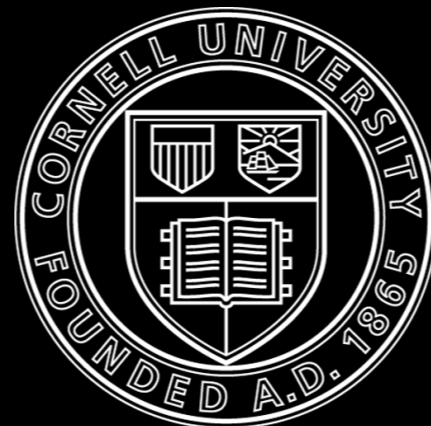


The Visual Anthropologist

Kavita Bala

Computer Science

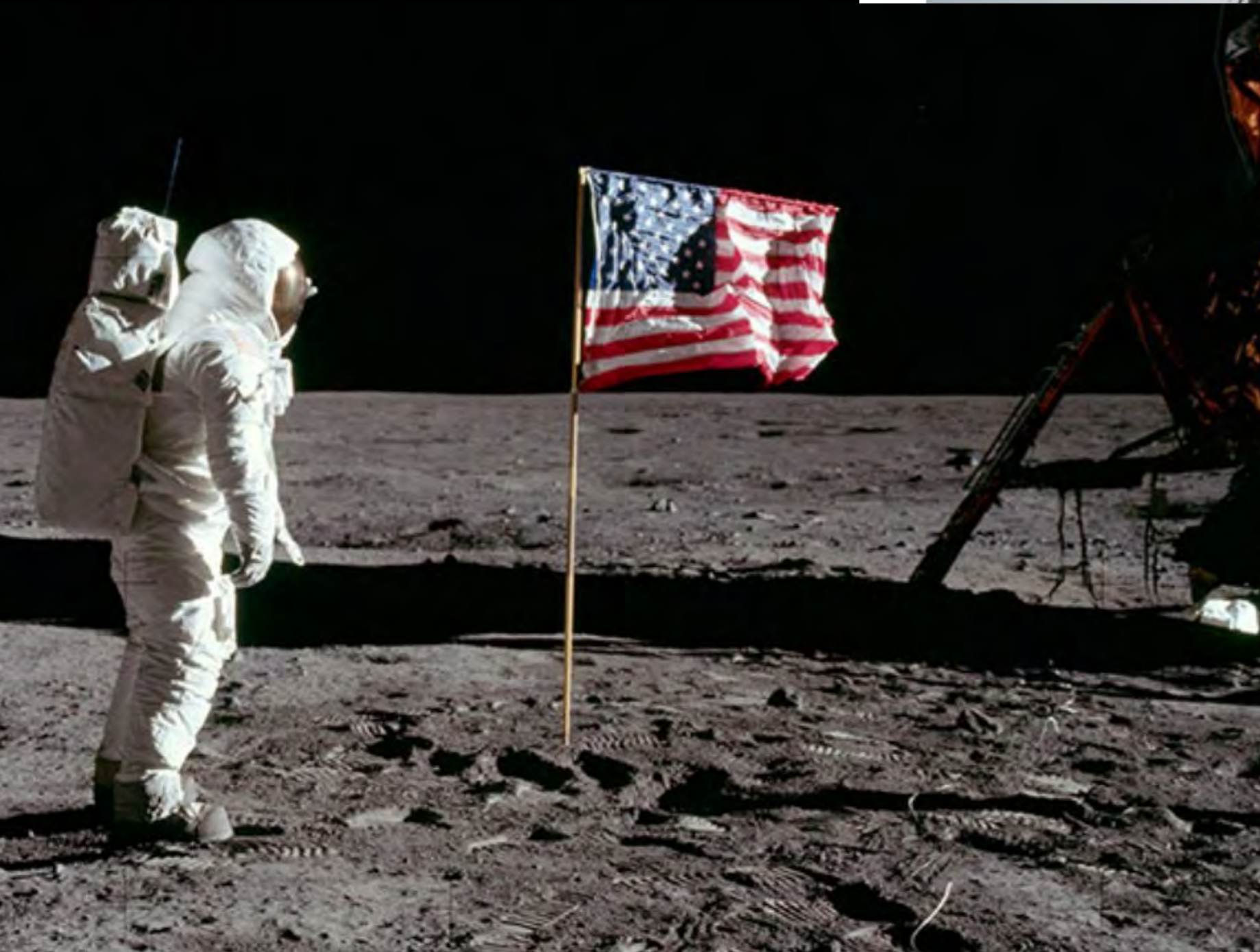


A picture is worth a thousand words





Four more years.



9





Our homes



Flickr



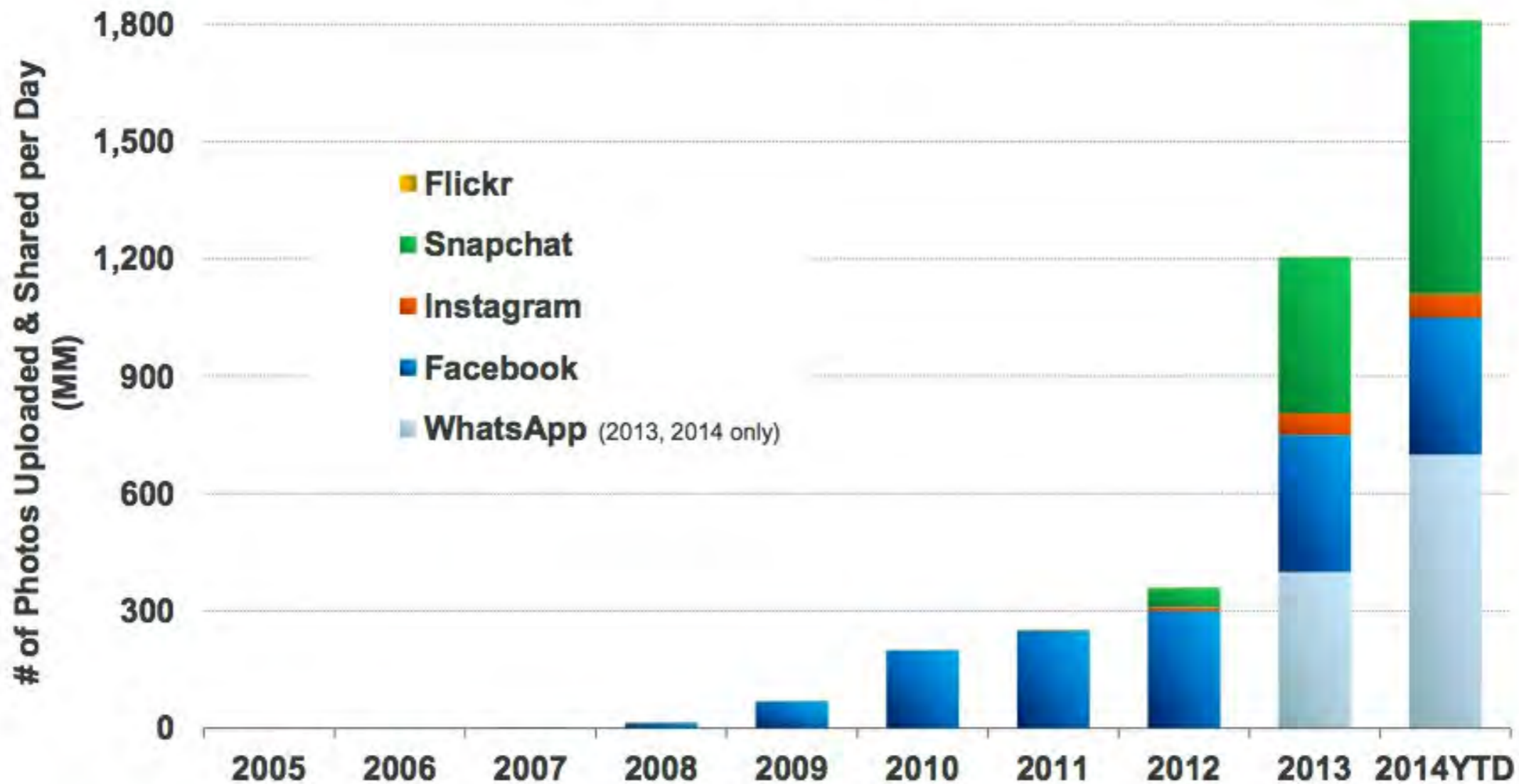
Our friends

Instagram



1.8 B + uploaded/shared images per day

Daily Number of Photos Uploaded & Shared on Select Platforms, 2005 – 2014YTD



Visual Anthropology

- What is the world made of?
- How do we live?
 - What do we wear?
 - What do we eat?
 - Who are our friends?
- How does our behavior change over time?

Why now?

- Confluence of data, compute, and deep learning
- At scale
 - Visual recognition
 - Visual search
 - Exploratory visual data analysis
 - Finding trends

Recognition and Scene Understanding



Recognition and Scene Understanding



Recognition and Scene Understanding



Wood

Rug

Paper

Scene Understanding



Scene Understanding



living room

Scene Understanding



Window

Ceiling

Wall

Sofa

Chair

Floor

Mat

living room

Scene Understanding



Glass

Window

Ceiling

Wall

Sofa

Fabric

Leather

Plastic

Fabric

Fabric

Wood
Floor

Chair

Fur
Mat

living room

Recognizing Materials



00: brick
01: carpet
02: ceramic
03: fabric
04: foliage
05: food
06: glass
07: hair
08: leather
09: metal
10: other
11: painted
12: paper
13: plastic
14: polishedstone
15: skin
16: tile
17: stone
18: water
19: wood

Previous: 40% for 10 classes
Ours: 85% accuracy for 20 classes

Beyond what people are good at



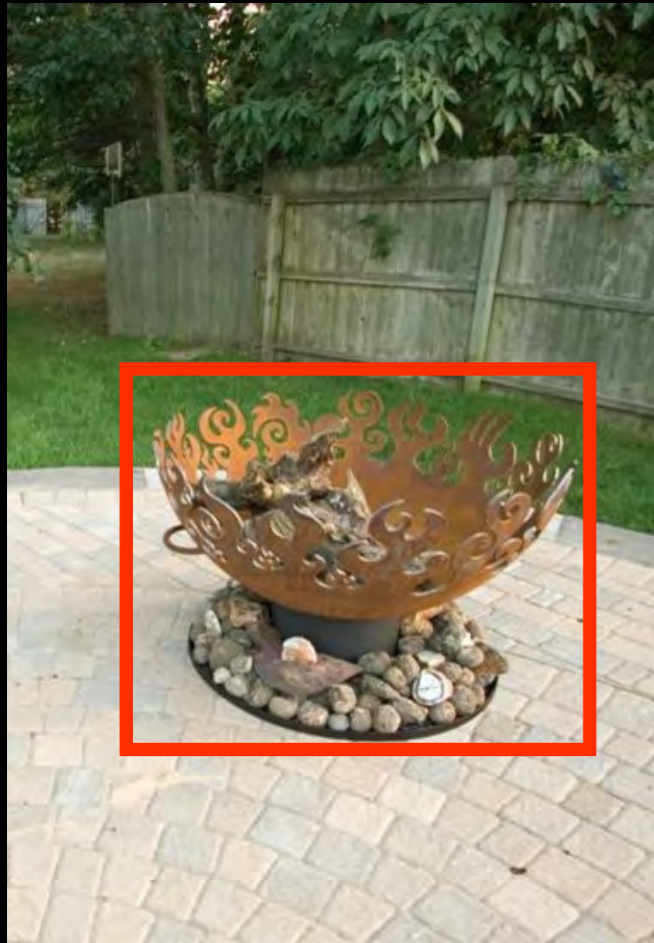
Beyond what people are good at

“What is this?”



Beyond what people are good at

“What is this?”



Category: Fire pit

Beyond what people are good at

“What is this?”



Name: "Great Bowl O'Fire
Sculptural Fire Bowl"

Category: Fire pit

Sold by: John T. Unger, LLC

Challenge

“What is this?”



Name: "Great Bowl O'Fire Sculptural Fire Bowl"

Category: Fire pit

Sold by: John T. Unger, LLC



Challenge

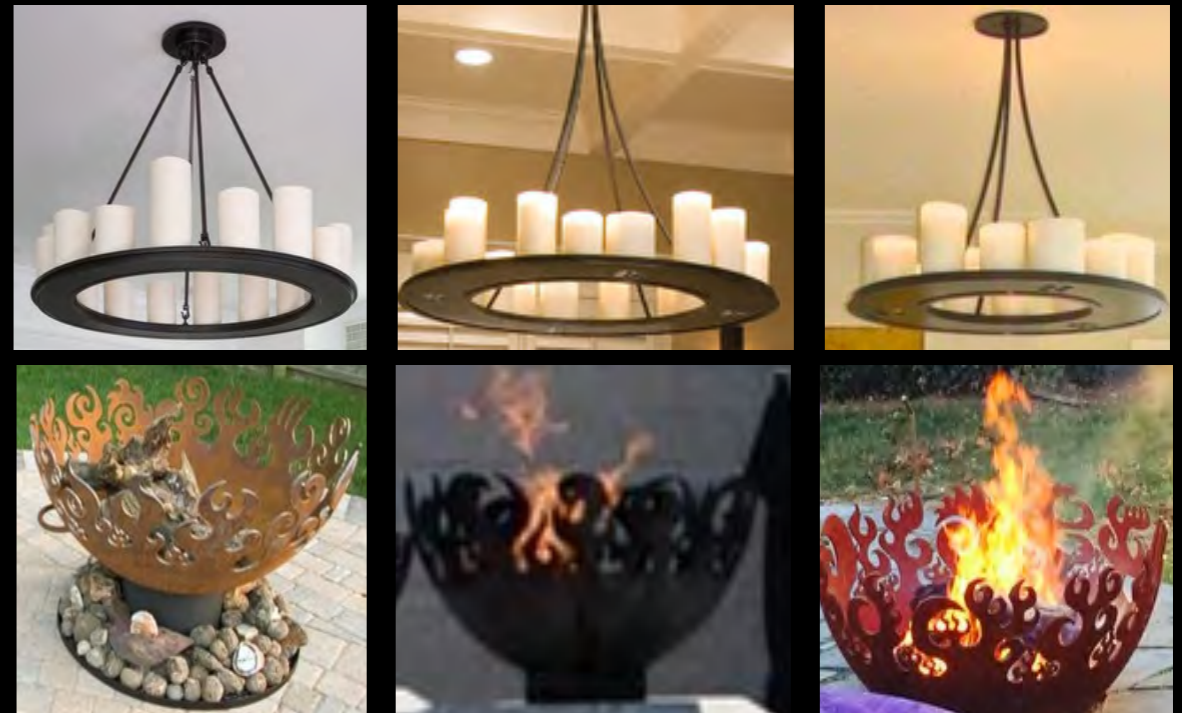


Determine whether these are the same product (different resolution, viewpoint, color, lighting, occlusions)

Problem

Problem

In context, “messy”, “in the wild”



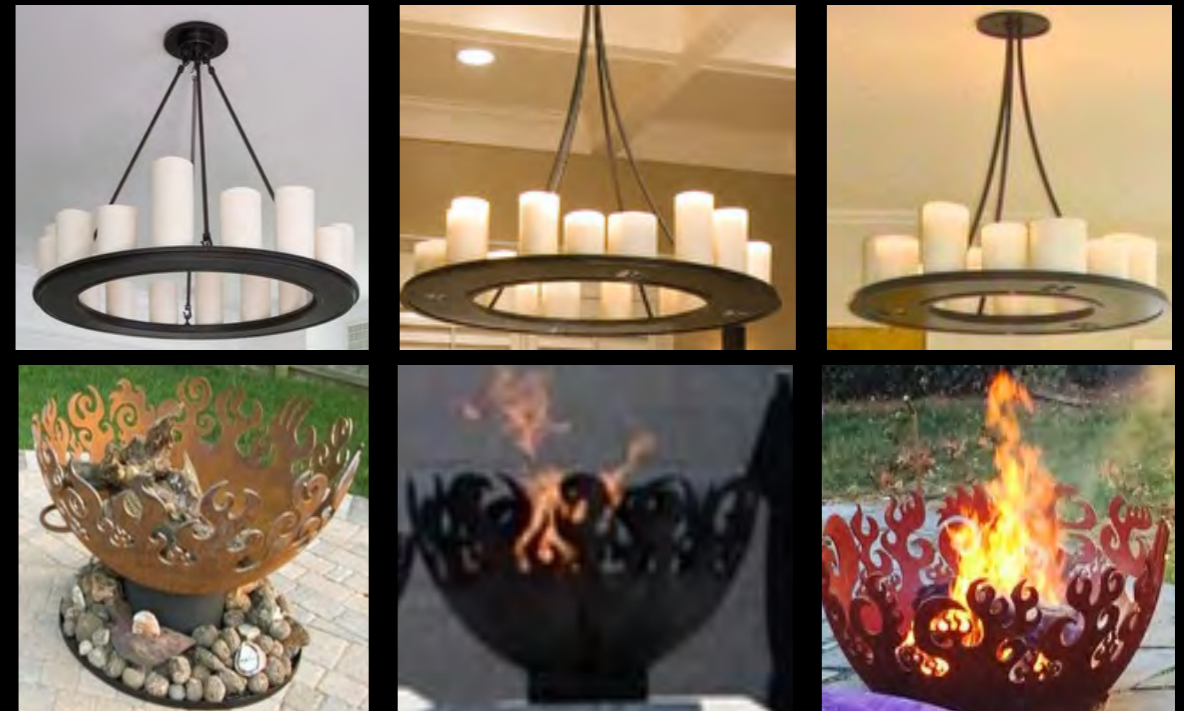
(Cropped from a scene photo)

Problem

Iconic, product or
“clean”



In context, “messy”, “in
the wild”



(From a product website)

(Cropped from a scene photo)

Problem: Find relation between these two kinds of images

Solution: Use deep learning to learn a joint embedding

Learn a projection into joint embedding

Iconic



In context



Learn a projection into joint embedding

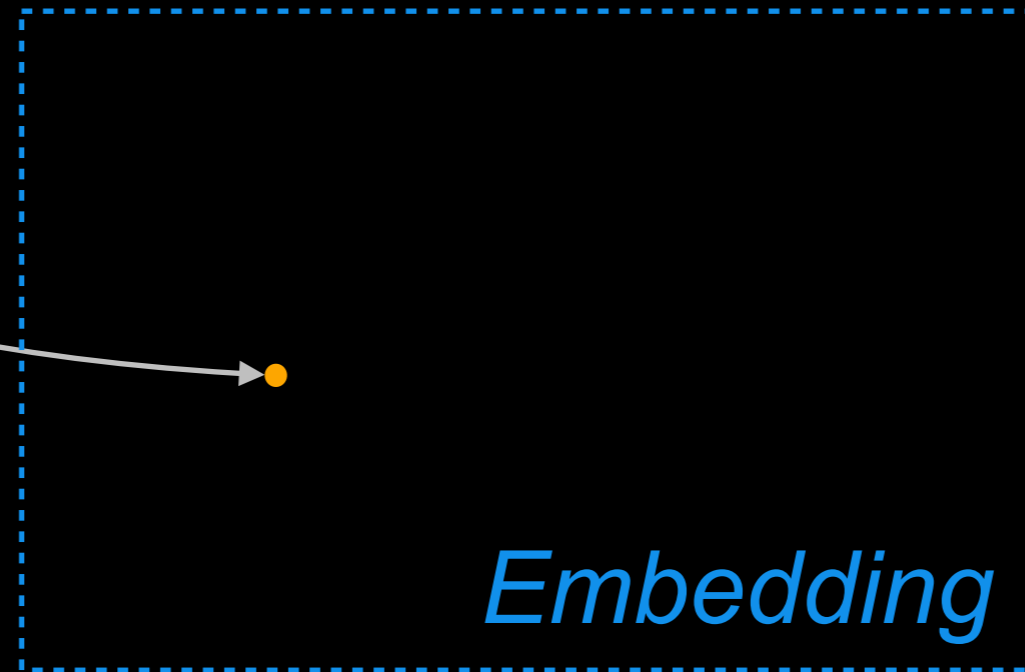
Iconic



In context



Project



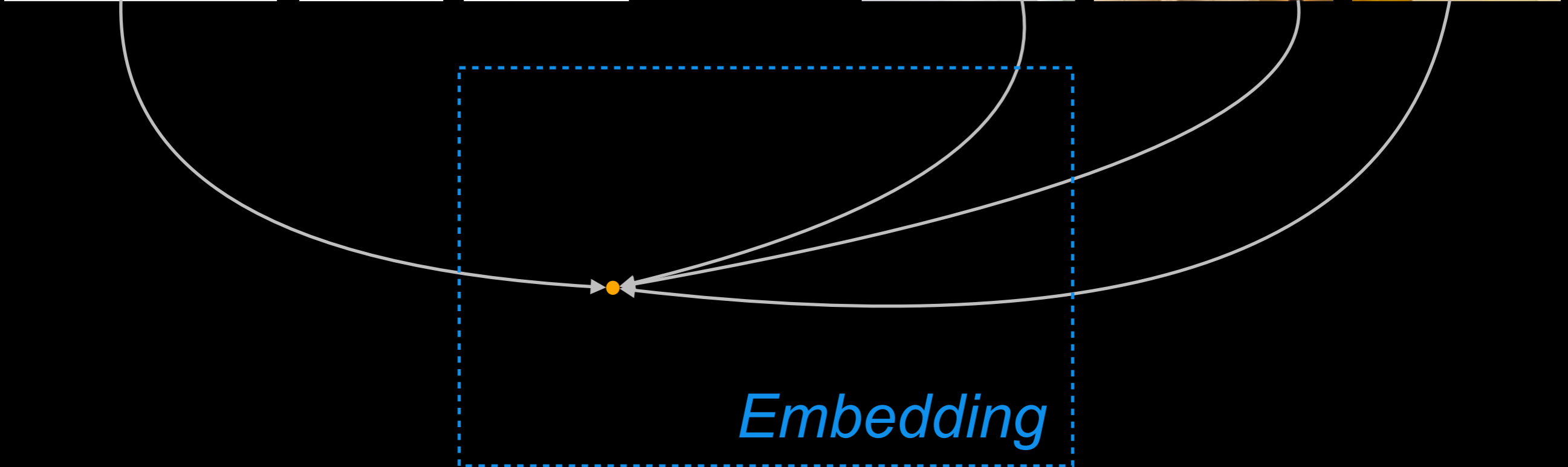
Embedding

Learn a projection into joint embedding

Iconic



In context

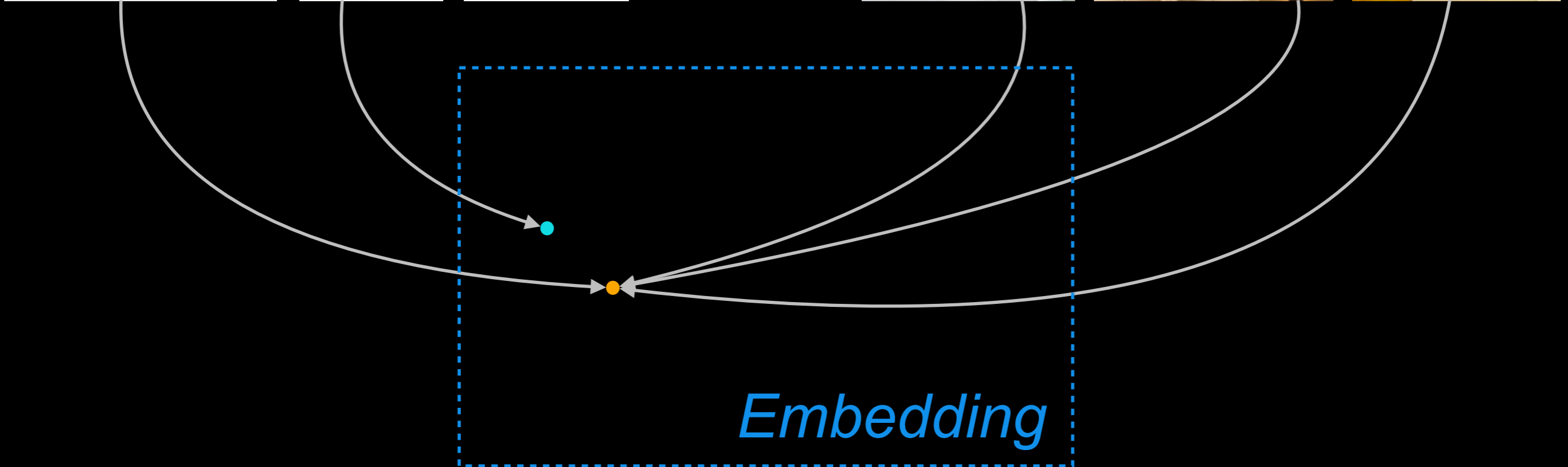


Learn a projection into joint embedding

Iconic



In context

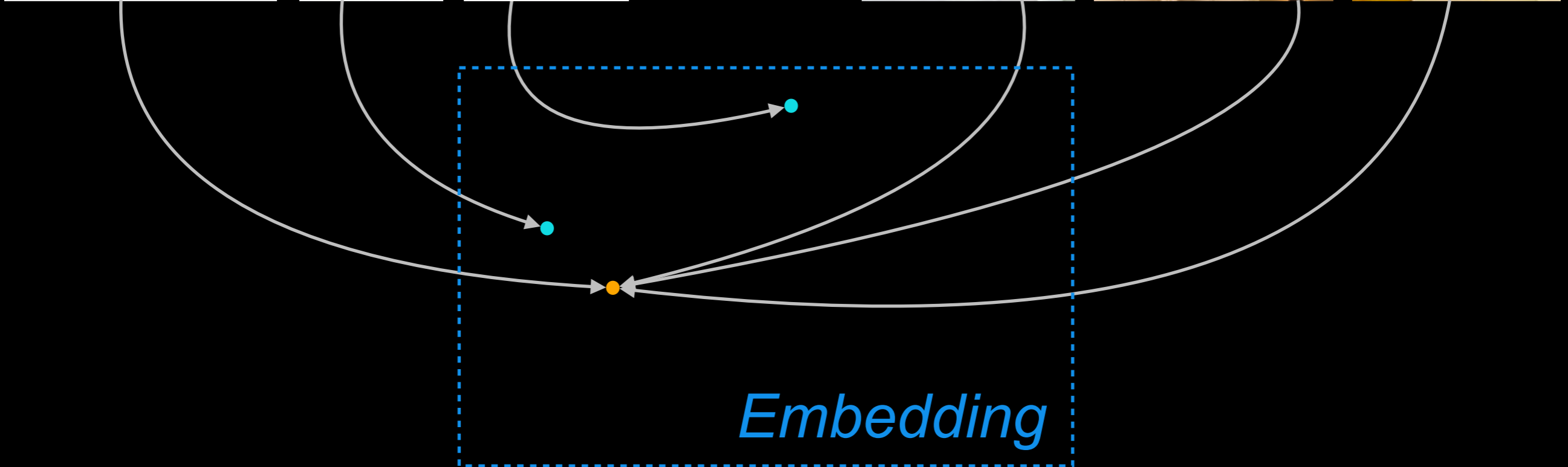


Learn a projection into joint embedding

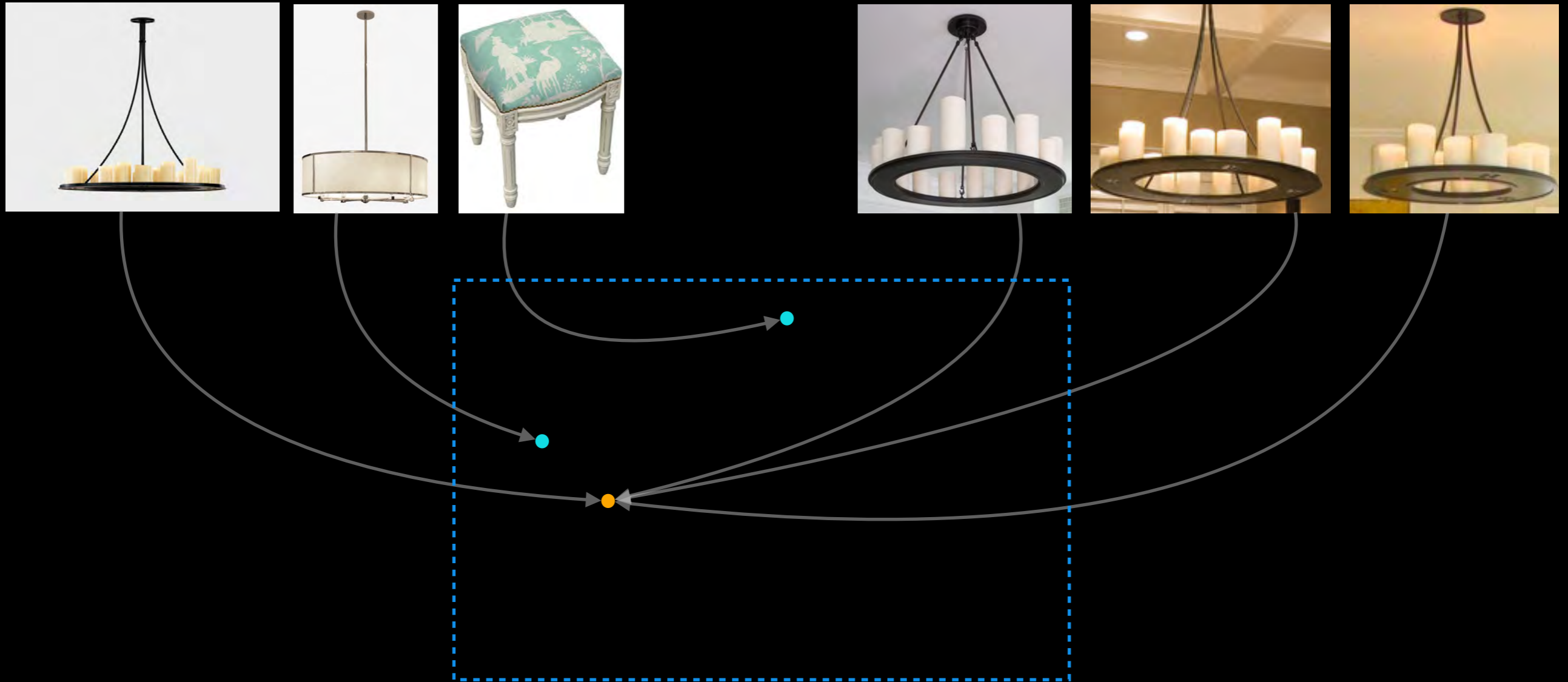
Iconic



In context

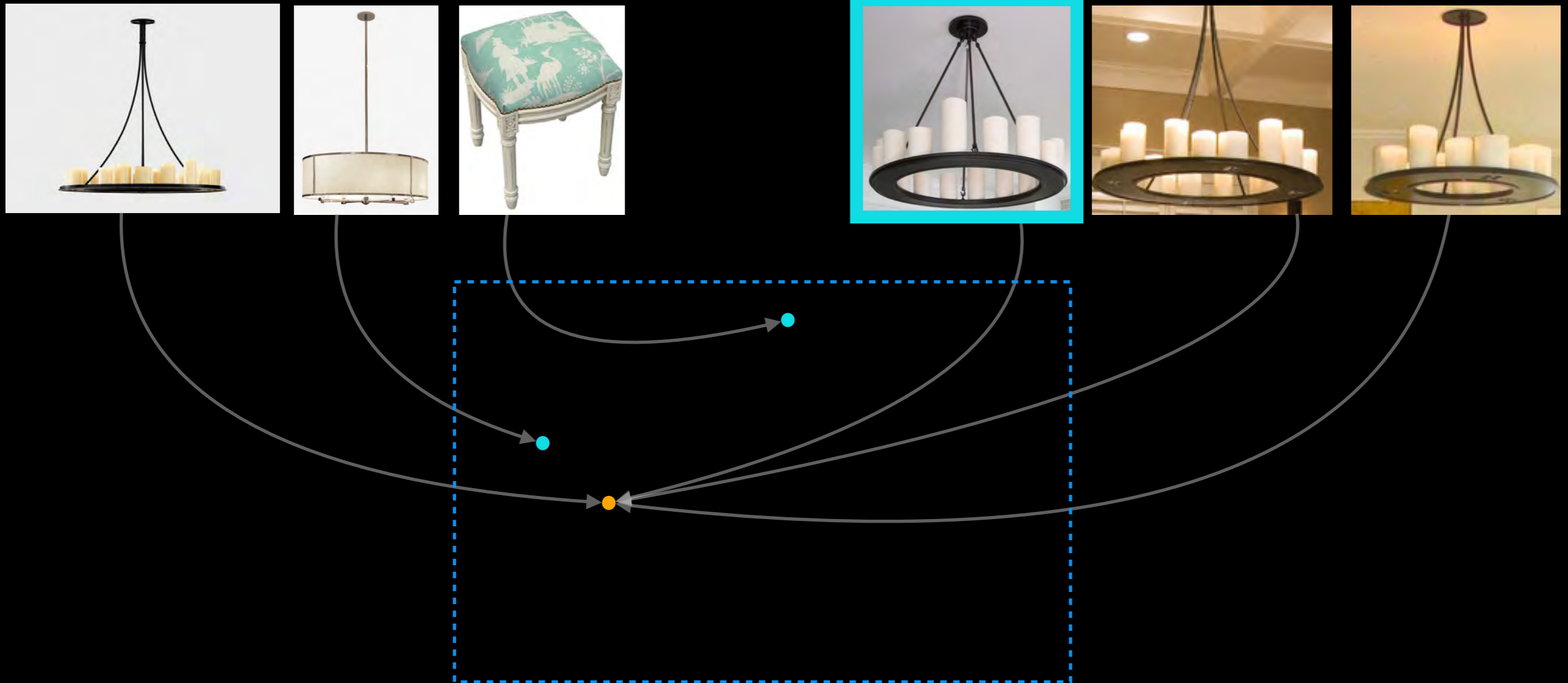


Search using the embedding



Search using the embedding

“What is it?”



Search using the embedding

“What is it?”



Retrieve

Project

Embedding

Name: Hemel Ring
Category: Hanging light
Sold by: Holly Hunt

Embedding (Projected, to 2D)



Embedding (Projected, to 2D)



Embedding Visualized in 2D: Chairs



Embedding Visualized in 2D: Chairs



Results: “What is it?”



“In the wild” query

Results: “What is it?”



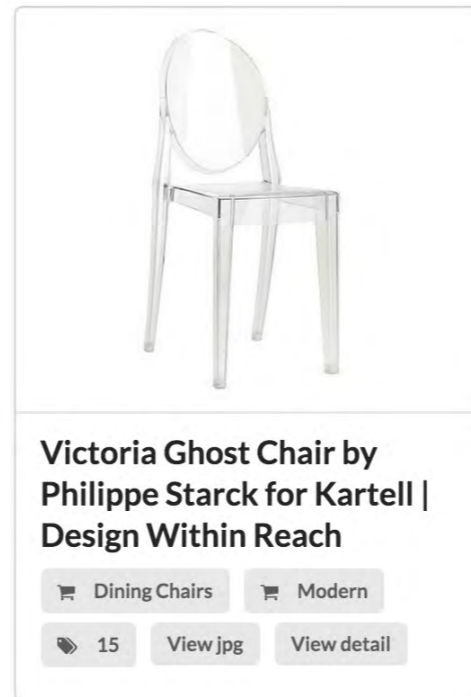
“In the wild” query

Results: "What is it?"

Query



Top 3 results:



0.019

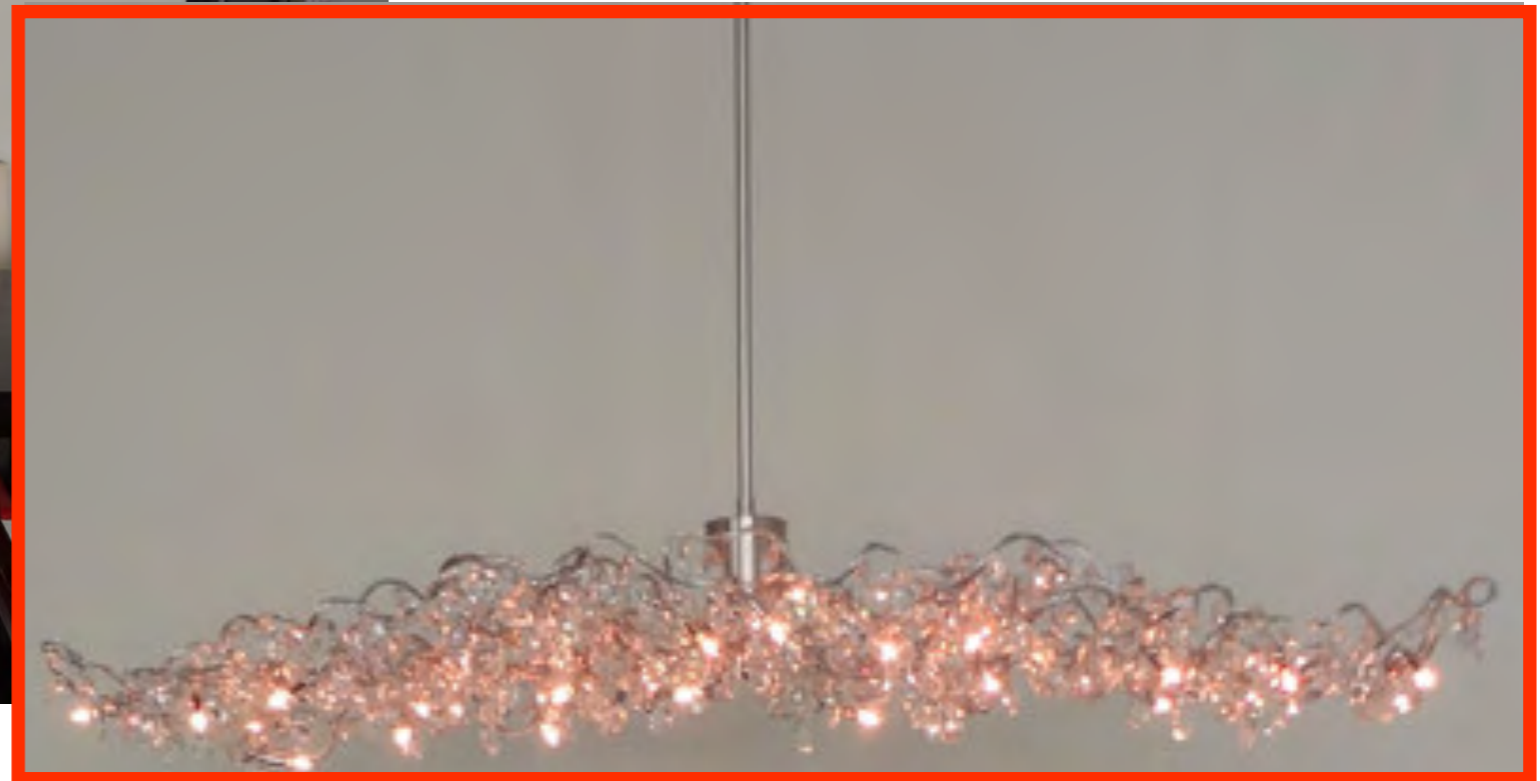
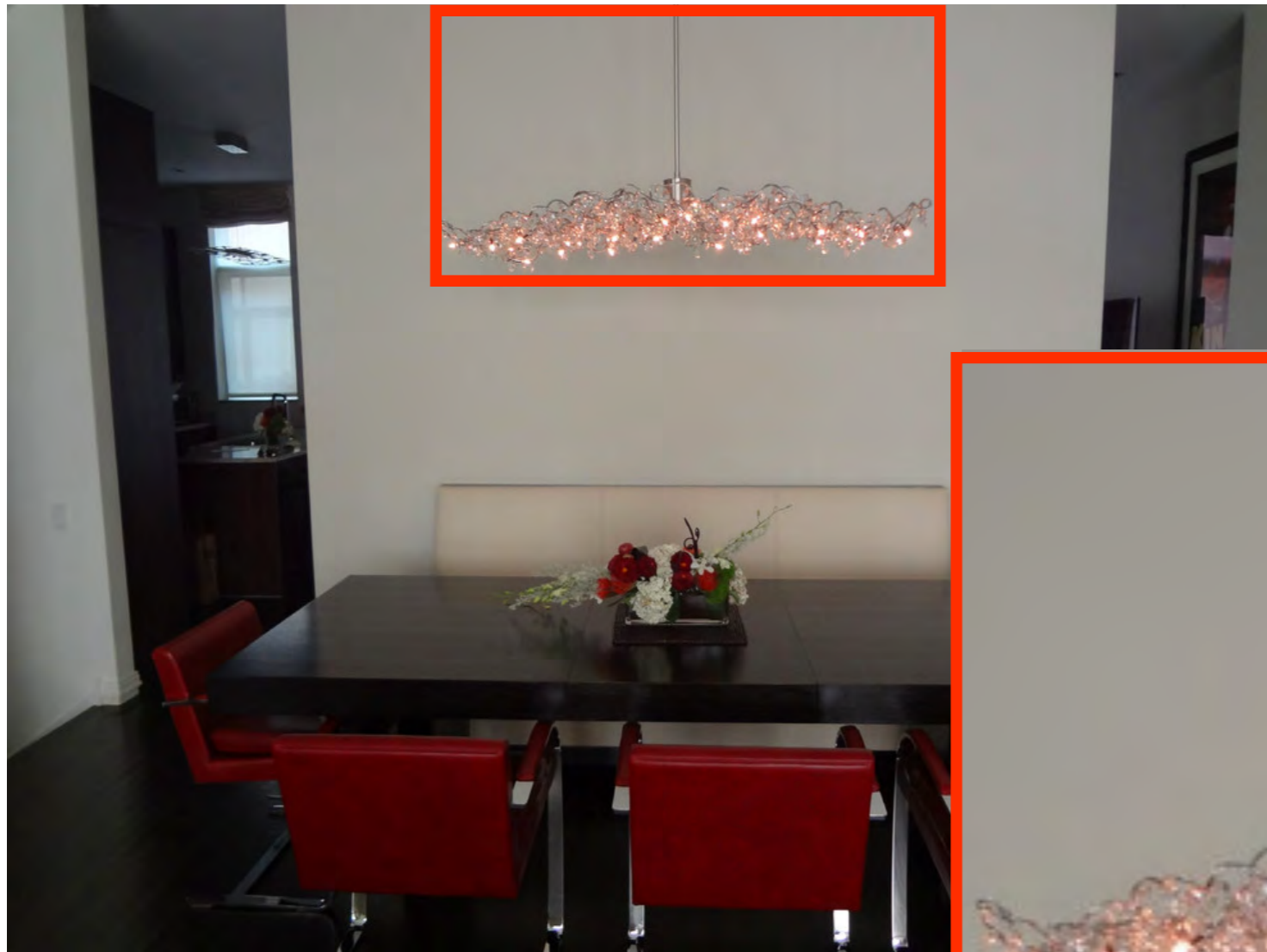


0.023



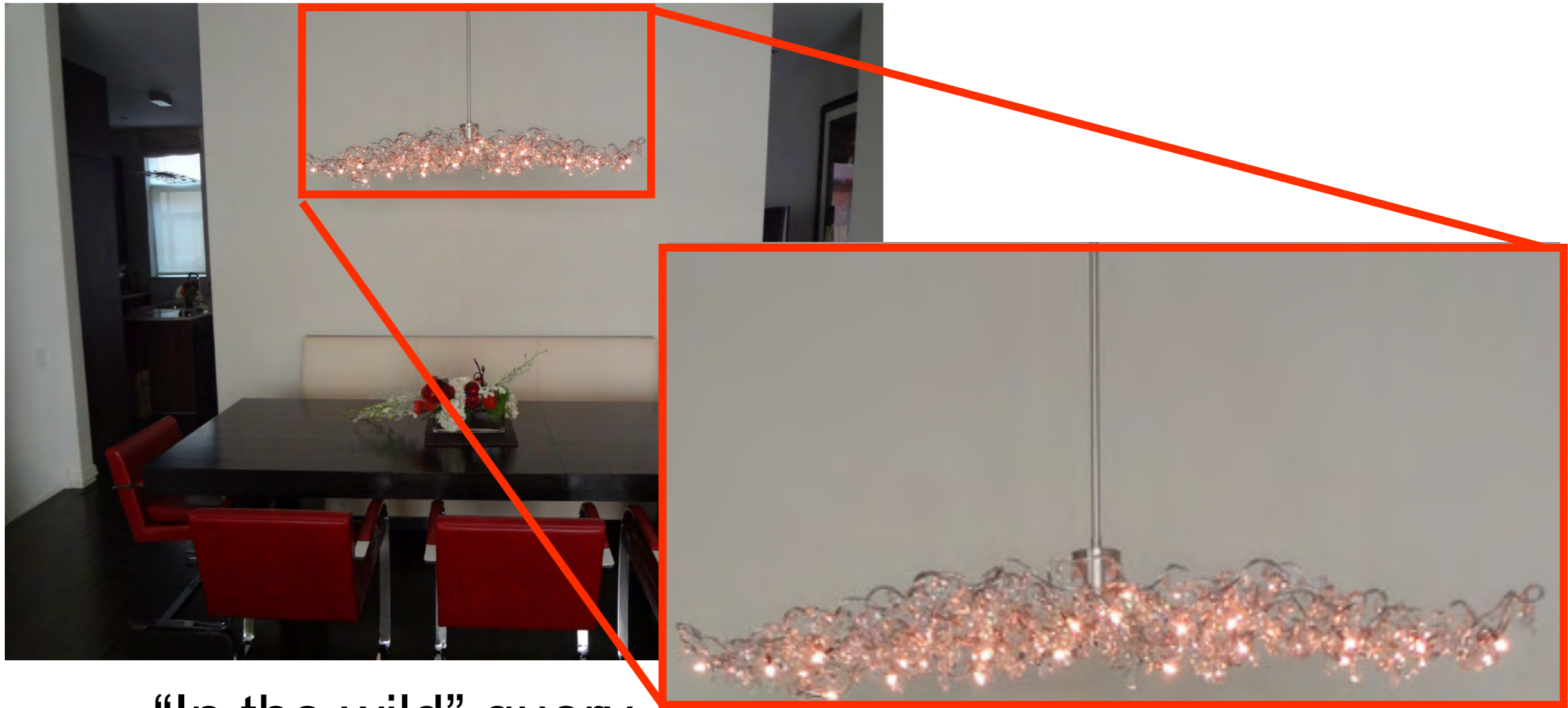
0.024

Results: “What is it?”



“In the wild” query

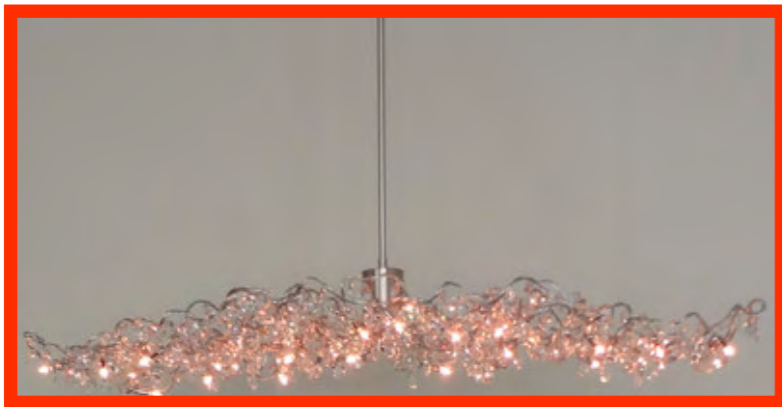
Results: “What is it?”



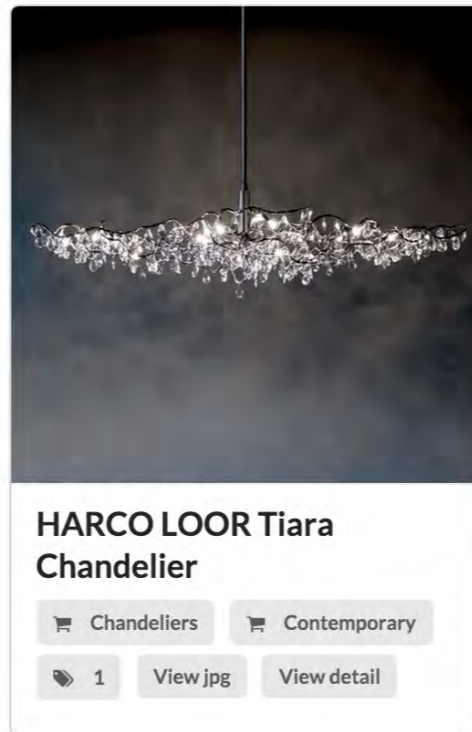
“In the wild” query

Results: "What is it?"

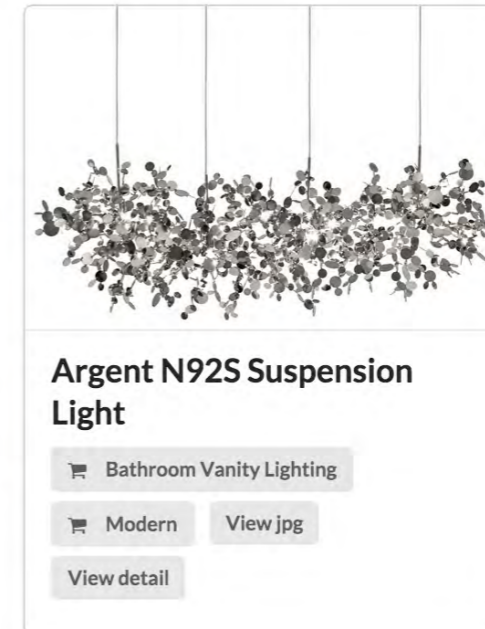
Query



Top 3 results:



0.035



0.038



0.041

RESULTS: "WHERE IS IT USED?"



"Maskros
Pendant
Lamp"

RESULTS: "WHERE IS IT USED?"



"Maskros
Pendant
Lamp"



Visual Anthropology

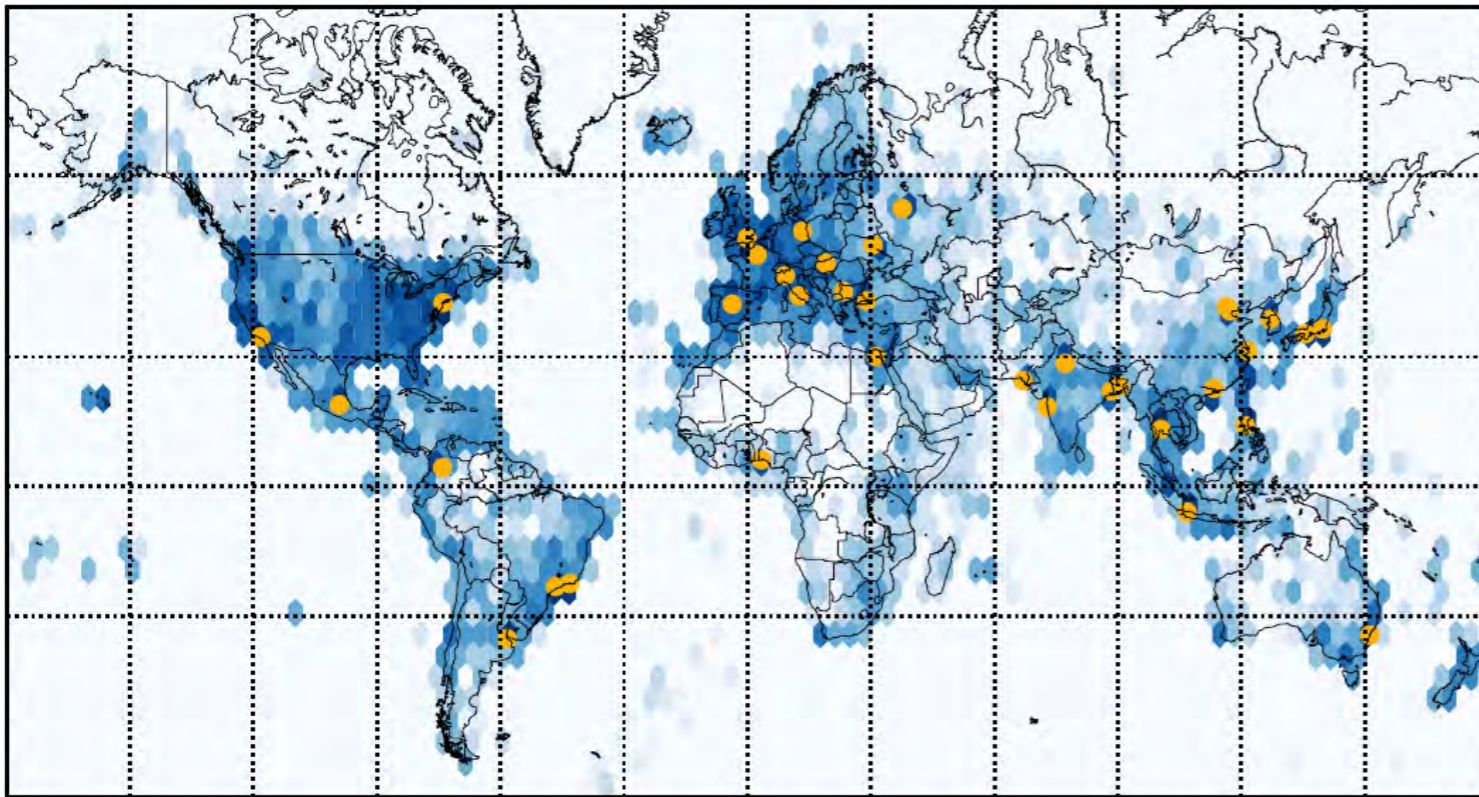
- What is the world made of?
- How do we live?
 - What do we wear?
 - What do we eat?
 - Who are our friends?
- How does our behavior change over time?



Instagram
Fast beautiful photo sharing



FACE++



Collect 8M people from 36 cities
27k annotated with clothing attributes - 12 attributes each



Stripes

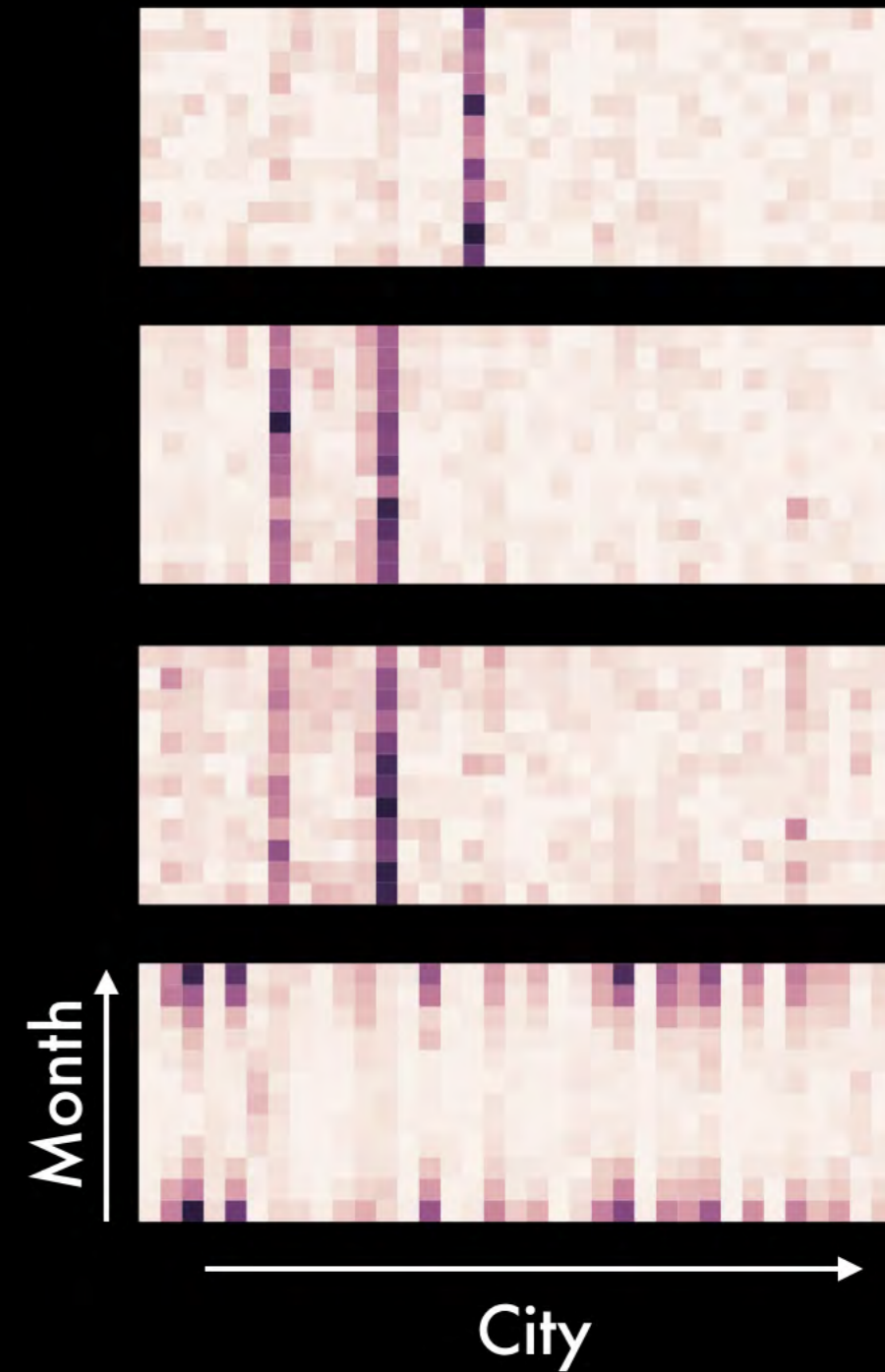
Jackets and Layers



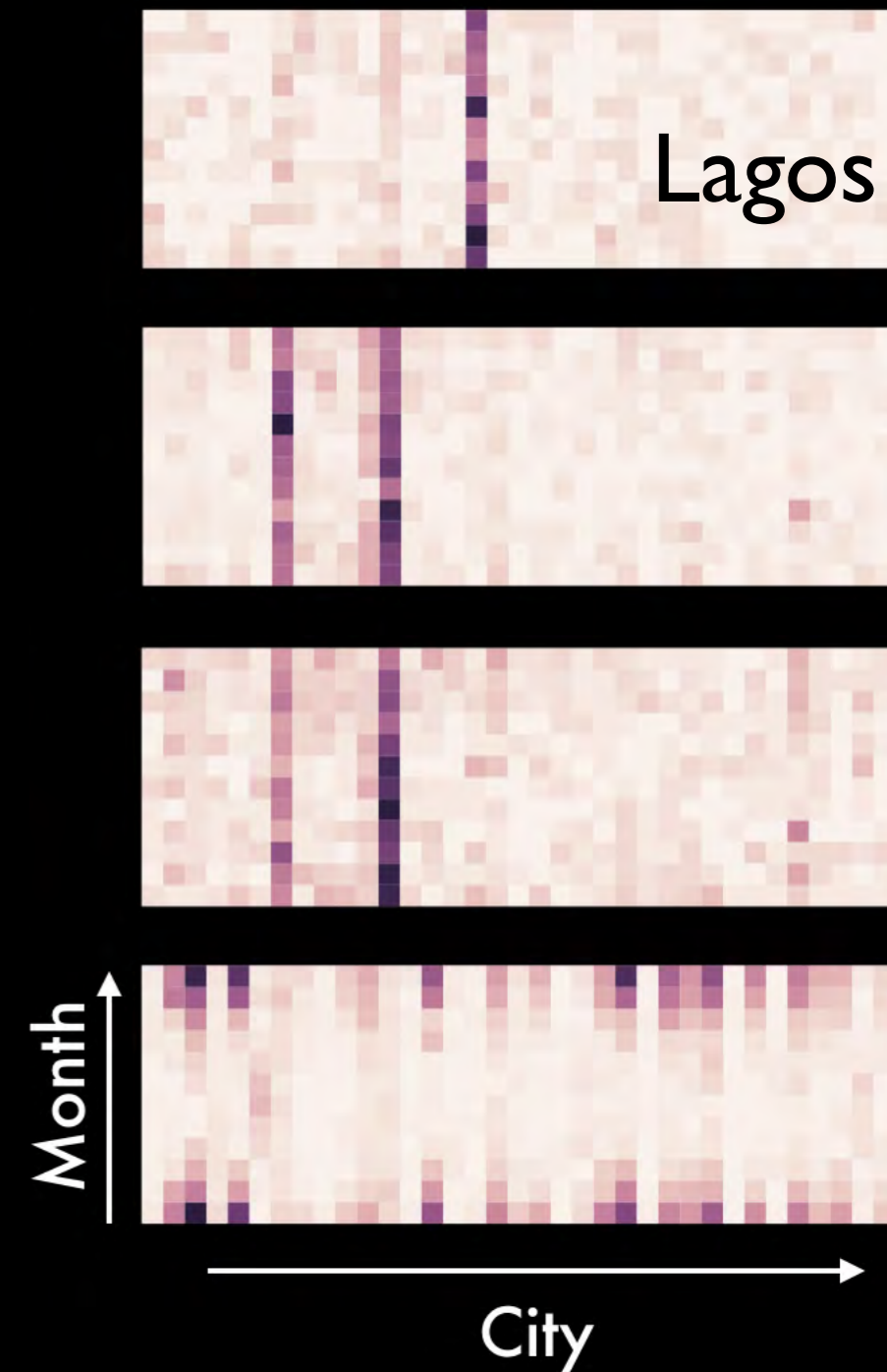
Plaids



Clusters that depend on place or time



Clusters that depend on place or time



Clusters that depend on place or time



Lagos



Cairo &
Jakarta



Month ↑



City →

Clusters that depend on place or time

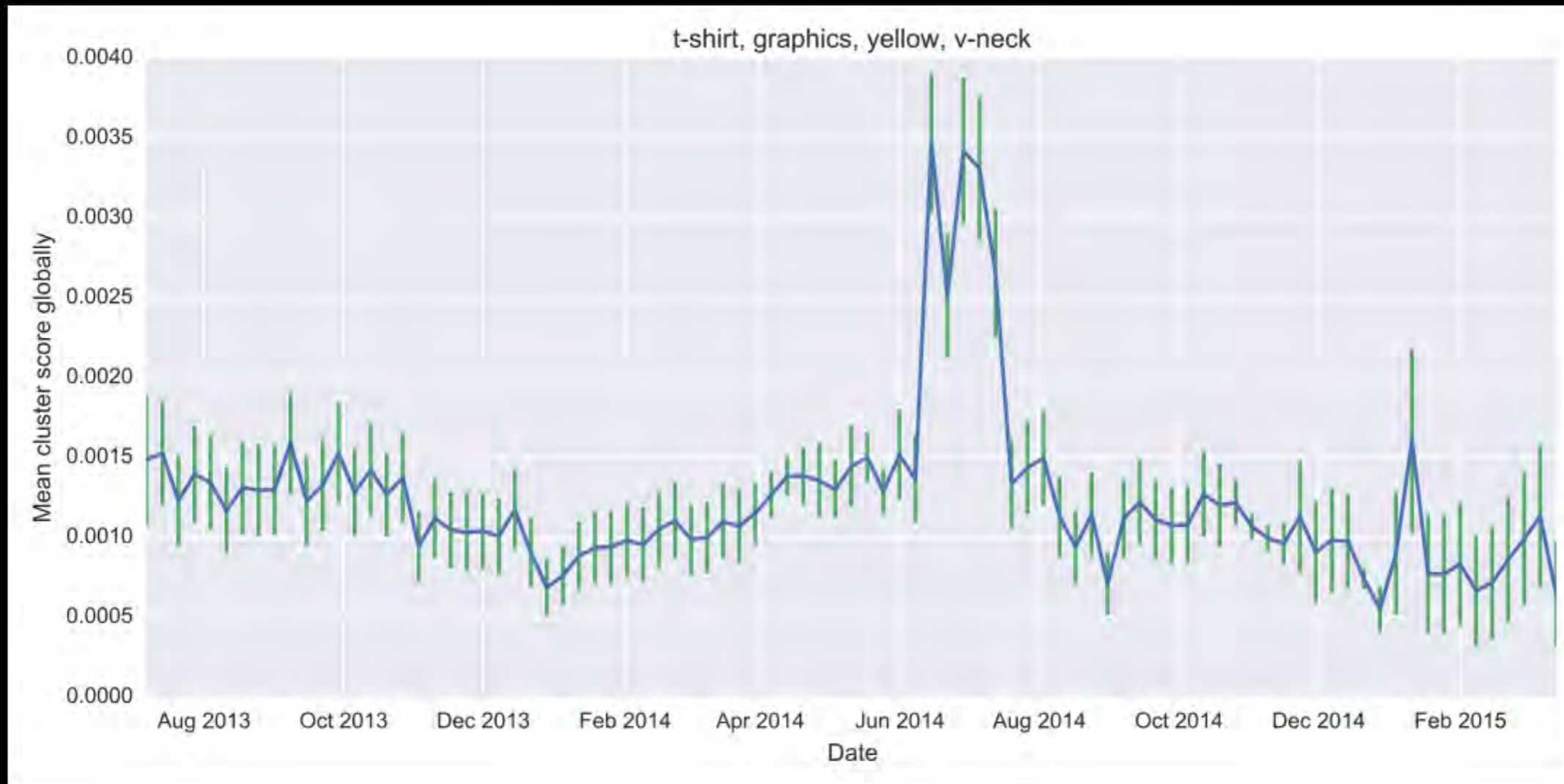


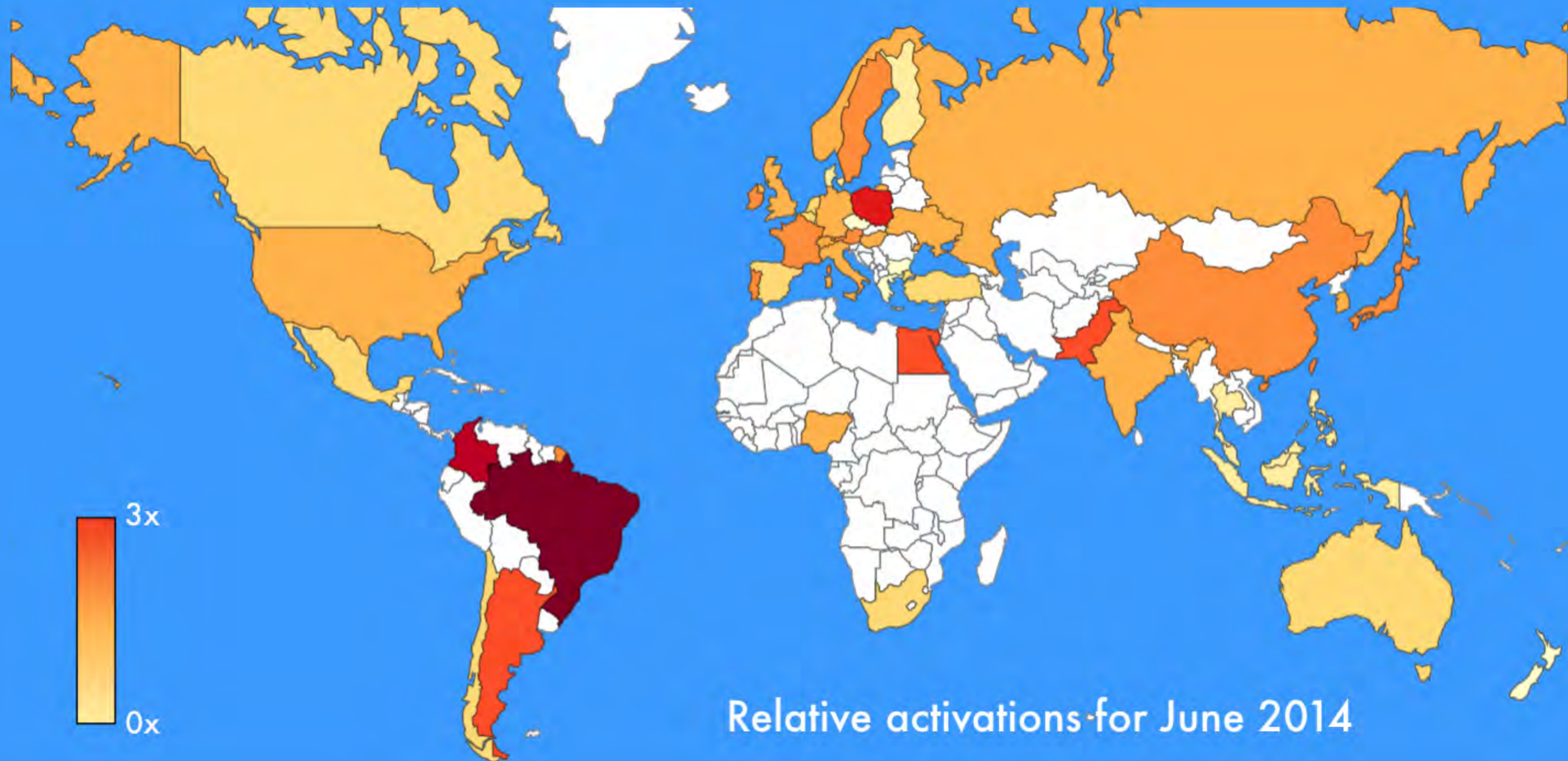
Clusters that don't depend on place or time



Exploratory Data Analysis

Find unique events





Relative activations for June 2014



Visual Anthropology

Collaborators



Sean Bell



Kevin Matzen



Noah Snavelly



Paul Upchurch