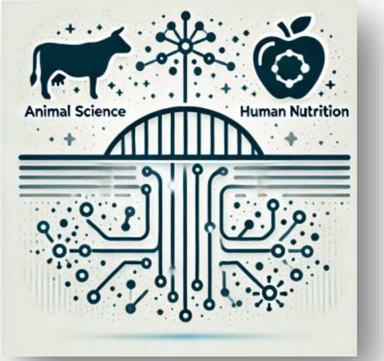
Bridging Animal and Human Nutrition Through INSIGHT: A Retrieved-Augmented Agentic Model

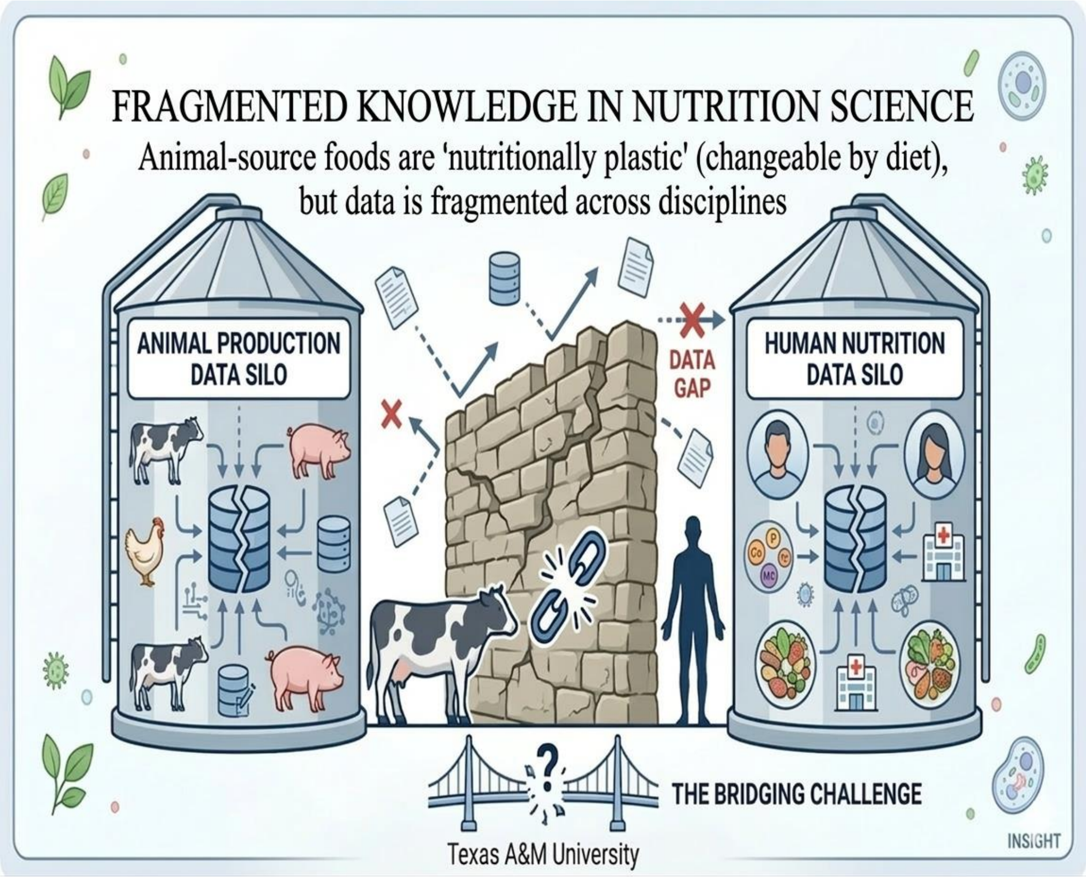
Robert Strong, Luis O. Tedeschi, Karun Kaniyamattam, Jian Tao, Dheeraj Mudireddy, Praneet Sai Madhu Surabhi, Nicole Geer, Ashley Wang

TEXAS A&M UNIVERSITY

Why is the literature hard to use?



- Evidence is spread across animal science, feed composition, food science, and human nutrition
- Findings sit in different species, journals, vocabularies, and experimental contexts.
- That fragmentation limits translation into actionable guidance.



Why Retrieval-Augmented Generation rather than a generic chatbot?

Problem with parametric only LLM

- hallucinations and fabricated citations
- knowledge boundaries from training data
- weak traceability for scientific claims

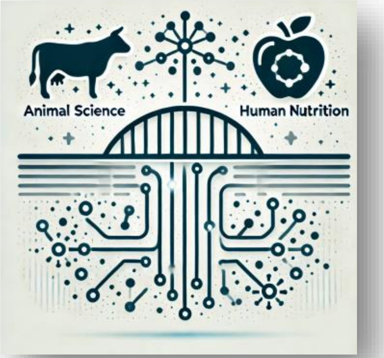
What RAG adds

- retrieved evidence from external documents
- better factual grounding and provenance
- freshness without relying only on model memory

INSIGHT goal

domain-specific, evidence-grounded synthesis


INTRODUCING INSIGHT: Intelligent System for Integrating Global Human & Animal Health Technology
GOAL: Domain-specialized synthesis of evidence with explict provenance (citations)
INSIGHT AI CORE
Tedeschi 2023
INSIGHT EVIDENCE BRIDGE
PROVENANCE SHIELD
[Ref. 12]
[Ref. 12]
[Ref. 12]
[Source 9]
[Source 8]
ANIMAL PRODUCTION & HEALTH DATA SILO
SYNTESIZED EVIDENCE (e.g., Reports, Models)
DOMAIN-SPECIFIC KNOWLEDGE GRAPHS
HUMAN NUTRITION & HEALTH DATA SILO
Texas A&M University
INSIGHT



Why is AI is needed?

~4,000

papers in the INSIGHT knowledge base

5

researchers on average

67.3

weeks for one systematic review

$141k–$180k

estimated labor cost per review

What is INSIGHT?

Truthful synthesis with provenance

- Cites source passages and exposes retrieval context

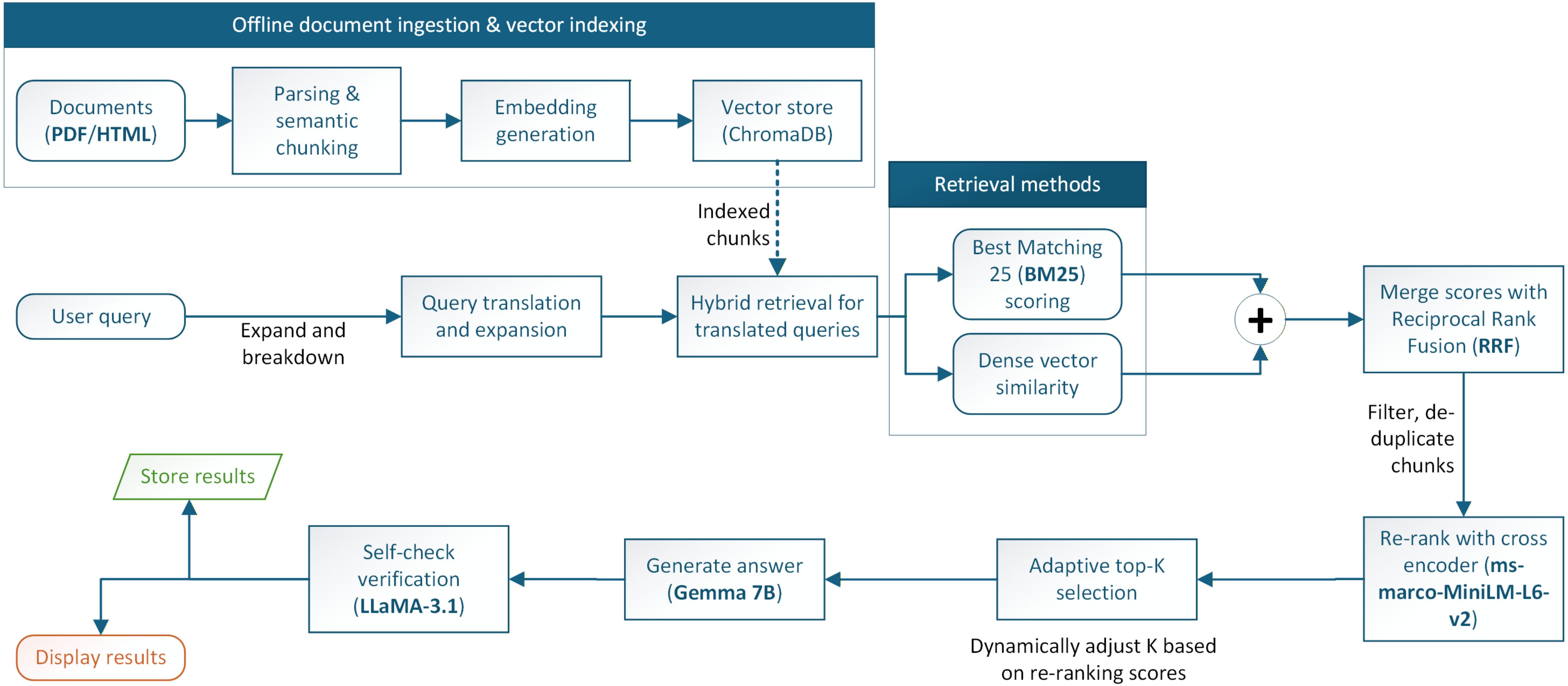
Agentic workflow

- Decides when to search, which tools to use, and when to escalate

Interoperability across domains

- Bridges animal science and human nutrition language

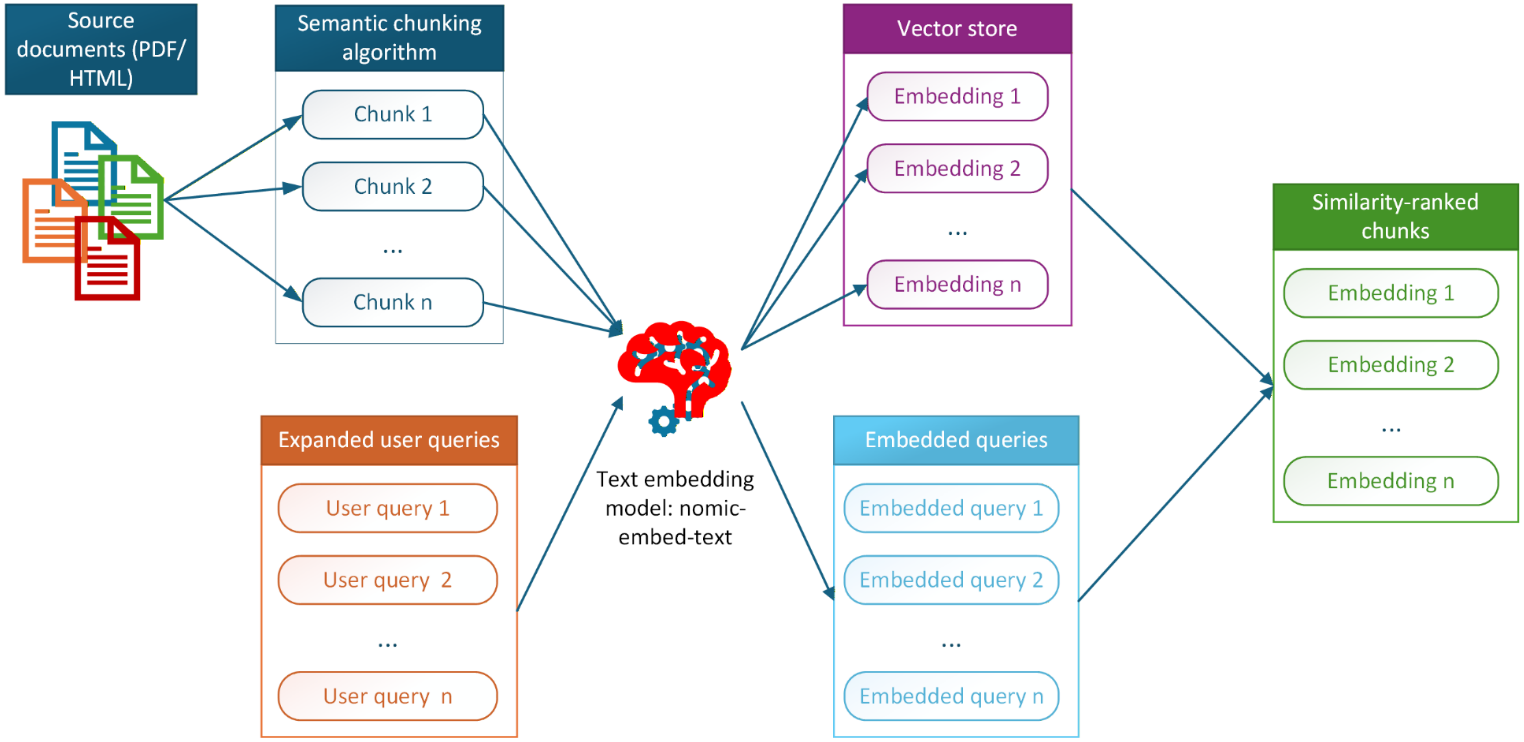
How the system works



Tedeschi et al. (2026; Frontiers in Nutr.)

Retrieval and indexing details

INSIGHT is engineered to improve recall without giving up evidence traceability



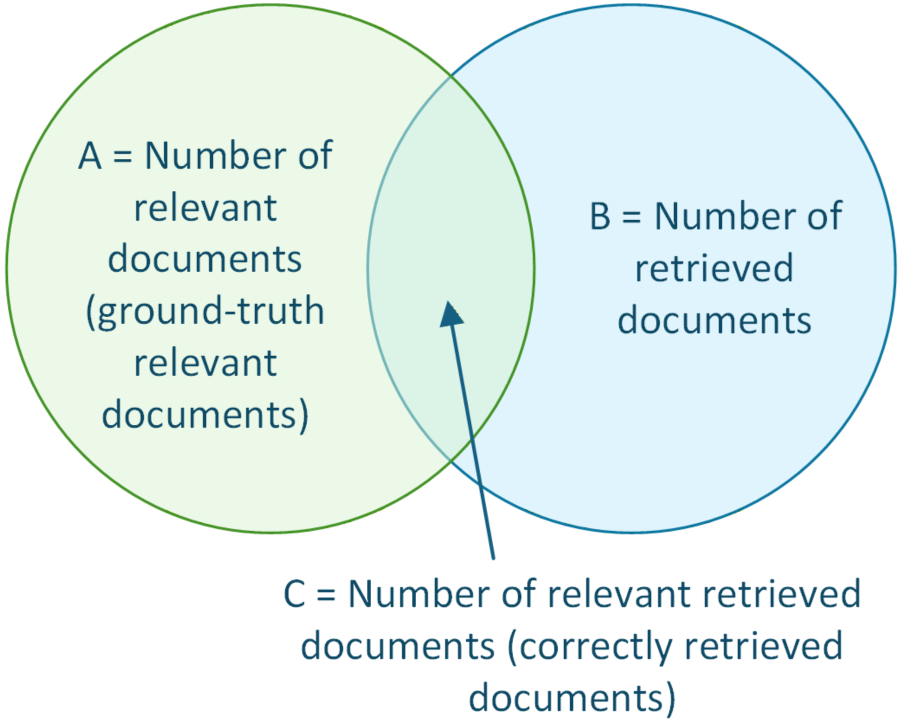
- Semantic chunking of extracted PDF/HTML content into 1024-token segments with 120-token overlap
- Text embeddings generated with Nomic Embed Text and stored in ChromaDB using Hierarchical Navigable Small World (HNSW) indexing
- Hybrid retrieval merges dense vector similarity with BM25 lexical matching via reciprocal rank fusion
- Cross-encoder reranking helps filter noisy candidates before final evidence selection

Tedeschi et al. (2026; Frontiers in Nutr.)

Evaluation of INSIGHT

- Documents were organized into ~10-paper topical groups so each question could be answered within a focused literature set
- Perplexity Deep Research generated 25 Q&A pairs per group and identified the ground-truth relevant papers
- Each question was run through INSIGHT, and document-level precision, recall, and F1 were computed

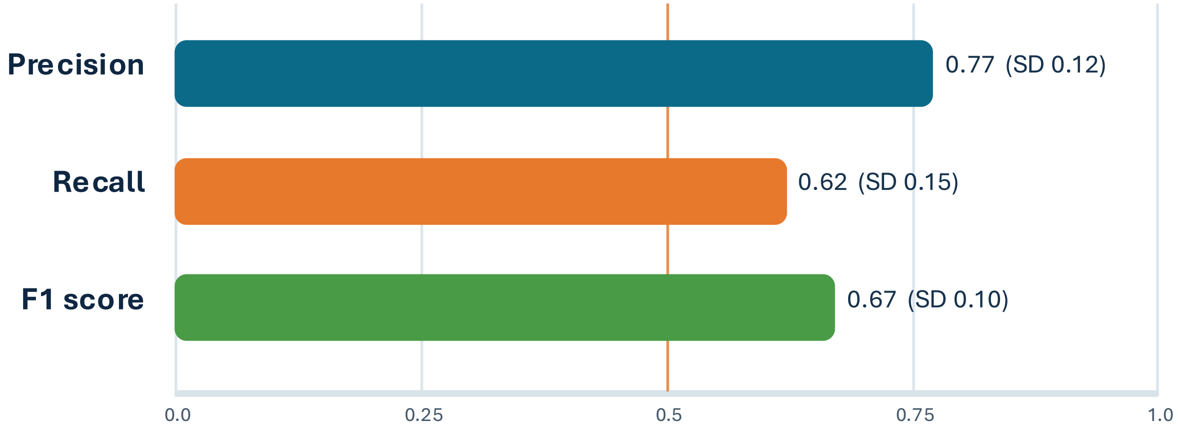
~4,000 papers in corpus → **282** papers sampled → **26** semantic groups → **614** total queries



Evaluation metrics: Precision (P) = C/B, Recall (R) = C/A, and F1 score (F1) = 2·P·R/(P + R)

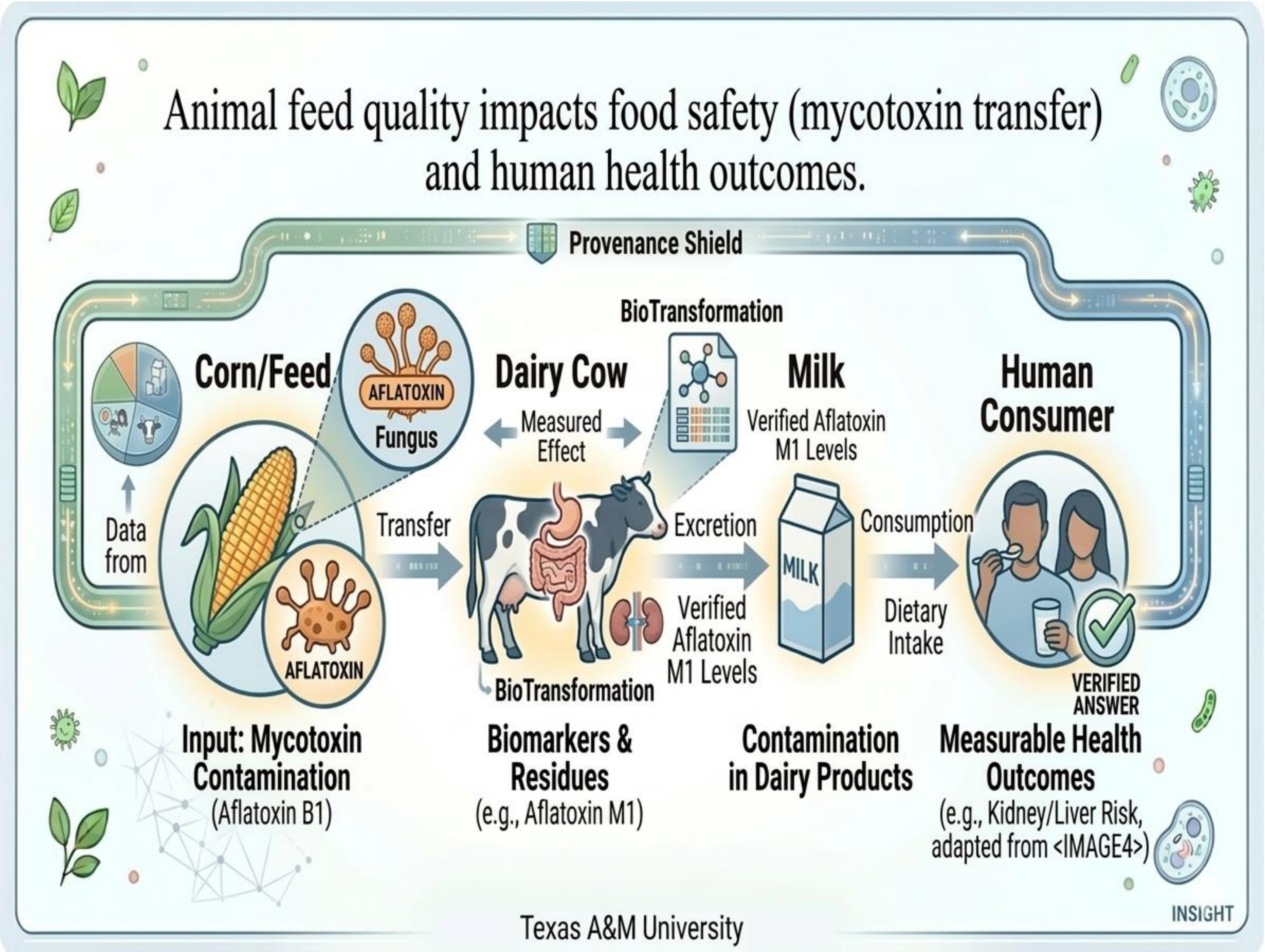
Evaluation results

- Strong precision means the retrieved evidence is usually relevant
- Lower recall means some relevant papers are still missed
- Performance changes across semantic domains, so hard topics remain hard topics



Case Study 1: The mycotoxin connection

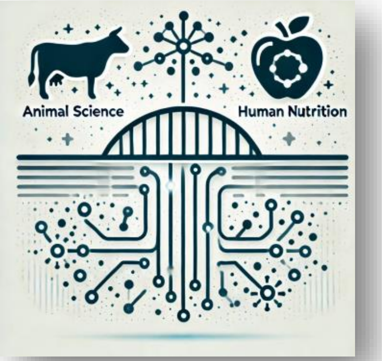
- Contaminated corn/feed (aflatoxin b1) is ingested by the animal
- INSIGHT tracks this flow to provide verified evidence on measurable health outcomes, such as kidney and liver risks



Case Study 2: Chromium supplementation

- Highlights how chromium propionate supplementation in feed leads to improved lean tissue deposition and higher growth efficiency, resulting in enriched beef for human consumption
- INSIGHT provides domain synthesis to verify how these specific animal feed interventions directly enhance the nutritional value of human food

Future Directions



What INSIGHT does well

- Gives transparent, document-grounded synthesis rather than unsupported narrative generation
- Performs reliably above baseline across a diverse, multi-group nutrition benchmark
- Supports cross-disciplinary discovery at the animal–human nutrition interface

What still needs improvement

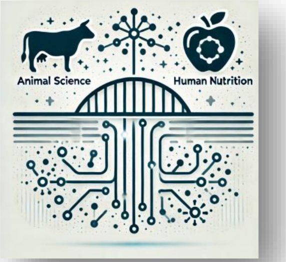
- Improve recall in semantically complex domains and with terminology variation
- Agentic AI-Knowledge Graph
- Expand corpus coverage and add broader expert evaluation beyond the current proof-of-concept benchmark

INSIGHT looks most valuable as an evidence-grounded assistant for researchers, practitioners, and policy work — not as a substitute for expert appraisal

Agentic AI + Knowledge Graphics

- **What it is**
  - AI agents using structured knowledge networks
- **Why it matters**
  - Reasoning, explainability, and multi-step tasks
- **Example flow**
  - Agent + KG ➔ Plans➔LLM➔Actions

Acknowledgments



Dr. David Hubbard (Texas A&M University Libraries, College Station, TX) for training graduate students on systematic literature search methods for building the corpus of the database for the INSIGHT RAG system

This material is based upon work supported by the **U.S. Department of Agriculture, Agricultural Research Service, under Agreement No. 58-3091-1-018**

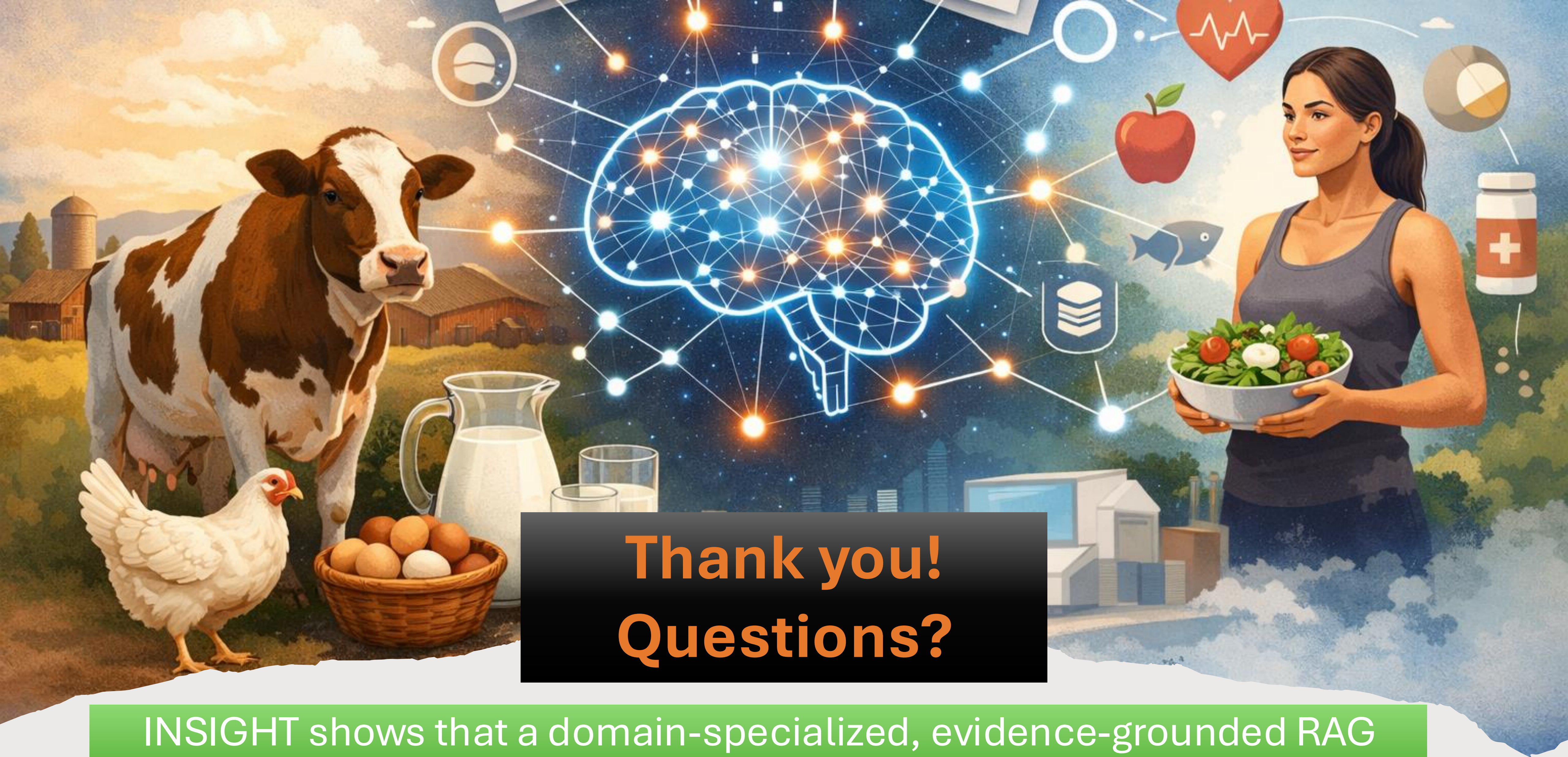
Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture

TEXAS A&M AGRILIFE

INSTITUTE FOR ADVANCING HEALTH THROUGH AGRICULTURE

USDA

U.S. Department of Agriculture



Thank you!
Questions?

INSIGHT shows that a domain-specialized, evidence-grounded RAG system can reliably retrieve relevant literature across animal and human nutrition while making its reasoning more transparent and traceable