

Attribute Labeling

Danna Gurari

The University of Texas at Austin

Fall 2019



<https://www.ischool.utexas.edu/~dannag/Courses/CrowdsourcingForCV/CourseContent.html>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Review

- Last week:
 - Scene classification applications
 - Scene classification datasets: key steps in creating them
 - Scene classification datasets: scaling up with *crowdsourcing* and *challenges*
- Assignments (Canvas)
 - Lab assignment 1 due yesterday
 - Reading assignment 3 due next week
 - Lab assignment 2 due in three weeks
- Questions?

Today's Topics

- Attribute labeling applications
- Attributes: dataset creation approaches
- Beyond binary classification: relative/indistinguishable pairs of attributes
- Lab: Connecting to Amazon Mechanical Turk

Today's Topics

- Attribute labeling applications
- Attributes: dataset creation approaches
- Beyond binary classification: relative/indistinguishable pairs of attributes
- Lab: Connecting to Amazon Mechanical Turk

Attribute Definition

Description

(as opposed to naming)



How would you describe this object?

Attribute Definition

Description

(as opposed to naming)



How would you describe this scene?

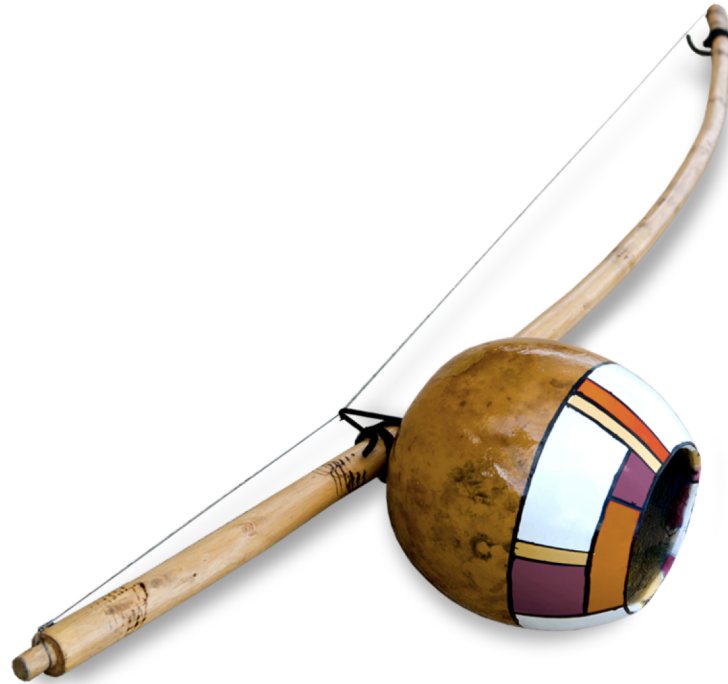
Attribute Definition

* Learning 30,000 objects equates to a person learning ~4.5 objects per day every day for 18 years

* Can be easier to “describe” than to “name” the unknown

Description

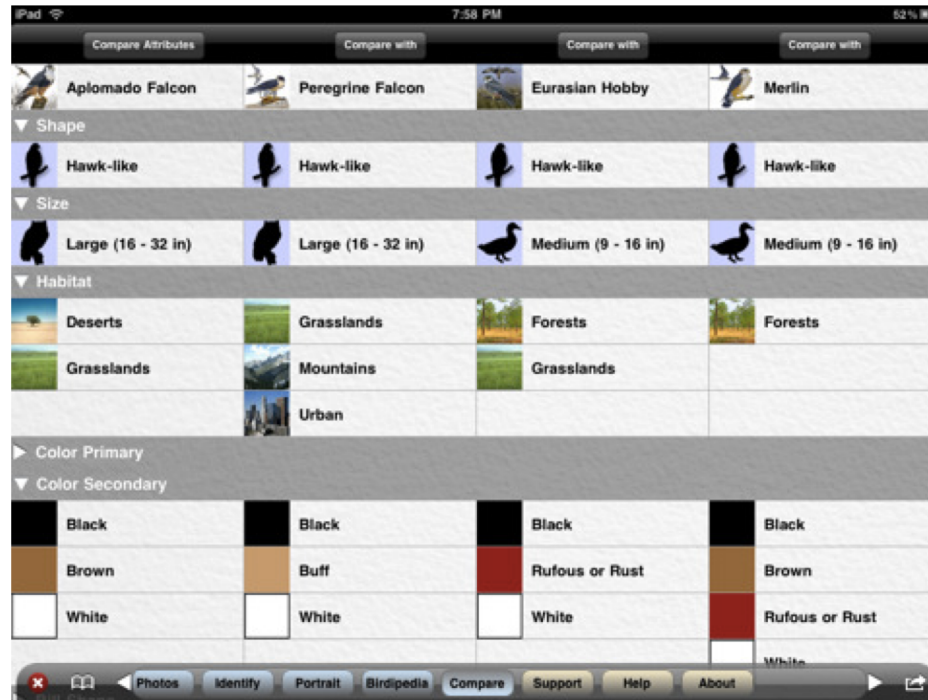
(as opposed to naming)



How would you describe this object?

Attribute: Recognition Applications

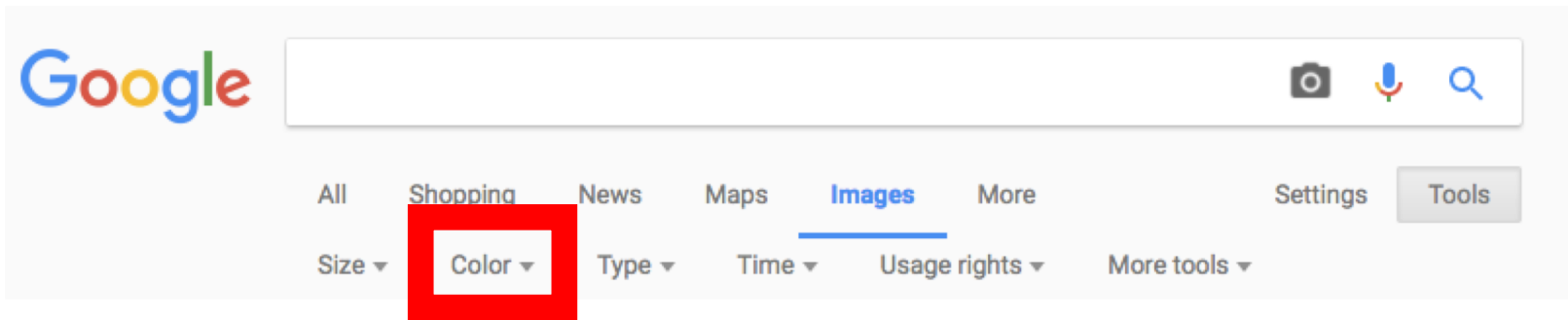
e.g., recognize objects with common knowledge instead of expert knowledge



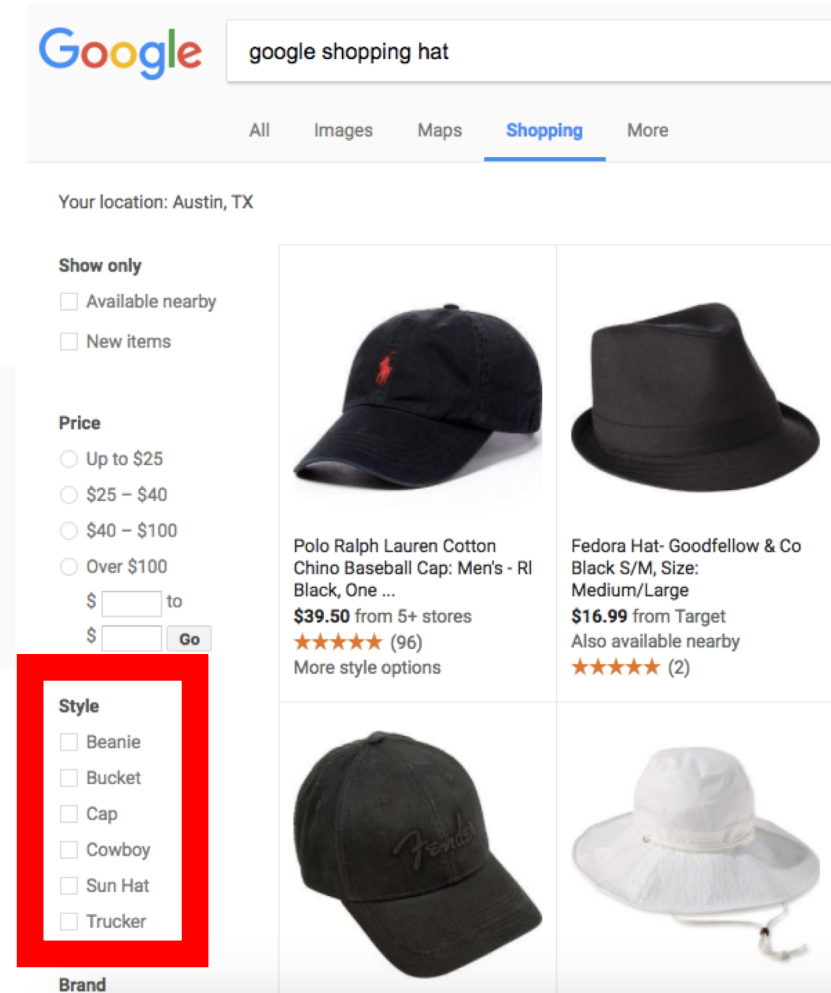
e.g., iBird: describe a bird to learn what type it is

Demo: https://www.youtube.com/watch?v=J1C-Q-z_np0

Applications: Expedite Search with Attributes



e.g., Image Search



e.g., Clothes Shopping

Additional Applications

- Recognize new objects with few/no examples; e.g., centaur



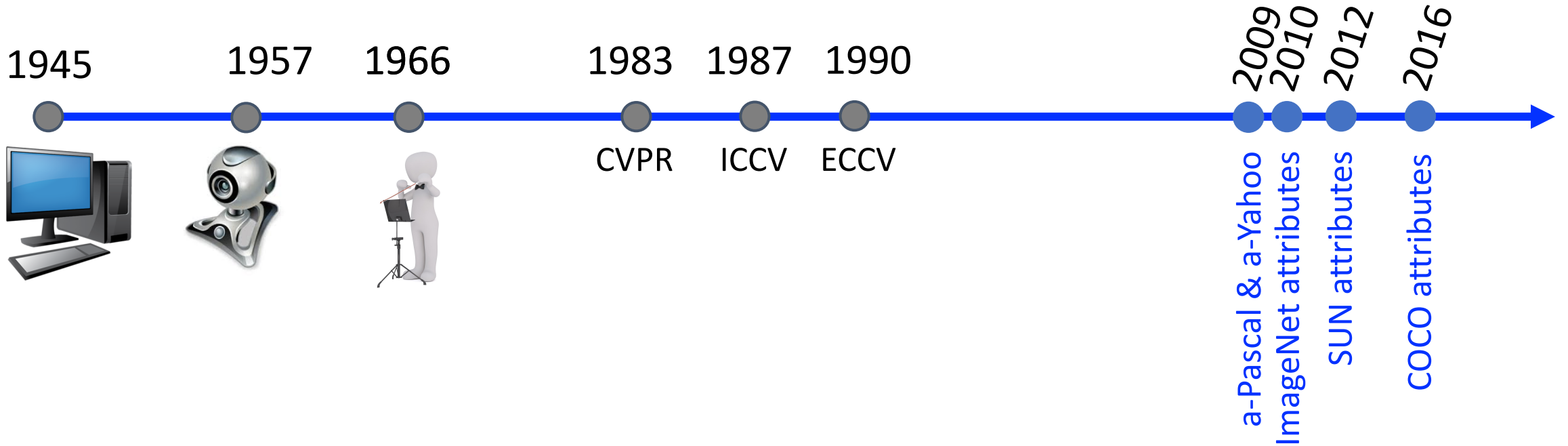
- Describe unusual aspects of a familiar object (intra-class variation); e.g.,



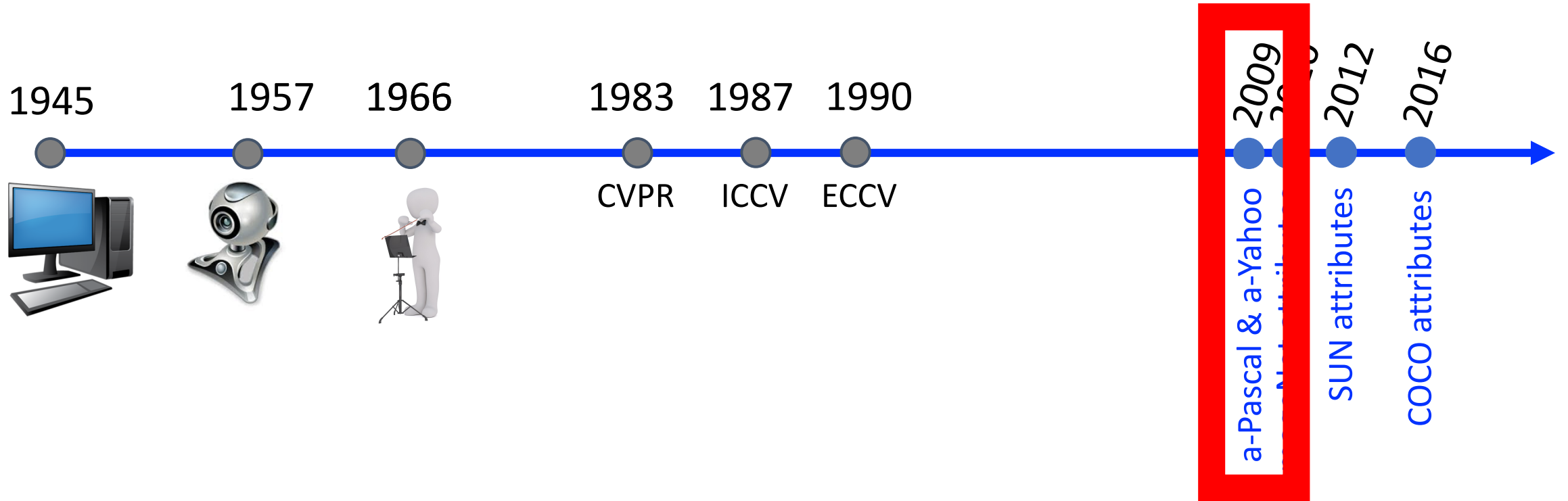
Today's Topics

- Attribute labeling applications
- **Attributes: dataset creation approaches**
- Beyond binary classification: relative/indistinguishable pairs of attributes
- Lab: Connecting to Amazon Mechanical Turk

Attribute Recognition Datasets



Attribute Recognition Datasets



Attribute Recognition Datasets: a-Pascal, a-Yahoo

1. Image Collection

- 12,000 VOC 2008 images
- Internet search on Yahoo!
for 12 object categories
- Objects are localized in
images with bounding boxes



Attribute Recognition Datasets: a-Pascal, a-Yahoo

1. Image Collection

- 12,000 VOC 2008 images
- Internet search on Yahoo! for 12 object categories
- Objects are localized in images with bounding boxes

2. Category Selection

- 64 attribute categories chosen by authors

1. **Shape attributes:** 2D and 3D properties such as “is 2D boxy”, “is 3D boxy”, “is cylindrical“, etc

2. **Part attributes:** parts that are visible, such as “has head”, “has leg”, “has arm”, “has wheel”, “has wing”, “has window”

3. **Material attributes:** describe what an object is made of, including “has wood”, “is furry”, “has glass”, “is shiny”

Attribute Recognition Datasets: a-Pascal, a-Yahoo

1. Image Collection

- 12,000 VOC 2008 images
- Internet search on Yahoo! for 12 object categories
- Objects are localized in images with bounding boxes

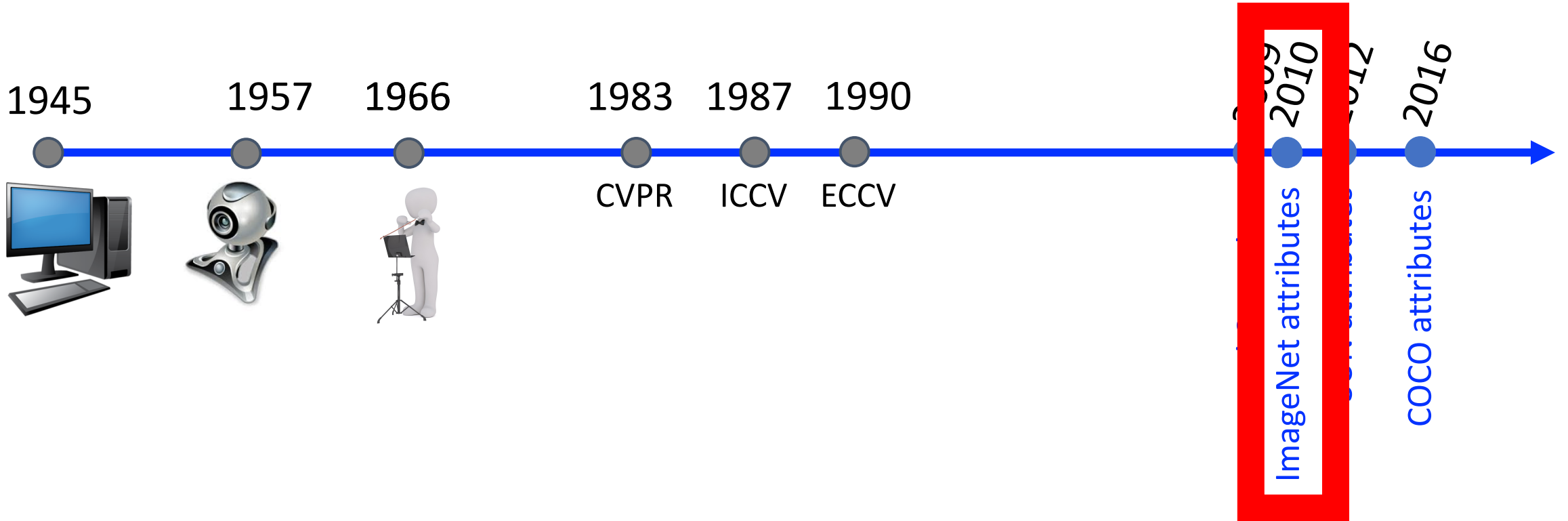
2. Category Selection

- 64 attribute categories chosen by authors

3. Human Labeling

- AMT crowd workers identify presence of each attribute

Attribute Recognition Datasets



Attribute Recognition Datasets: ImageNet

1. Image Collection

- Candidate images are all ImageNet images for which objects are localized in images with bounding boxes
- Include images in a “synset” for which the attribute is contained in the synset’s name or definition

Attribute Recognition Datasets: ImageNet

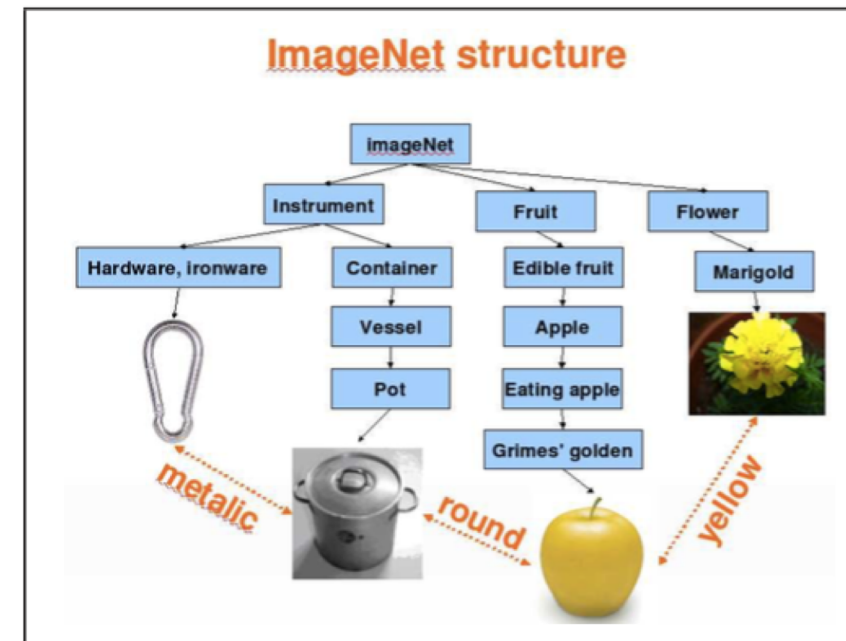
1. Image Collection

- Candidate images are all ImageNet images for which objects are localized in images with bounding boxes
- Include images in a “synset” for which the attribute is contained in the synset’s name or definition

2. Category Selection

- 20 categories:
 - (1) 8 colors
 - (2) furry, long, metallic, rectangular, rough, round, shiny, smooth, spotted, square, striped, wet, vegetation, wooden

Aim is to identify *visual* connections between objects



Attribute Recognition Datasets: ImageNet

1. Image Collection

- Candidate images are all ImageNet images for which objects are localized in images with bounding boxes
- Include images in a “synset” for which the attribute is contained in the synset’s name or definition

2. Category Selection

- 20 categories:
 - (1) 8 colors
 - (2) furry, long, metallic, rectangular, rough, round, shiny, smooth, spotted, square, striped, wet, vegetation, wooden

3. Human Labeling

- AMT crowd workers identify presence of each attribute for 106 images per HIT

Attribute Recognition Datasets: ImageNet

1. Task Design

Instructions:

- Color attribute: is a significant part of the object (at least 25%) that color?
- Other attributes: would person describe the object as a whole using that attribute?

Interface:

- Label one attribute




2. Crowdsourcing Platform

amazon mechanical turk™
Artificial Artificial Intelligence

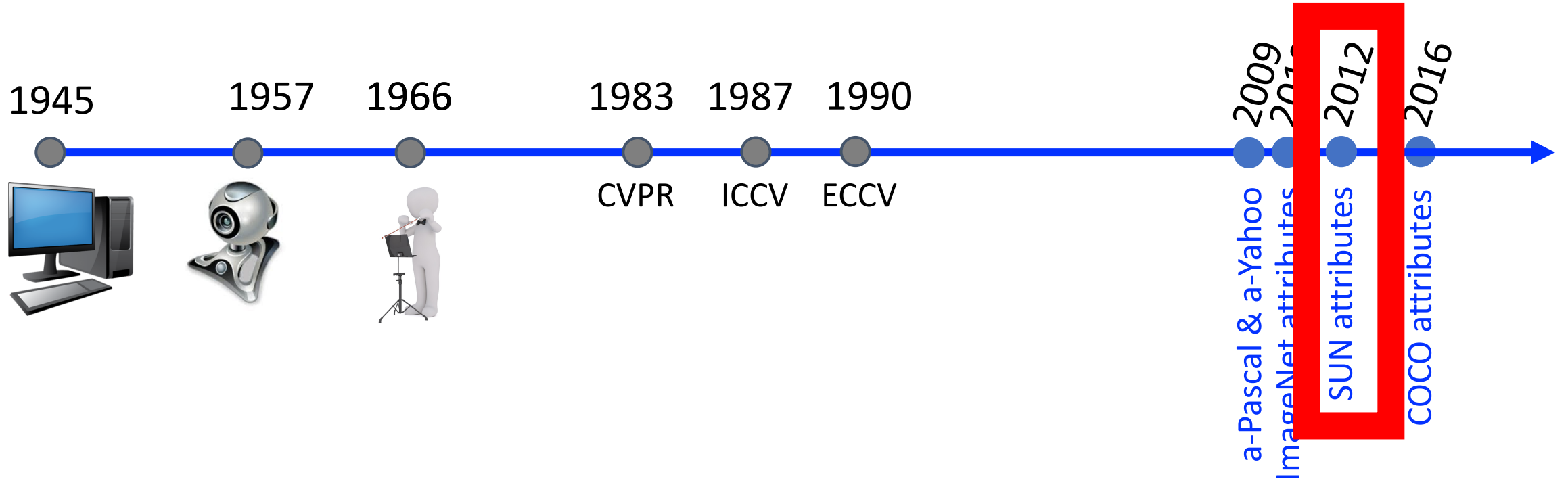
3. Quality Control

- 3 workers label each image and it is kept only if 3 workers agree that the attribute is present
- “Honeypot”: 6 known images per included (unspecified what was done)

Attribute Recognition Datasets: ImageNet

<p>metallic</p>	<p>fork (.72), transporter (.56), roller coaster (.49), stick (.41), wheel (.38), police van (.37), keyboard (.34), sail (.31), bridge (.31), building (.28), ski (.25), bowhead (.25)</p> 
<p>rectangular</p>	<p>police van (.90), transporter (.84), cabinet (.61), marimba (.50), window (.44), varietal (.42), flag (.38), bridge (.38), kummel (.31), pot (.29), generic (.28), pool table (.26)</p> 
<p>yellow</p>	<p>egg yolk (1.00), sunflower (.86), omelet (.70), kedgeree (.64), flan (.61), tostada (.48), succotash (.42), pizza (.35), zabaglione (.26), ravigote (.25), curry (.23), casserole (.21)</p> 

Attribute Recognition Datasets



Attribute Recognition Datasets: SUN

1. Image Collection

- 20 scenes from each of the
717 SUN scene categories

Attribute Recognition Datasets: SUN

1. Image Collection

- 20 scenes from each of the 717 SUN scene categories

2. Category Selection

- Discover *attribute types* from image descriptions by AMT workers: material, object & envelope, surface property, affordance, spatial

- Choose *discriminative* attributes offered by AMT workers for the 5 types

- Authors removed and added some categories resulting in 102 categories

Which attributes distinguish the scenes on the *left* from the scenes on the *right*?



rock, warm, barren, natural |

Attribute Recognition Datasets: SUN

1. Image Collection

- 20 scenes from each of the 717 SUN scene categories

2. Category Selection

- Discover *attribute types* from image descriptions by AMT workers: material, object & envelope, surface property, affordance, spatial
- Choose *discriminative* attributes offered by AMT workers for the 5 types
- Authors removed and added some categories resulting in 102 categories

3. Human Labeling

- AMT crowd workers identify presence of each attribute for 48 images per HIT

Attribute Recognition Datasets: SUN



1. Task Design

Instructions:

Scene Attribute Labeling When you mouse over one of the images, a larger version of that image will appear in the box below.

Click on the scenes below that contain the following lighting or material:

camping Either an actual camp site, or scene in wilderness suitable enough for humans to make a tent and/or sleep.

Example Scene Example Scene

These HITs are reviewed before being approved or rejected. [For further instructions Click Here!](#)

This task can be very subjective. If you are not sure about which images should be selected, please *SKIP THIS HIT* or email us to ask for clarification. There are more HITs with less subjective attributes.

Interface:



Images continued down the page ... ↓

Attribute Recognition Datasets: SUN

1. Task Design

Instructions:

Scene Attribute Labeling

Click on the scenes below that contain the following lighting or material:

camping *Either an actual camp site, or scene in wilderness suitable enough for humans to make a tent and/or sleep.*



Example Scene Example Scene

When you mouse over one of the images, a larger version of that image will appear in the box below.

These HITs are reviewed before being approved or rejected.

[For further instructions Click Here!](#)

This task can be very subjective. If you are not sure about which images should be selected, please ***SKIP THIS HIT*** or email us to ask for clarification. There are more HITs with less subjective attributes.

Interface:



Images continued down the page ... ↓

Scene Attribute Labeling

Click on the scenes below that contain the following lighting or material:

camping *Either an actual camp site, or scene in wilderness suitable enough for humans to make a tent and/or sleep.*



Example Scene Example Scene

When you mouse over one of the images, a larger version of that image will appear in the box below.



These HITs are reviewed before being approved or rejected.

[For further instructions Click Here!](#)

This task can be very subjective. If you are not sure about which images should be selected, please ***SKIP THIS HIT*** or email us to ask for clarification. There are more HITs with less subjective attributes.

Attribute Recognition Datasets: SUN

1. Task Design

(grid of 48 images)

Instructions:

Scene Attribute Labeling When you mouse over one of the images, a larger version of that image will appear in the box below.

Click on the scenes below that contain the following lighting or material:

camping Either an actual camp site, or scene in wilderness suitable enough for humans to make a tent and/or sleep.

 Example Scene

 Example Scene



These HITs are reviewed before being approved or rejected.

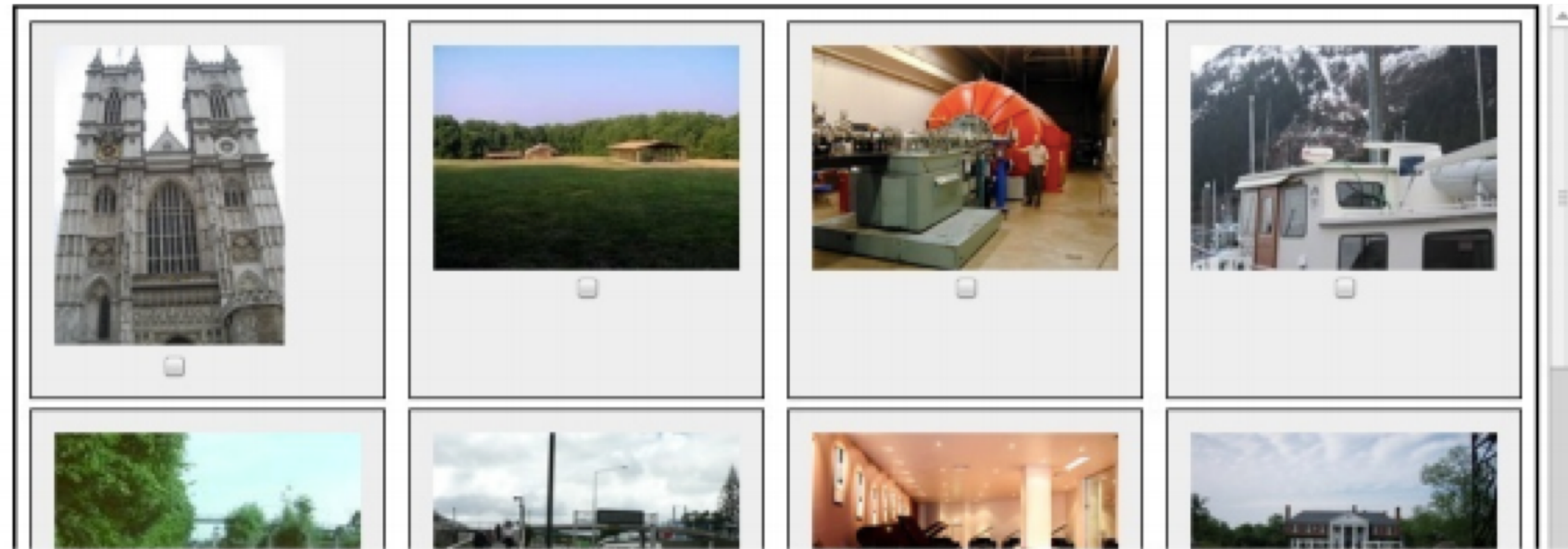
[For further instructions Click Here!](#)

This task can be very subjective. If you are not sure about which images should be selected, please *SKIP THIS HIT* or email us to ask for clarification. There are more HITs with less subjective attributes.

Interface:



Images continued down the page ... ↓



Images continued down the page ... ↓

Attribute Recognition Datasets: SUN

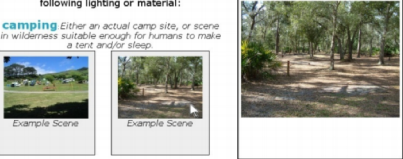
1. Task Design

Instructions:

Scene Attribute Labeling When you mouse over one of the images, a larger version of that image will appear in the box below.

Click on the scenes below that contain the following lighting or material:

camping Either an actual camp site, or scene in wilderness suitable enough for humans to make a tent and/or sleep.




Example Scene Example Scene Example Scene

These HITs are reviewed before being approved or rejected.
[For further instructions Click Here!](#)

This task can be very subjective. If you are not sure about which images should be selected, please *SKIP THIS HIT* or email us to ask for clarification. There are more HITs with less subjective attributes.

Interface:



Images continued down the page ... ↓

2. Crowdsourcing Platform



3. Quality Control

- Workers must pass a quiz
- List of “suspicious” workers created: any worker whose average # of labels or work time exceeds one standard deviation from average of all workers
- “Bad” workers filtered: found by sampling work from “suspicious workers” and from random sampling
- “Good” workers cultivated
- Relabel all images twice w/ 28 trusted workers; use majority vote of 3

Attribute Recognition Datasets: SUN

1. Image Collection

- 20 scenes from each of the 717 SUN scene categories

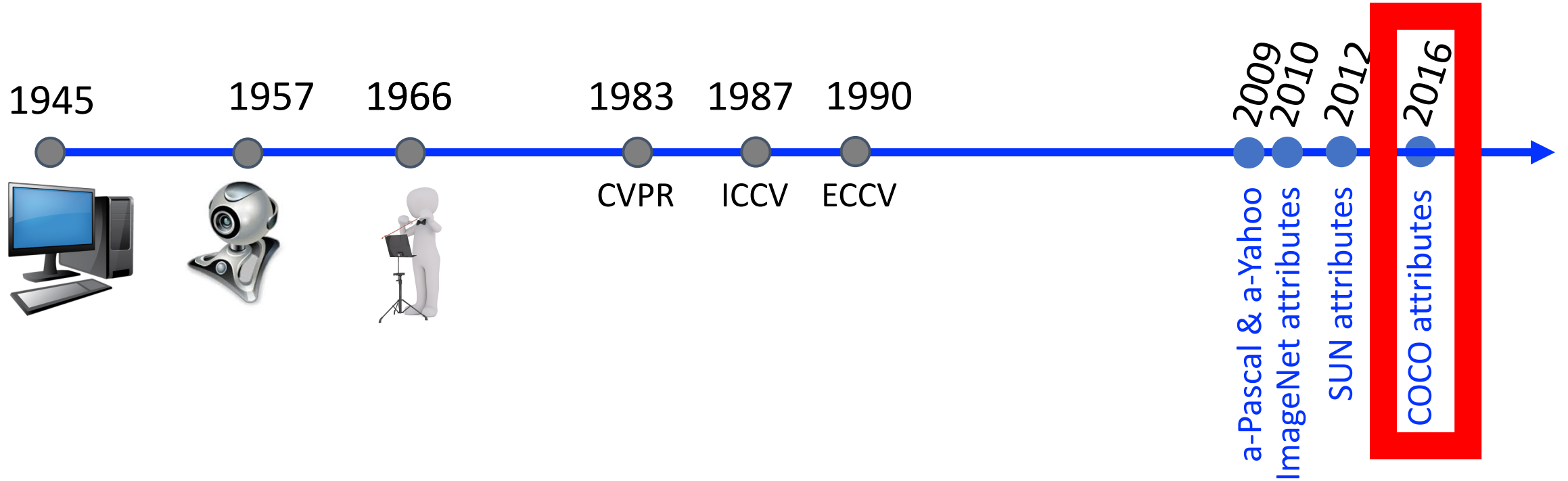
2. Category Selection

- Discover *attribute types* from image descriptions by AMT workers: material, object & envelope, surface property, affordance, spatial
- Choose *discriminative* attributes offered by AMT workers for the 5 types
- Authors removed and added some categories resulting in 102 categories

3. Human Labeling

- AMT crowd workers identify presence of each attribute for 48 images per HIT

Attribute Recognition Datasets



Attribute Recognition Datasets: COCO

1. Image Collection

- Subset of COCO images
which have objects localized
with bounding boxes

Attribute Recognition Datasets: COCO

1. Image Collection

- Subset of COCO images which have objects localized with bounding boxes


2. Category Selection

- Attributes solicited from AMT crowd workers
- Sample of candidate adjectives mined from a New York Times Corpus were suggested to workers
- Attributes discovered for object categories in COCO
- Authors manually refined list to 196 attributes


Attribute Recognition Datasets: COCO

Task:

Picture A



Picture B



What is different about the *person* in **Picture A** compared to the *person* in **Picture B**?

In each picture, the *person* is outlined in white.
Please provide at least three new descriptive words.

Words other people thought of:

- bored
- thinking
- female
- male
- playing
- watching
- standing
- young

New descriptive words:

smiling
surfing

Submit New Words

Attribute Recognition Datasets: COCO

1. Image Collection

- Subset of COCO images which have objects localized with bounding boxes

2. Category Selection

- Attributes solicited from AMT crowd workers
- Sample of candidate adjectives mined from a New York Times Corpus were suggested to workers
- Attributes discovered for object categories in COCO
- Authors manually refined list to 196 attributes

3. Human-in-the-Loop Labeling

- AMT crowd workers identify presence of each attribute for 48 images per HIT

Attribute Recognition Datasets: COCO

1. Task Design

Instructions:

Instructions

Examine the outlined Dog in each image. Click on all of the descriptive **attributes** that apply to the Dog in the image.

In each picture, the Dog is outlined in white. If you hover your cursor over the image, the outline will disappear.



This person Not this person

If none of the **attributes** apply, check None.

Interface:

Dogs **Attributes**



Image 1

- enjoying
- loving
- sleeping
- sniffing
- hunting
- hanging
- sneezing
- smiling / sniffing
- stretching
- coughing
- None

- barking
- talking
- walking
- leaping
- grazing
- cleaning
- socializing
- drinking
- learning

Image 2

Attribute Recognition Datasets: COCO

1. Task Design

Instructions:

Instructions

Examine the outlined Dog in each image. Click on all of the descriptive attributes that apply to the Dog in the image.

In each picture, the Dog is outlined in white. If you hover your cursor over the image, the outline will disappear.



This person. Not this person.

If none of the attributes apply, check None.

Interface:

Dogs



Attributes

Image 1

- enjoying
- loving
- looking
- sniffing
- hunting
- hanging
- sneezing
- smelling / sniffing
- stretching
- coughing
- None

Image 2

- drinking
- holding
- talking
- walking
- listening
- grazing
- cleaning
- socializing
- driving
- learning

Instructions

Examine the outlined Dog in each image. Click on all of the descriptive attributes that apply to the Dog in the image.

In each picture, the Dog is outlined in white. If you hover your cursor over the image, the outline will disappear.



This person. Not this person.

If none of the attributes apply, check None.

In order to receive a Qualification for this type of HIT, please complete an Image or Object Annotation Qualification Quiz HIT. If you pass the quiz, you will be granted the qualification.

Attribute Recognition Datasets: COCO


1. Task Design

Instructions:

Instructions

Examine the outlined Dog in each image. Click on all of the descriptive attributes that apply to the Dog in the image.

In each picture, the Dog is outlined in white. If you hover your cursor over the image, the outline will disappear.




This person Not this person

If none of the attributes apply, check None.

Interface:

Dogs





Attributes

Image 1

- enjoying
- loving
- feeding
- floating
- hunting
- hanging
- skating
- smelling / sniffing
- stretching
- cooking
- None
- thinking
- holding
- talking
- skiing
- listening
- grazing
- looting
- socializing
- driving
- learning

Image 2

Dogs (10 images)



Attributes

Image 1

- enjoying
- loving
- feeding
- floating
- hunting
- hanging
- skating
- smelling / sniffing
- stretching
- cooking
- None
- thinking
- holding
- talking
- skiing
- listening
- grazing
- looting
- socializing
- driving
- learning

(20 attributes)

Image 2

Attribute Recognition Datasets: COCO

1. Task Design

Instructions:

Instructions

Examine the outlined Dog in each image. Click on all of the descriptive **attributes** that apply to the Dog in the image.

In each picture, the Dog is outlined in white. If you hover your cursor over the image, the outline will disappear.



If none of the **attributes** apply, check None.

Interface:

Dogs




Attributes

Image 1

- enjoying
- loving
- leaping
- flailing
- hunting
- hanging
- sneezing
- smelling / sniffing
- stretching
- coughing
- None
- hissing
- holding
- talking
- asking
- salivating
- sneezing
- sneezing
- socializing
- drinking
- learning

Image 2

2. Crowdsourcing Platform



Artificial Artificial Intelligence

(\$0.10 per HIT)

3. Quality Control

- Workers must pass a quiz
- List of “suspicious” workers created: any worker whose average disagreement with trusted annotations exceeds one standard deviation from average disagreement of all workers
- “Bad” workers filtered: found by verifying work of “suspicious workers”
- Use majority vote label of 3 people

Attribute Recognition Datasets: COCO

1. Image Collection

- Subset of COCO images which have objects localized with bounding boxes

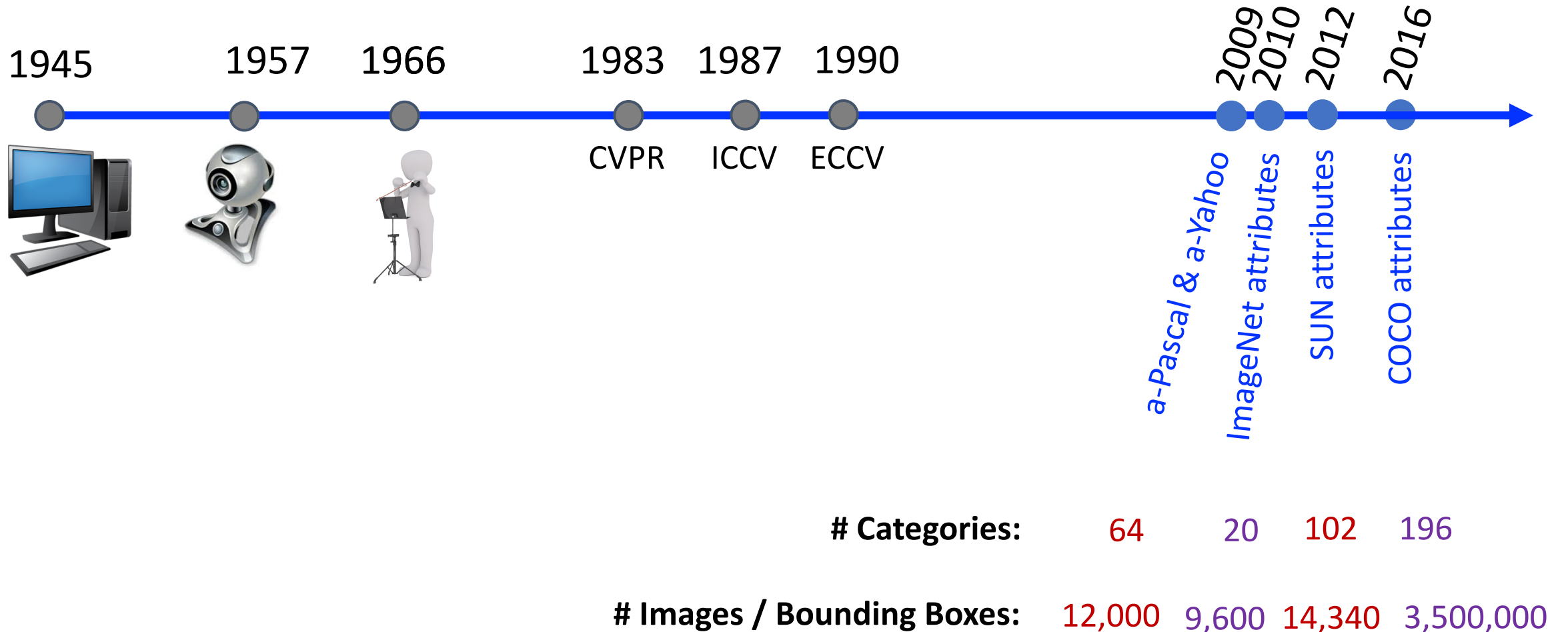
2. Category Selection

- Attributes solicited from AMT crowd workers
- Sample of candidate adjectives mined from a New York Times Corpus were suggested to workers
- Attributes discovered for object categories in COCO
- Authors manually refined list to 196 attributes

3. Human-in-the-Loop Labeling

- AMT crowd workers identify presence of each attribute for 48 images per HIT

Attribute Recognition Datasets



Attribute Recognition Datasets: Summary

- Key steps in creating dataset:

1. Image Collection

2. Category Selection

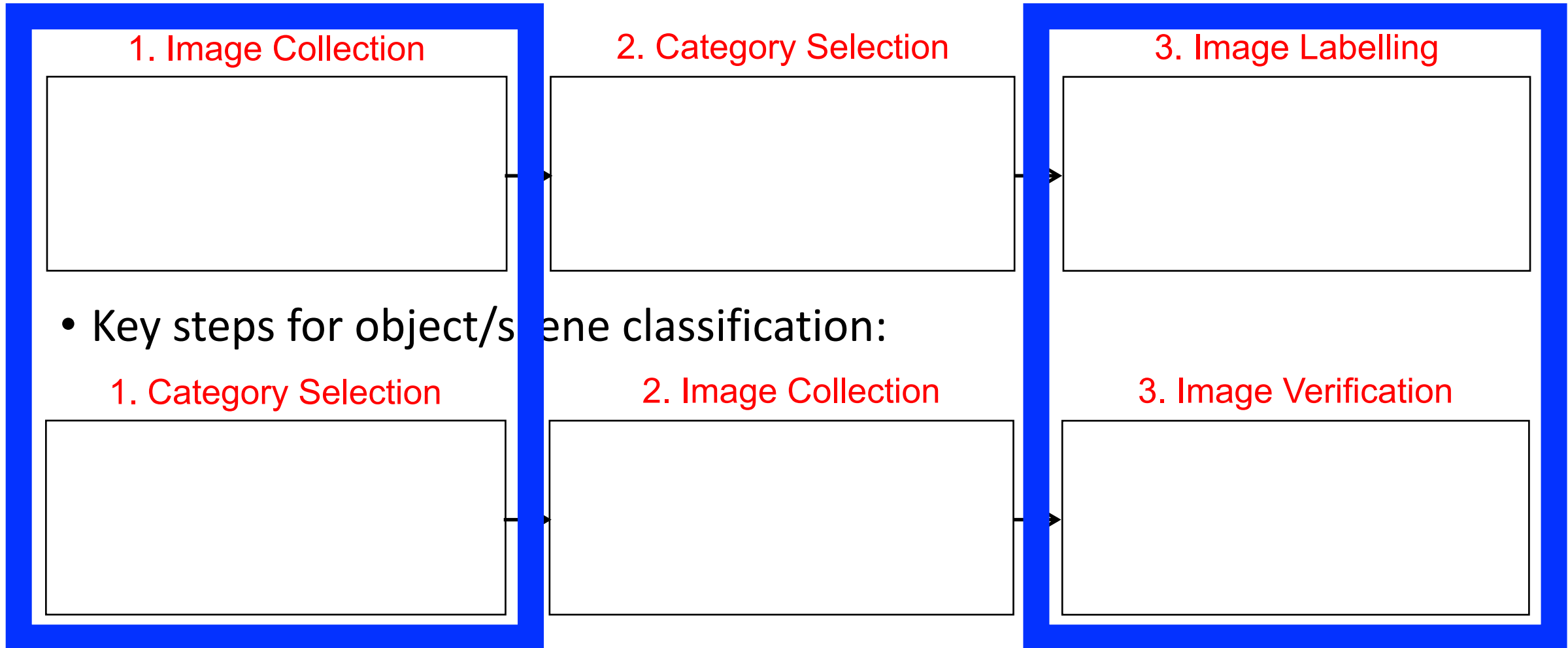
3. Image Labelling



How is this process different from the process used for object recognition and scene classification?

Attribute Recognition Datasets: Summary

- Key steps in creating dataset:



Today's Topics

- Attribute labeling applications
- Attributes: dataset creation approaches
- **Beyond binary classification: relative/indistinguishable pairs of attributes**
- Lab: Connecting to Amazon Mechanical Turk

Discussion: Challenges of Attribute Labeling

What is the shape of this flag? *



Is this drinkable? *



Is the person on the right taller than the person on the left? *



What is the color of these shoes? *



Is this person smiling? *



1. What makes each task difficult?
2. What label do you agree on for each task and why?

Relative Attributes (Rather Than Categorical)

Has a *spectrum* of strengths; e.g.,

- Expressions (smiling, surprised)
- Shapes (flat, boxy)
- Material properties (metallic, furry),
- Functions (suitable, drinkable)

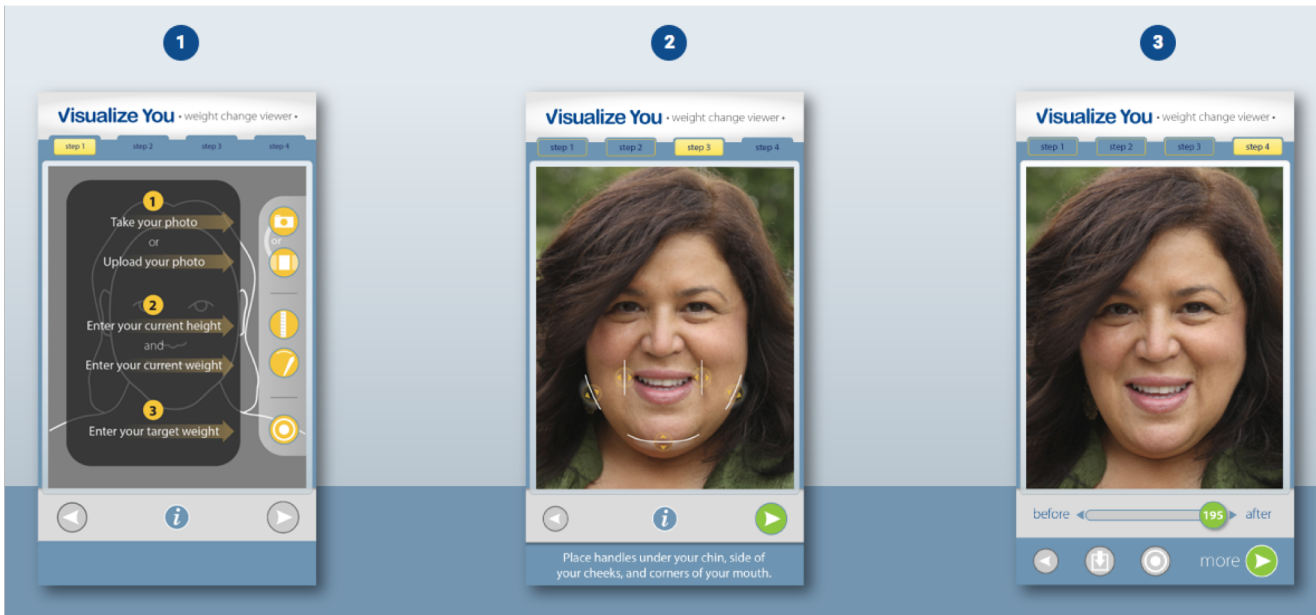


Relative Attributes for Shoe Shopping

The image shows a web interface for "Whittle Search" designed for finding shoes based on relative attributes. On the left, a reference image of a yellow and grey sneaker is shown with the text "Find Shoes like the one below". To the right, a control panel titled "Select a Range of the Attribute Strengths on Sliders below" features three horizontal sliders. The first slider is for "BrightColored", the second for "Feminine", and the third for "Sporty". Each slider has "Less" and "More" labels and a central handle. Below the sliders, a small image of a pair of sandals is shown, indicating the current selection. To the right of the sliders are three attribute labels: "BrightColored X", "Feminine X", and "Sporty X", each with a close button. Below the control panel, a text prompt reads "Give feedback using images below as references | Indicate more/less of an attribute than the reference image". The bottom section of the interface displays a grid of 14 different shoe images for selection.

Demo: <https://www.youtube.com/watch?v=3A6YkHn6OU0>

Relative Attributes for Altering Appearance



e.g., simulate weight loss/gain
www.visualizeyourweight.com



e.g., simulate aging and different lifestyles
<http://www.mastersingerontology.com/top-25-incredible-age-progression-tools-online.html>

Relative Attributes for Finding Criminals



Please compare the subject in the lower video to the subject in the top video.
For example if the subject in the bottom video is taller than the subject

Attribute	Annotation	Certainty
Age	Older	100%
Bottom subject is OLDER than the top		
Hair Colour	Same	100%
Subjects have roughly the SAME hair colour		
Hair Length	Longer	100%
Bottom subject has LONGER hair than the top		
Height	Taller	100%
Bottom subject is TALLER than the top		
Figure	Same	100%
Subjects both have roughly the SAME figure		
Neck Length	Same	100%
Subjects have roughly the SAME length neck		
Neck Thickness	Thinner	100%
Bottom subject has a THINNER neck than the top		
Shoulder Shape	Same	100%
Subjects have roughly the SAME shoulder shape		
Chest	Same	100%
Subjects have roughly the SAME size chest		
Arm Length	Longer	100%
Bottom subject has a LONGER arms than the top		

e.g., Biometrics: “the suspect is *taller* than him”
[D. Reid, M. Nixon, IJCB 2011]

Today's Topics

- Attribute labeling applications
- Attributes: dataset creation approaches
- Beyond binary classification: relative/indistinguishable pairs of attributes
- **Lab: Connecting to Amazon Mechanical Turk**