

# RAG & SEARCH



THE #1 AI TRAINING  
CONFERENCE



Amy Hodler

David Hughes



SAN FRANCISCO | OCT 28-30

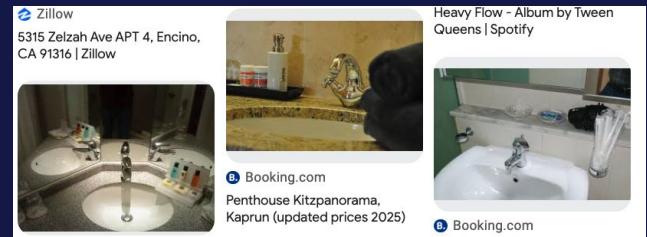
# BEYOND TEXT: BUILDING & EVALUATING MULTIMODAL GRAPH RAG



# AI Doesn't "Get It"



**Google search** gave us  
a bunch of **sink** photos



**Claude** thought it was a  
**joke about T-Rex** having  
short arms  
Both missed the toothbrush

# LLMs & GraphRAG

## We're Supposed to Give Us the Answers!



**Today's GraphRAG** is good at adding guardrails but **misses**:

Latent image data and Interwoven data types  
 $(1+1 = 3)$

Associative detail

Lacks Interpretability

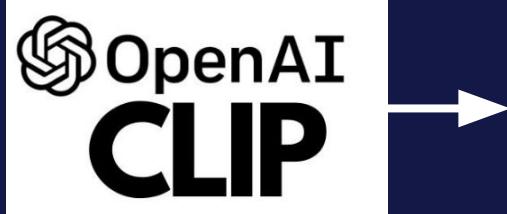
# Multimodal GraphRAG provides Associative Intelligence



**mmGraphRAG breaks down components that can be explored (texture, spatial placement, etc) individually or together**

**Blending of semantic context and data (text, visual, etc) enables reasoning across multiple levels of abstraction and associations**

# The Role of Images in RAG

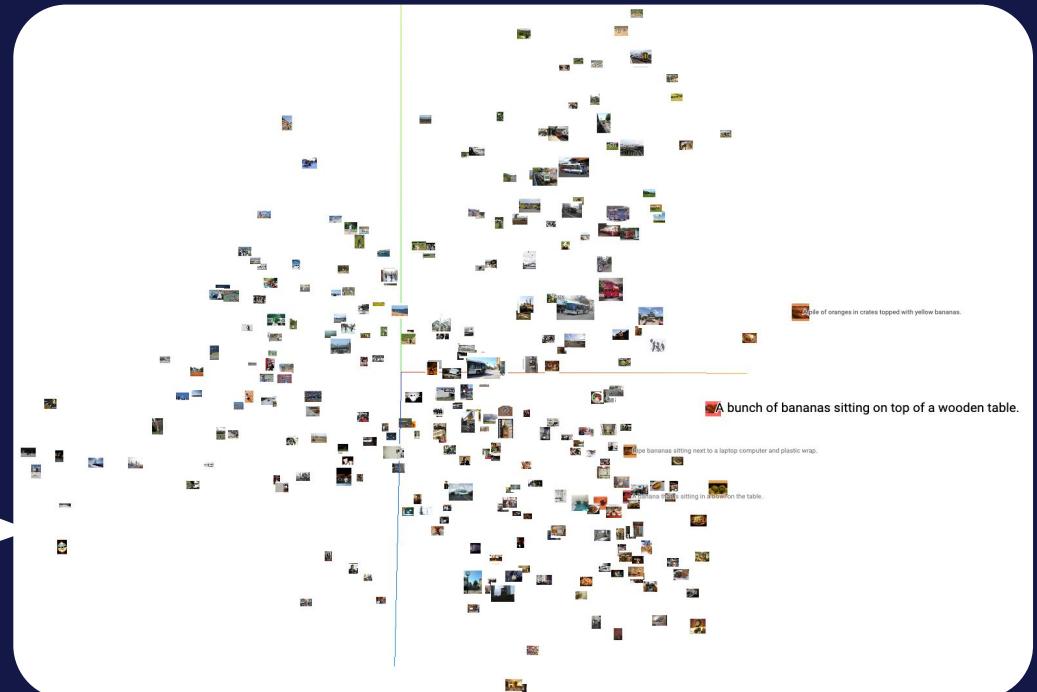


A toy dinosaur...

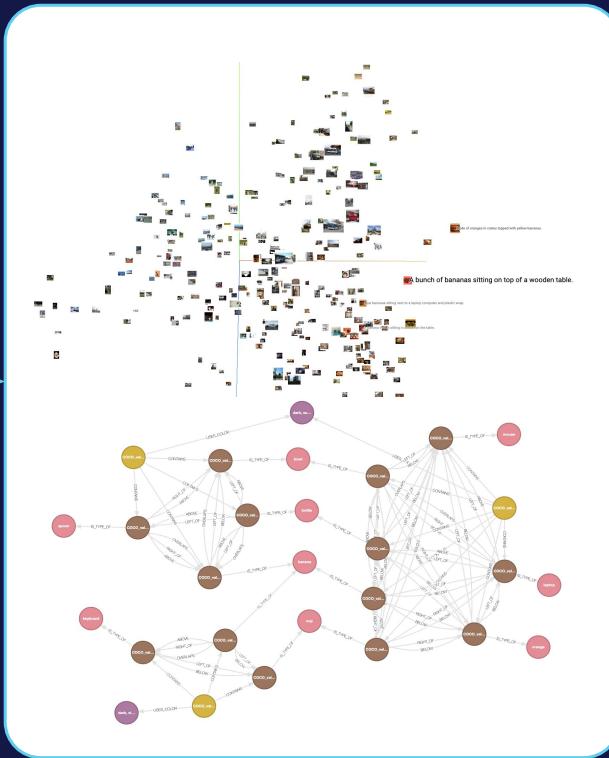
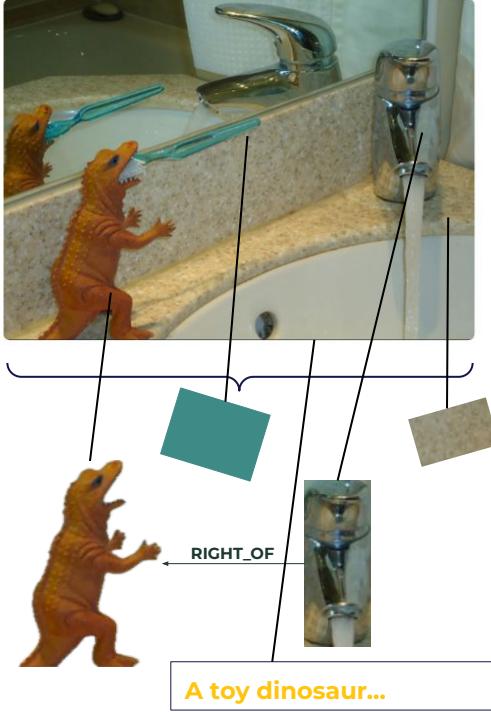
# The Role of Images in RAG



A toy dinosaur...



# Multimodal Search: Semantics + Reasoning!?



Search Images

Find images with basins

Search

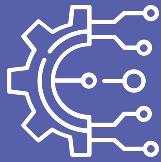
Image with sink

These images contain sinks located in bathrooms

Image with sink

# Next Several Mins...

## Advancing Graph RAG



Demonstrate how  
**mmGraphRAG** works



Highlight the integration of  
**vision and audio models** with  
graphs



Introduce **BAML** for AI  
Development



**David  
Hughes**



**Amy  
Hodler**

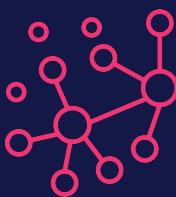
# Multimodal Graph RAG

# mmGraphRAG in a Nutshell



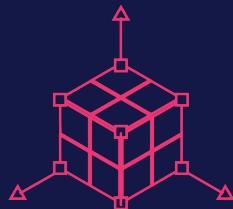
## Decomposition

Break down images/audio:  
create subcomponents



## Graph

Components and subs are loaded into a graph



## Embedding Space

Space with anchors created for global search in the graph



## Multimodal Context

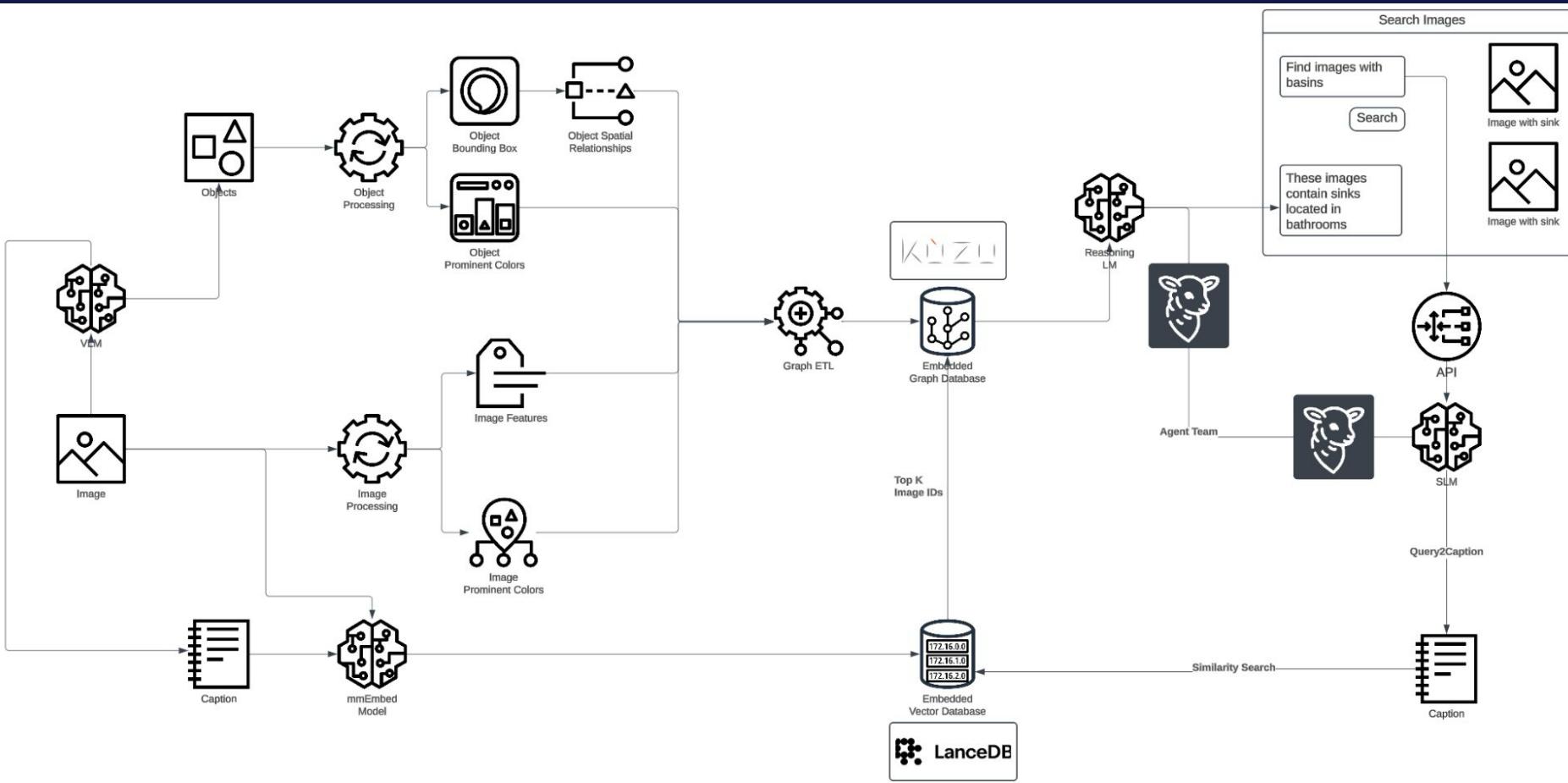
Provide blended results to LLMs for full context of text and images

Generate captions

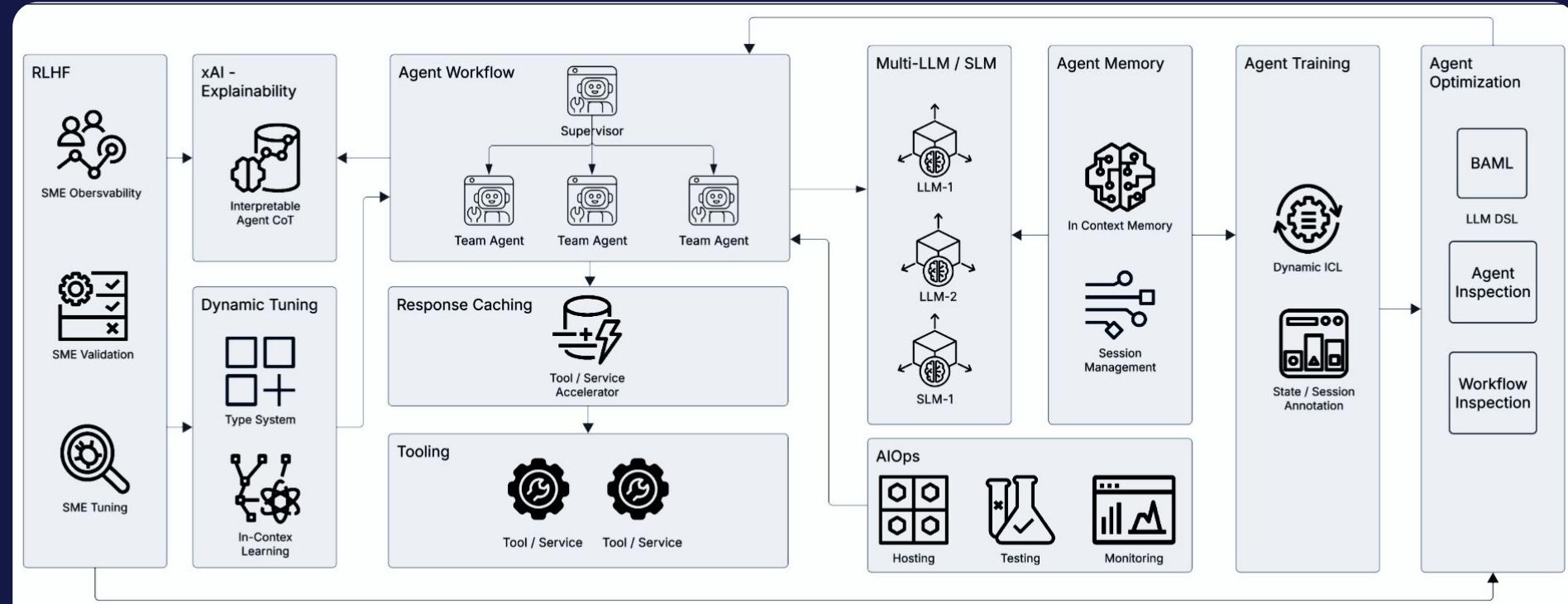
Preserve context

Basis of semantic search

Associative Intelligence



# (:Agents)-[:ORCHESTRATED\_IN]→(:Lamb)



git.new/ODSC-East-25

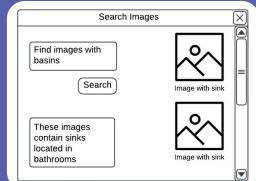
## **Follow Along**

### Notebook for Associative Search

DAVID

# Architectural Layers

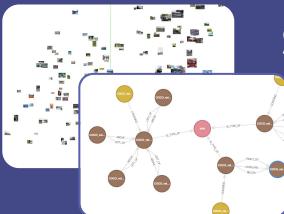
Data Sources  
Presentation  
Apps



## Query and Explain

- Vector similarity searches on images
- Graph queries retrieve contextual details on similar images
- LLMs generate explanations linking results to user queries

Semantics



## Semantic and Graph Construction

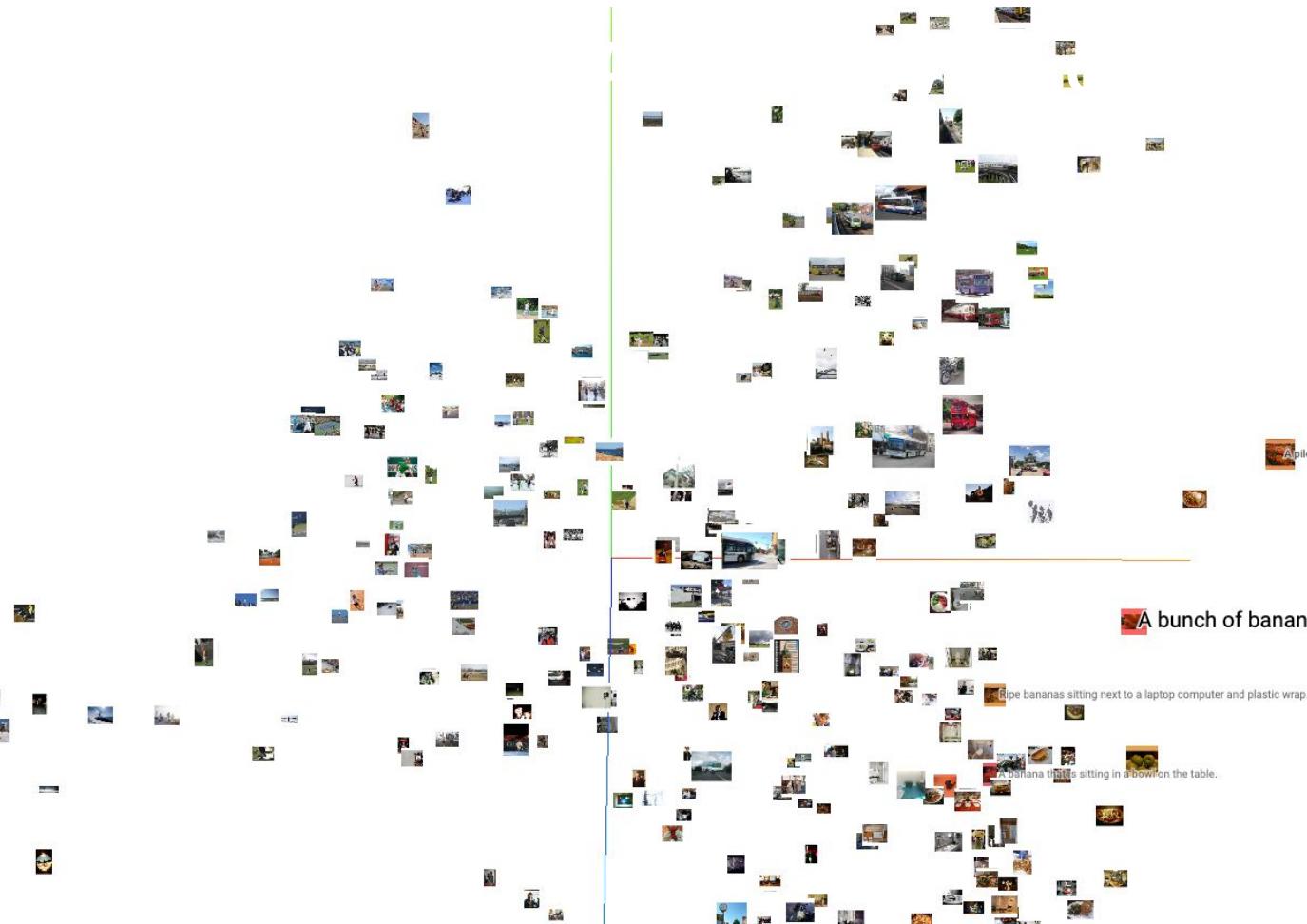
- Load hypervectors & relationships into vector & graph DB
- Connect nodes representing image features, metadata, & text-based context



## Data Extraction and Embedding

- Visual features extracted using pretrained models
- Features converted into hypervectors for high-dimensional encoding

# The Semantic & Graph Layer



 A pile of oranges in crates topped with yellow bananas.

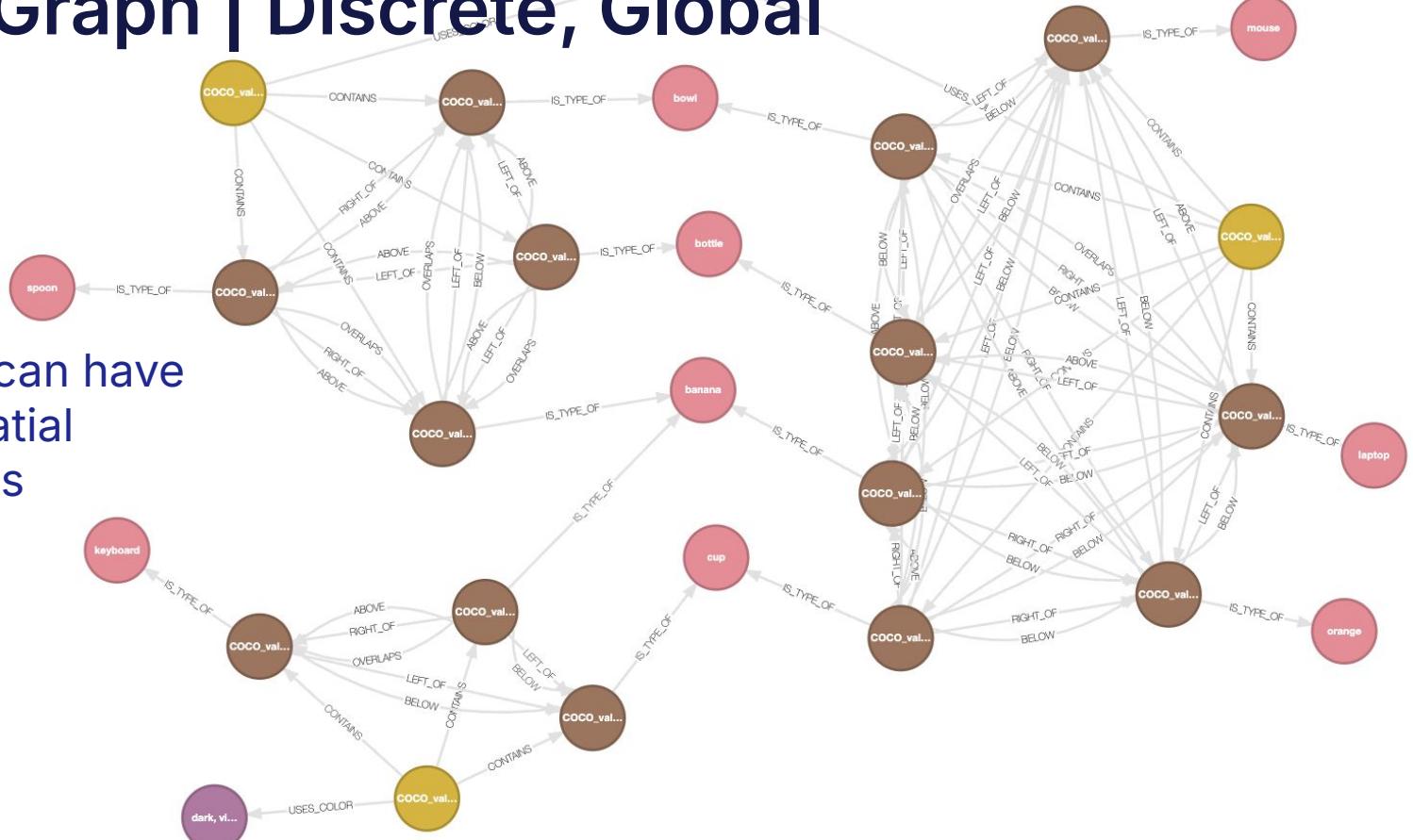
 A bunch of bananas sitting on top of a wooden table.

 Ripe bananas sitting next to a laptop computer and plastic wrap.

 Banana thins sitting in a bowl on the table.

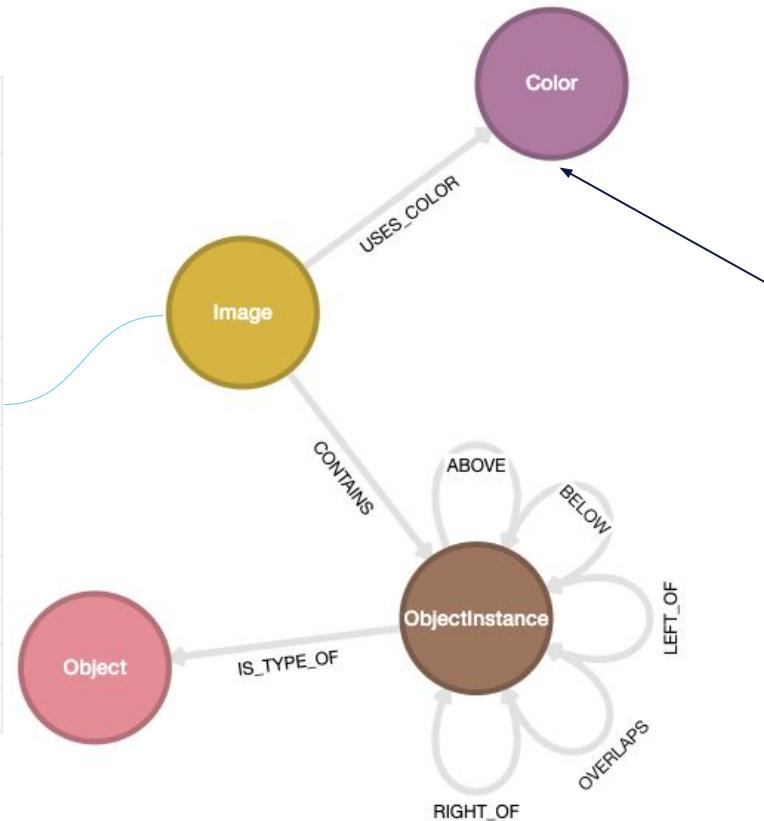
# Image Graph | Discrete, Global

Each node can have  
multiple spatial  
relationships



# Image Graph

<b>id</b> <small>PK</small>	COCO_val2014_000000000208.jpg
<b>name</b>	COCO_val2014_000000000208.jpg
<b>caption</b>	A toy dinosaur standing on a sink next to a running faucet.
<b>complexity_value</b>	12759
<b>complexity_description</b>	highly complex
<b>texture_value</b>	42.90502166748047
<b>texture_description</b>	smooth
<b>sharpness_value</b>	1603.90087890625
<b>sharpness_description</b>	sharp
<b>lightness_value</b>	124.84366607666016
<b>lightness_description</b>	balanced
<b>pattern</b>	geometric patterns



E.g.  
color present,  
major color,  
minor color,  
dominant color  
score

An aerial photograph of several ships on the ocean. In the bottom left foreground, a large aircraft carrier is visible. To its right, a smaller ship is moving away, leaving a white wake. Further back and to the left, another ship is moving towards the viewer. The water is a deep blue.

# High Fidelity Reasoning *within Visual Components*

Are these ships  
moving together or  
apart?

Are there similar  
marked ships nearby?

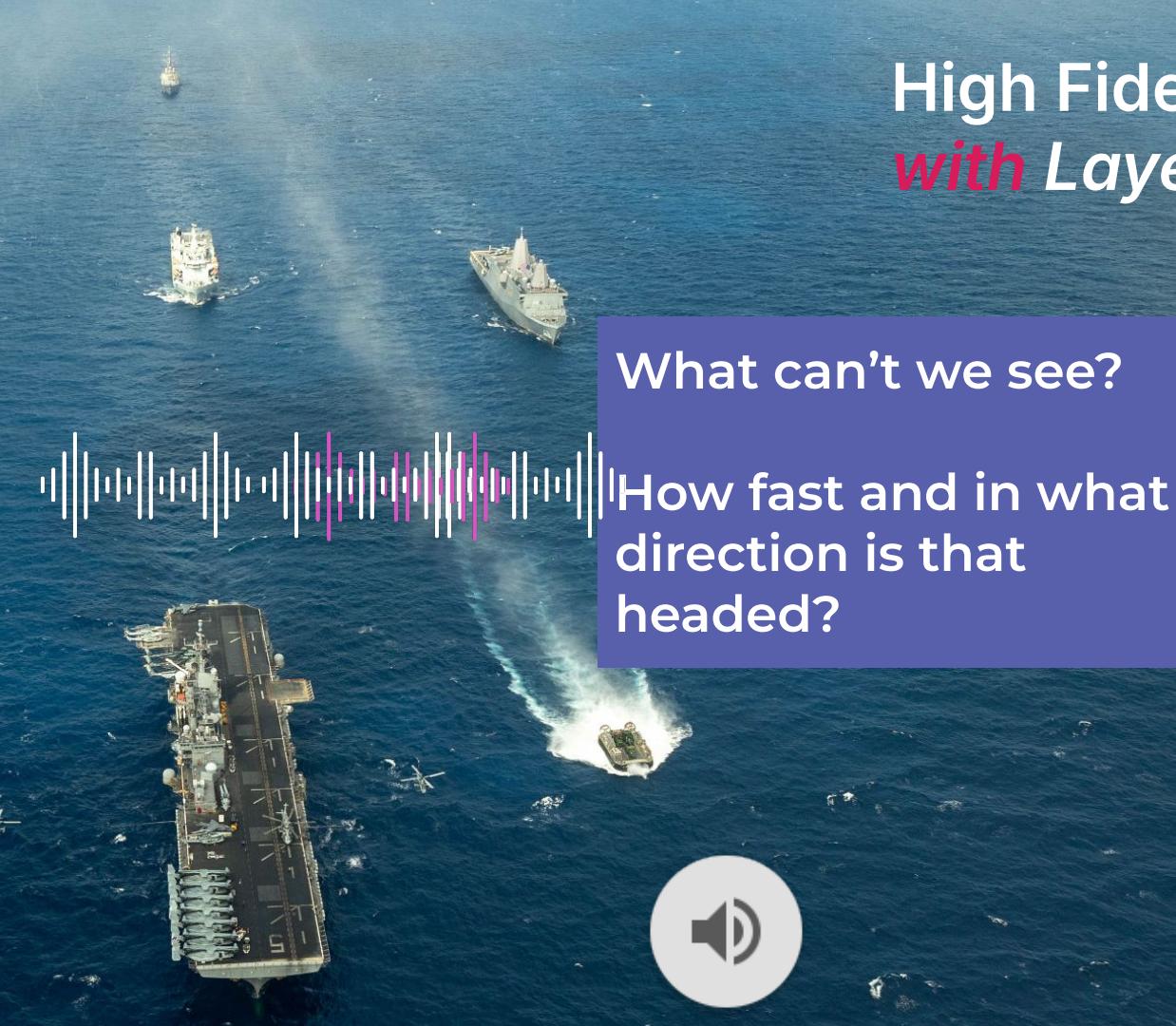
Contextual  
Understanding

Precision Object  
Recognition

Nuanced Similarity

Graph Reasoning

Explainability

An aerial photograph of several ships on the ocean. In the foreground, a large aircraft carrier is positioned vertically. To its right, a smaller boat is moving away from the viewer, creating a white wake. Further back, two more ships are visible: a white supply vessel and a grey naval ship. A prominent feature is a white soundwave graphic on the left side of the slide, composed of vertical bars of varying heights.

# High Fidelity Reasoning *with Layered Audio*

What can't we see?

How fast and in what direction is that headed?



Temporal / Spatial

Directionality & Speed

Off Stage  
Object Segmentation

Contextual  
Understanding

Precision Object  
Recognition

Nuanced Similarity

Graph Reasoning

Explainability

## Source Image Description:

Mediterranean Transit

The amphibious assault ship USS Bataan, dock landing ship USS Carter Hall, amphibious transport dock ship USS Mesa Verde and guided missile destroyer USS Arleigh Burke transit in formation with the British primary casualty receiving ship RFA Argus in the Mediterranean Sea, Feb. 25, 2024. The Bataan Amphibious Ready Group will be transferred to NATO command in the eastern Mediterranean to train with elements of Allied Maritime Command's Standing NATO Maritime Group 2 and the Turkish navy.



# High Fidelity Reasoning with Layered Audio

"name": "helicopter",

"start\_time": 0.0,

"duration": 67.83,

"location": "overlap",

"reasoning": "The helicopter sound is present throughout the entire audio clip, indicating it is located near the recording device.",

"travel\_of\_direction": "going\_forward",

"reasoning": "**The significant Doppler shift (pitch rising then falling)**

**starting immediately indicates the helicopter is moving quickly relative to the observer**, initially approaching and then passing by or moving away. While the specific direction relative to the image is hard to pinpoint from the audio alone, the rapid pitch change suggests a forward motion relative to the audio sensor's perspective.",

"speed": 50.0,

"reasoning": "The rapid onset and prominent Doppler shift in the audio suggest a high speed for the helicopter as it passes close to the observer. A speed of 50 m/s (approx 100 knots) is a reasonable estimate for a helicopter in flight, consistent with the audio characteristics."

},

e": "Columbus Circle & Central Park South".

ption": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.",

ocation": "Columbus Circle, Manhattan, New York City (40.7684° N, 73.9818° W)",

ation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the en from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.",

sis": {

": "Eight police cars with sirens passing from right to left, direction confirmed using stereo and doppler shift",

nt": 8,

siren\_1", "location": "right", "direction": "right\_to\_left"},

siren\_2", "location": "right", "direction": "right\_to\_left"},

siren\_3", "location": "right", "direction": "right\_to\_left"},

siren\_4", "location": "right", "direction": "right\_to\_left"},

siren\_5", "location": "right", "direction": "right\_to\_left"},

siren\_6", "location": "right", "direction": "right\_to\_left"},

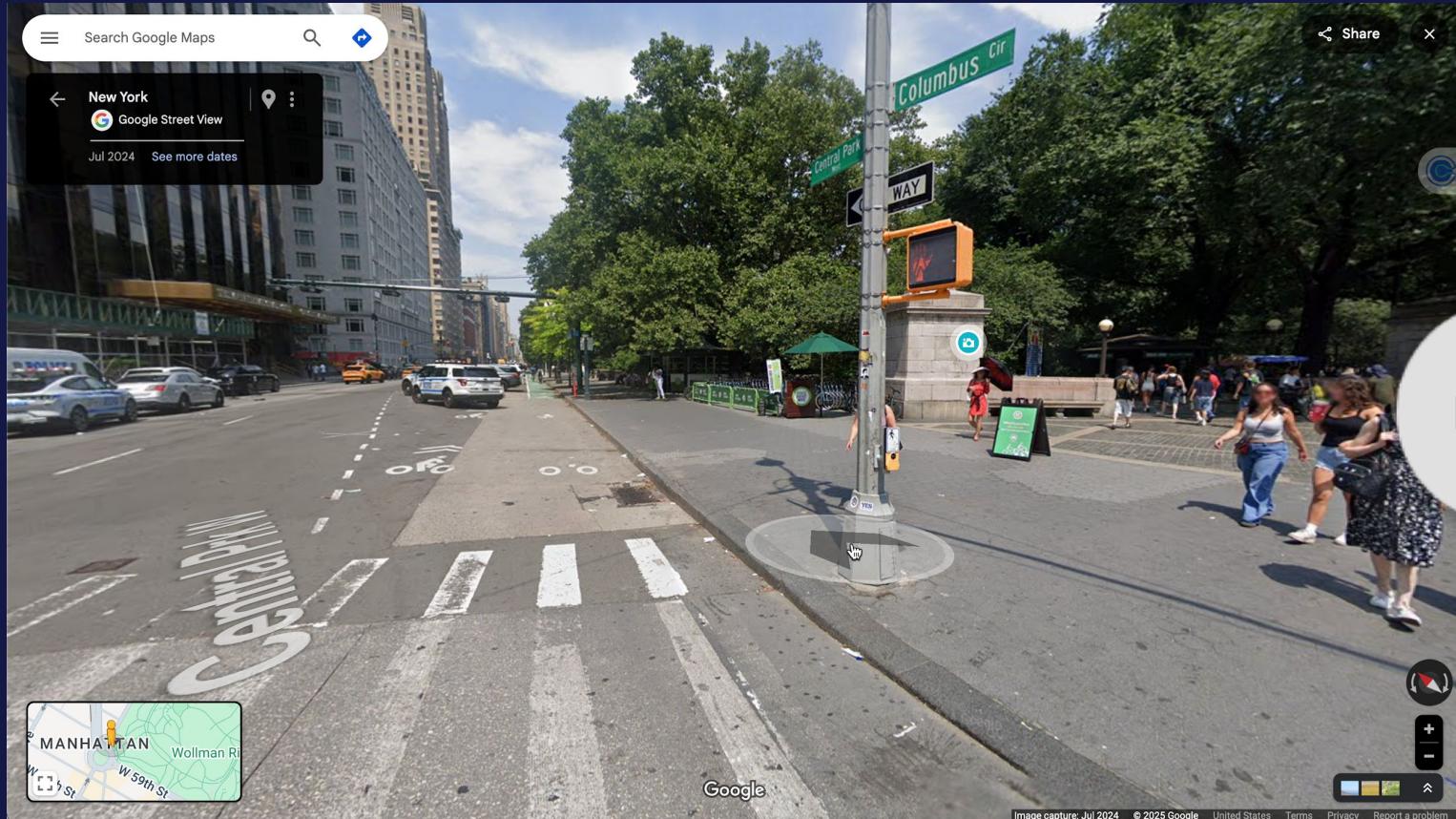
siren\_7", "location": "right", "direction": "right\_to\_left"},

siren\_8", "location": "right", "direction": "right\_to\_left"}

cts\_detected": [



# One More Example for Fun!



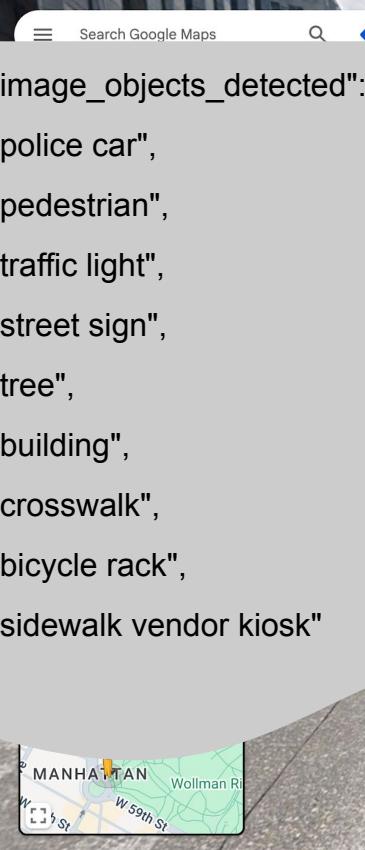
Google

# One More Example for Fun!

```
"image_objects_detected": [  
    "police car",  
    "pedestrian",  
    "traffic light",  
    "street sign",  
    "tree",  
    "building",  
    "crosswalk",  
    "bicycle rack",  
    "sidewalk vendor kiosk"  
]
```



# One More Example for Fun!

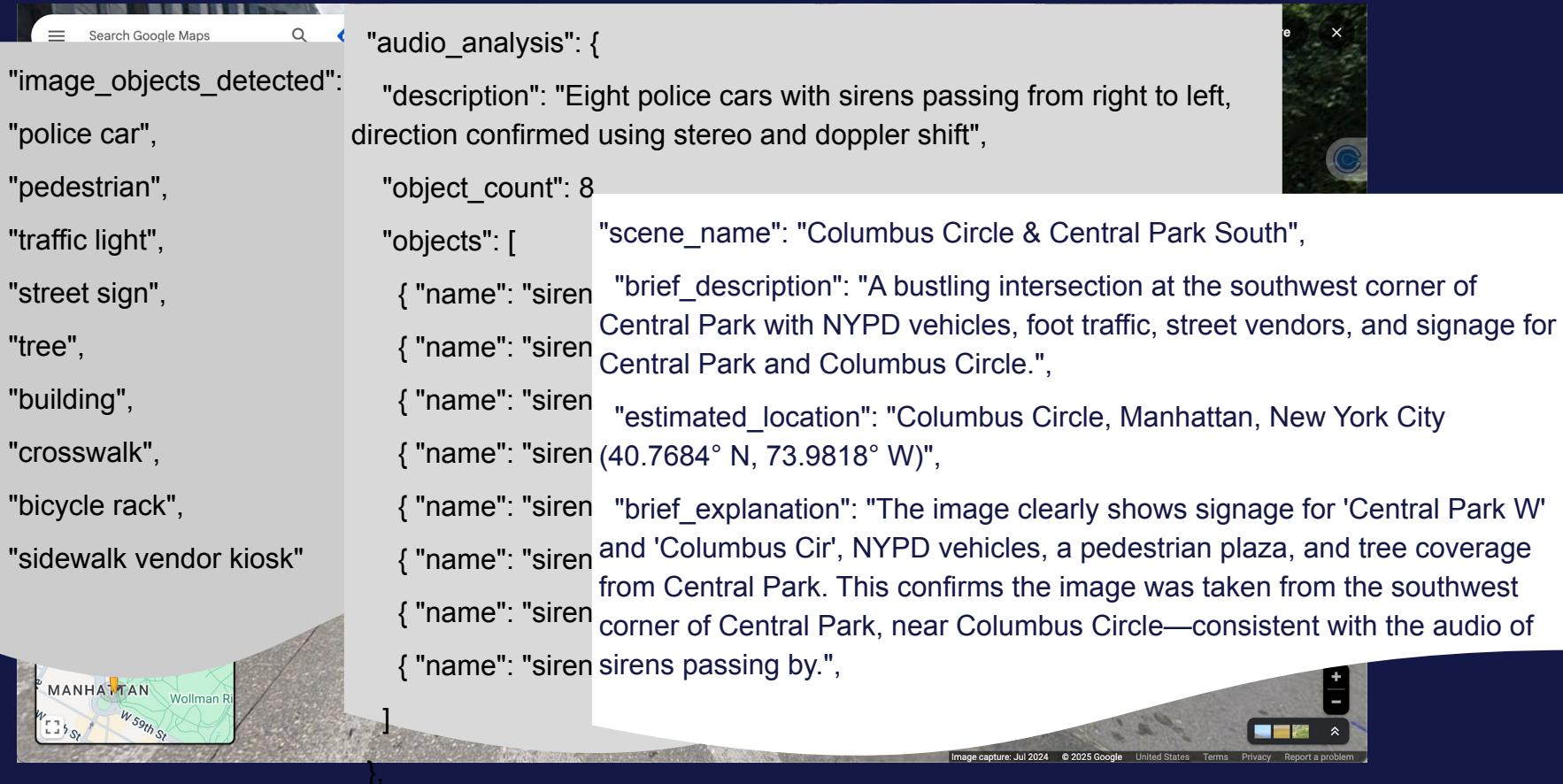


```
  "audio_analysis": {  
    "description": "Eight police cars with sirens passing from right to left,  
    direction confirmed using stereo and doppler shift",  
    "object_count": 8,  
    "objects": [  
      { "name": "siren_1", "location": "right", "direction": "right_to_left" },  
      { "name": "siren_2", "location": "right", "direction": "right_to_left" },  
      { "name": "siren_3", "location": "right", "direction": "right_to_left" },  
      { "name": "siren_4", "location": "right", "direction": "right_to_left" },  
      { "name": "siren_5", "location": "right", "direction": "right_to_left" },  
      { "name": "siren_6", "location": "right", "direction": "right_to_left" },  
      { "name": "siren_7", "location": "right", "direction": "right_to_left" },  
      { "name": "siren_8", "location": "right", "direction": "right_to_left" }  
    ]  
  },  
  "image_objects_detected": [  
    "police car",  
    "pedestrian",  
    "traffic light",  
    "street sign",  
    "tree",  
    "building",  
    "crosswalk",  
    "bicycle rack",  
    "sidewalk vendor kiosk"  
  ]  
}
```



Image capture: Jul 2024 © 2025 Google United States Terms Privacy Report a problem

# One More Example for Fun!



A screenshot of a Google Maps street view interface. The map shows a bustling intersection at the southwest corner of Central Park and Columbus Circle in Manhattan. The view includes parts of Central Park, the Columbus Circle monument, and surrounding city streets. A small watermark in the bottom right corner of the map area says "COLUMBUS CIRCLE".

```
  "audio_analysis": {  
    "description": "Eight police cars with sirens passing from right to left, direction confirmed using stereo and doppler shift",  
    "object_count": 8,  
    "objects": [  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbus Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"},  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbus Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"},  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbus Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"},  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbus Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"},  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbs Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"},  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbs Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"},  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbs Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"},  
      { "name": "siren", "brief_description": "A bustling intersection at the southwest corner of Central Park with NYPD vehicles, foot traffic, street vendors, and signage for Central Park and Columbus Circle.", "estimated_location": "Columbs Circle, Manhattan, New York City (40.7684° N, 73.9818° W)", "brief_explanation": "The image clearly shows signage for 'Central Park W' and 'Columbus Cir', NYPD vehicles, a pedestrian plaza, and tree coverage from Central Park. This confirms the image was taken from the southwest corner of Central Park, near Columbus Circle—consistent with the audio of sirens passing by.", "name": "siren"}],  
  }},  
  "image_objects_detected": [  
    "police car",  
    "pedestrian",  
    "traffic light",  
    "street sign",  
    "tree",  
    "building",  
    "crosswalk",  
    "bicycle rack",  
    "sidewalk vendor kiosk"  
  ]  
}
```

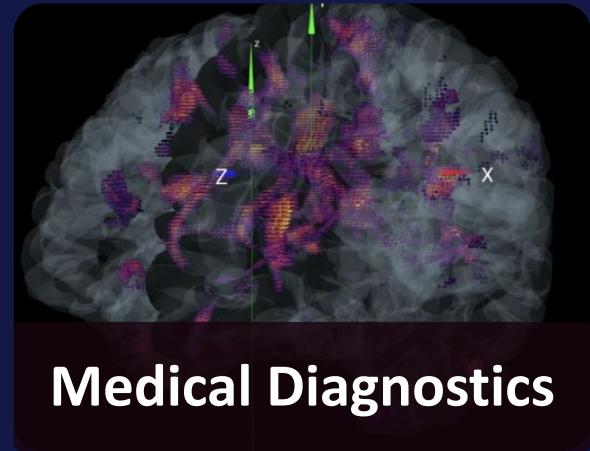
Image capture: Jul 2024 © 2025 Google United States Terms Privacy Report a problem



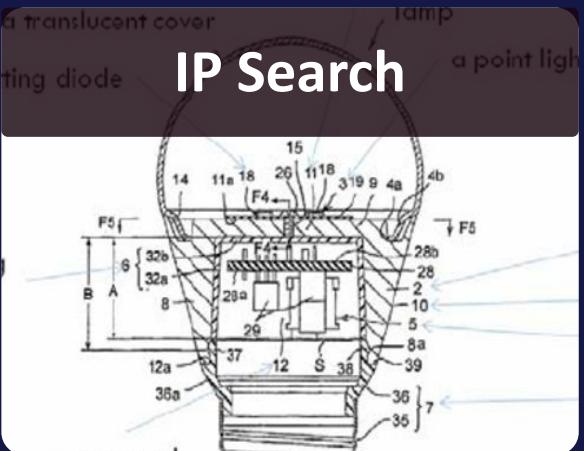
# Geospatial Analysis



# Security



# Medical Diagnostics



# IP Search



# Observation

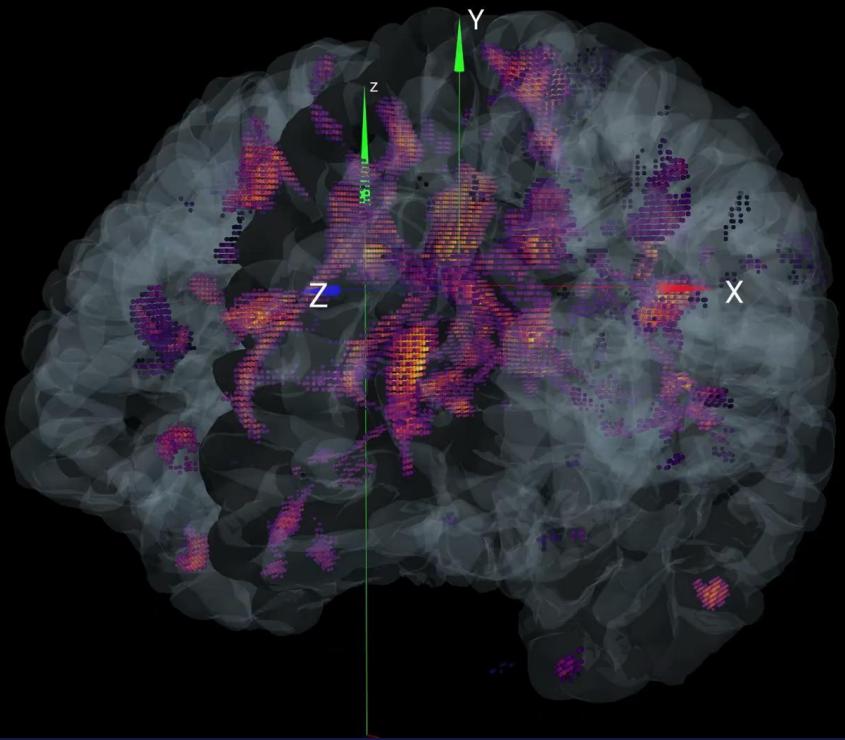


# Predictive Maintenance

# Future Directions

# BrainGraph

## Using Latent Data and Graph Analytics

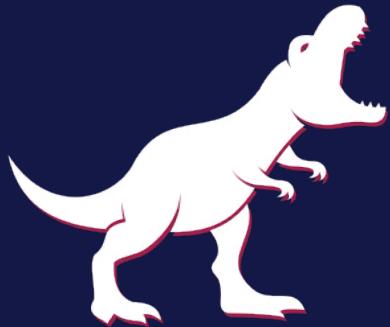


Medical images consist of 3D pixels called voxels.  
**Voxels are nodes in a graph with neighbors.**



Communities of voxels can represent anatomical structures, or abnormalities like tumors. **Evolution in the graph** can represent disease progression or treatment response.

# Reach Out!



GraphGeeks.org

Join our vendor neutral community for  
graph practitioners & enthusiasts

Amy Hodler



# Please Reach Out



## Q&A

David Hughes

