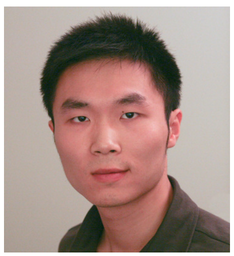
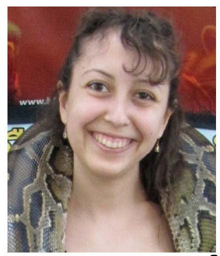
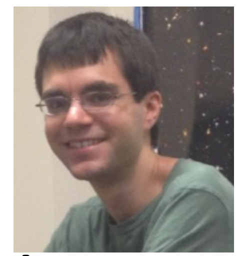
MICHIGAN

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

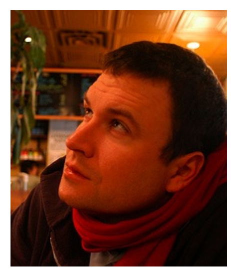
STANFORD HCI GROUP

Scalable Multi-Label Annotation

Jia Deng   Olga Russakovsky   Jonathan Krause,

Michael Bernstein   Alexander Berg   Li Fei-Fei


Multi-label annotation

Data Item

Labels

Table	Chair	Horse	Dog	Cat	Bird
+	+	-	-	-	-

Task: Crowdsource object labels for images.

Application: Benchmarking, training, modeling

Generalization:

- musical attributes of songs
- actions in movies
- sentiments in documents

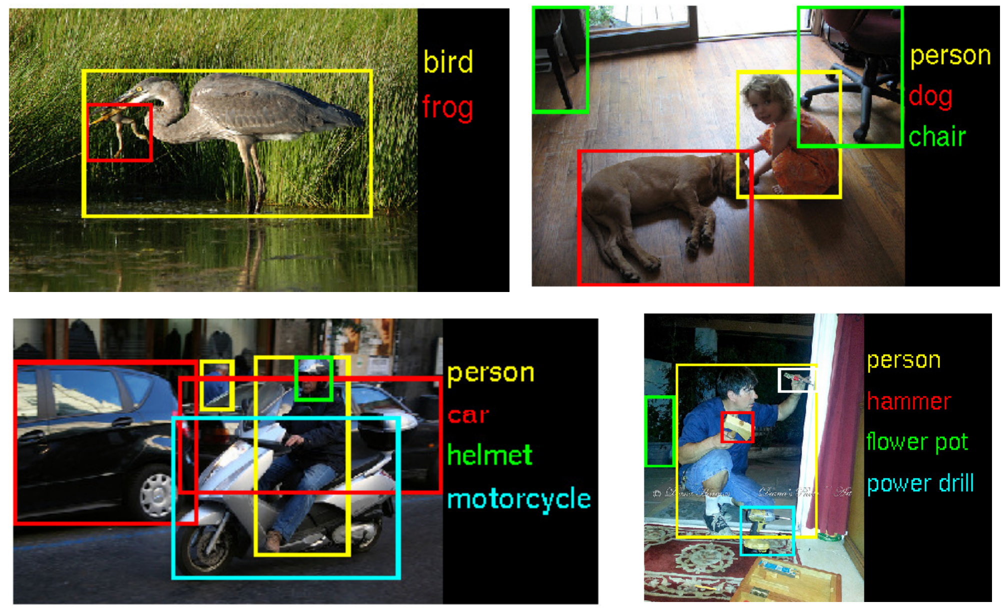
Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

IMAGENET

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Large-Scale Visual Recognition Challenge ILSVRC 2010-2014

Current focus:
200 Category Detection
(~100,000 fully labeled images)



Decaf / Caffe
a Berkeley Vision Project

UNIVERSITY OF OXFORD

The New York Times

NEC Laboratories America
Relentless passion for innovation

NUS National University of Singapore

NYU

facebook

Google

DNNresearch

University of Amsterdam

EUVISION technologies

Clarifai

Microsoft Research

IBM Research

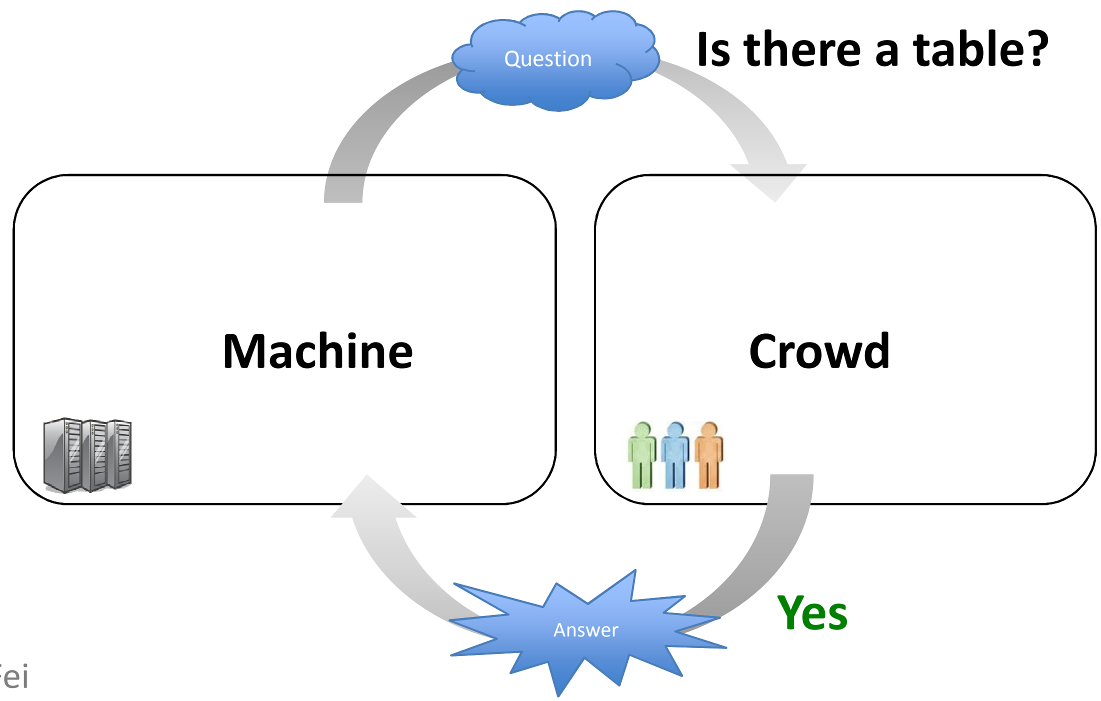
Andrew Howard Consulting

CloudCV: Object Detection

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

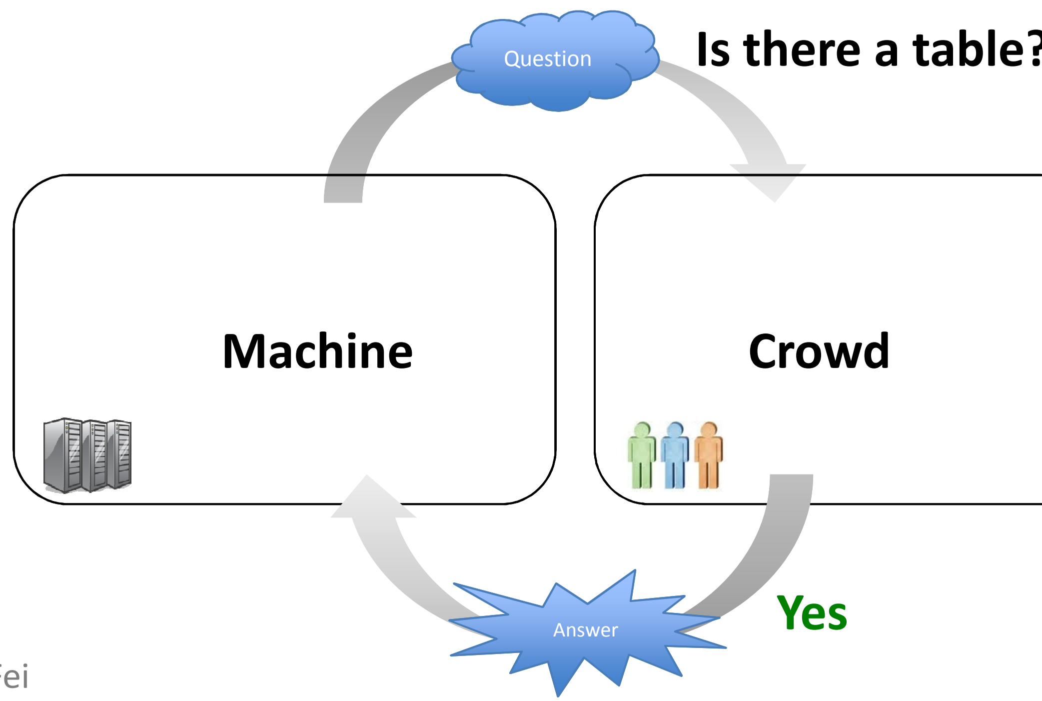
Table	Chair	Horse	Dog	Cat	Bird
?	?	?	?	?	?



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

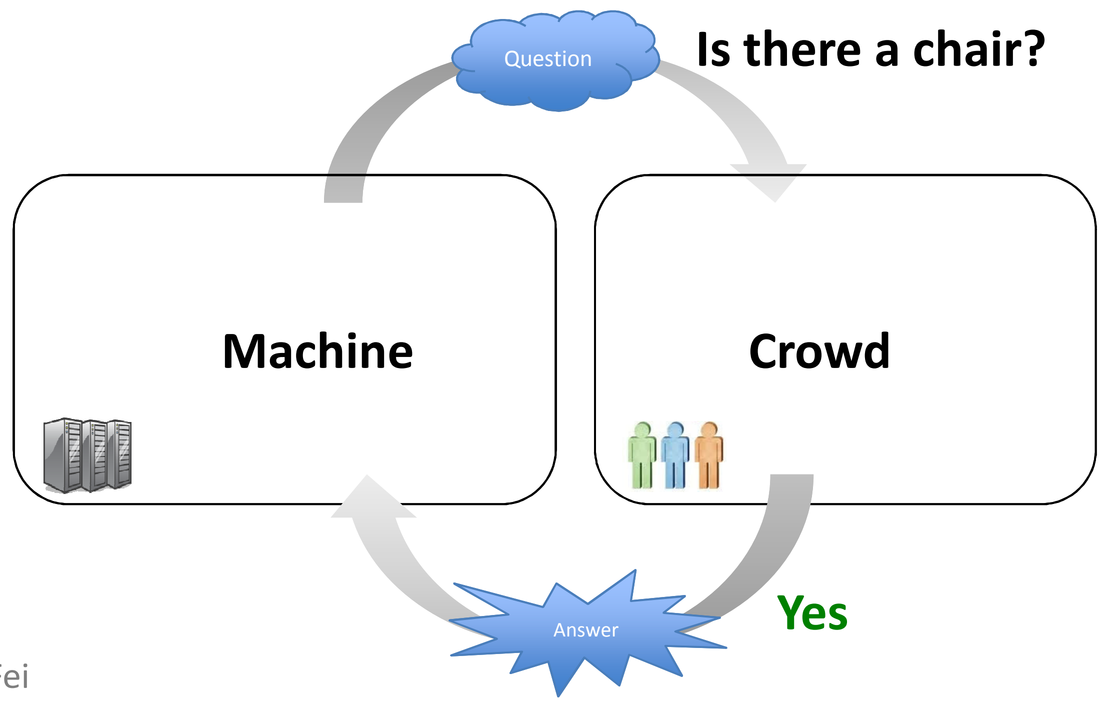
Table	Chair	Horse	Dog	Cat	Bird
+	?	?	?	?	?



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

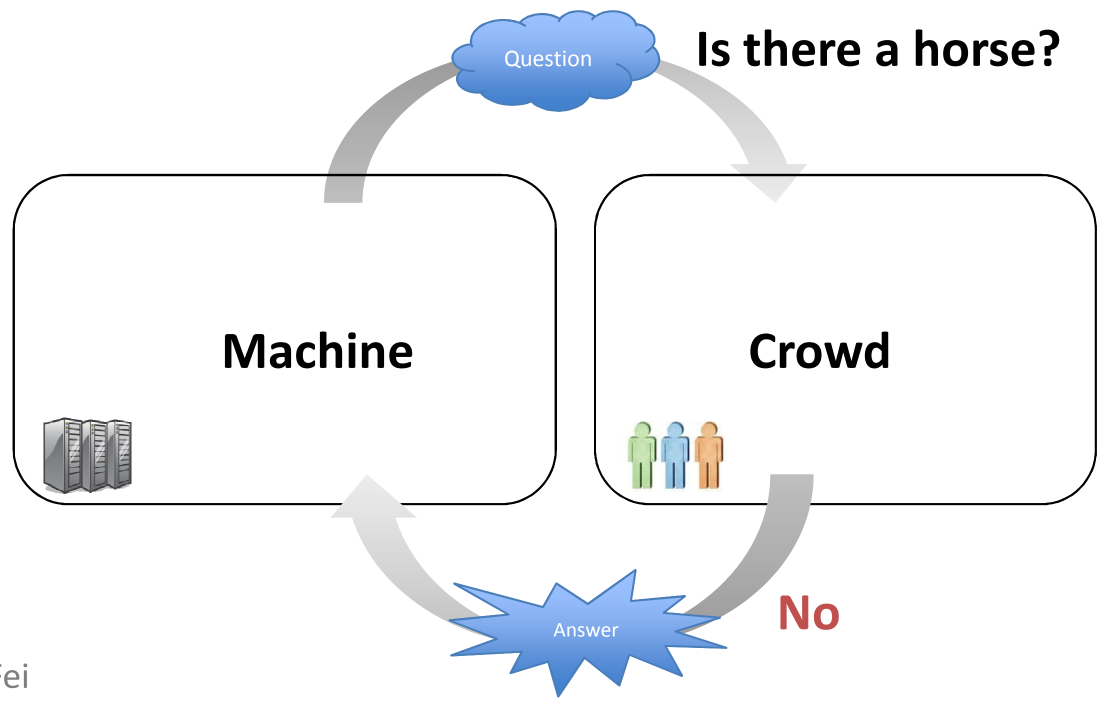
Table	Chair	Horse	Dog	Cat	Bird
+	+	?	?	?	?



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

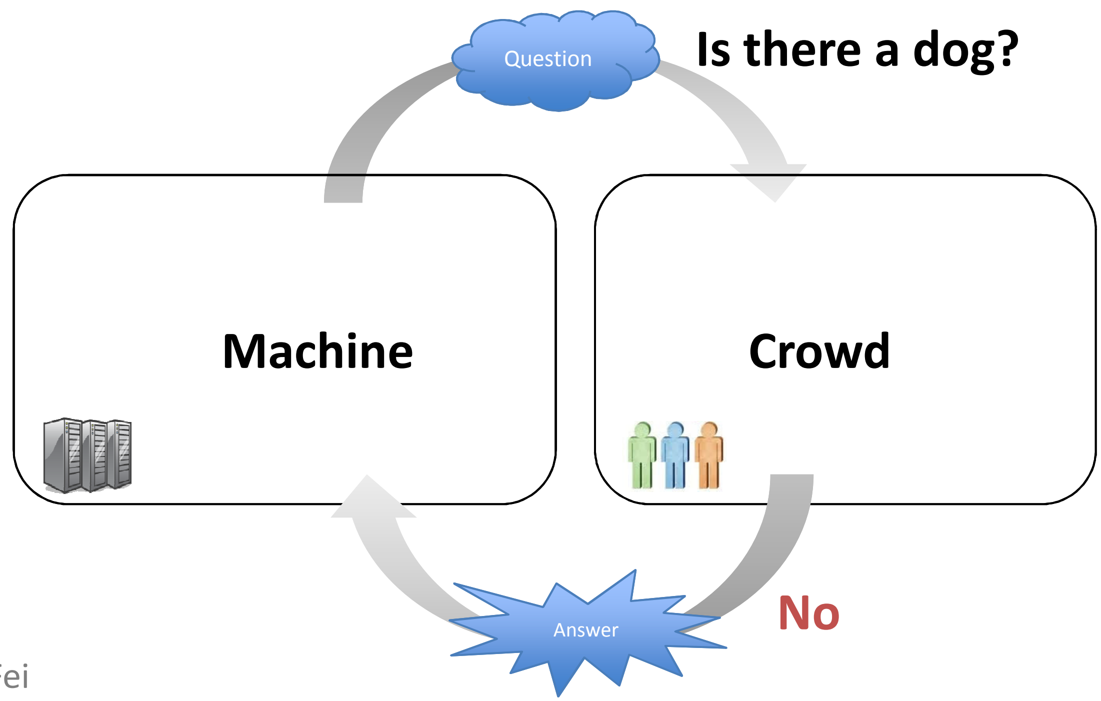
Table	Chair	Horse	Dog	Cat	Bird
+	+	-	?	?	?



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

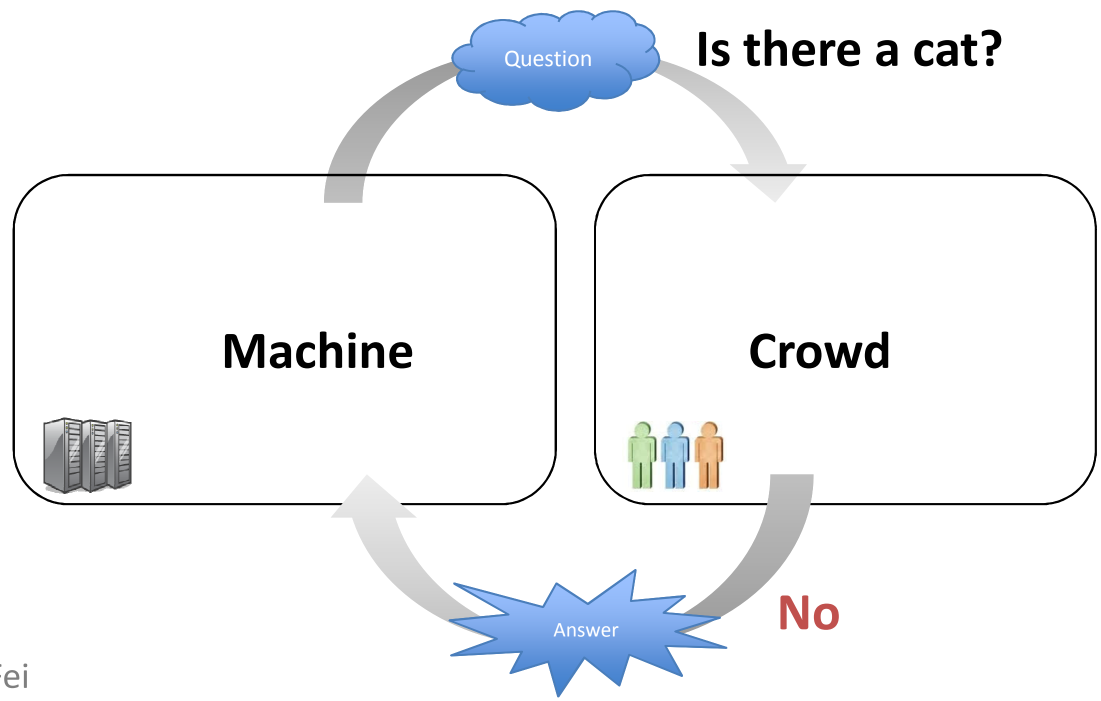
Table	Chair	Horse	Dog	Cat	Bird
+	+	-	-	?	?



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

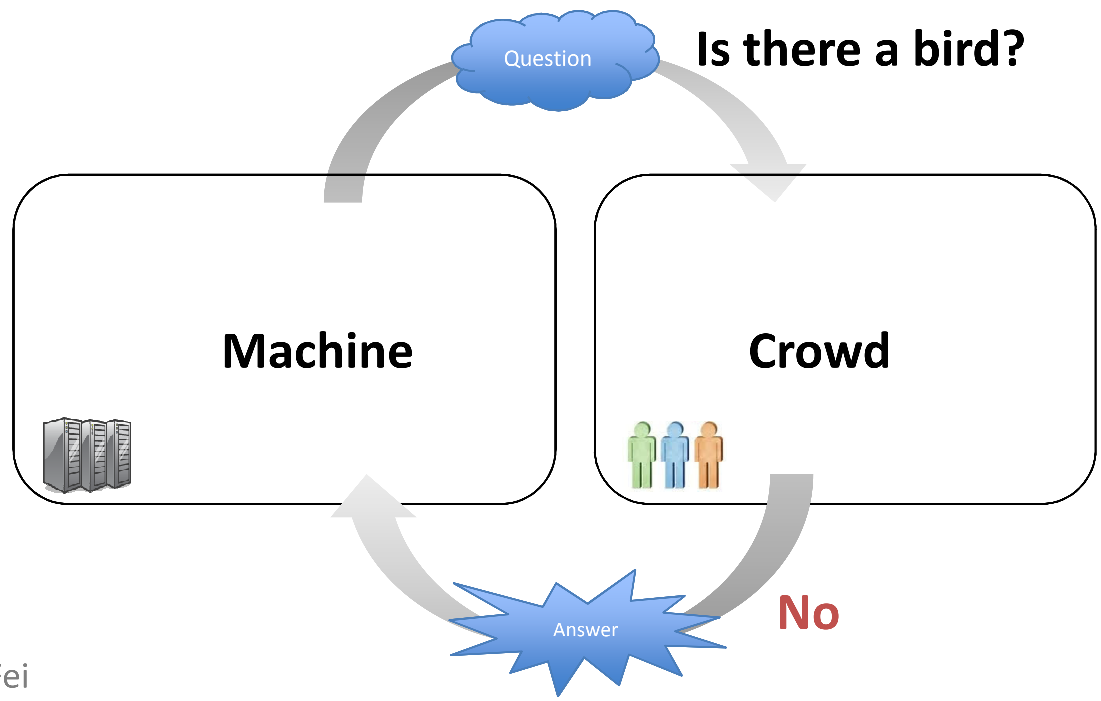
Table	Chair	Horse	Dog	Cat	Bird
+	+	-	-	-	?



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

Table	Chair	Horse	Dog	Cat	Bird
+	+	-	-	-	-



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Naïve approach: ask for each object

Cost: O(NK) for N images and K objects

	Table	Chair	Horse	Dog	Cat	Bird
	+	+	-	-	-	-
	+	-	-	-	+	-
	+	+	-	-	-	-

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Hierarchy

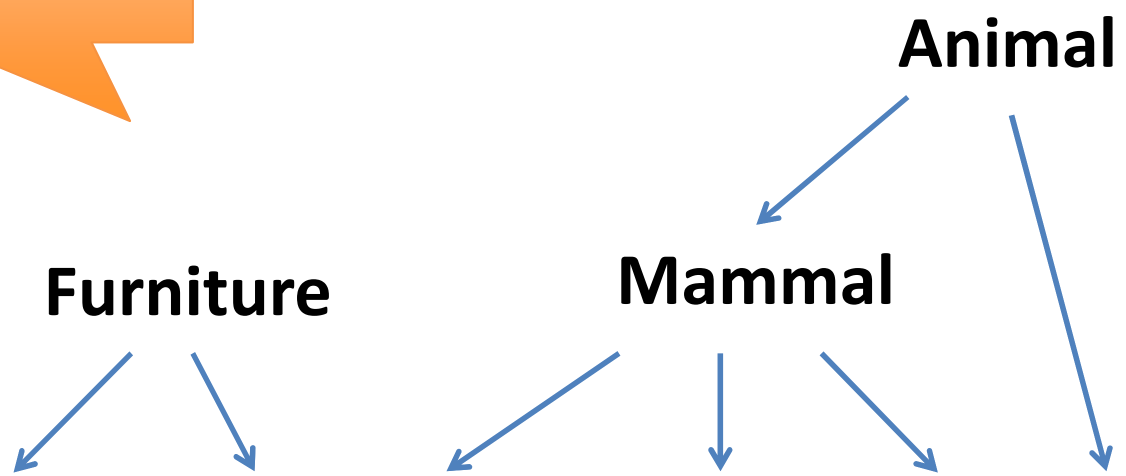


Table	Chair	Horse	Dog	Cat	Bird
+	+	-	-	-	-

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Hierarchy

Animal

Mammal

Furniture

	Table	Chair	Horse	Dog	Cat	Bird
	+	+	-	-	-	-
	+	-	-	-	+	-
	+	+	-	-	-	-

Correlation

Sparsity

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Better approach: exploit label structure

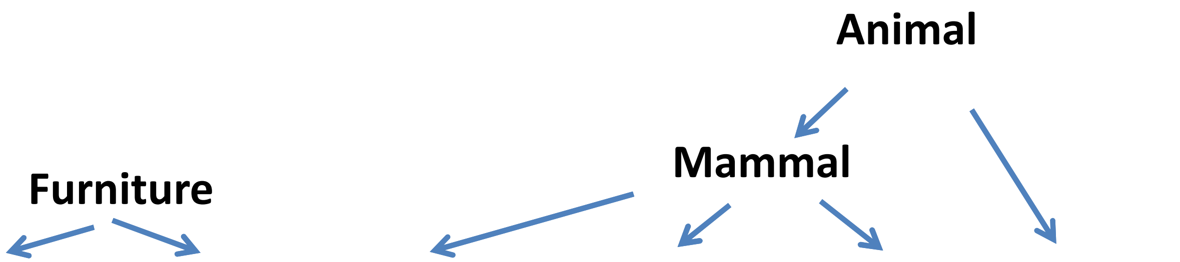
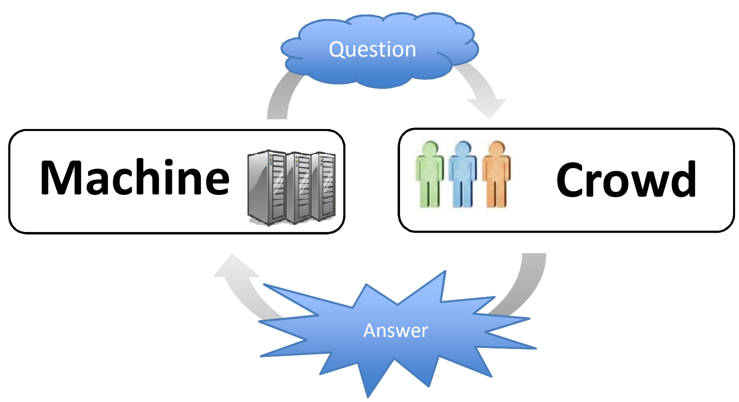


Table	Chair	Horse	Dog	Cat	Bird
?	?	?	?	?	?



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

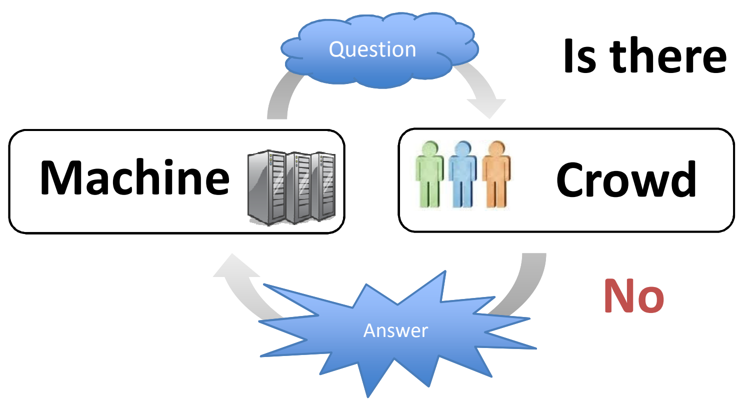
Better approach: exploit label structure

Animal

Mammal

Furniture

Table	Chair	Horse	Dog	Cat	Bird
?	?	?	?	?	?



Is there an animal?

No

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Better approach: exploit label structure

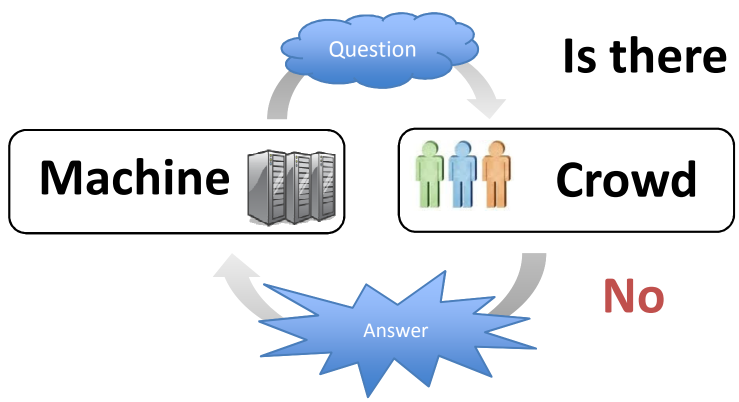
Animal

Mammal

Furniture

Table	Chair	Horse	Dog	Cat	Bird
?	?	-	-	-	-

Is there an animal?



No

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Better approach: exploit label structure

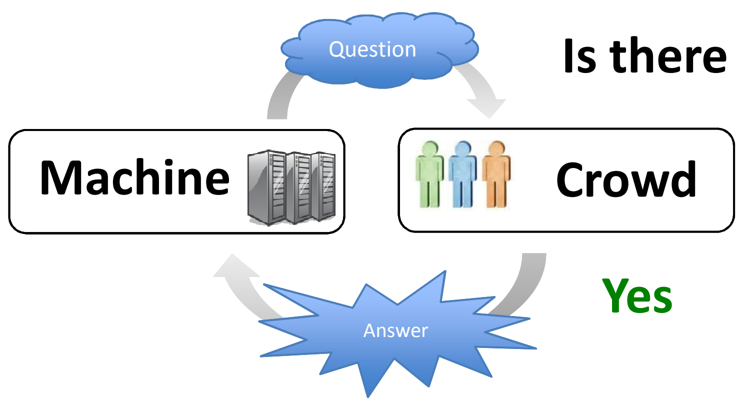
Animal

Furniture

Mammal

Table	Chair	Horse	Dog	Cat	Bird
?	?	-	-	-	-

Is there furniture?



Yes

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

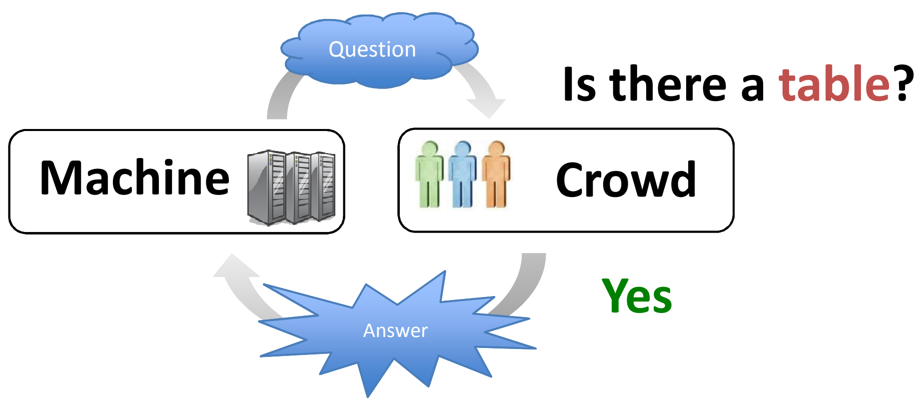
Better approach: exploit label structure

Animal

Furniture

Mammal

Table	Chair	Horse	Dog	Cat	Bird
?	?	-	-	-	-



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

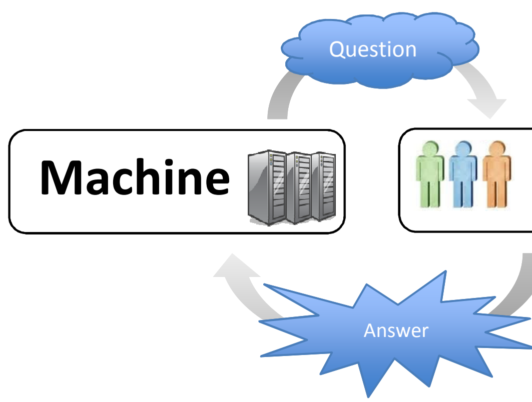
Better approach: exploit label structure

Animal

Furniture

Mammal

Table	Chair	Horse	Dog	Cat	Bird
+	?	-	-	-	-



Is there a chair?

Yes

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Better approach: exploit label structure

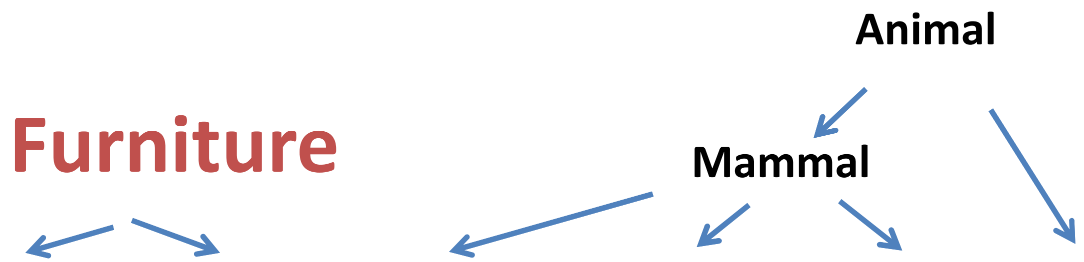
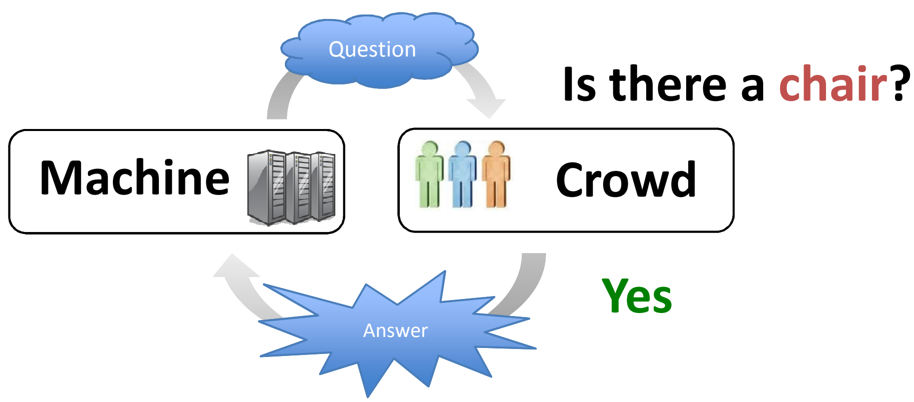


Table	Chair	Horse	Dog	Cat	Bird
+	+	-	-	-	-



Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Selecting the Right Question

Goal:

Get as much utility (new labels) as possible,
for as little cost (worker time) as possible,
given a desired level of accuracy

Accuracy constraint

- User-specified accuracy threshold, e.g., 95%
- Majority voting assuming uniform worker quality

  [GAL: Sheng, Provost, Ipeirotis KDD '08]
- Might require only one worker, might require several based on the task

Cost: worker time (time = money)

expected human time to get an answer with 95% accuracy

Question (is there …)	Cost (second)
a thing used to open cans/bottles	14.4
an item that runs on electricity (plugged in or using batteries)	12.6
a stringed instrument	3.4
a canine	2.0

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Utility: expected # of new labels

Table	Chair	Horse	Dog	Cat	Bird
?	?	?	?	?	?

Is there a table?

Yes

Table	Chair	Horse	Dog	Cat	Bird
+	?	?	?	?	?

No

Table	Chair	Horse	Dog	Cat	Bird
-	?	?	?	?	?

utility = 1

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Utility: expected # of new labels

Table	Chair	Horse	Dog	Cat	Bird
?	?	?	?	?	?

Is there a table?

Yes

Table	Chair	Horse	Dog	Cat	Bird
+	?	?	?	?	?

No

Table	Chair	Horse	Dog	Cat	Bird
-	?	?	?	?	?

utility = 1

Table	Chair	Horse	Dog	Cat	Bird
?	?	?	?	?	?

Is there an animal?

Pr(Y) = 0.5

Table	Chair	Horse	Dog	Cat	Bird
?	?	?	?	?	?

Pr(N) = 0.5

Table	Chair	Horse	Dog	Cat	Bird
?	?	-	-	-	-

utility = 0.5 * 0 + 0.5 * 4 = 2

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Selecting the Right Question

Pick the question with the most labels per second

Query: Is there a...	Utility (num labels)	Cost (worker time in secs)	Utility-Cost Ratio (labels per sec)
mammal with claws or fingers	12.0	3.0	4.0
living organism	24.8	7.9	3.1
mammal	17.6	7.4	2.4
creature without legs	5.9	2.6	2.3
land or avian creature	20.8	9.5	2.2

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei

Results

- Dataset: 20K images from ImageNet Challenge 2013.
- Labels: 200 basic categories (dog, cat, table…), 64 internal nodes in hierarchy

Results

- Dataset: 20K images from ImageNet Challenge 2013.
- Labels: 200 basic categories (dog, cat, table…), 64 internal nodes in hierarchy
- Setup:
  - 50-50 training test split
  - Estimate parameters on training, simulate on test
  - Future work: online estimation

Results: accuracy

Annotating 10K images with 200 objects

Accuracy Threshold per question (parameter)	Accuracy (F1 score) Naïve approach	Accuracy (F1 score) Our approach
0.95	99.64 (75.67)	99.75 (76.97)
0.90	99.29 (60.17)	99.62 (60.69)

Results: cost

Annotating 10K images with 200 objects

Accuracy **Threshold** per question (parameter)	**Cost saving** (our approach compared to naïve approach)
0.95	**3.93**x
0.90	**6.18**x

Results: cost

Annotating 10K images with 200 objects

Accuracy **Threshold** per question (parameter)	**Cost saving** (our approach compared to naïve approach)
0.95	**3.93**x
0.90	**6.18**x

6 times more labels per second

Conclusions

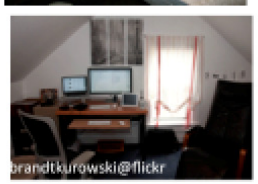
Speeds up crowdsourced multi-label annotation by exploiting the structure and distribution of labels.
Could be a bargain for you!

Animal

Mammal

Hierarchy

Furniture

	Table	Chair	Horse	Dog	Cat	Bird
	+	+	-	-	-	-
	+	-	-	-	+	-
	+	+	-	-	-	-

Sparsity

Correlation

Deng, Russakovsky, Krause, Bernstein, Berg, Fei-Fei