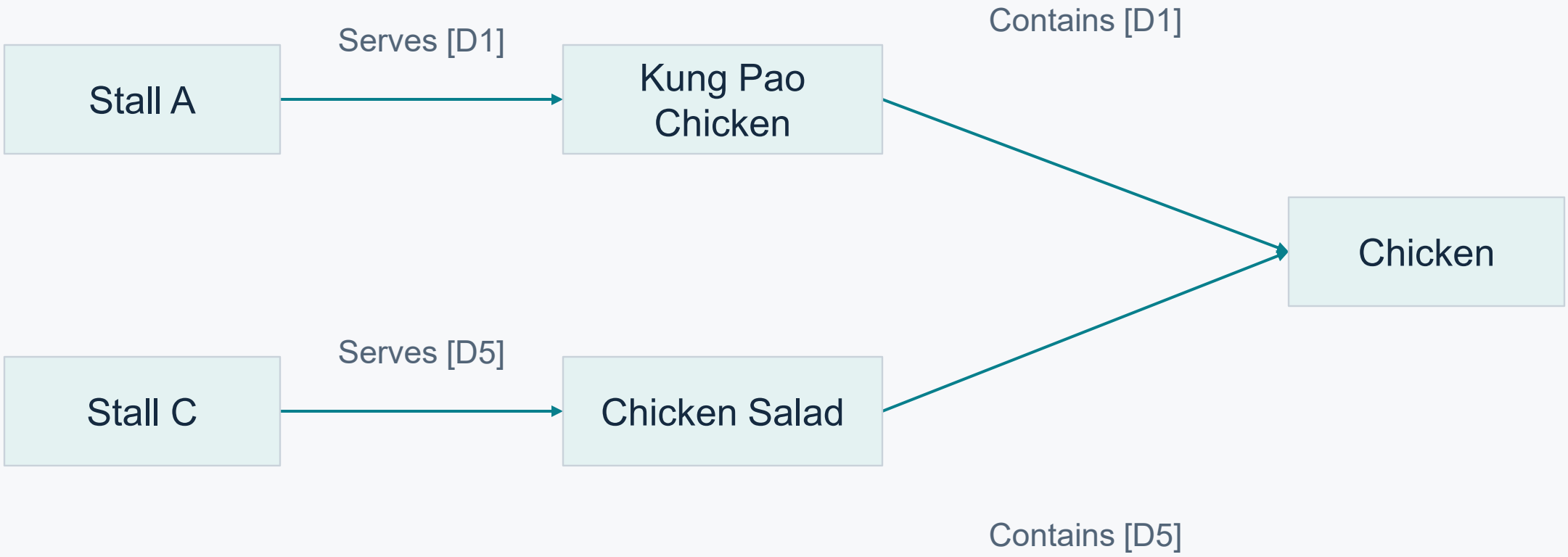
Teaching simulation: use the same campus dining records for all questions.

Tiny campus dining corpus

D1	Dish record	Stall A: Kung Pao Chicken contains chicken and peanuts.
D2	Dish record	Stall B: Mapo Tofu contains tofu and beef.
D3	Student review	Stall A has long waits at the evening peak.
D4	Operations	Stall A has many evening-peak orders.
D5	Dish record	Stall C: Chicken Salad contains chicken and lettuce.
D6	Student review	Stall C’s Chicken Salad has a mild taste.

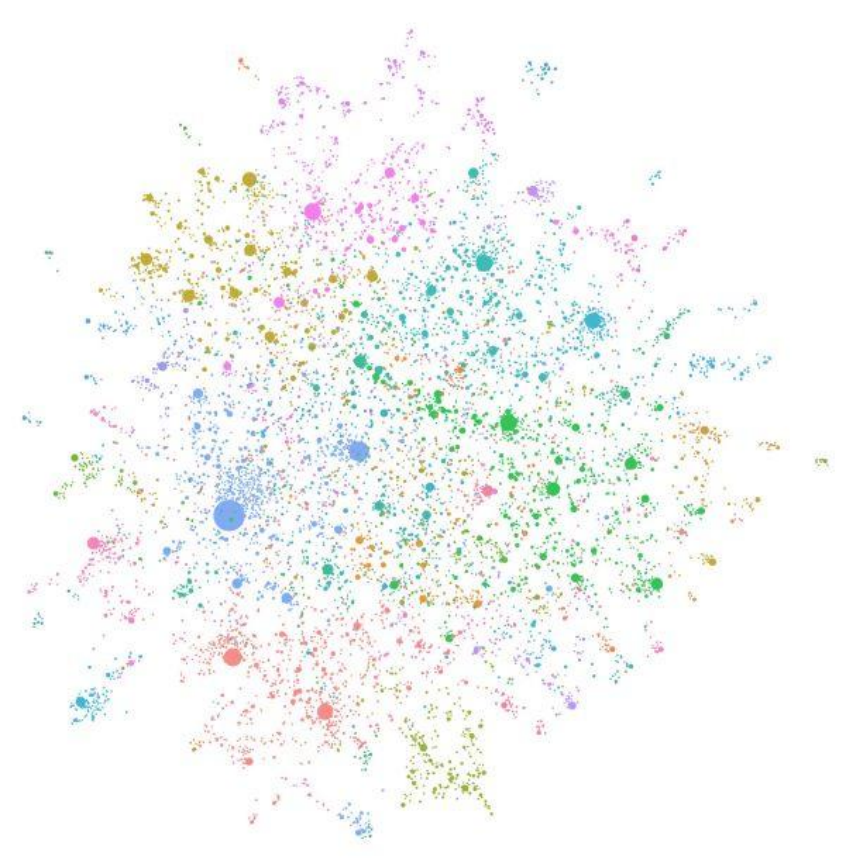
D1–D6 are invented for teaching; no real campus or merchant data.

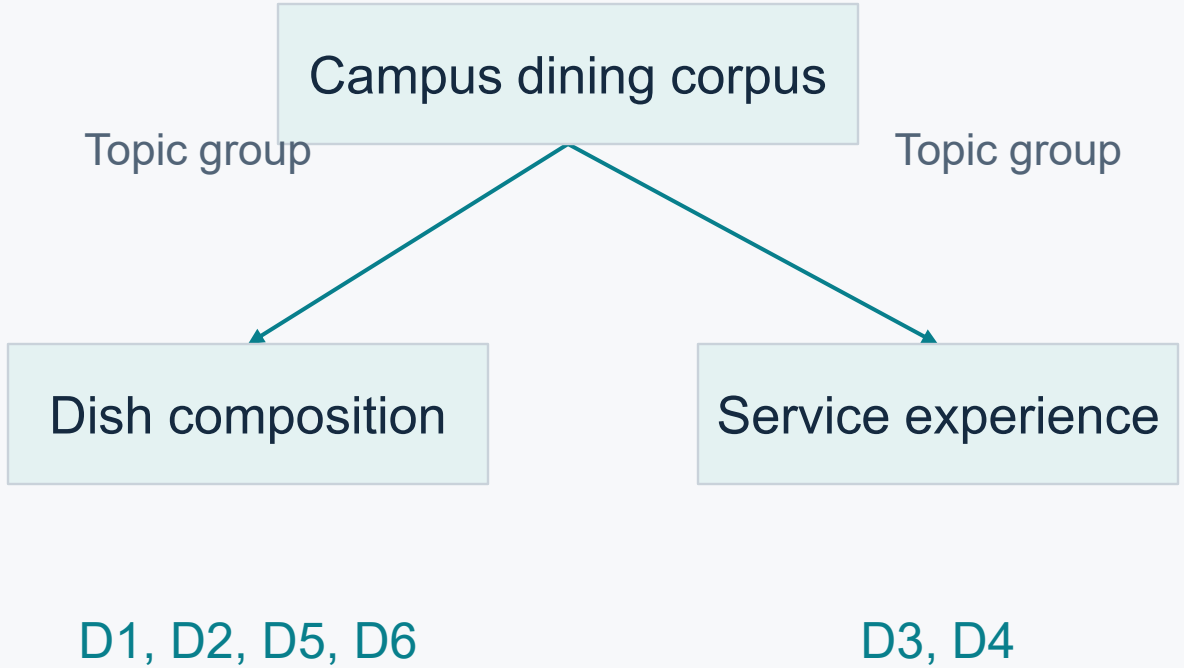
Entities connect separate records



GraphRAG extracts entities and relations using LLMs. Teaching example: D1 and D5.

Graph communities and topic levels





Communities reflect graph links; generated community names still need review.

Batch aggregation in global search

Query: Which topics appear across the corpus?



Based on official documentation. The answers below are teaching examples.

Evidence and inference in an answer

Source records

D3: Long evening-peak waits at Stall A
D4: Many evening-peak orders at Stall A

Supported description

Both records concern Stall A’s evening peak

Inference to test

More orders caused the longer waits

D3 and D4 occurring together does not establish a causal relationship.

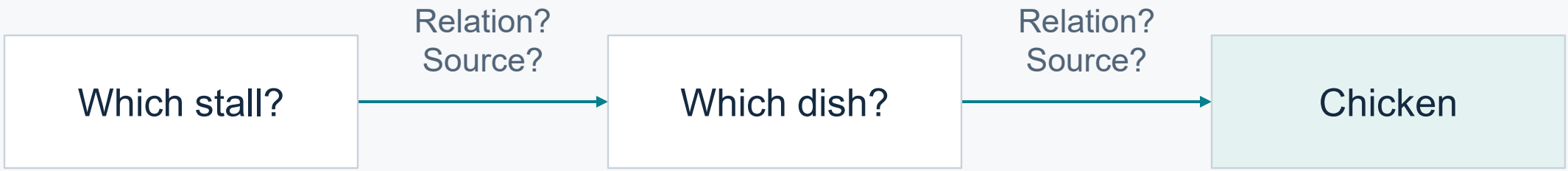
GraphRAG findings and limits

Evaluation aspect	Article reports	Conditions to preserve
Comprehensiveness & diversity	Better than compared baselines	Depends on questions and corpus
Human verification	Source leads are provided	Check each citation and claim
Faithfulness to sources	Similar to the baseline	Does not imply hallucination removal

The February 2024 article reports an early comparison, not a universal accuracy gain.

Activity A: complete evidence paths

Query: Which stalls serve dishes containing chicken?

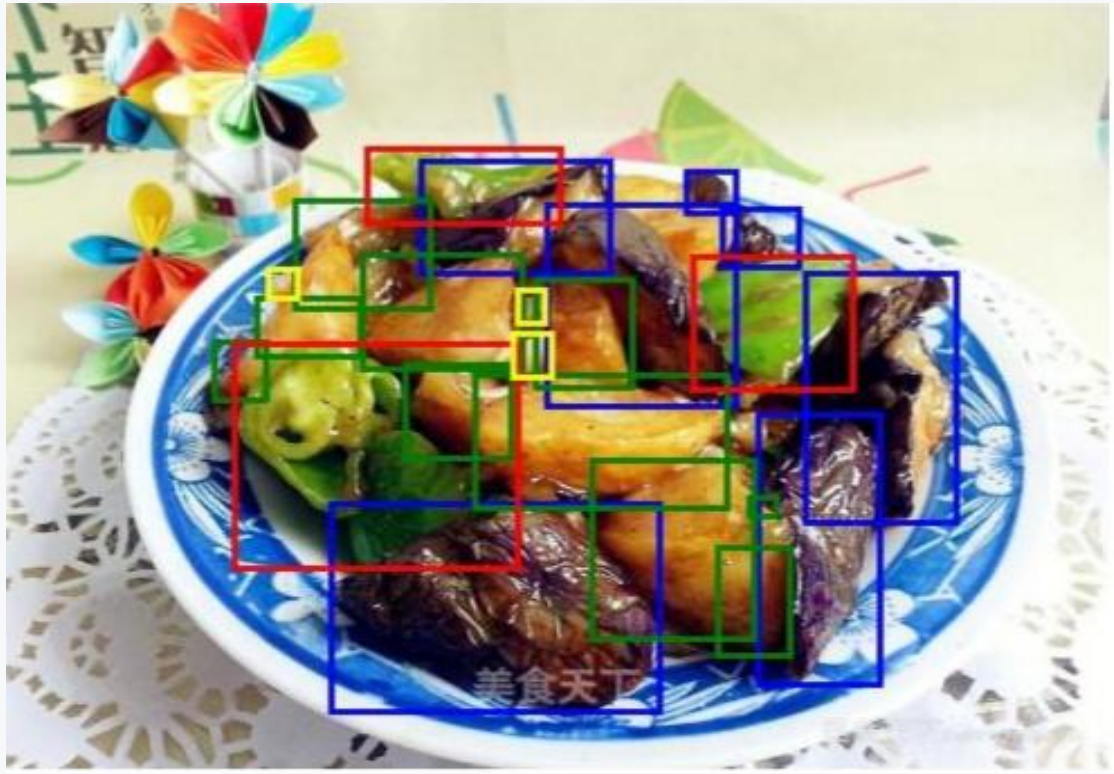


Submit two paths, a sourced answer and one evidence limit

Test D7: Stall A's Kung Pao Chicken has no peanuts. How should you revise the graph?

2m alone, 5m group graph, 5m peer test and revision.

What text and images reveal



Text may mention

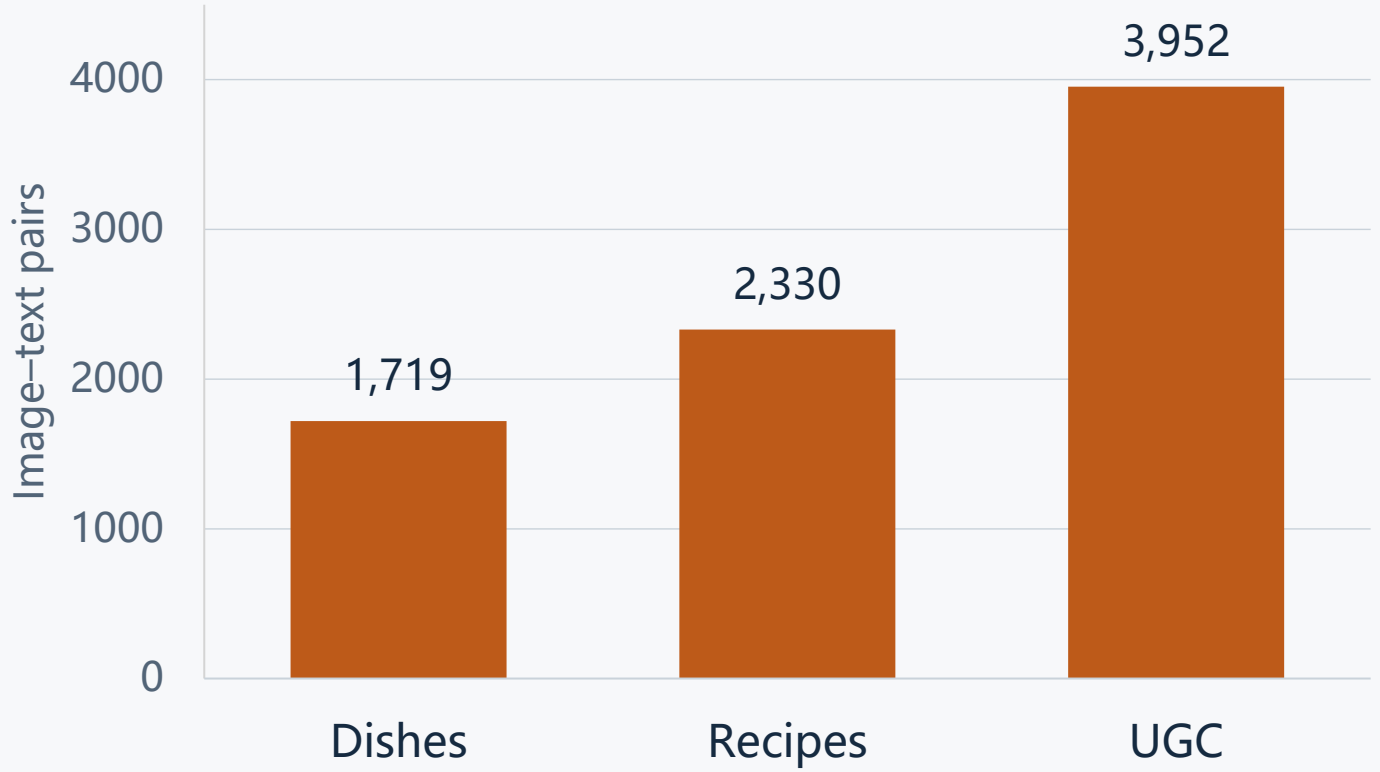
Ingredients, seasonings and cooking steps

Image evidence adds

Visible ingredients and their locations

Image excerpt: recipe example in Meituan Fig. 2.1; original annotations retained.

CMIngre sources and annotation



Annotation level

Ingredient name
Ingredient bounding box

Normalize synonyms

Songhua egg = preserved egg

Source reports: 8,001 image–text pairs, 429 ingredients, 95,290 bounding boxes.

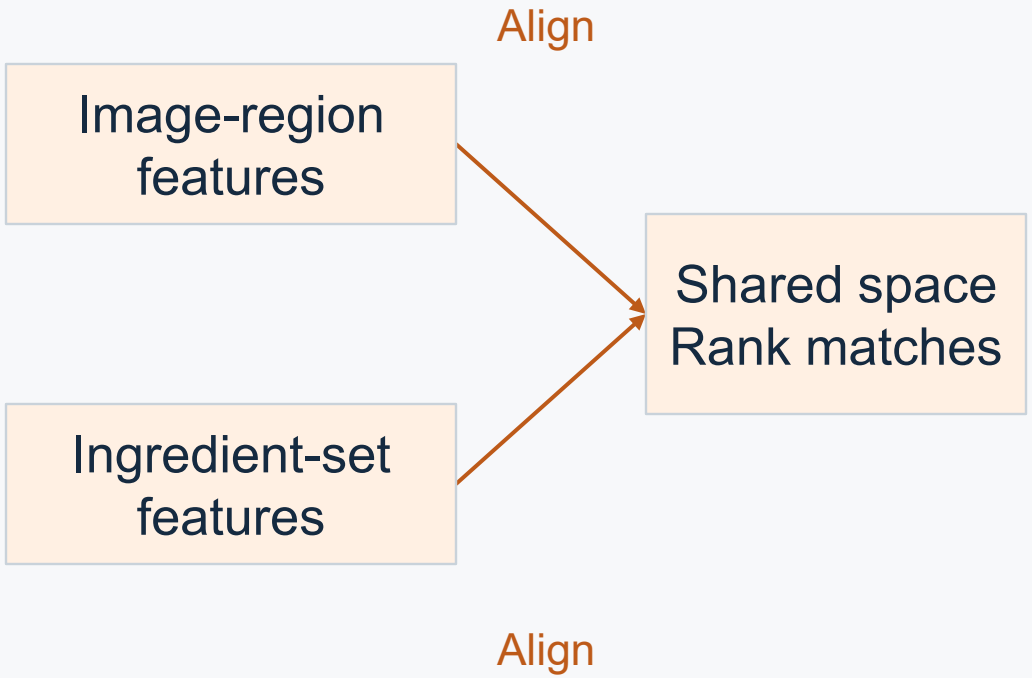
Detection and cross-modal retrieval

Task 1: ingredient detection



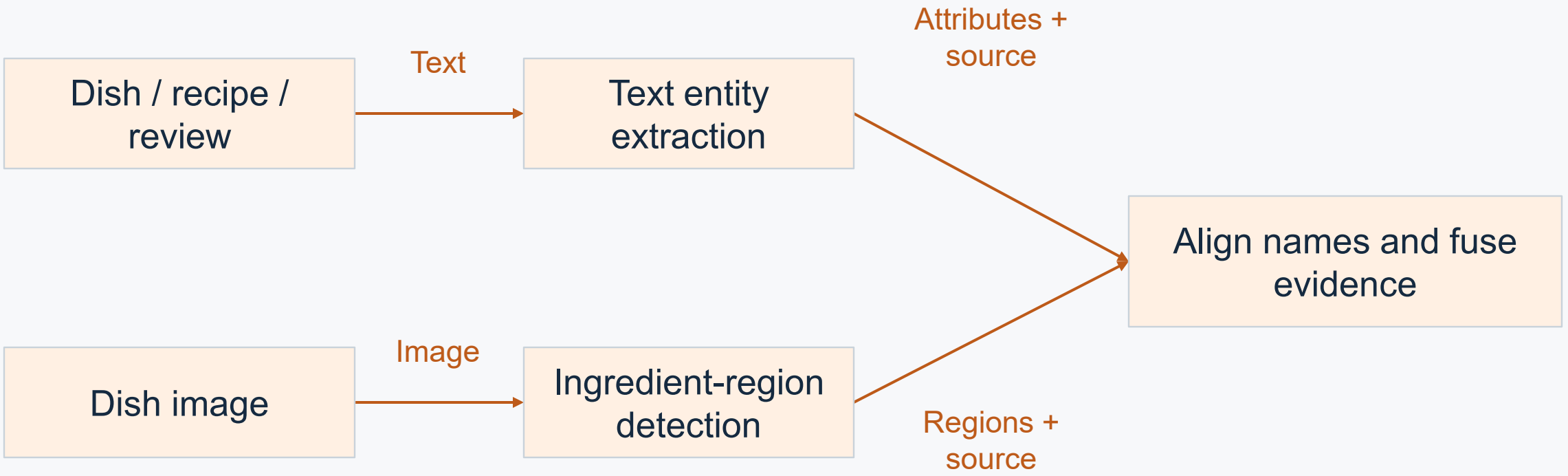
Output: class + location

Task 2: cross-modal retrieval



Detection experiments merge 429 fine-grained labels into 67 higher-level classes.

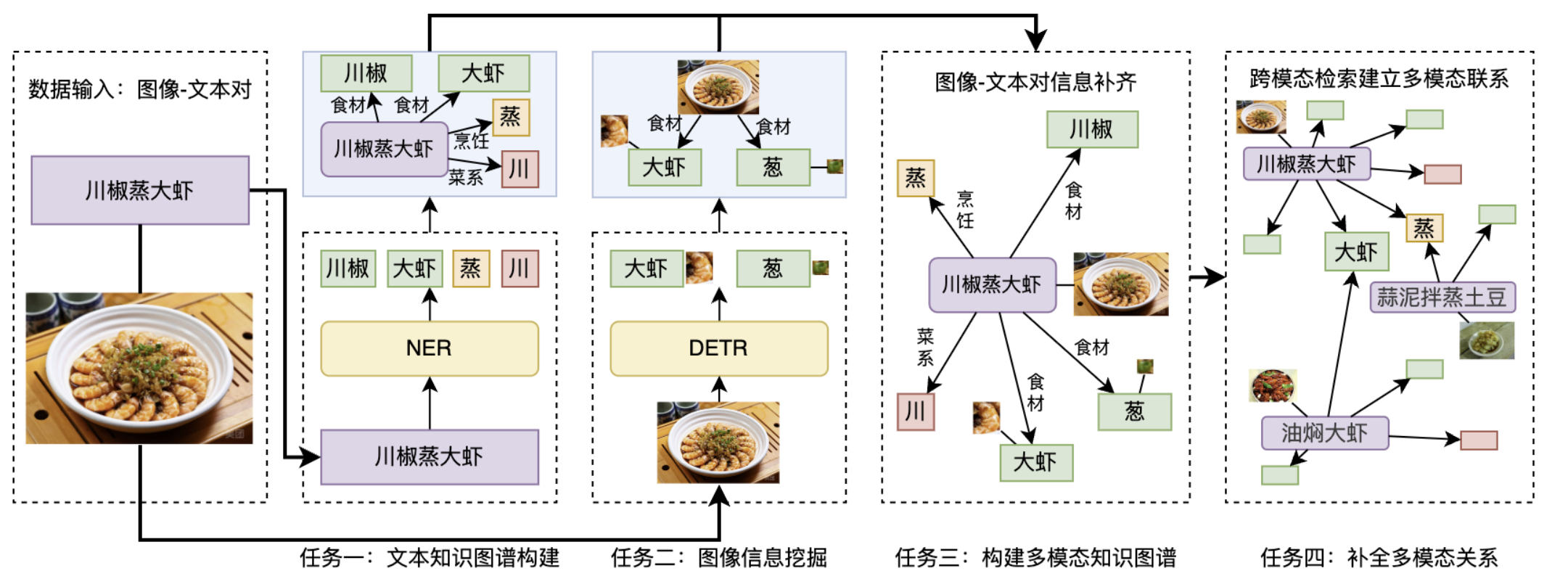
Extract and fuse multimodal knowledge



Teaching recommendation: preserve and review conflicts

Redrawn from Meituan. Preserve sources and review image–text disagreement.

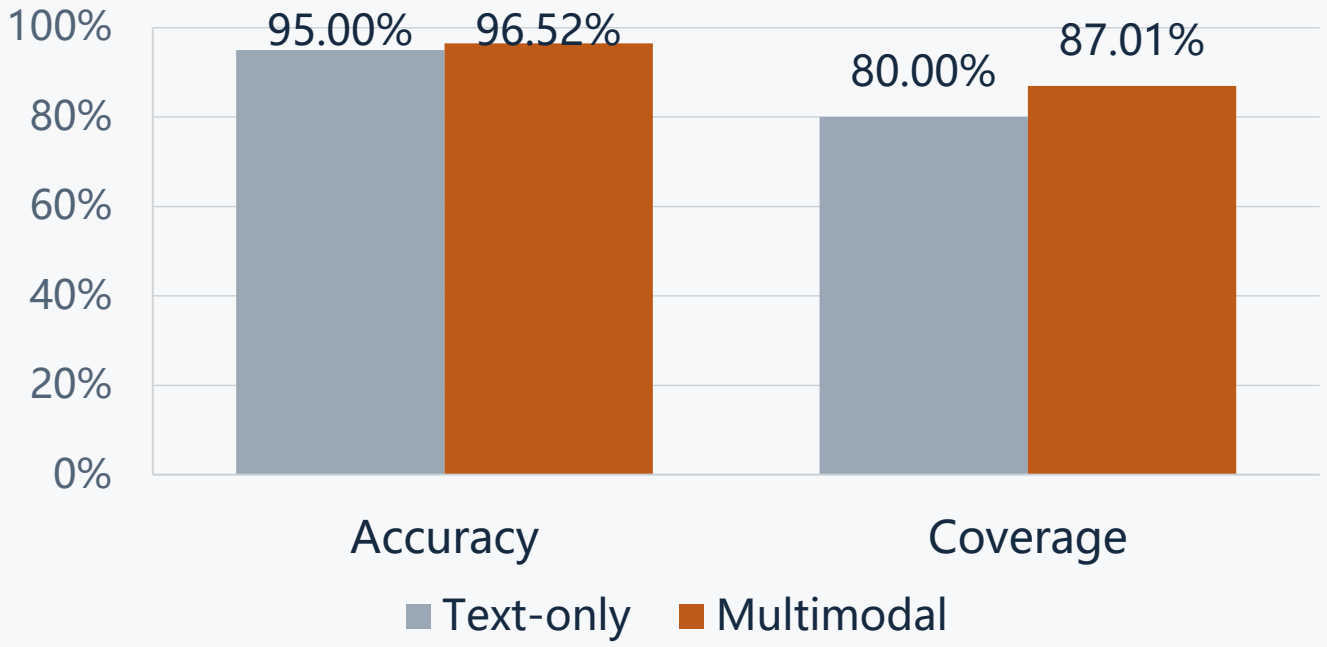
Cross-modal links in the original pipeline



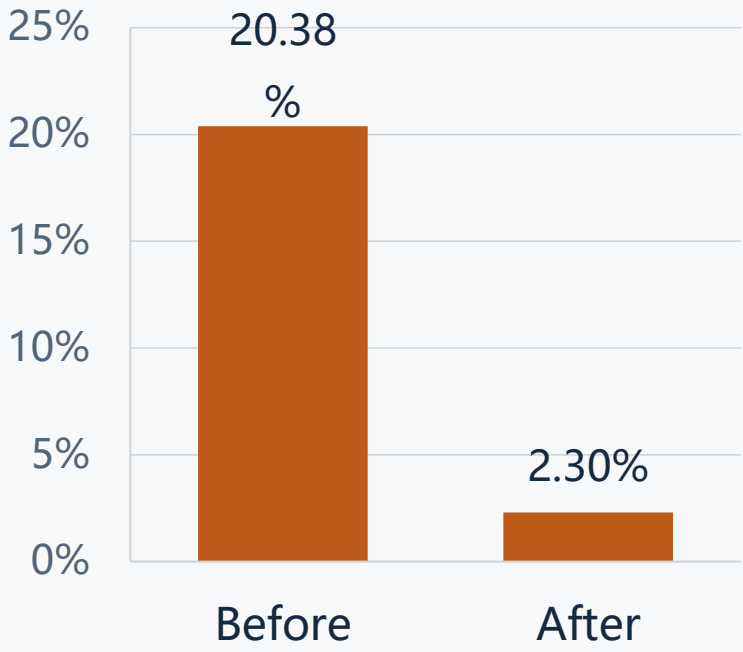
Original: Meituan Fig. 5.1. Which step adds image evidence to the text? Original Chinese figure retained.

Reported business effects

Attribute knowledge (higher is better)



Identification error (lower is better)



Left: attribute knowledge. Right: downstream product identification. Different measures cannot form one score.

Visible ingredients and unknowns

Candidate conclusion	Evidence status?	What is missing?
A: D1 records peanuts in Kung Pao Chicken	Supported / inference / needs review	Give the record ID
B: No peanuts visible, so the dish has none	Supported / inference / needs review	Explain limits of visual evidence
C: Chicken detected, so its weight is known	Supported / inference / needs review	Separate classification from measuring quantity

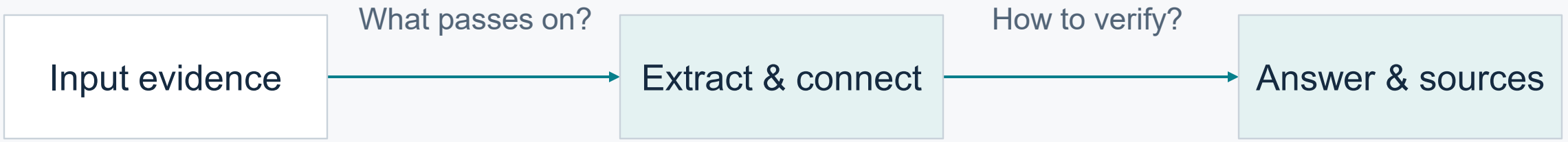
Not seeing an ingredient does not establish that the dish is free of it.

Where the two articles fit

Dimension	GraphRAG	Meituan dish knowledge graph
Input	Text documents	Images and text
Core process	Entities, relations, communities, summaries	Ingredient detection, matching, fusion
Role of graph	Organize evidence for answers	Add attributes and connect dishes
Typical output	Sourced answers and topic summaries	Ingredient knowledge and identification

Activity B: transfer to a new case

Scenario: build QA for a campus plant image–text archive



Submit 6 nodes, 5 named relations, a query and an exception

Rubric: evidence 30, relations 25, process 25, limits 20

5m individual design, then 4m peer review. This is a proposed joint system.

Exit check and assignment

Exit check	Draw a source-to-claim path and state its limitation
Assignment	Choose a multimodal case: one graph plus 150 English words
Must include	Source IDs, relation names, normal result, missing/conflict exit

Use paper, PowerPoint or a text table; no paid platform needed

Sources and further reading

Microsoft Research article	GraphRAG: Unlocking LLM discovery…
GraphRAG documentation	Indexing and global search
Meituan technical article	Building a dish graph with multimodal extraction
Research dataset	CMIngre: annotations and retrieval examples
GraphRAG research paper	From Local to Global…

Course follows the 2024 articles; official docs clarify the workflow. Accessed 2026-10-18.