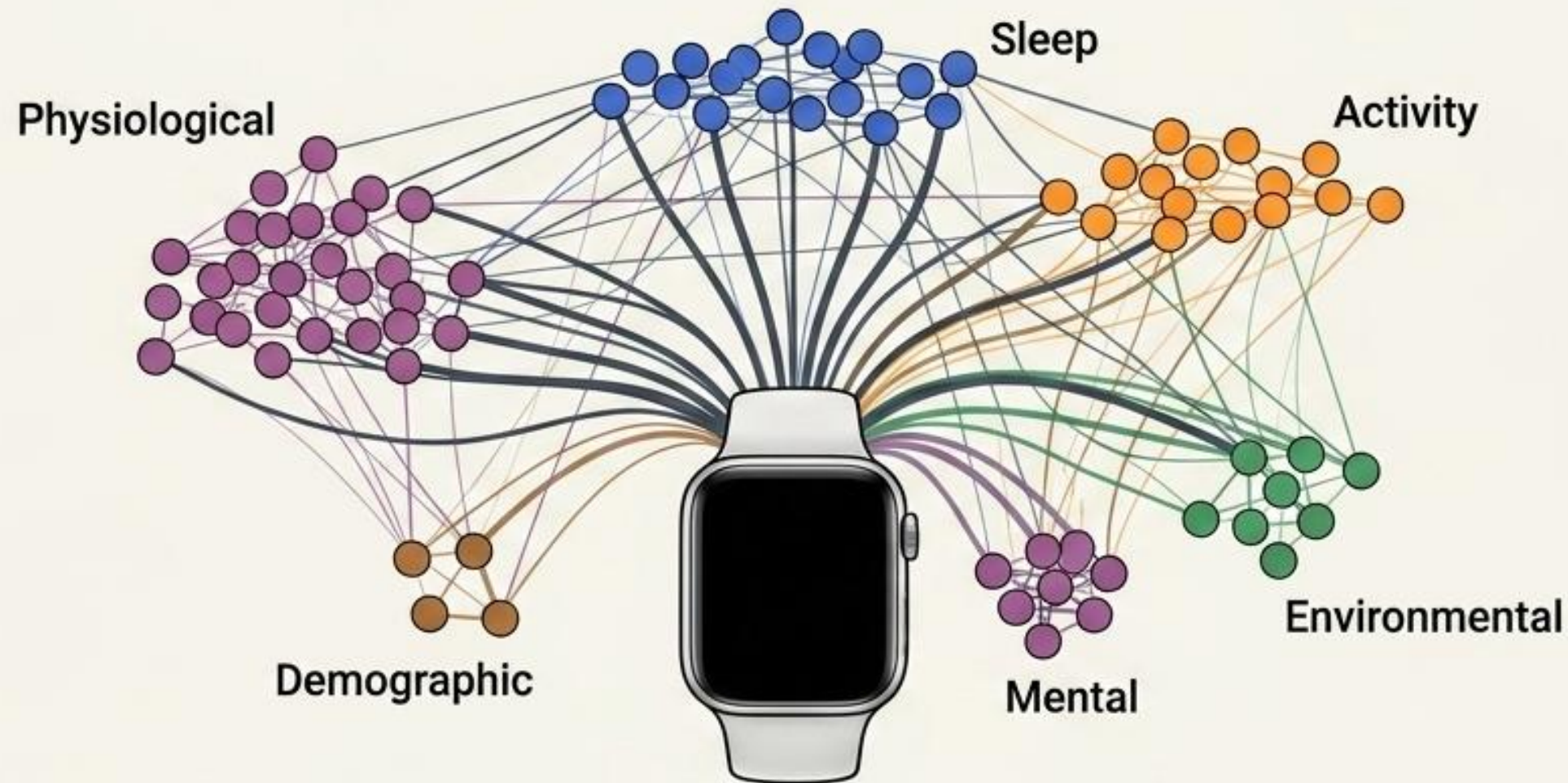


# Wearable As Graph: Query-Conditioned Context Retrieval



**A dynamic, graph-based framework enabling contextualized LLM reasoning over long-term, multimodal wearable data.**

Lu, Abbasian, Rahmani  
CMU & UC Irvine

## The Data Scale

Modern wearables generate continuous, multi-modal time-series data over months or years.

## The Data Scale

Modern wearables generate continuous, multi-modal time-series data over months or years.

## The LLM Limit

Feeding all historical data exceeds context windows, skyrocketing compute costs and inference time.

## The LLM Limit

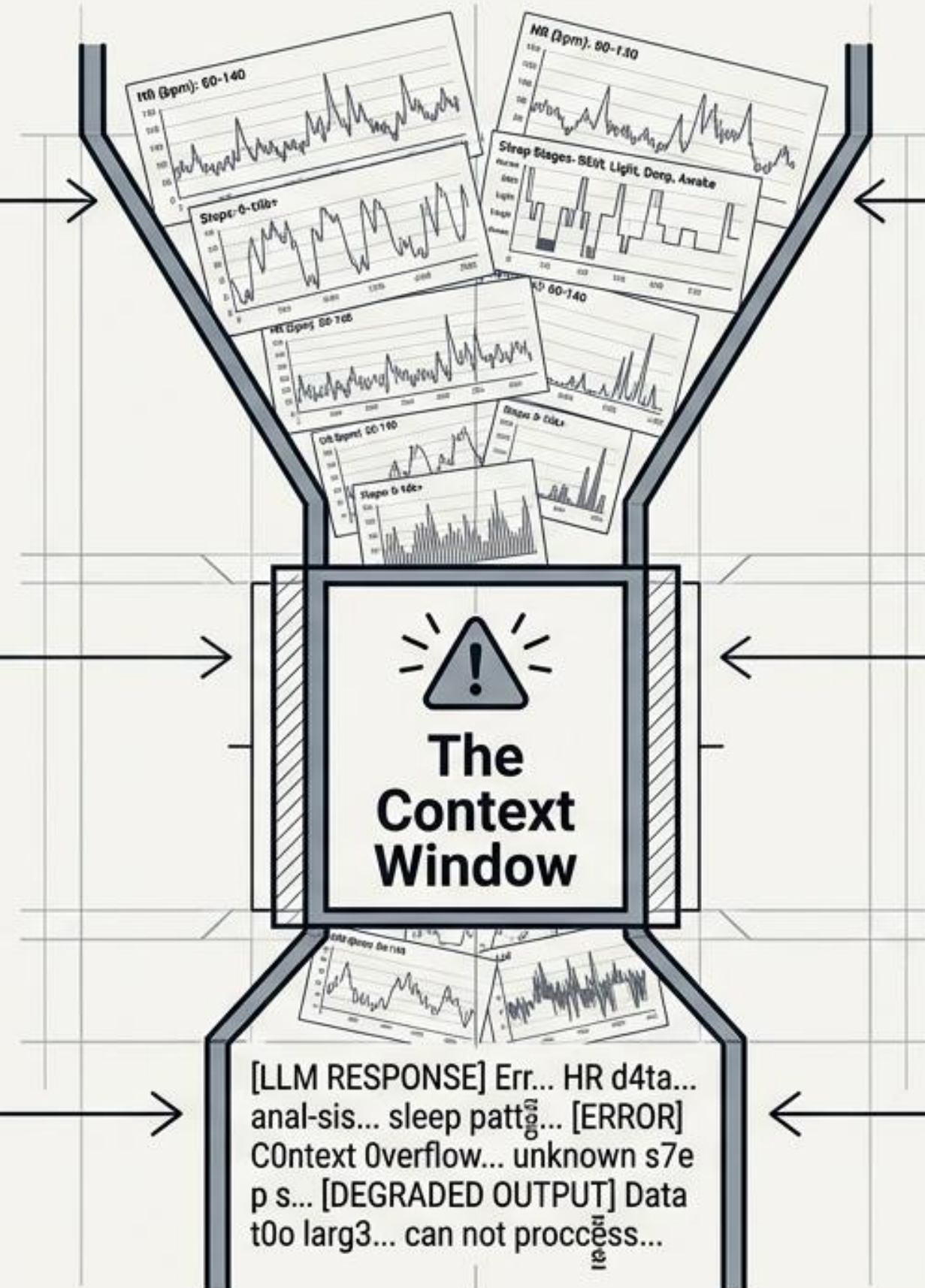
Feeding all historical data exceeds context windows, skyrocketing compute costs and inference time.

## The Quality Drop

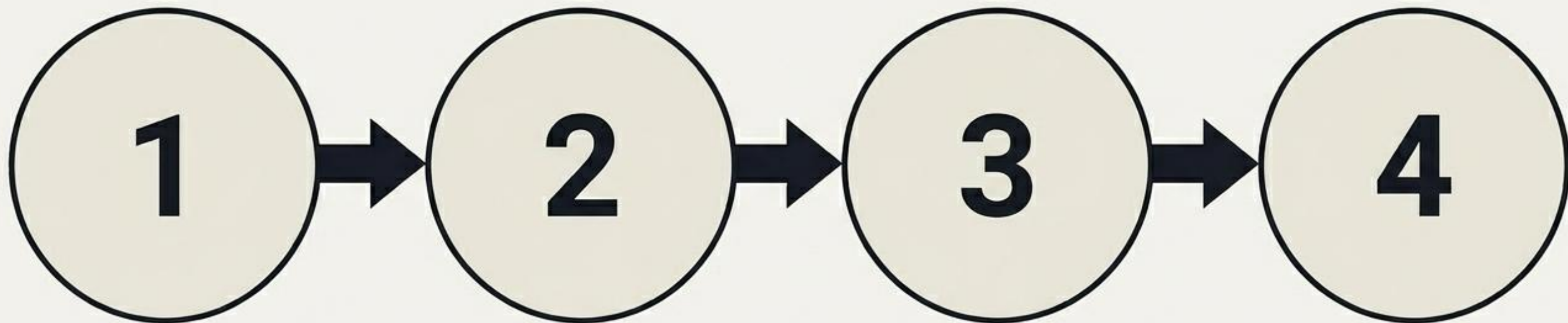
Massive, uncurated context causes LLMs to suffer from degradation, severely hurting analytical reasoning.

## The Quality Drop

Massive, uncurated context causes LLMs to suffer from degradation, severely hurting analytical reasoning.



	Standard Baseline	Standard RAG	WAG (Our Approach)
Context Handling	Raw data only	Isolated text chunks	Interconnected Subgraphs
Relationship Awareness	None	Keyword-matched local data	Global + Local dynamic relationships
Temporal Sensitivity	Static	Static text retrieval	Dynamically re-weighted via recent anomalies



## Query Parsing

Extracts target metrics, time window, and an exploratory openness score.

## Graph Construction

Merges universal medical knowledge with personalized health data streams.

## Dynamic Retrieval

Adjusts edge weights based on long-term memory (Global) and short-term reactivity (Local).

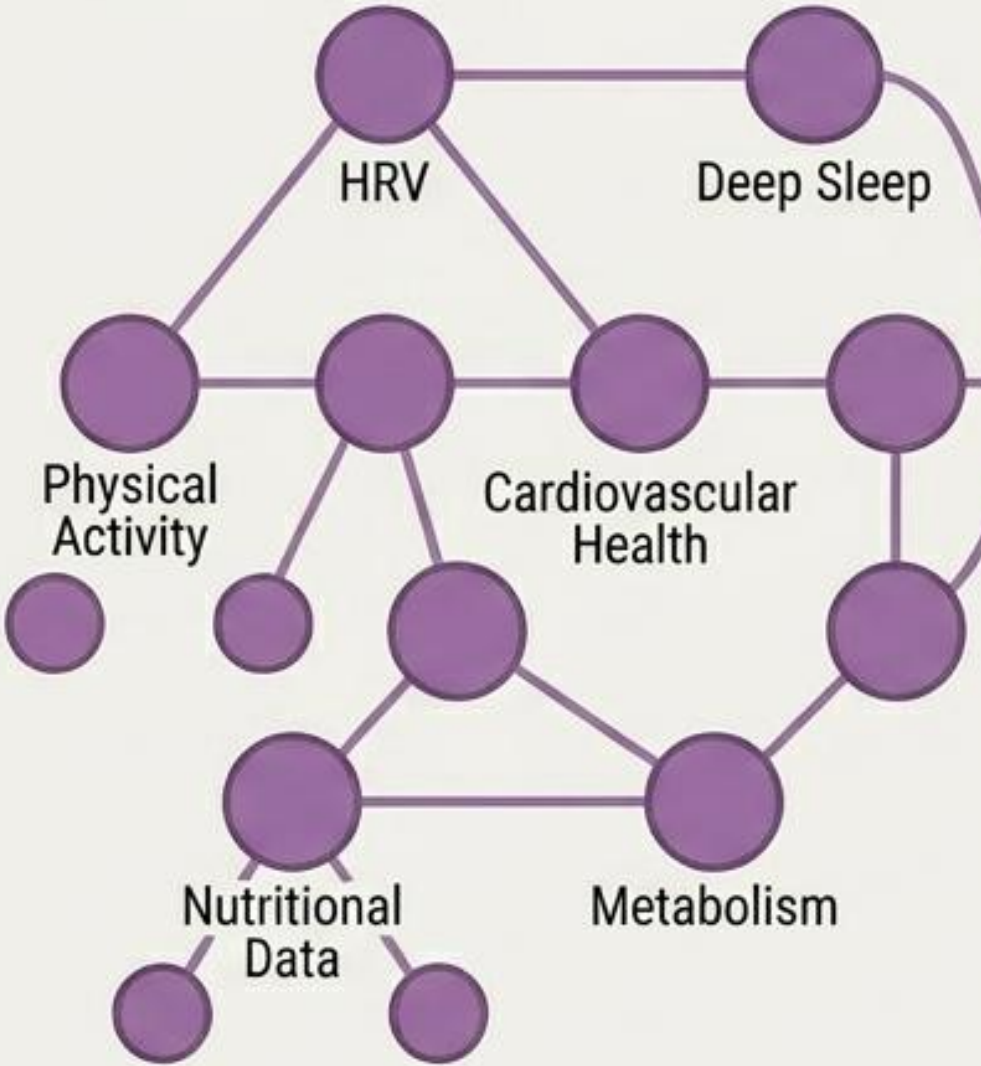
## Augmented Generation

Feeds the highly tailored, query-conditioned subgraph into the LLM for grounded reasoning.

# Graph Construction Synthesis

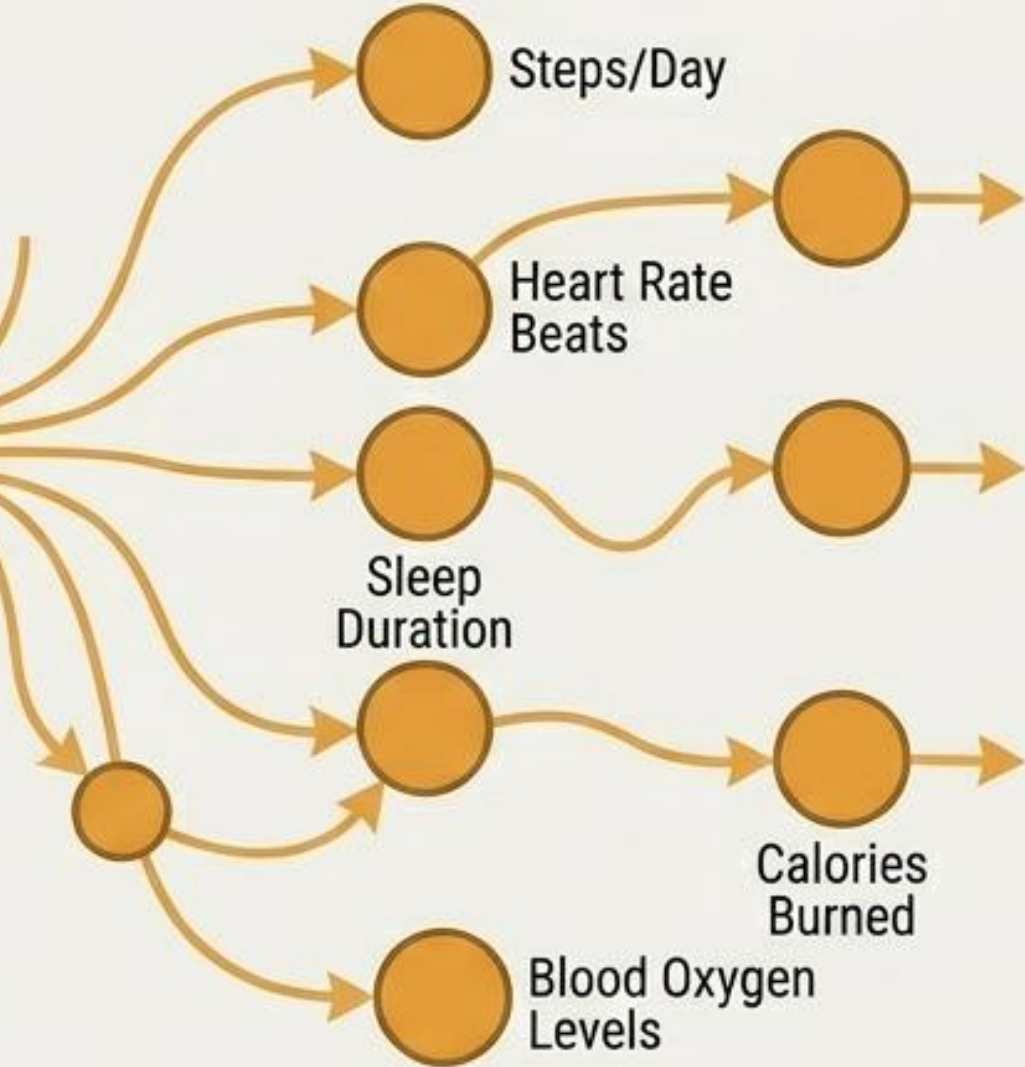
## General Medical Knowledge

WAG maps established medical truths from Web, DBs, and scientific literature (e.g., HRV impacts Deep Sleep).

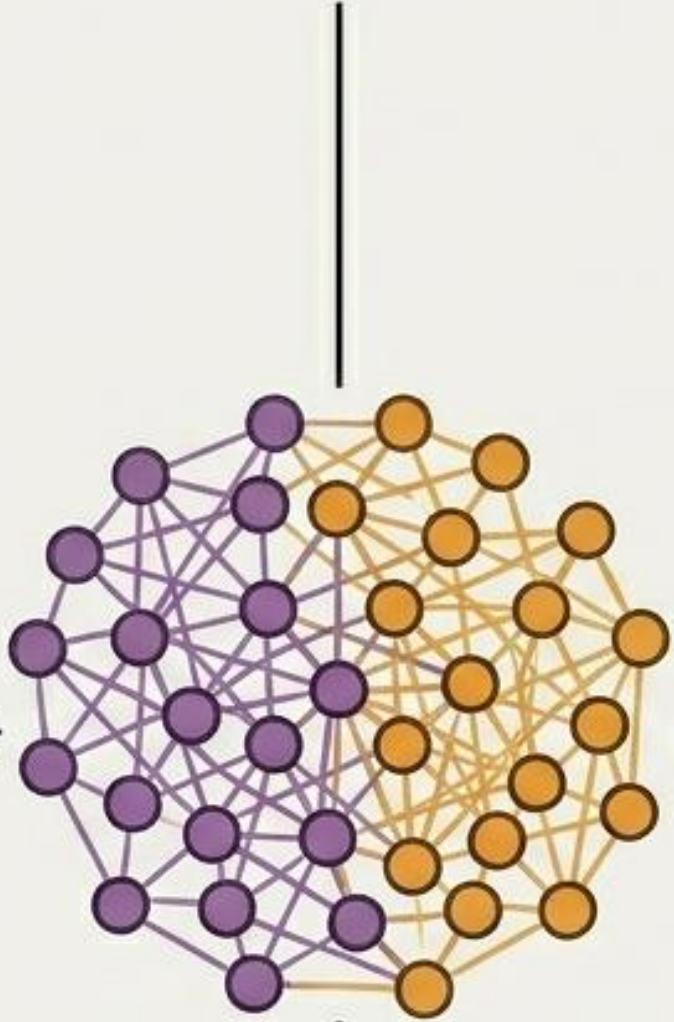


## Personal Data Streams

Individual user datasets from wearables and lifelogging.



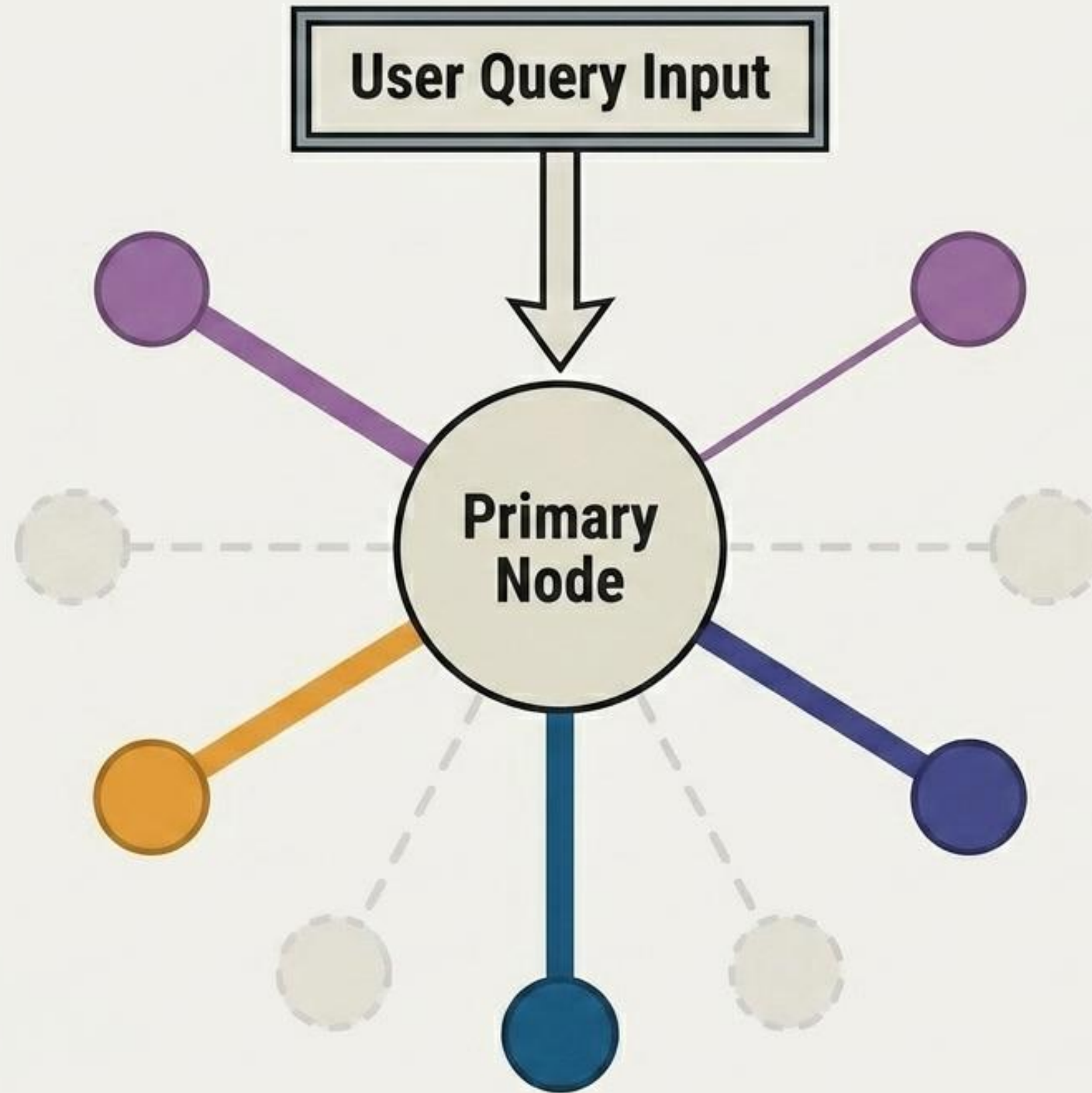
**The Personalized Knowledge Graph (PKG)**  
Universal medical truths are continuously conditioned by the user's actual physiological baselines.



## The Trigger

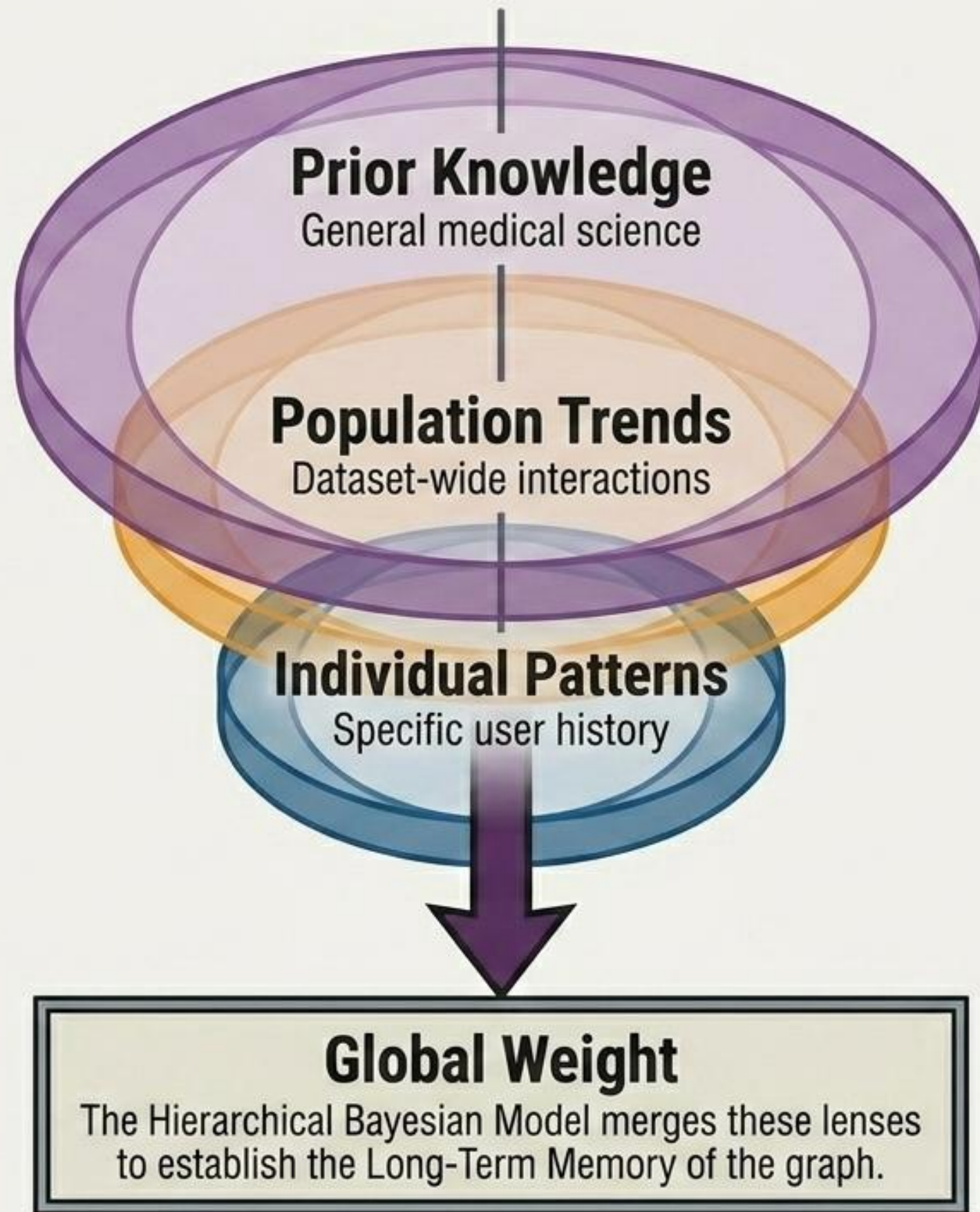
Every query is mathematically parsed for three variables:

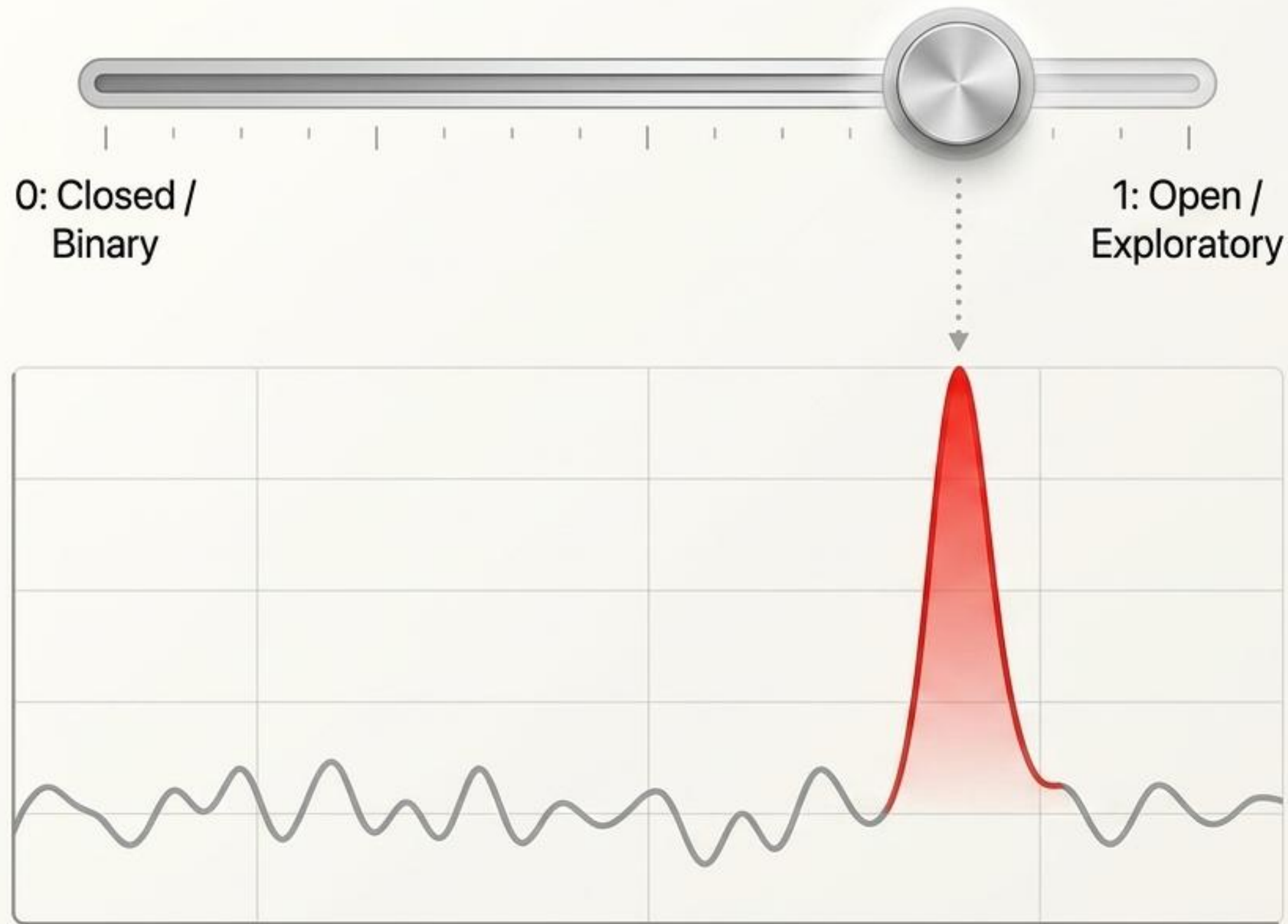
- Target Metrics
- Time Window
- Query Openness



## The Reaction

WAG dynamically redraws the map of what matters right now. It mathematically strengthens connections to relevant secondary metrics and fades irrelevant noise.





## The Openness Dial

Phrasing matters.  
Exploratory queries (e.g.,  
"Why am I stressed?")  
crank up the openness  
dial.

This amplifies the graph's  
sensitivity to recent  
anomalies (Z-scores),  
pulling in a broader set of  
related symptoms for the  
LLM to analyze.

**Long-Term Wisdom**  
Global Modeling

**Short-Term Reactivity**  
Local Modeling

$$\text{Final Weight} = (1 - \beta) * \text{Global} + \beta * \text{Local}$$



### **The Synthesis**

By combining deep historical medical baselines with hyper-sensitive recent anomaly detection, WAG constructs the perfect, custom-tailored context window—eliminating noise and hallucination risks.

## 1. Input

Why are my circadian rhythm patterns off over the past 30 days?

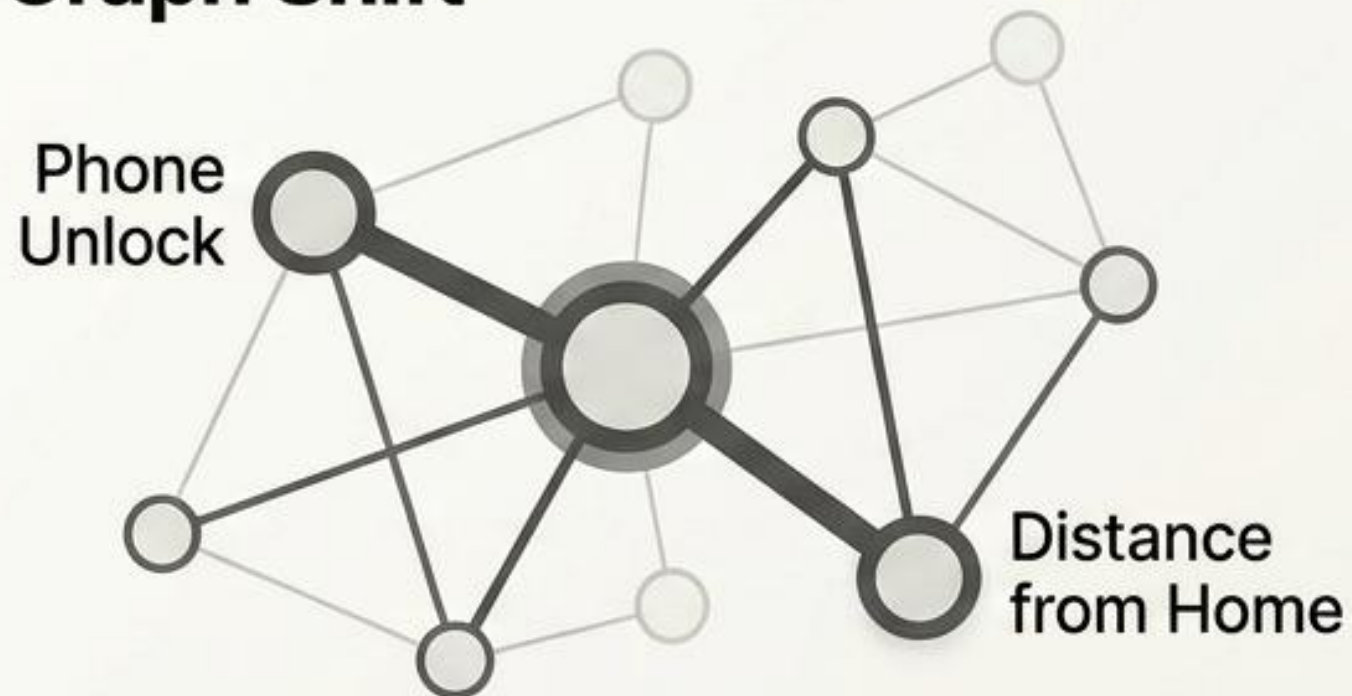
## 2. Extraction

Circadian Rhythm

Openness = 0.8

Window = 30 Days

## 3. Graph Shift



## 4. Insight

“ Because the query was highly exploratory, the graph expanded its search, identifying that late-night screen time and travel were the root causes of the physiological anomaly. ”



**Globem**



**Lifesnaps**

# 10,341

Data-Grounded Queries  
Evaluated



**IFH Affect**



**Pmdata**

## Dual-Evaluation Methodology



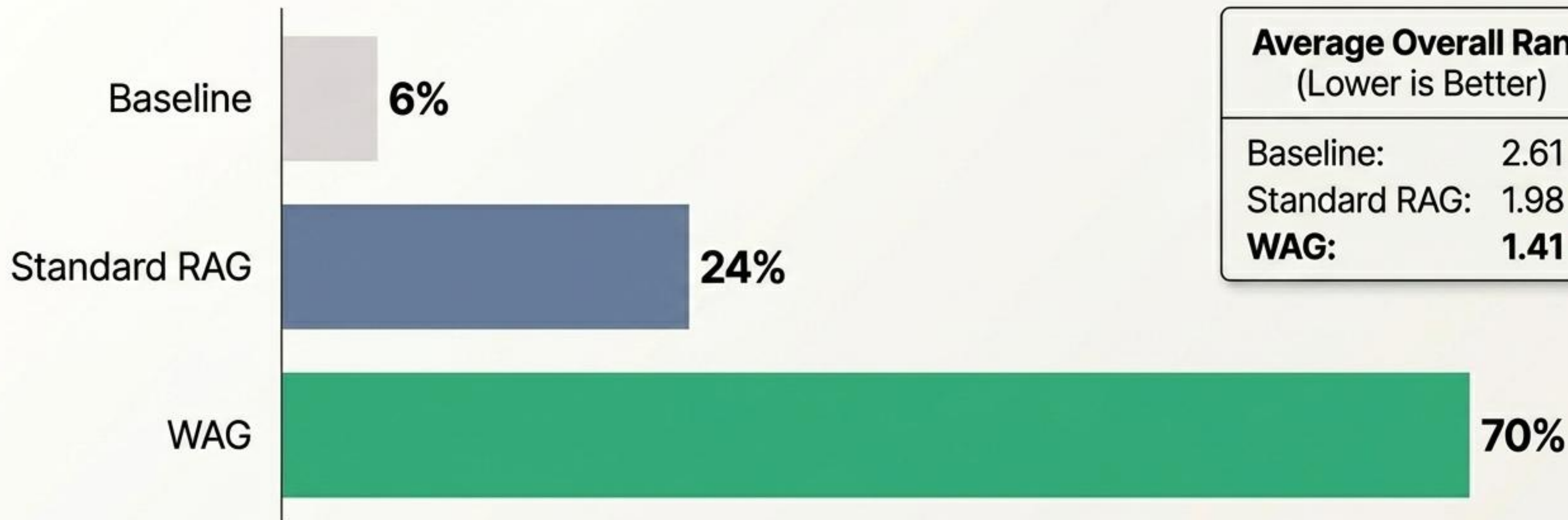
Advanced LLM  
Evaluation



Human Domain  
Experts

Multi-Dimensional Grading: Insightfulness, Relevance, Groundedness, Personalization, Clarity, Safety.

# Absolute Win Rate



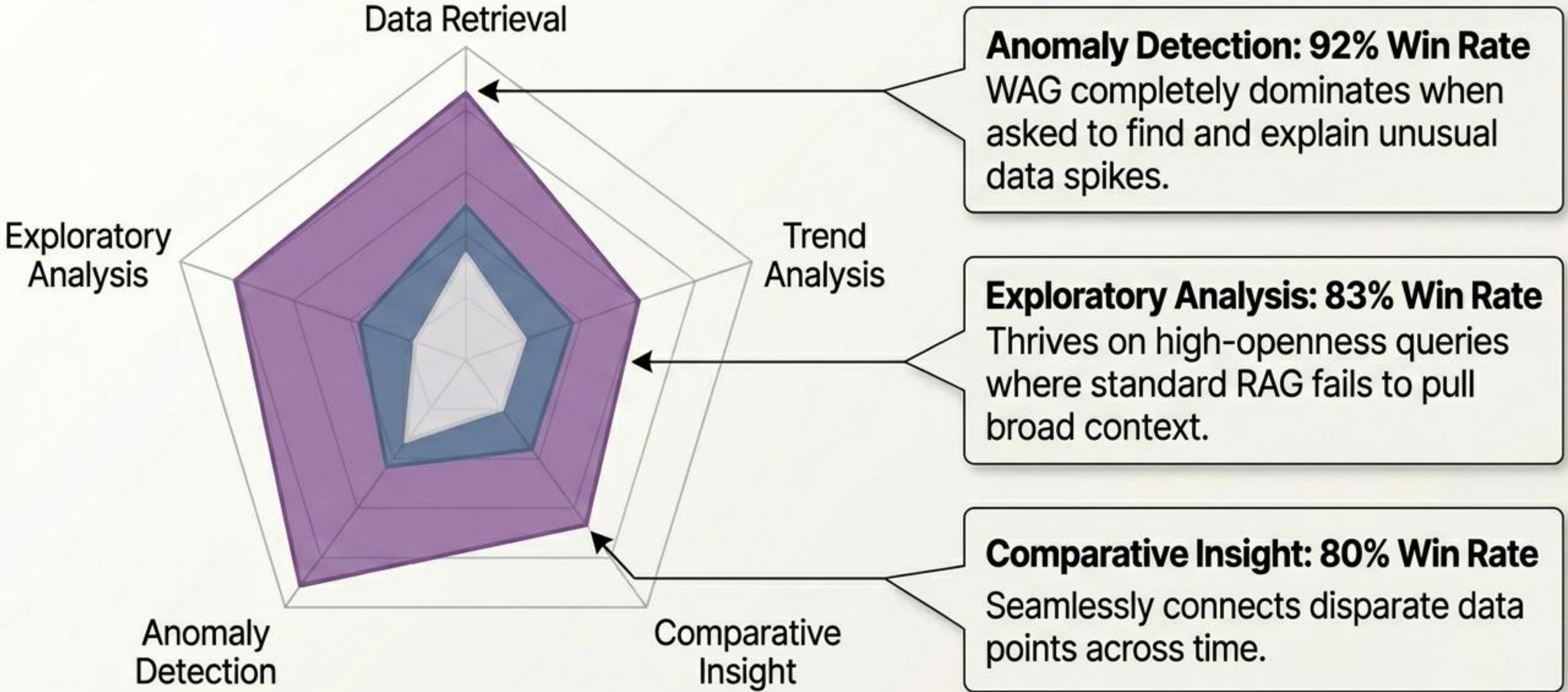
**Average Overall Rank**  
(Lower is Better)

Baseline:	2.61	↓
Standard RAG:	1.98	↓
<b>WAG:</b>	<b>1.41</b>	▼

## The Verdict

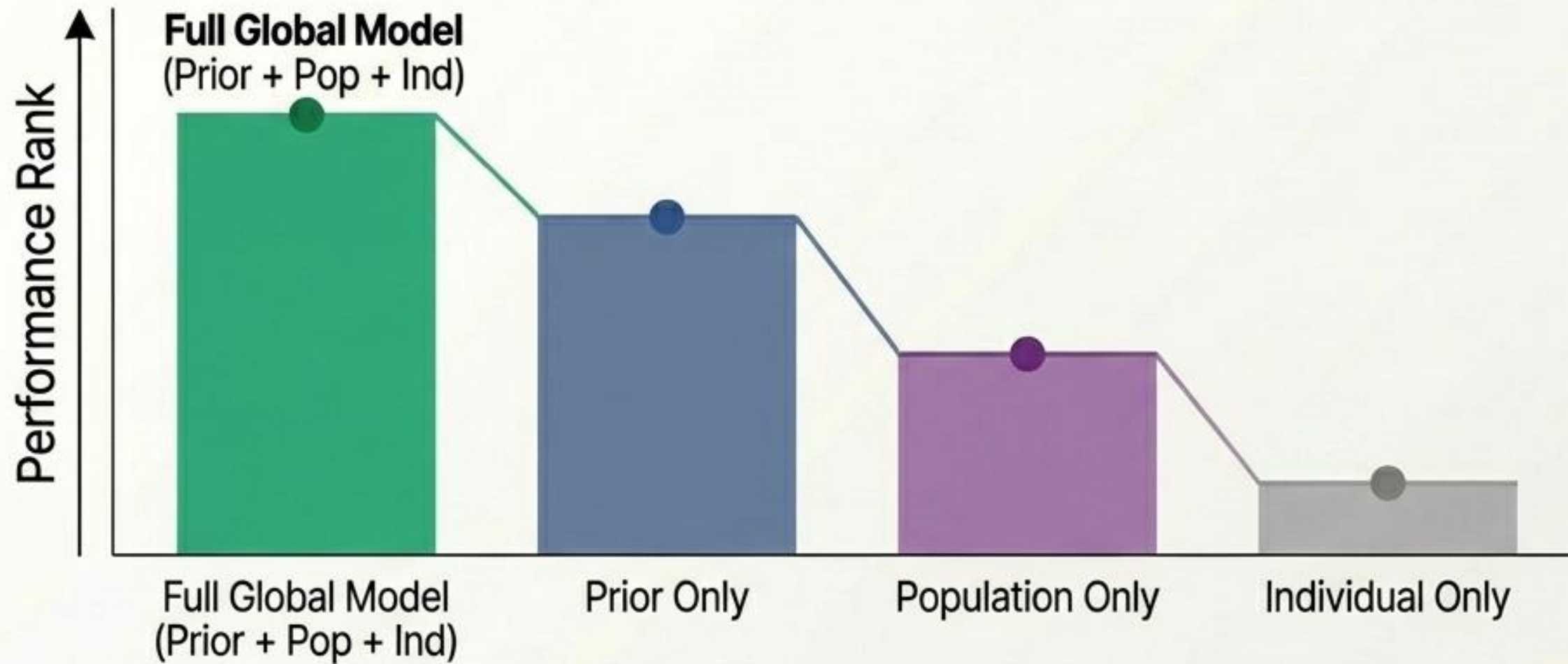
WAG represents a **massive 56% improvement** over **standard** RAG methodologies, proving that flat text retrieval is insufficient for multimodal health data.

# WAG Dominates on Complex Analytical Queries

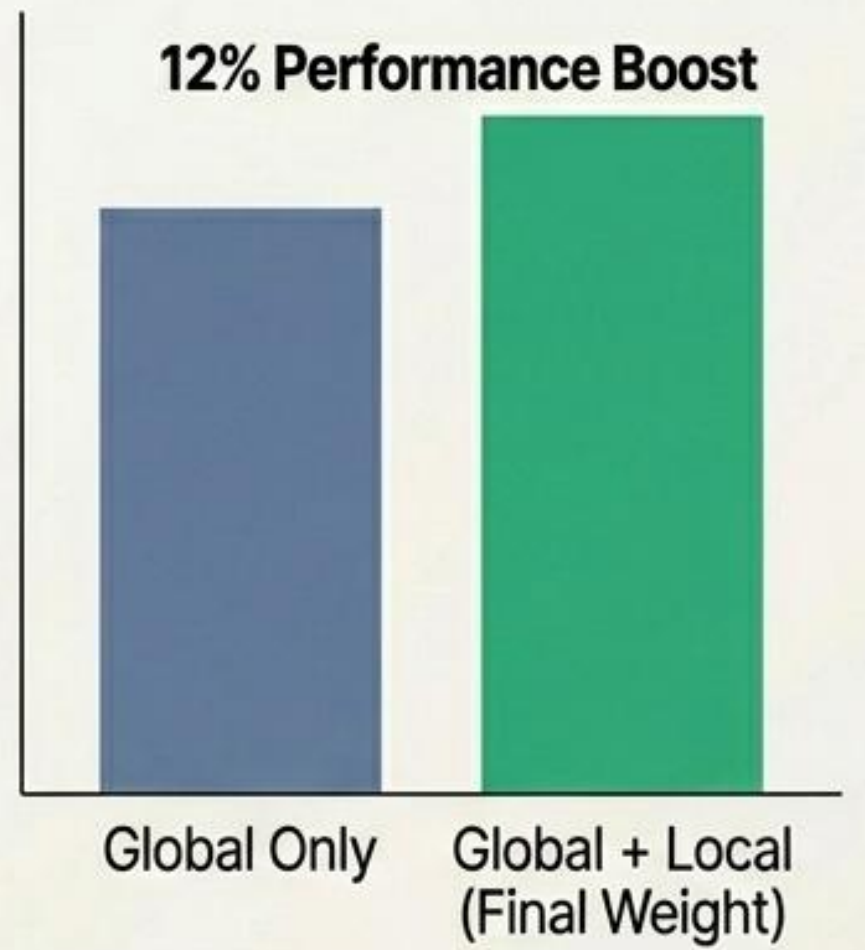


# Ablation Performance Analysis

## Component Contribution to Performance Rank



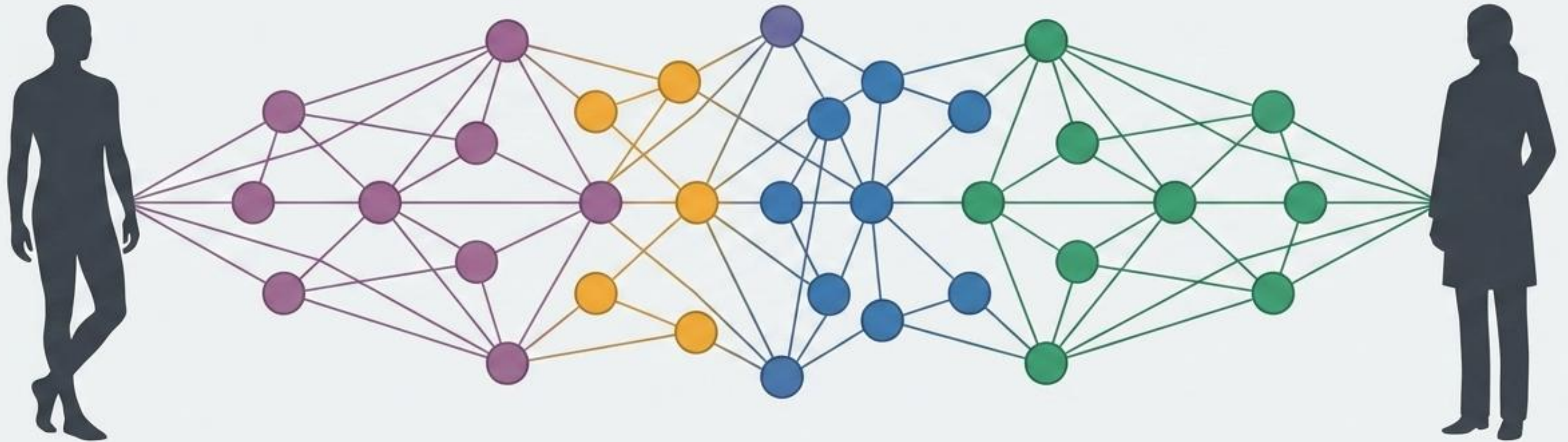
## Global vs. Global + Local



### Every Piece Matters

Combining Prior, Population, and Individual patterns mathematically outperforms any single-source weighting strategy. Adding short-term reactivity (Local) provides a consistent **12% performance boost** over relying on long-term relationships alone.

# The Future of Explainable Health AI



## Co-Management

WAG serves as a flexible, dynamic interface where clinical expertise and lived human experience co-exist in the exact same mathematical space.

## The Result

Deeply personalized, highly grounded LLM analysis without hallucination—unlocking the true predictive power of wearable ecosystems.