

RAG VS. MODEL TRAINING:

What Academic Technology Teams **Need to Know**

For academic technology teams evaluating AI solutions, understanding the difference between retrieval and model training is important. Retrieval-Augmented Generation (RAG) lets an AI application use information from external sources to answer a question without building that information into the model through training.

THE SIMPLE DISTINCTION

RAG brings information to the model for a response. Training changes the model itself.

HOW RAG WORKS AT QUERY TIME

- 1 ASK**
A user asks a question or gives the AI a task.
- 2 FIND**
The system searches connected sources for relevant information.
- 3 ADD CONTEXT**
Relevant information is sent to the AI with the user's question.
- 4 ANSWER**
The AI uses the retrieved information to help create its response.
- 5 TRACE THE SOURCE**
Citations, links, or source details can help users check where information came from.

WHAT RAG DOES NOT DO

IT DOESN'T CHANGE THE MODEL

Retrieved information helps answer the current question. The model itself stays the same.

IT DOESN'T AUTOMATICALLY 'LEARN' THE CONTENT

A source does not become part of the model simply because it was retrieved for a response.

IT DOESN'T DETERMINE HOW DATA IS HANDLED

RAG describes how information is retrieved and supplied to an AI system. Data handling, retention, logging, and downstream use depend on the application and provider.

RAG VS. MODEL TRAINING

COMPARISON POINT	RAG / RETRIEVAL	MODEL TRAINING
HOW IS CONTENT USED?	Used to help answer the current question or task.	Used to teach or adjust the model.
WHEN IS CONTENT USED?	When a user asks a question.	During a separate process used to build or update the model.
DOES THE MODEL CHANGE?	No. Retrieval leaves the model unchanged.	Yes. Training changes the model based on training data.
HOW IS IT KEPT CURRENT?	Sources can be updated without retraining the model.	Changing what the model has learned requires additional training.

WHY THE DISTINCTION MATTERS TO ACADEMIC TECHNOLOGY TEAMS

INSTITUTIONAL CONTROL

Universities need visibility into how AI platforms access and use institutional and licensed content, including retention and downstream use.

GOVERNANCE

Retrieval can preserve access controls, source attribution, and links back to authoritative systems. Those controls are part of the application design, not the model itself.

ARCHITECTURE & ACCESS

A retrieval service can sit between an AI application and institutional or licensed sources, applying authentication, access rules, filtering, and source selection before sending information to the AI.

FRESHNESS

Because the knowledge stays outside the model, sources can be updated independently. The model does not need to be retrained every time the knowledge base changes.

HOW RAG SUPPORTS ACADEMIC INTEGRITY

A retrieval-based service can make university library databases available within an AI workflow without using that content to retrain the model. It can verify access rights, return relevant research, and preserve attribution and links to the original source.

By limiting the LLM's data sources to the university's licensed library databases, retrieval supports factual accuracy and traceable, verifiable citations. This benefits end users, the library, and the institution's investment in academic technology.