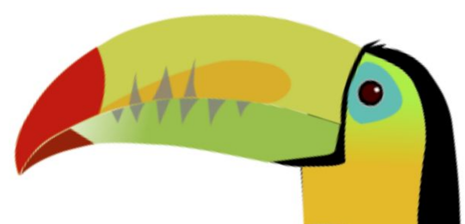
Fine Grained Visual Category Recognition and Perceptual Embedding

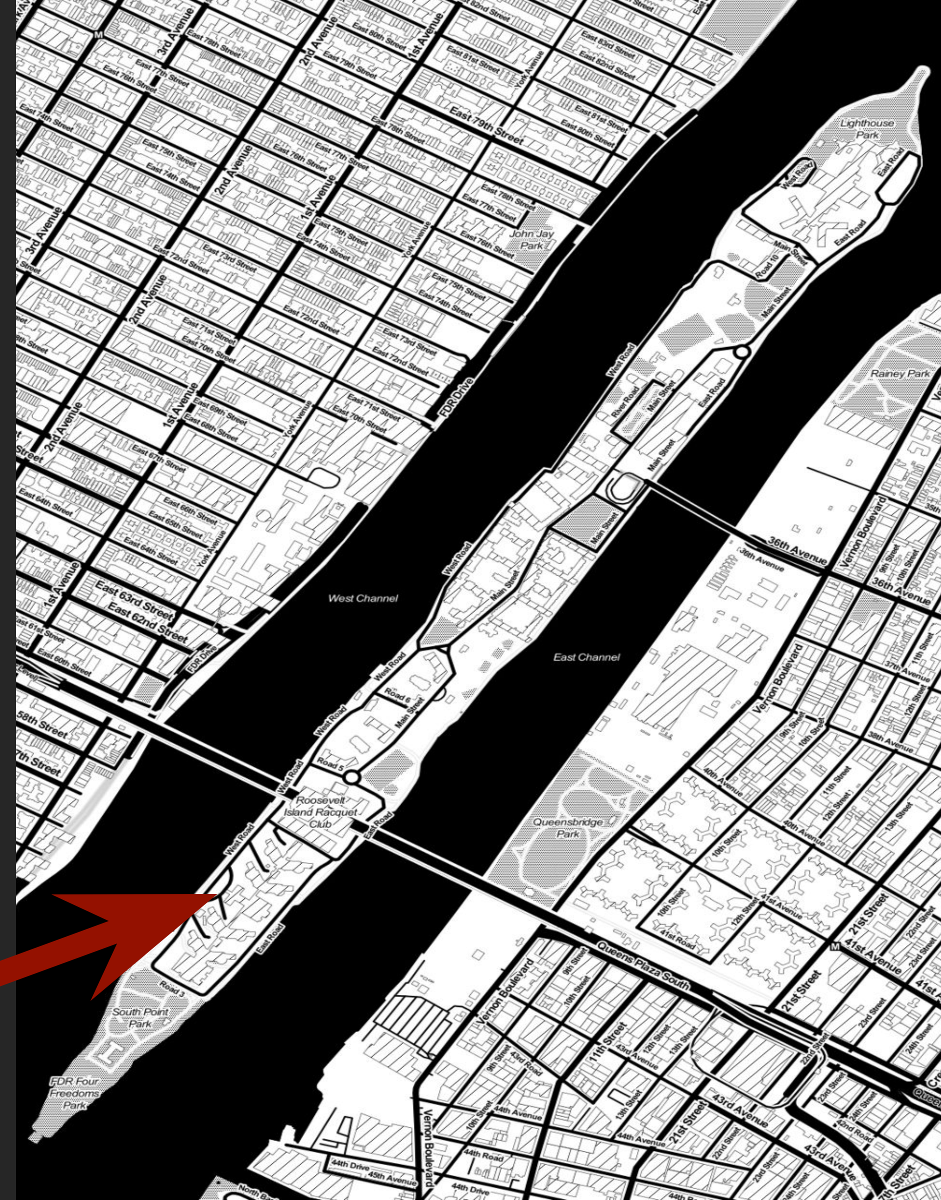
Serge Belongie



Cornell Tech is the technology-focused campus of Cornell University

Founded in 2012

Future Location

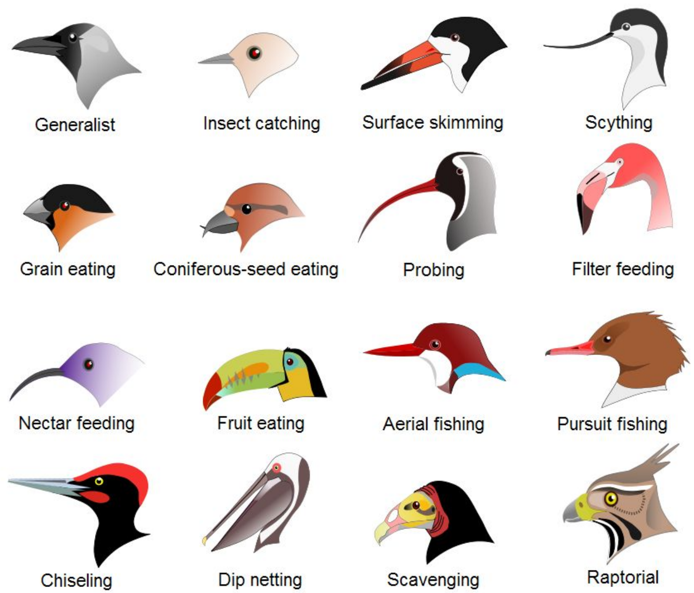


FUTURE CAMPUS (2017)

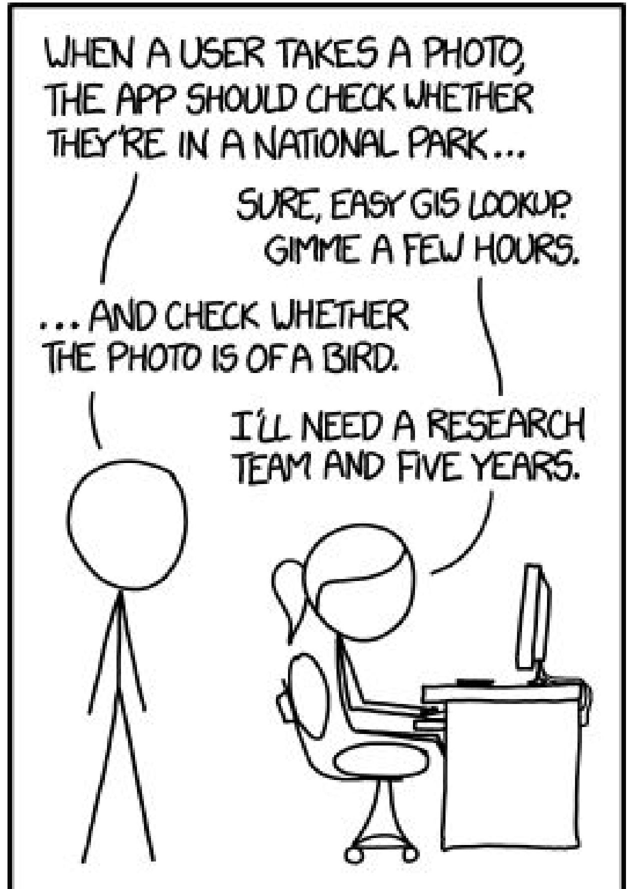
Part I
Fine Grained Category Recognition with Humans in the Loop

What Is Visipedia?

- A user-generated encyclopedia of visual knowledge
- An effort to associate articles with large quantities of well-organized, intuitive visual concepts



http://en.wikipedia.org/wiki/Bird



IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

http://xkcd.com/1425

code.flickr.com

Flickr | Flickr Blog | @flickr | @flickrapi | Developer Guidelines | API | Jobs

Search

Posted on October 20, 2014 by Rob Hess, Clayton Mellina, and Friends

← Previous

Introducing: Flickr PARK or BIRD

Zion National Park Utah by Les Haines (cc) BY

OR

Secretary Bird by Bill Gracey (cc) BY-NC-ND

tl;dr: Check it out at parkorbird.flickr.com!

Motivation

- People will willingly label or organize certain images if:
  - They are interested in a particular subject matter
  - They have the appropriate expertise

Ring-tailed lemur

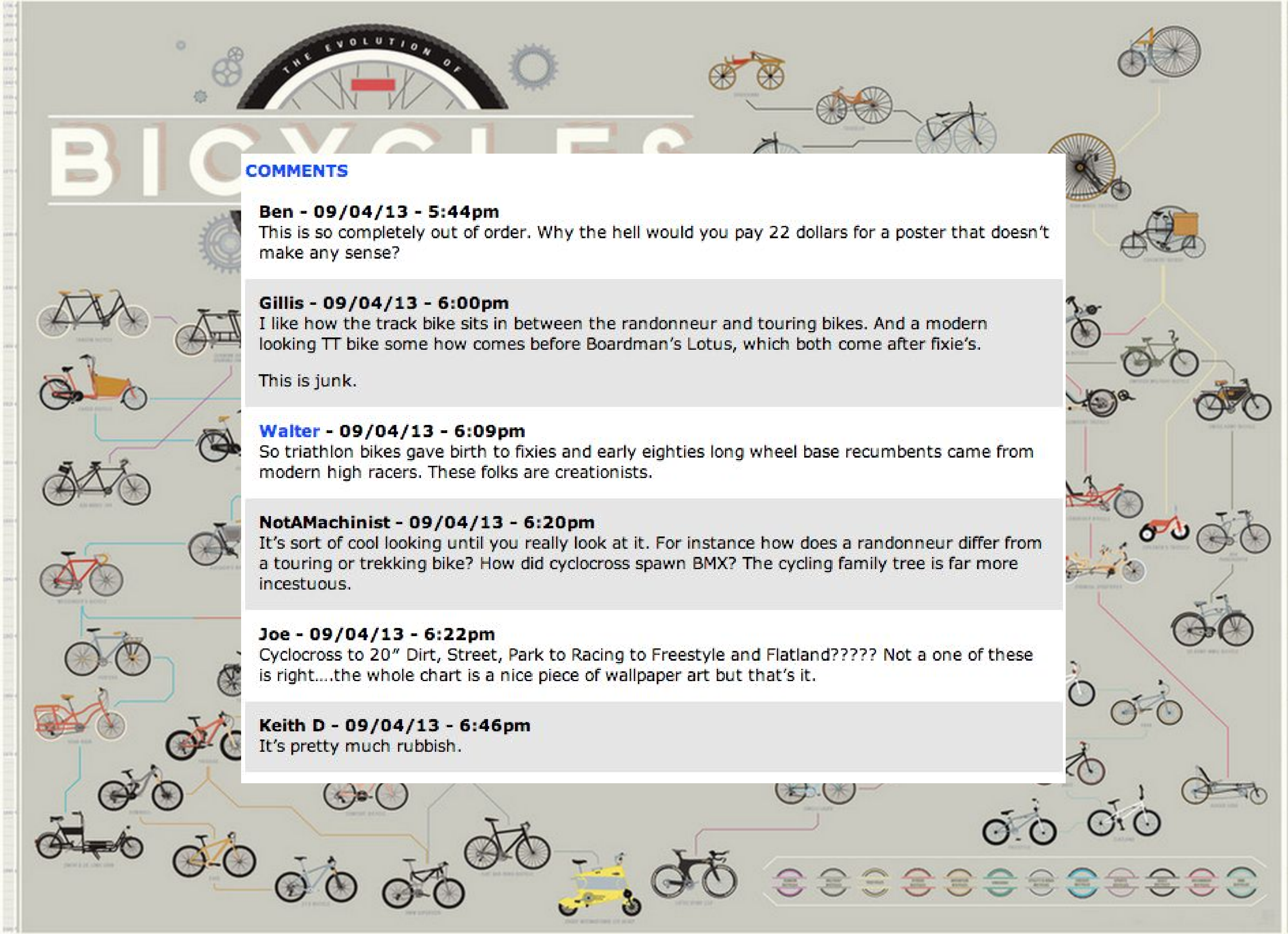
Thruxton Jackaroo

A DUBIOUSLY ACCURATE 233 YEAR HISTORY OF CYCLING

posted by Saris - September 4, 2013 - 5pm EDT

While we're not particularly certain about some of the claims in this bicycle family tree (e.g. freeride bikes spawned downhill bikes, which gave birth to 29ers?), we're certain you'll appreciate the artwork. You can head to their website to buy your own copy and laugh at the implication that big wheels evolved into recumbents for only $22.

[BikeRumor.com]



COMMENTS

Ben - 09/04/13 - 5:44pm
This is so completely out of order. Why the hell would you pay 22 dollars for a poster that doesn't make any sense?

Gillis - 09/04/13 - 6:00pm
I like how the track bike sits in between the randonneur and touring bikes. And a modern looking TT bike some how comes before Boardman's Lotus, which both come after fixie's.

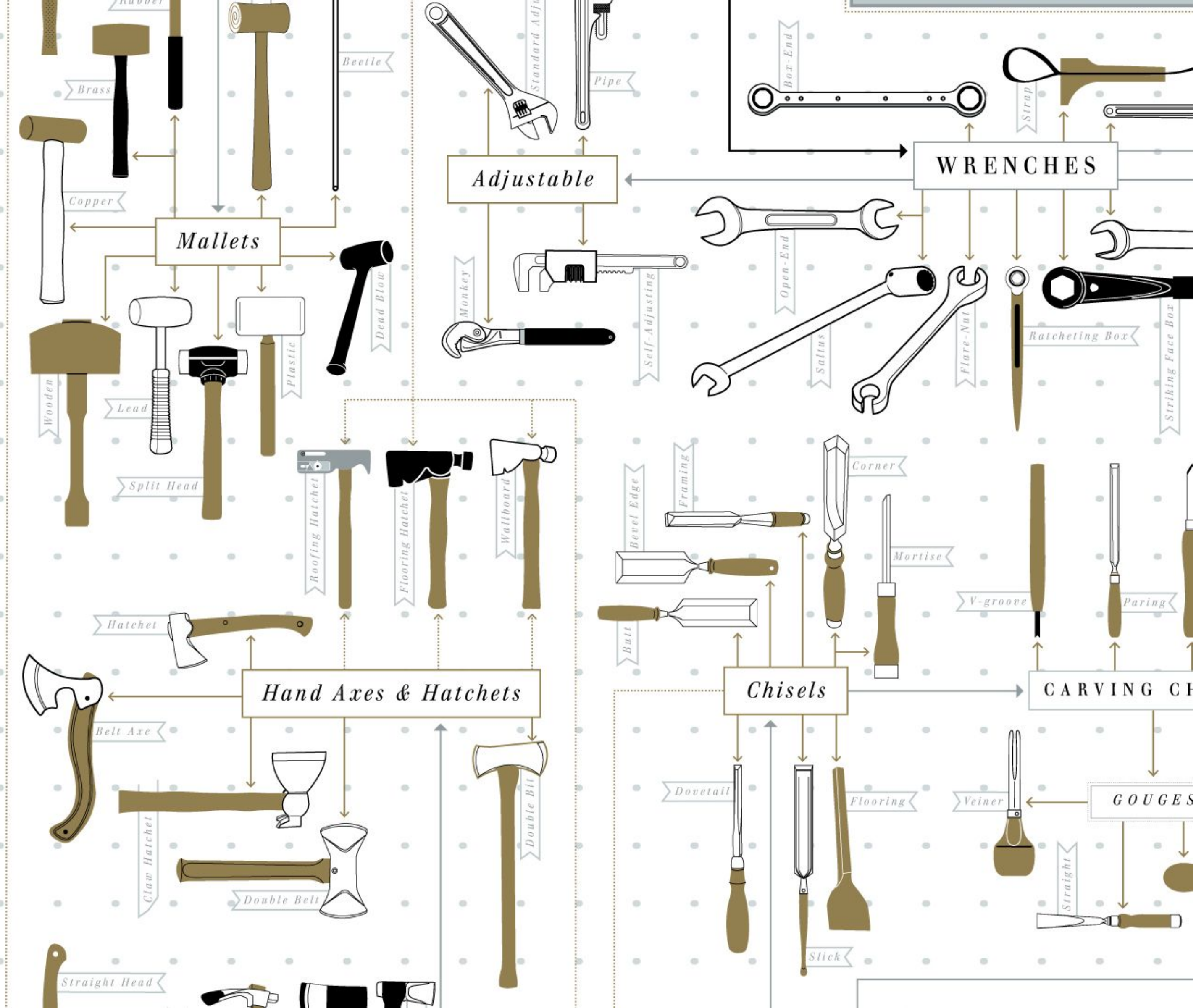
This is junk.

Walter - 09/04/13 - 6:09pm
So triathlon bikes gave birth to fixies and early eighties long wheel base recumbents came from modern high racers. These folks are creationists.

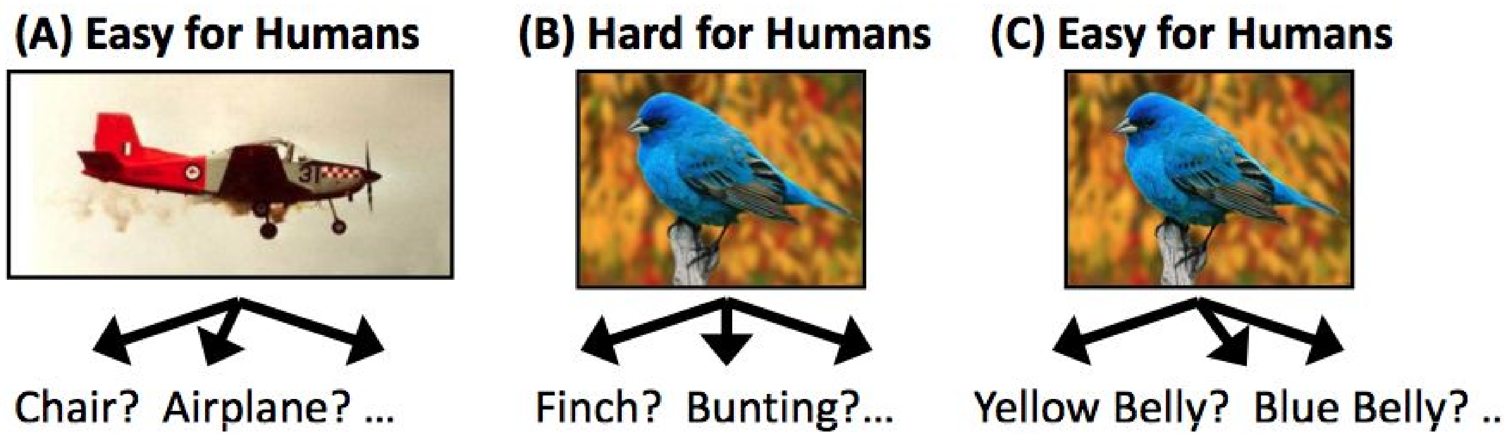
NotAMachinist - 09/04/13 - 6:20pm
It's sort of cool looking until you really look at it. For instance how does a randonneur differ from a touring or trekking bike? How did cyclocross spawn BMX? The cycling family tree is far more incestuous.

Joe - 09/04/13 - 6:22pm
Cyclocross to 20" Dirt, Street, Park to Racing to Freestyle and Flatland????? Not a one of these is right....the whole chart is a nice piece of wallpaper art but that's it.

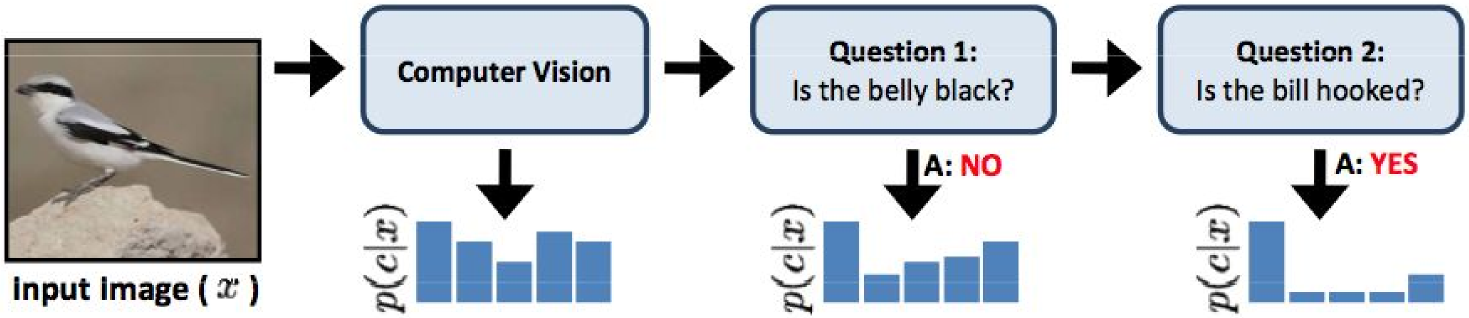
Keith D - 09/04/13 - 6:46pm
It's pretty much rubbish.


Brass
Beetle
Copper
Mallets
Dead Blow
Wooden
Lead
Plastic
Split Head
Pipe
Adjustable
Monkey
Self-Adjusting
Box-End
Strap
WRENCHES
Open-End
Saltus
Flare-Nut
Ratcheting Box
Striking Face Box
Roofing Hatchet
Flooring Hatchet
Wallboard
Hatchet
Hand Axes & Hatchets
Belt Axe
Claw Hatchet
Double Belt
Double Bit
Straight Head
Framing
Bevel Edge
Corner
Mortise
Butt
V-groove
Paring
Chisels
CARVING
Dovetail
Flooring
Veiner
Straight
Slick


THE CHARTED COLLECTION OF CONTEMPORARY FOOTWEAR
Gladiator
Loafer
Sneaker
High Heel
Sandal
Lace-Up
T-Strap
Mary Jane
Ankle Strap
Boot
Slingback
Low Heel
D'orsay
Wedge
Wedge Sandal
Cork Wedge Sandal
Wedge Peep Toe Pump
Wedge Pump
Knee-High Wedge Boot
Peep Toe Ankle Strap
Platform Wedge
Slingback Espadrille Wedge
Slingback Wedge
Slingback Peep Toe Wedge
Gladiator Heel Sandal
T-Strap Platform Pump
T-Strap Platform Peep Toe Pump
Wedge Ankle Bootie
Knee-High Heel Boot


(A) Easy for Humans
(B) Hard for Humans
(C) Easy for Humans
Chair? Airplane? ...
Finch? Bunting?...
Yellow Belly? Blue Belly? ...

Visual 20 Questions



Algorithm 1 Visual 20 Questions Game

$U^0 \leftarrow \emptyset$
for $t = 1$ to 20 **do**
  $j(t) = \max_k \mathrm{I}(c; u_k | x, U^{t-1})$
  Ask user question $u_{j(t)}$, and $U^t \leftarrow U^{t-1} \cup u_{j(t)}$.
end for
Return class $c^* = \max_c p(c|x, U^t)$

antedeepluvian

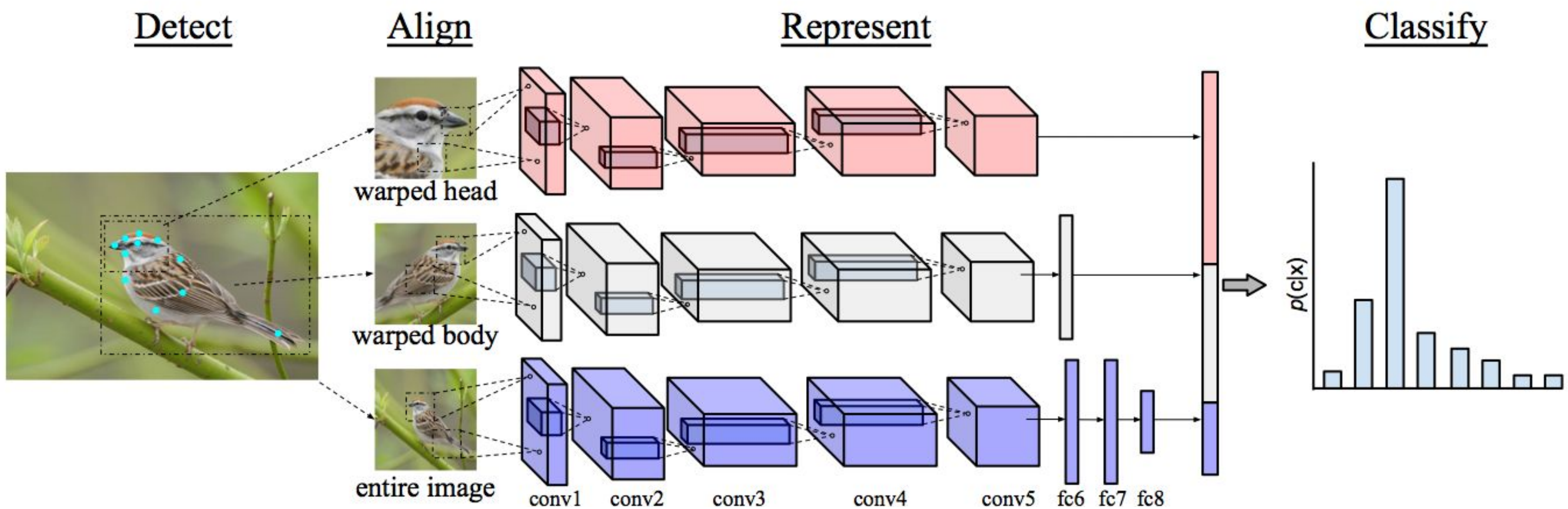
an·te·deep·lu·vi·an

ˌan(t)ēdēpˈlo͞ovēən/

adjective

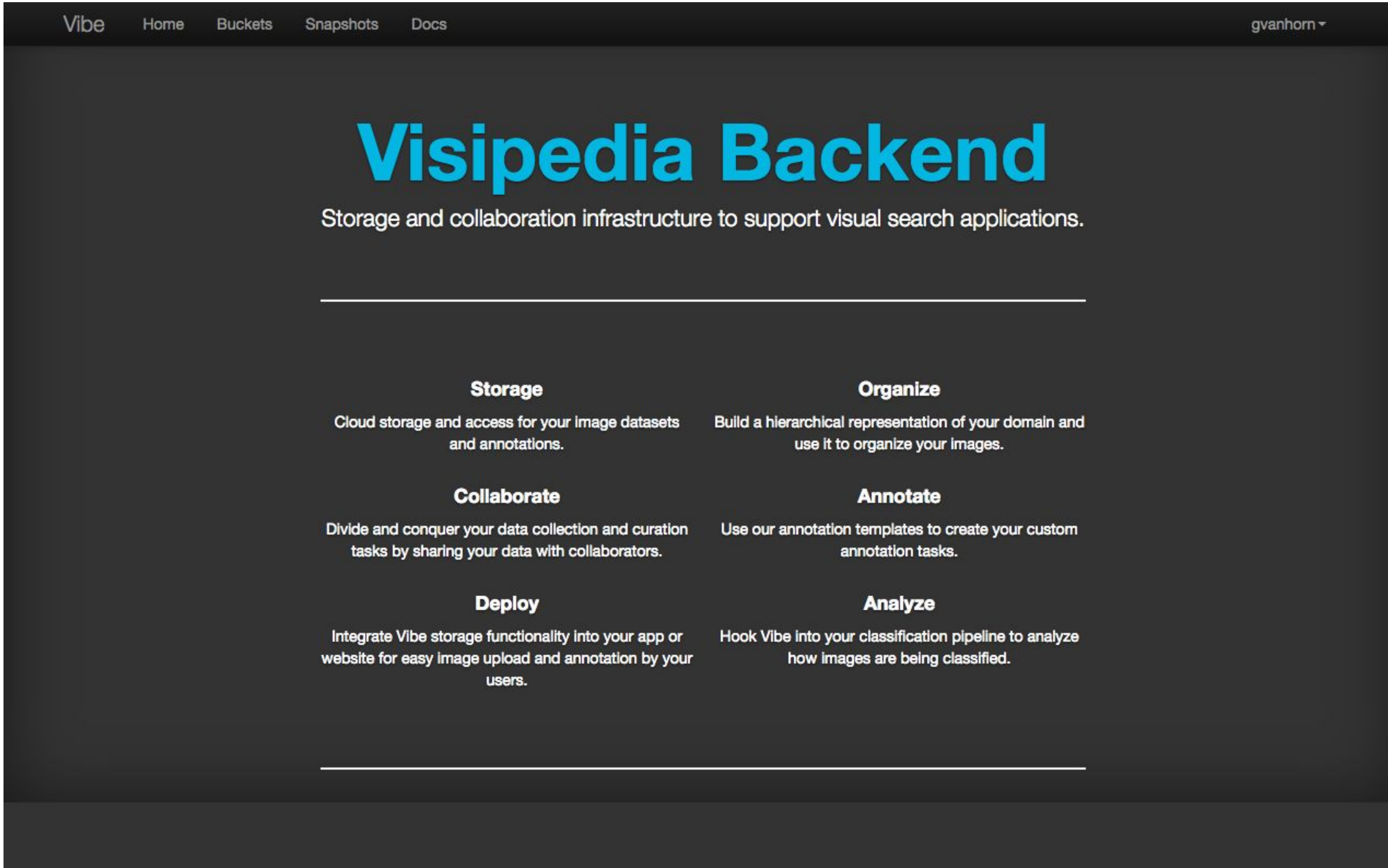
1. before the flood of deep learning papers
2. “Histograms of vector quantized filter responses are *antedeepluvian* features.”

Pose Normalized Deep ConvNets

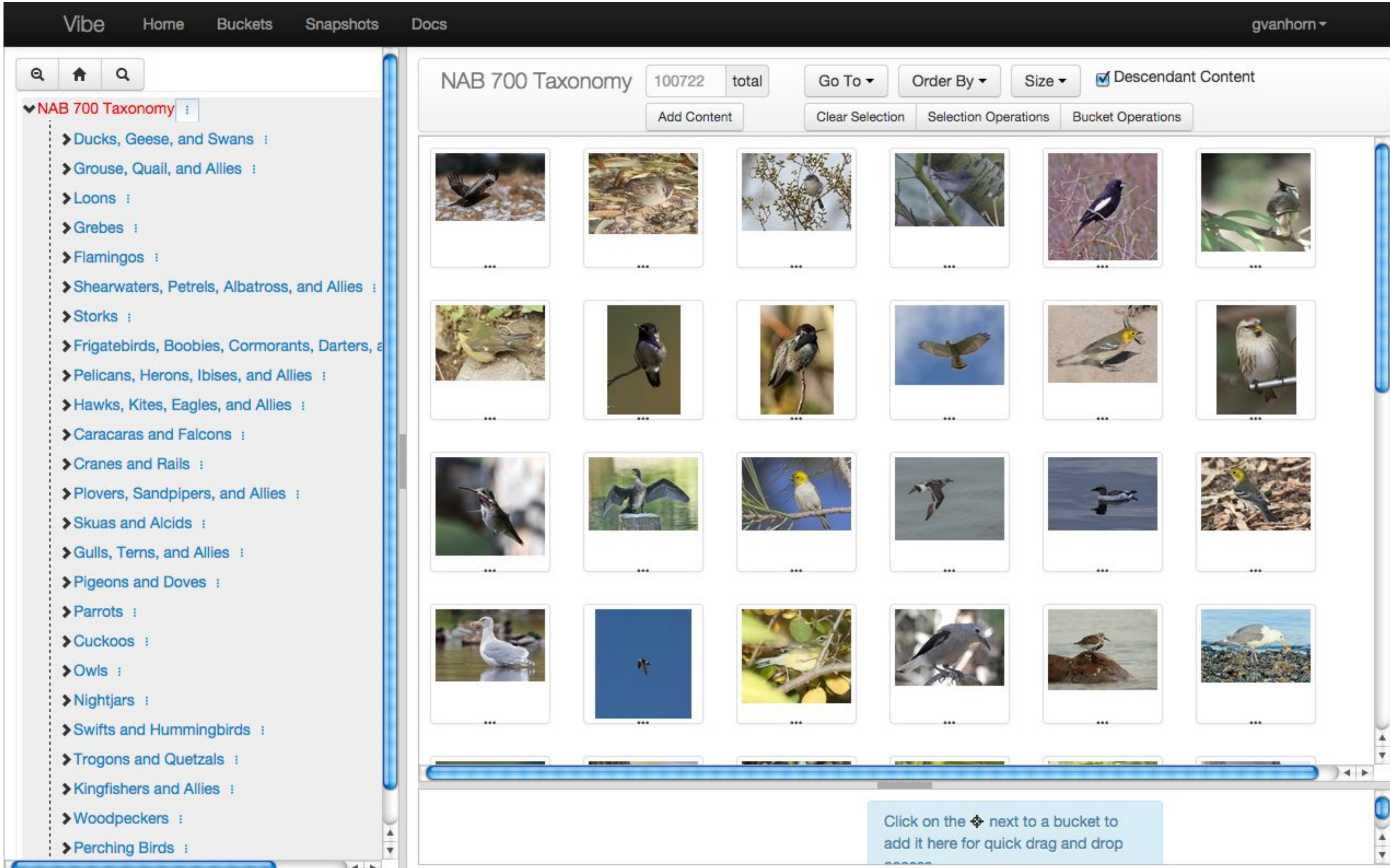


[Van Horn, Branson, Perona, Belongie BMVC 2014]


CCUB NABirds
700
Try out on a new dataset for fine-grained recognition, featuring 550 of North America's most common birds. The full dataset will be available in the fall. Join the competition today and download the "taster" dataset!
http://birds.cornell.edu/nabirds
CCUB NABirds includes:
• More than 700 visual categories, organized taxonomically
• Photos curated in collaboration with domain experts
• Data organized in a researcher-friendly, widely-used PASCAL VOC format
For more information contact: Ryan Farrell (farrell@eecs.berkeley.edu)
TheCornellLab of Ornithology



amazon web services™


Vibe
Home
Buckets
Snapshots
Docs
gvanhorn
NAB 700 Taxonomy
Ducks, Geese, and Swans
Grouse, Quail, and Allies
Loons
Grebes
Flamingos
Shearwaters, Petrels, Albatross, and Allies
Storks
Frigatebirds, Boobies, Cormorants, Darters, a
Pelicans, Herons, Ibises, and Allies
Hawks, Kites, Eagles, and Allies
Caracaras and Falcons
Cranes and Rails
Plovers, Sandpipers, and Allies
Skuas and Alcids
Gulls, Terns, and Allies
Pigeons and Doves
Parrots
Cuckoos
Owls
Nightjars
Swifts and Hummingbirds
Trogons and Quetzals
Kingfishers and Allies
Woodpeckers
Perching Birds
NAB 700 Taxonomy
100722
total
Go To
Order By
Size
Descendant Content
Add Content
Clear Selection
Selection Operations
Bucket Operations
Click on the ✥ next to a bucket to
add it here for quick drag and drop


Vibe
Home
Buckets
Snapshots
Docs
gvanhorn
Close
Tree View
Leaf Statistics
Total Nodes: 1674
Total Leaves: 929
Threshold
Sort Order:
Original Order
Reverse
0
50
100
150
200
250
300
350
400
450
500
550
600
Black-bellied Whistling-Duck 169
Fulvous Whistling-Duck 44
Greater White-fronted Goose 162
Snow Goose (White morph) 177
Snow Goose (Blue morph) 118
Ross's Goose 155
Brant 149
Cackling Goose 166
Canada Goose 153
Mute Swan 205
Trumpeter Swan 171
Tundra Swan 190
Muscovy Duck 91
Wood Duck (Breeding male) 261
Wood Duck (Female/Eclipse male) 186
Gadwall (Breeding male) 226
Gadwall (Female/Eclipse male) 109
American Wigeon (Breeding male) 190
American Wigeon (Female/Eclipse male) 109
Eurasian Wigeon (Breeding male) 154
Eurasian Wigeon (Female/Eclipse male) 24
American Black Duck 121
Mallard (Breeding male) 242
Mallard (Female/Eclipse male) 278
Mottled Duck 139
Blue-winged Teal (Male) 187
Blue-winged Teal (Female/juvenile) 121
Cinnamon Teal (Male) 211
Cinnamon Teal (Female/juvenile) 108
Northern Shoveler (Breeding male) 250
Northern Shoveler (Female/Eclipse male) 150

Merlin About Identify Contact

Have a bird photo?

Help us test Merlin Bird *Photo* ID

Start Bird Photo ID >

The Cornell Lab of Ornithology and Visipedia are collaborating to develop computer vision technology to identify birds in photos. Help test our new tool!

Note: Merlin Bird *Photo* ID does not work on tablets or mobile devices at this time. For best results, please use a computer with a recent version of Chrome or Safari.

http://merlin.allaboutbirds.org/photo-id

Merlin About Identify Contact

Select your photo.

Browse

Or drag and drop your file into this box.

1. Upload one image at a time.
2. Upload a jpeg or png image.
3. Image must be less than 10 MB.

Merlin About Identify Contact

Crop the bird by clicking and dragging a box.

Next

Merlin About Identify Contact

Click on the bill tip.

Next

Merlin About Identify Contact

Click on the eye. If both eyes are visible, click on the side of the head that is more visible.

Next

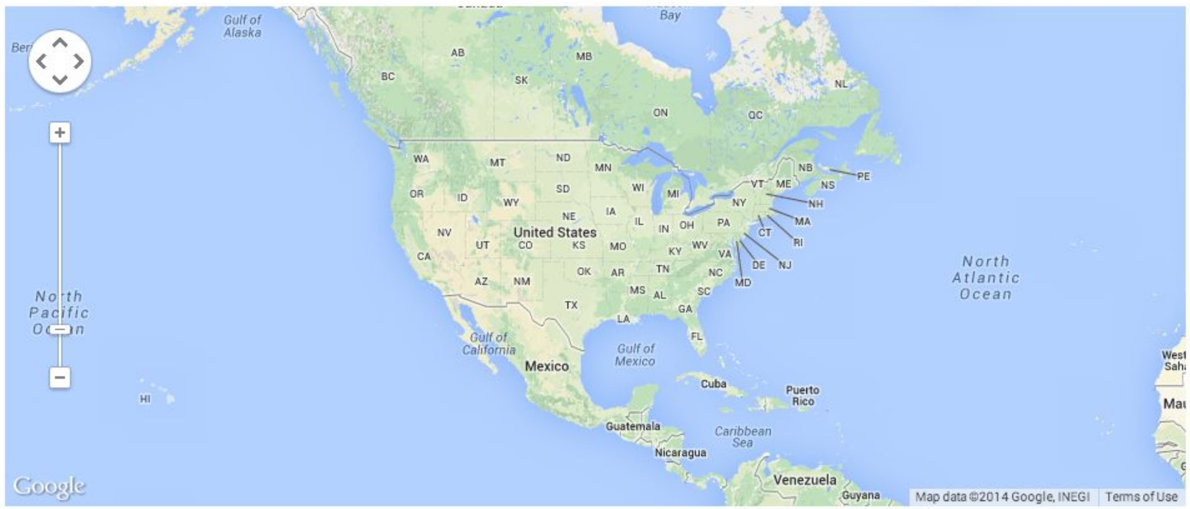
Merlin About Identify Contact

Where did you see the bird?

1. Search or click on the map to place a marker where the photo was taken.
2. Then, click Next to confirm the location.

Next

Search by City and State or Zip Code



When did you see the bird?

August 6

Next

Merlin About Identify Contact

Creating list of possible birds...

Merlin About Identify Contact

Dryden, NY, USA

Aug 6

Black-capped Chickadee

Christopher L. Wood

Christopher L. Wood

Christopher L. Wood

This Is My Bird

165,000 images uploaded
(since June 2015)

Can't find your bird?

Merlin considers the 400 most familiar species in North America. The bird you're seeing might not be in our database yet. Browse species.

ID Another Bird

http://merlin.allaboutbirds.org/photo-id

Part II
Learning about Similarity from Human and Machine Expertise

Our goals:

We want to **pull out** humans' intuitive notion of **perceptual similarity!**

 $\approx$  $\not\approx$ 

How can we combine **machine** and **human** expertise?
How can we **efficiently ask humans** about their knowledge?

[Wilber et al. ICCV 2015]

We can't directly measure similarity

"These two foods taste similar."

Disagree 1 2 3 4 5 Agree

We can't directly measure similarity

"These two foods taste similar."

Disagree 1 2 3 4 5 Agree

We can't directly measure similarity

“These two foods taste similar.”

Disagree 1 2 3 4 5 Agree

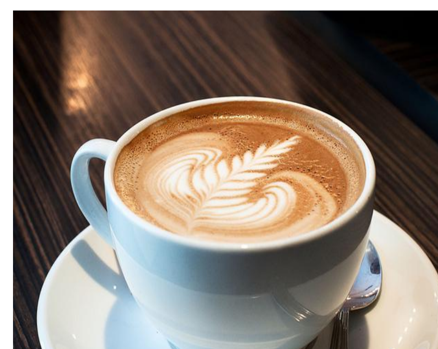
How could we learn how humans **perceive taste**?

"Pepperoni pizza **tastes similar** to other pizza"

≈

"Pepperoni pizza **tastes different** from coffee"

≉



Our design: **Grid questions**

To collect constraints efficiently, we ask:

“Please select the **four foods** that taste similar to the one on the left.”

Our design: **Grid questions**

To collect constraints efficiently, we ask:

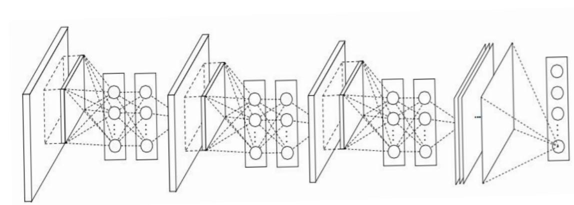
"Please select the **four foods** that taste similar to the one on the left."

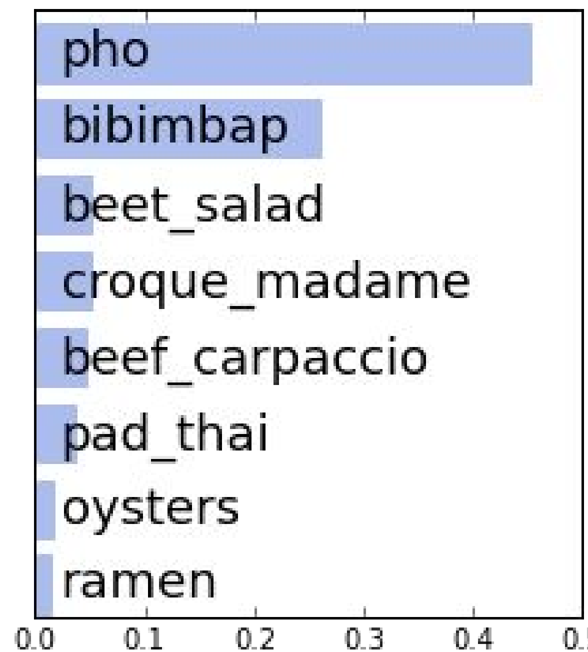
Our design: **Grid questions**

Selected points are more similar to the reference food than unselected points.

Machine similarity: Recognition is not enough.

Current **deep learning** approaches are great at recognizing objects, but they can't understand food taste without help.





Machine similarity: Recognition is not enough

Both of these are **pepperoni pizzas**, but they taste very different!

Machine similarity: Recognition is not enough

Is this **guacamole** or **wasabe**? This is not apparent from visual appearance.

Machine similarity: Taxonomies may be imprecise

This is a **wasabi Kit-Kat bar**. Where does it fit into the food taxonomy?

A perfect model can’t help us.

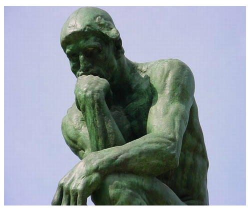
Machine similarity: Taxonomies may be imprecise

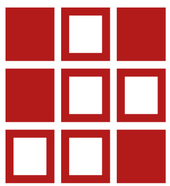
This is a **Berliner**. Berliner hobbyists will refuse to call this a "Jelly Doughnut", because it does not have a hole.

Stochastic Neighbor and Crowd Kernel (SNaCK)

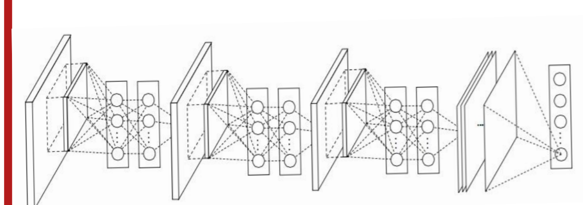
Our SNaCK embedding algorithm combines the expert's **triplet constraints** with a **visual similarity kernel**.

Human Expertise





Deep Learning



Our Algorithm

Concept Embedding

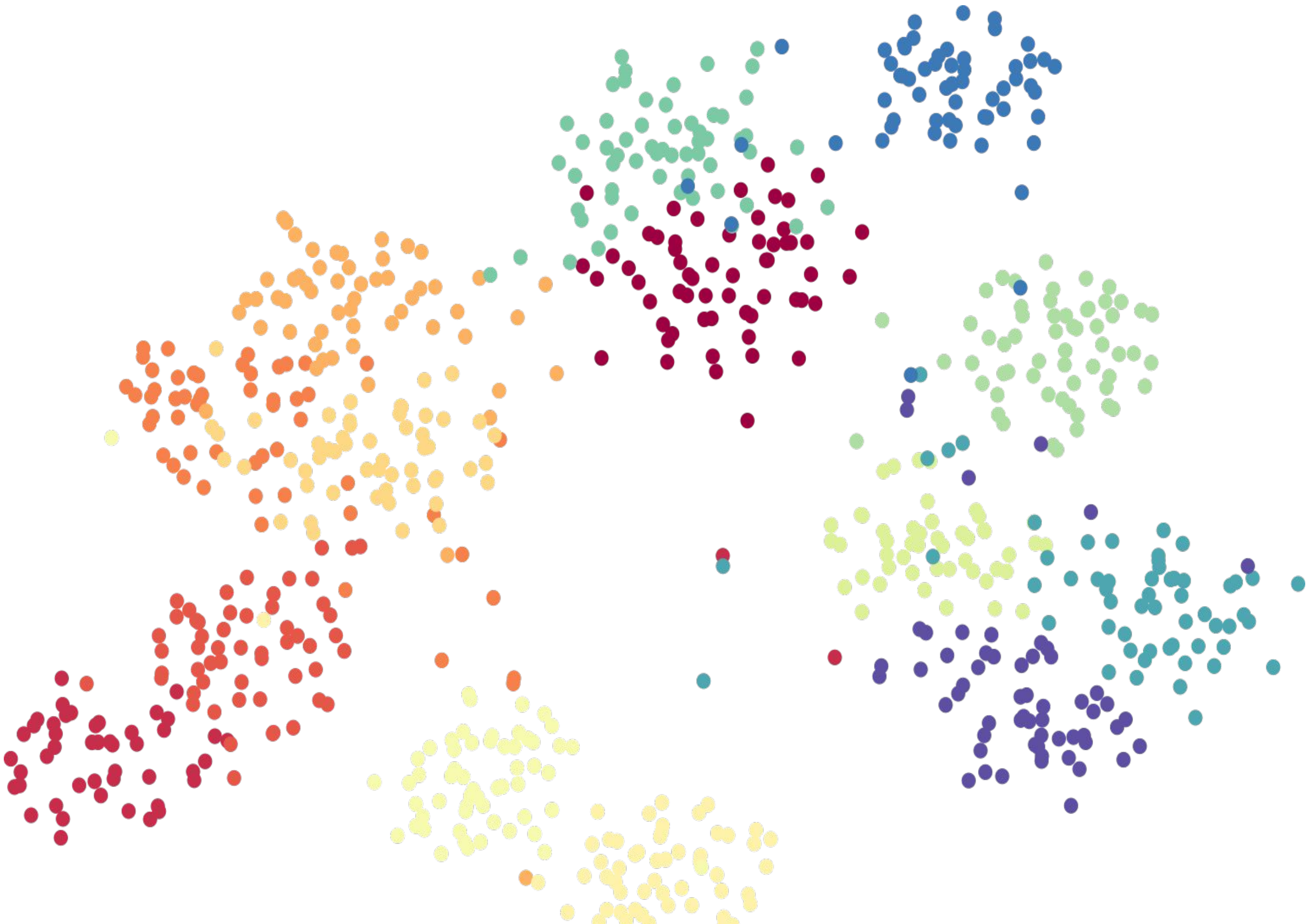
Embedding example: Food-10k

Explore: http://tinyurl.com/snack-embeddings

Embedding example: Food-10k

Explore: http://tinyurl.com/snack-embeddings

Embedding example: CU Birds 200 “Birdlets” subset

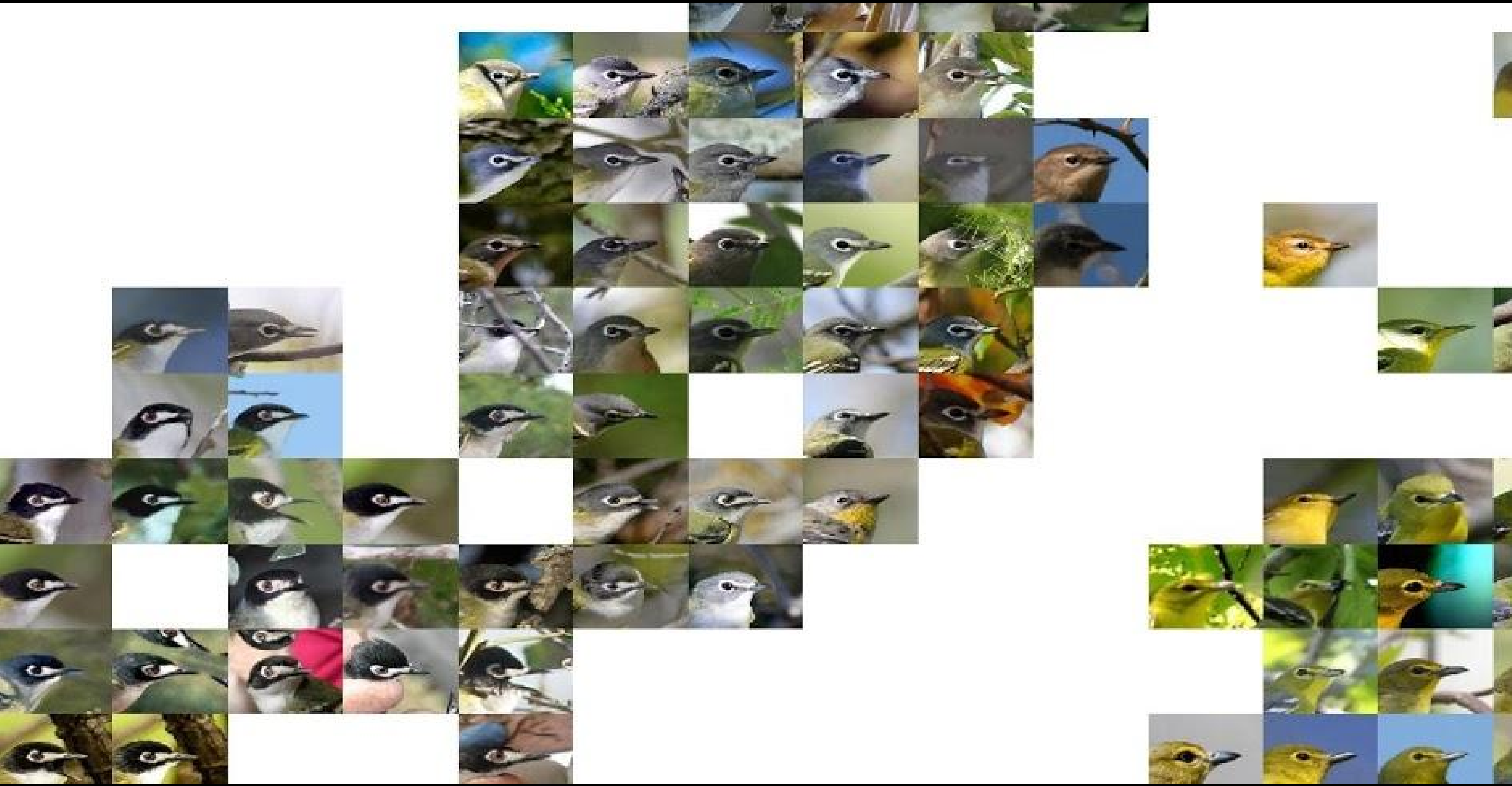


Explore: http://tinyurl.com/snack-embeddings

Embedding example: CU Birds 200 “Birdlets” subset

Explore: http://tinyurl.com/snack-embeddings

Embedding example: CU Birds 200 “Birdlets” subset



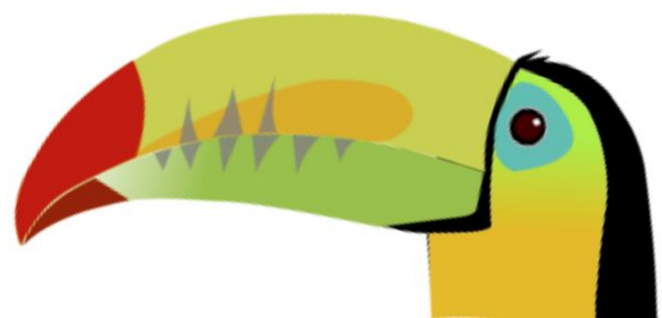
Explore: http://tinyurl.com/snack-embeddings

What’s Next for Visipedia?

- More taxa: flowers, trees, terrestrial mammals, amphibians, seashells, mushrooms, bugs, ...
- Partnership with Macaulay Digital Archive
- Visipedia.org Schema & Data Sharing Best Practices
- Suite of digital field guide apps
- CV/ML research for the most challenging taxa
- FGVC workshops: datasets and competitions

Thank You

- Cornell: Michael Wilber, Jessie Barry, Scott Haber
- UCSD: Iljung Sam Kwak, David Kriegman
- Caltech: Steve Branson, Grant Van Horn, Pietro Perona
- BYU: Ryan Farrell
- Google Focused Research Award
- Jacobs Technion-Cornell Institute



RESEARCH

Security and Privacy (Shmatikov, Juels, Pass, Ristenpart)

Computer Vision (Belongie, Snavely, Zelnik, Zabih)

Social/Mobile Computing (Estrin, Naaman, Dell)

tech.cornell.edu/programs

PhD (CS, ECE, InfoSci…)

Postdocs

Runway Postdocs

Masters: Connective Media
Health Tech
MBA (1 year)
CS MEng (1 year)
LLM (1 year)
ORIE (1 year)