Image Searches, Abstraction, Invariance

36-350: Data Mining
8 September 2008

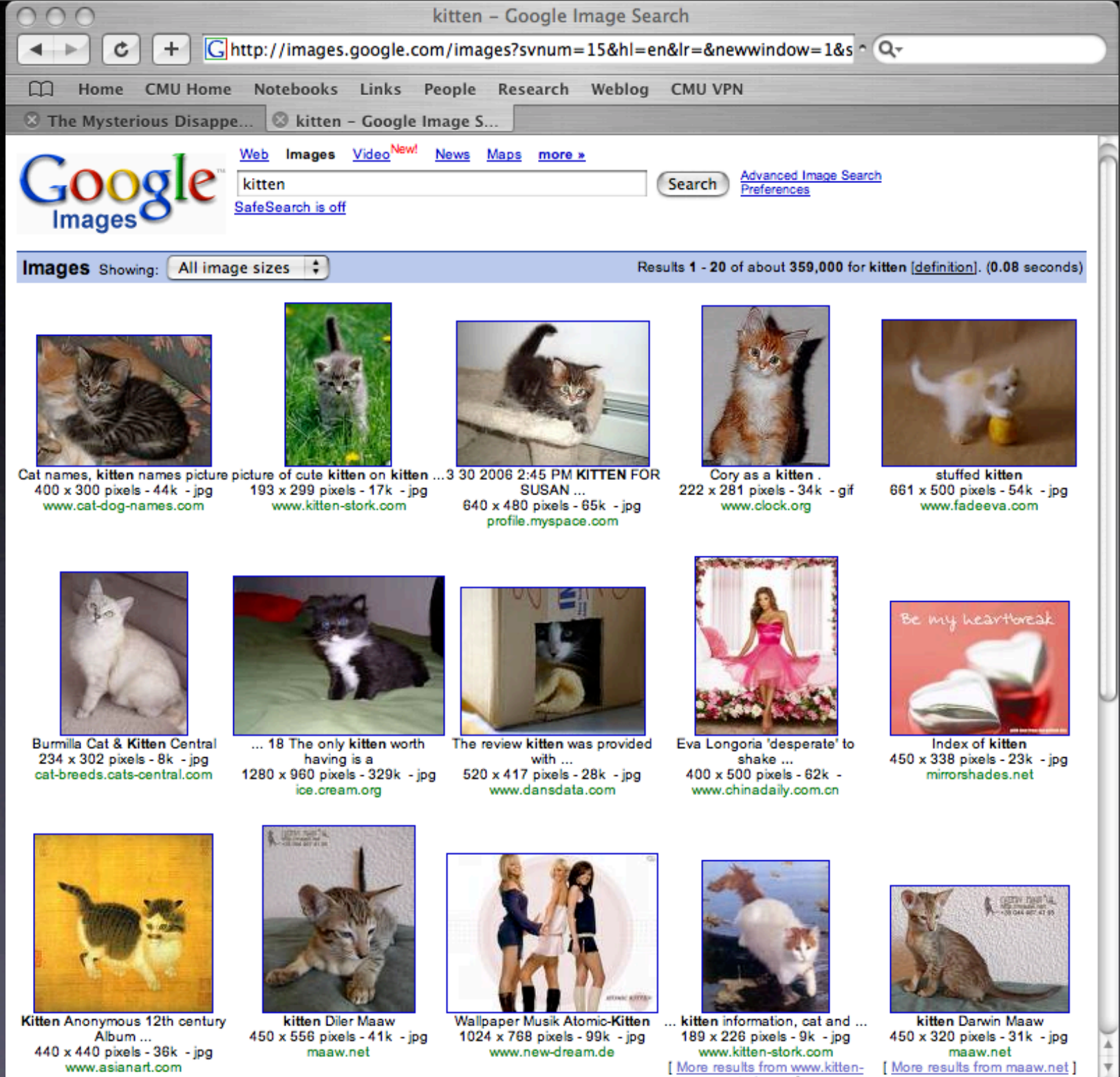
- Medical: x-rays, brain imaging, histology ("do these look like cancerous cells?")
- Satellite imagery
- Fingerprints
- Finding illustrations for lectures...

Searching for Images by Searching for Text

- Assume there's text accompanying the images ("annotation")
- Search those text records with the query phrase
- Take images which appear close to the query phrase on highly-ranked records
- This how Google does it

Sometimes this works perfectly...

...and sometimes it doesn't; depends on the text!

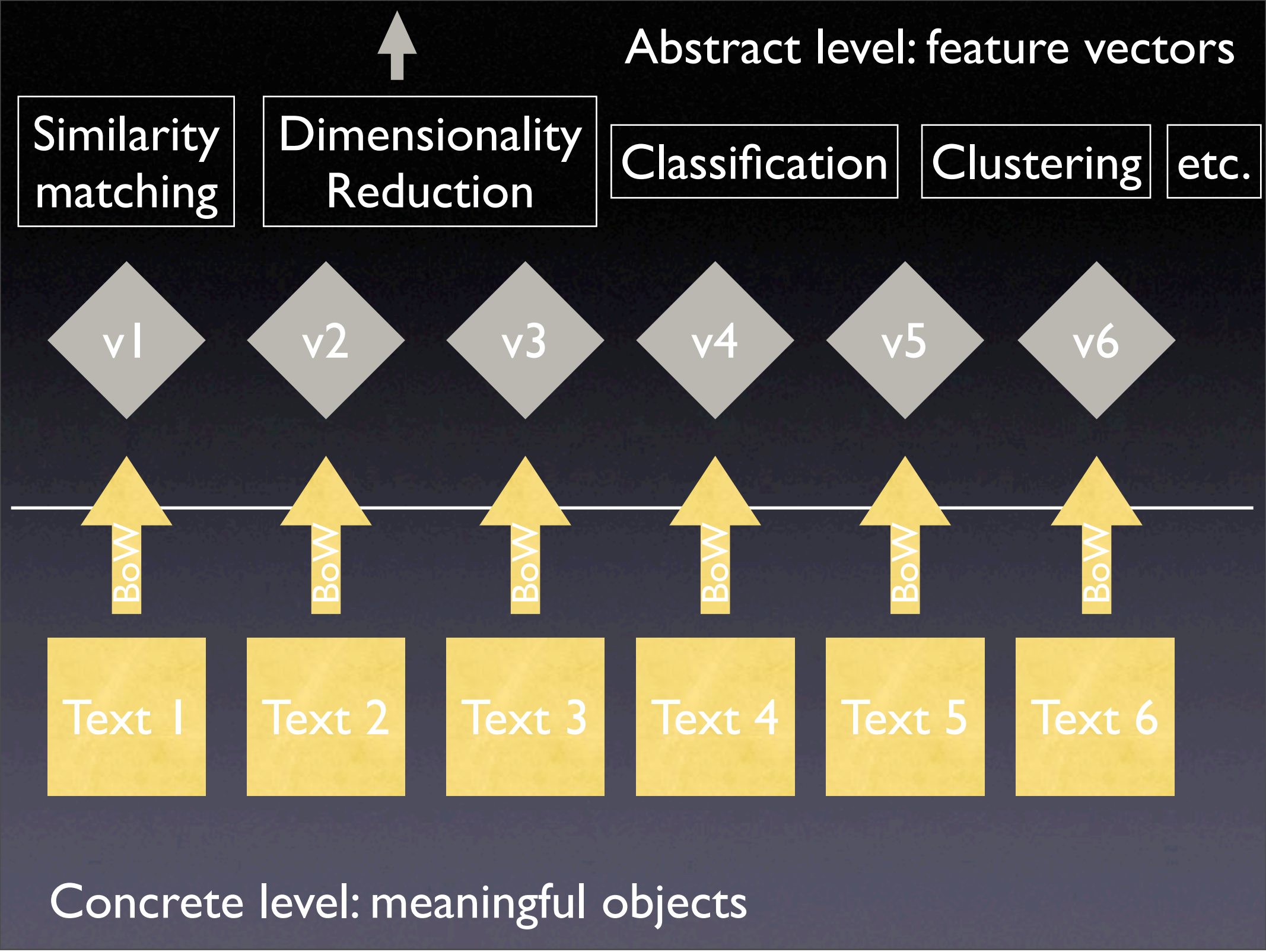


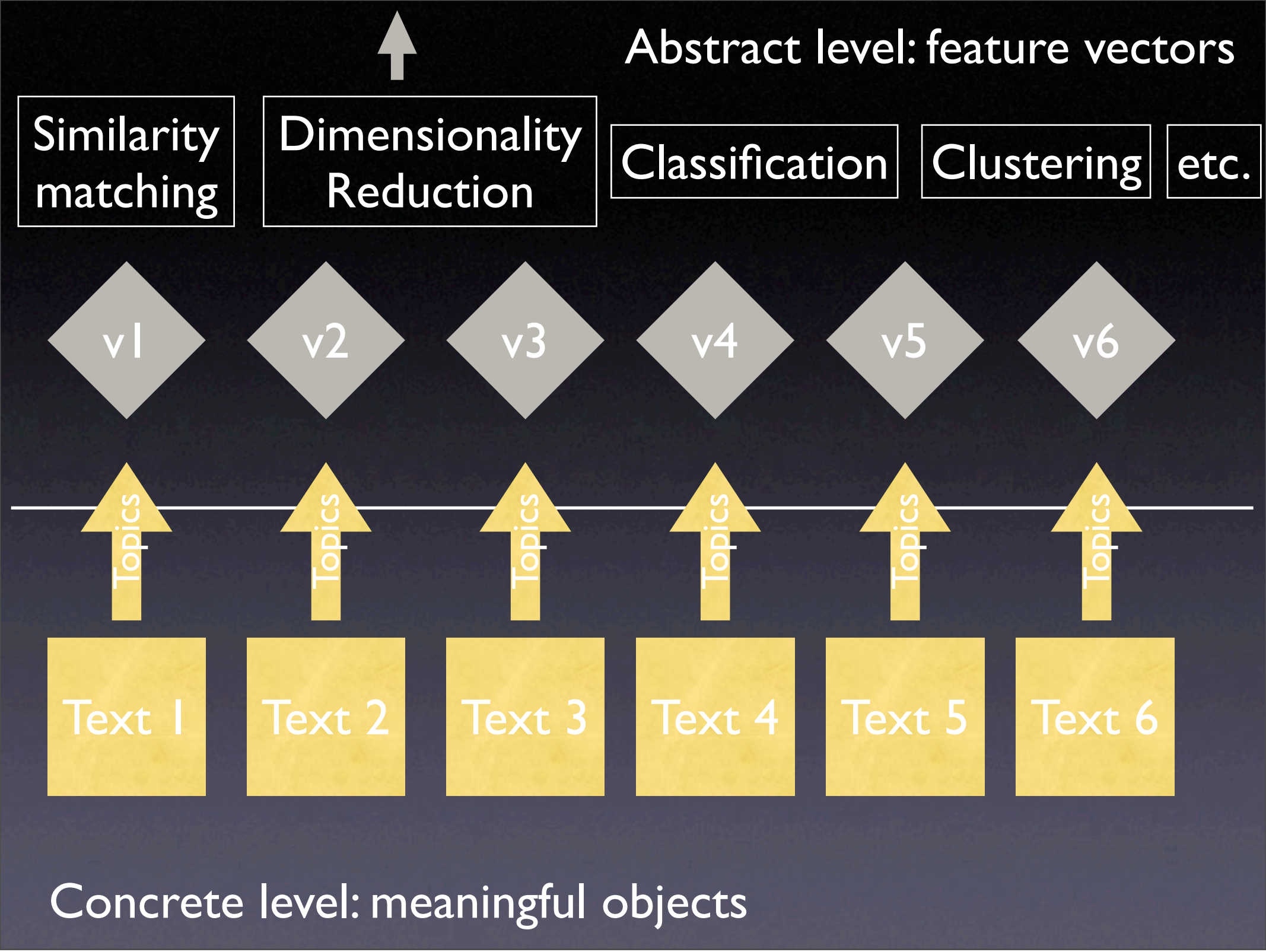
Searching for images by representing images

- For text, we only cared about *features*, and only worked with *feature vectors*
- Define numerical features for images and everything carries over
- Abstraction

Abstraction

- Remove some of the details but keep others
  - Kept details = features
- Then act on abstracta
- Hopes:
  - Simplifies problem
  - Lets you treat many problems similarly


Abstract level: feature vectors
Similarity matching
Dimensionality Reduction
Classification
Clustering
etc.
v1
v2
v3
v4
v5
v6
BoW
BoW
BoW
BoW
BoW
BoW
Text 1
Text 2
Text 3
Text 4
Text 5
Text 6
Concrete level: meaningful objects


Abstract level: feature vectors
Similarity matching
Dimensionality Reduction
Classification
Clustering
etc.
v1
v2
v3
v4
v5
v6
Topics
Topics
Topics
Topics
Topics
Topics
Text 1
Text 2
Text 3
Text 4
Text 5
Text 6
Concrete level: meaningful objects

Abstract level: feature vectors

Similarity matching

Dimensionality Reduction

Classification

Clustering

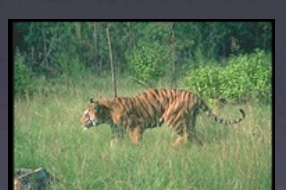
etc.

v1 v2 v3 v4 v5 v6

Bitmap Bitmap Bitmap Bitmap Bitmap Bitmap





Pic. 1 Pic. 2 Pic. 3 Pic. 4 Pic. 5 Pic.6

Concrete level: meaningful objects

Abstract level: feature vectors

Similarity matching

Dimensionality Reduction

Classification

Clustering

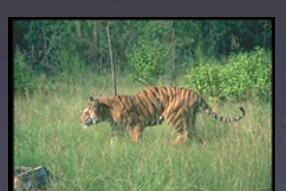
etc.

v1 v2 v3 v4 v5 v6

Bag of colors

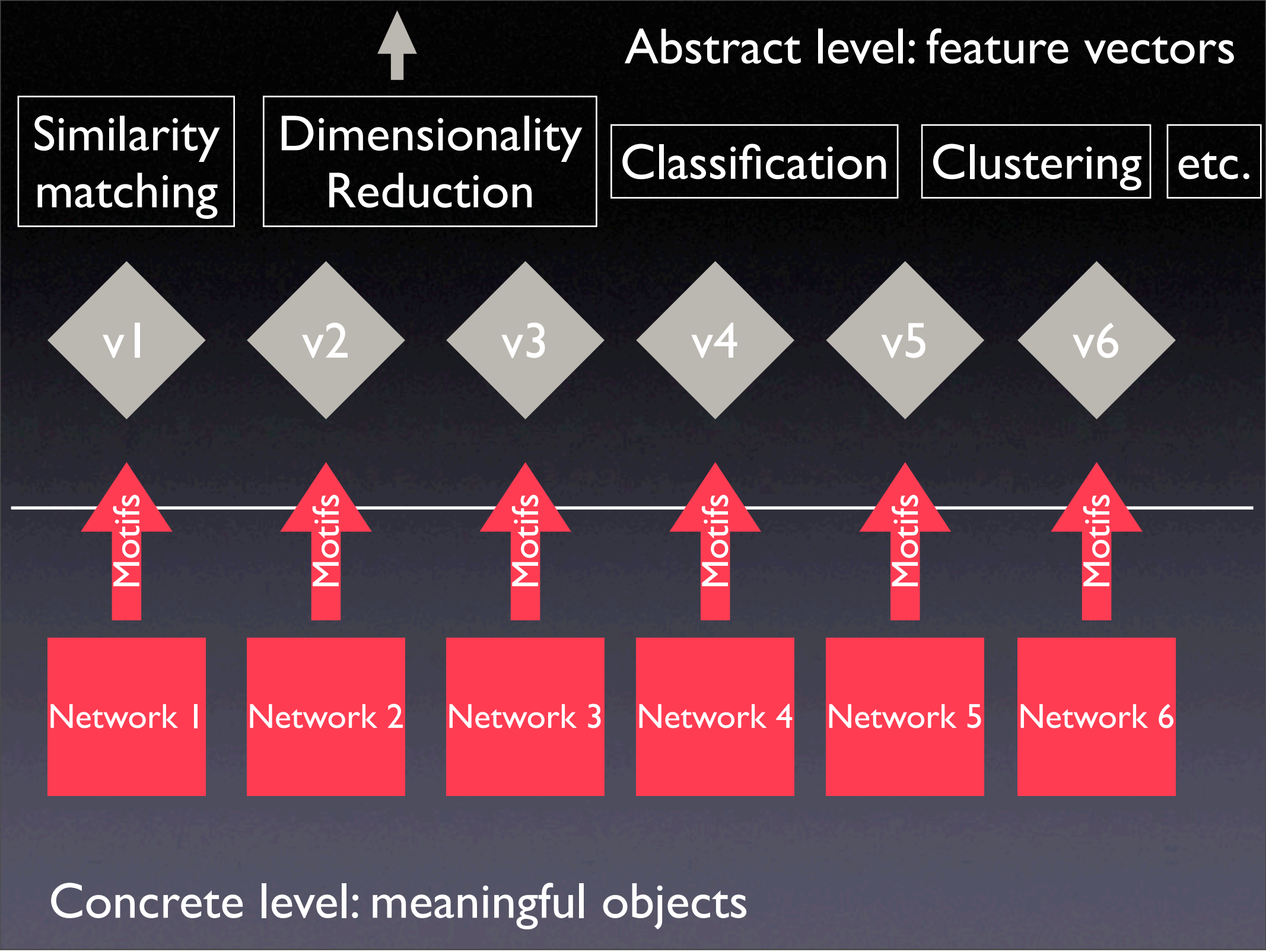




Pic. 1 Pic. 2 Pic. 3 Pic. 4 Pic. 5 Pic.6

Concrete level: meaningful objects


Abstract level: feature vectors
Similarity matching
Dimensionality Reduction
Classification
Clustering
etc.
v1
v2
v3
v4
v5
v6
Motifs
Motifs
Motifs
Motifs
Motifs
Motifs
Network 1
Network 2
Network 3
Network 4
Network 5
Network 6
Concrete level: meaningful objects

- Need to find right (relevant) representation
- Representation = concrete/abstract interface
  - Go read *The Sciences of the Artificial*!
- Great methods at the abstract level generally fail if the representation is bad
  - missing what's relevant
  - including what's irrelevant
  - comparing apples to platypi
    - both multicellular sexually-reproducing carbon-based lifeforms...
- A lot of your work will be designing representations

Abstract level: feature vectors

Similarity matching

Dimensionality Reduction

Classification

Clustering

etc.

v1 v2 v3 v4 v5 v6

BoW BoW Topics Bitmap Bag of colors Motifs

Text 1 Text 2 Text 3

Pic. 1

Pic. 2

Social Network

Concrete level: meaningful objects

flower1

flower2

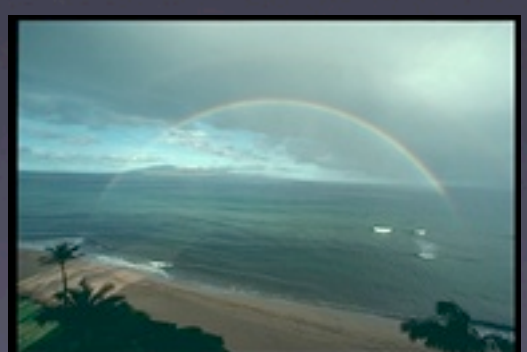
flower3

tiger1

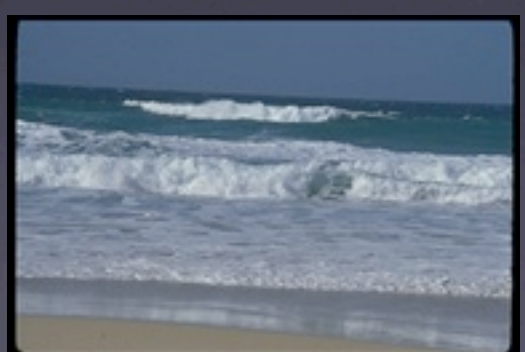
tiger2

tiger3

ocean1



ocean2



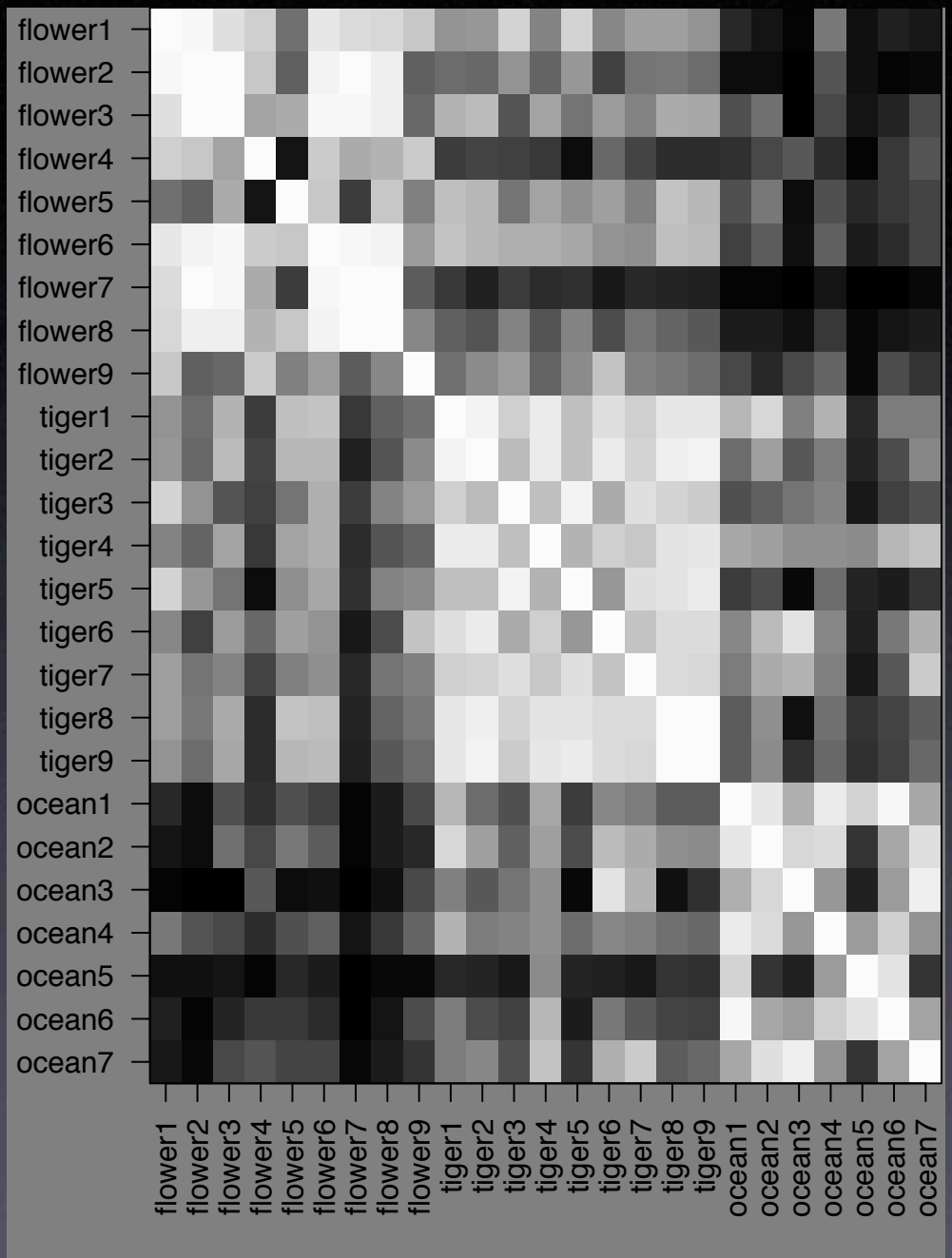
ocean3

Euclidean Distance of Images

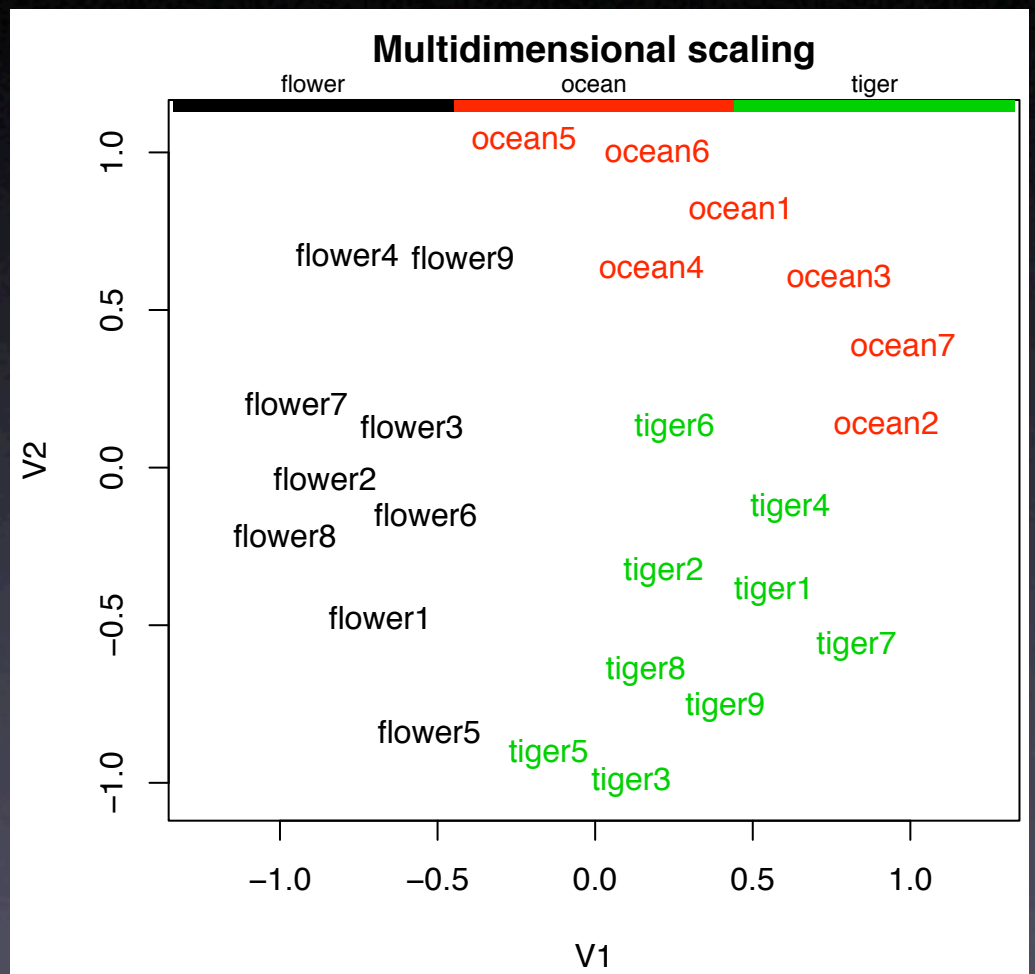
- Image is MxN pixels, each with 3 color components, so a 3MN vector
- Euclidean distance possible, and OK for some kinds of noise-removal
- but hopeless even at grouping flower1 with flower2
- or slight changes in perspective, lighting...

Bag of Colors

- "If it works, try it some more"
- For each possible color, count how many pixels there are of that color
- Use Euclidean distance on color-count vectors
- Too many colors, so *quantize* them down to a manageable number (like stemming, or combining synonyms)



Distances between images



MDS plot of images

Representation and Invariance

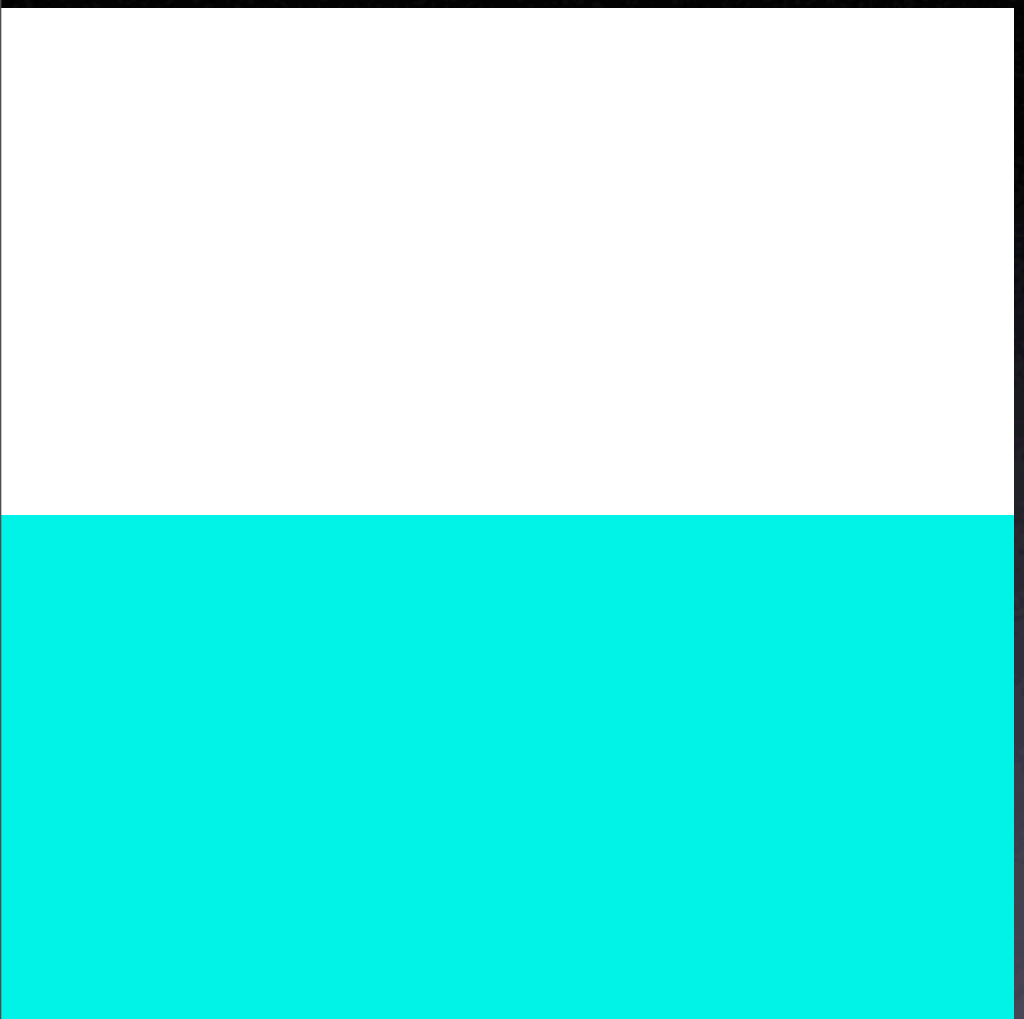
- Invariances of a representation = how can we change the underlying object without changing the representation?
- What differences does the representation ignore?

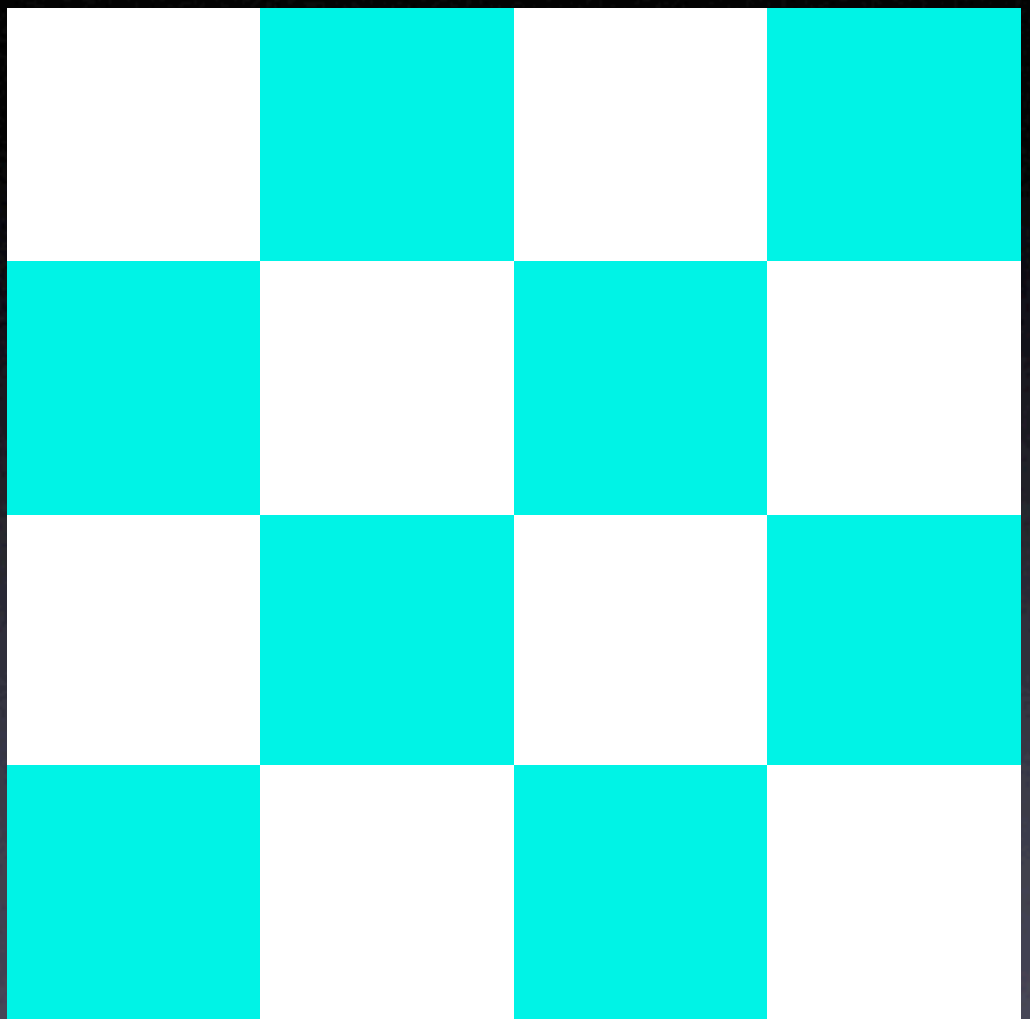
Invariants of bags of words

- Punctuation and word order
- Universal words (exact count of “the”, “of”, “to”, ...), if using inverse document frequency
- Word-endings, if using stemming
- Grammar, context, word proximity ...
  - “Send lawyers, guns and money” vs. “Sending the Guns’ lawyers for the money”

Invariants of bags of colors

- Small changes in orientation, pose, some rotations
- Small amounts of color noise or weird colors
- Texture





Same color counts, different textures

Non-invariants

- Lighting, shadows
- Occlusion, 3D effects
- Blurring
  - There are good ways to deal with blur (from astronomy)
- but full vision is *very, very hard*

- Breaking an invariance is easy
  - e.g., add features for textures
  - or sub-divide the image and do color-counts on each part
- Adding invariances is hard
  - often need to go back to scratch and chose a different representation