

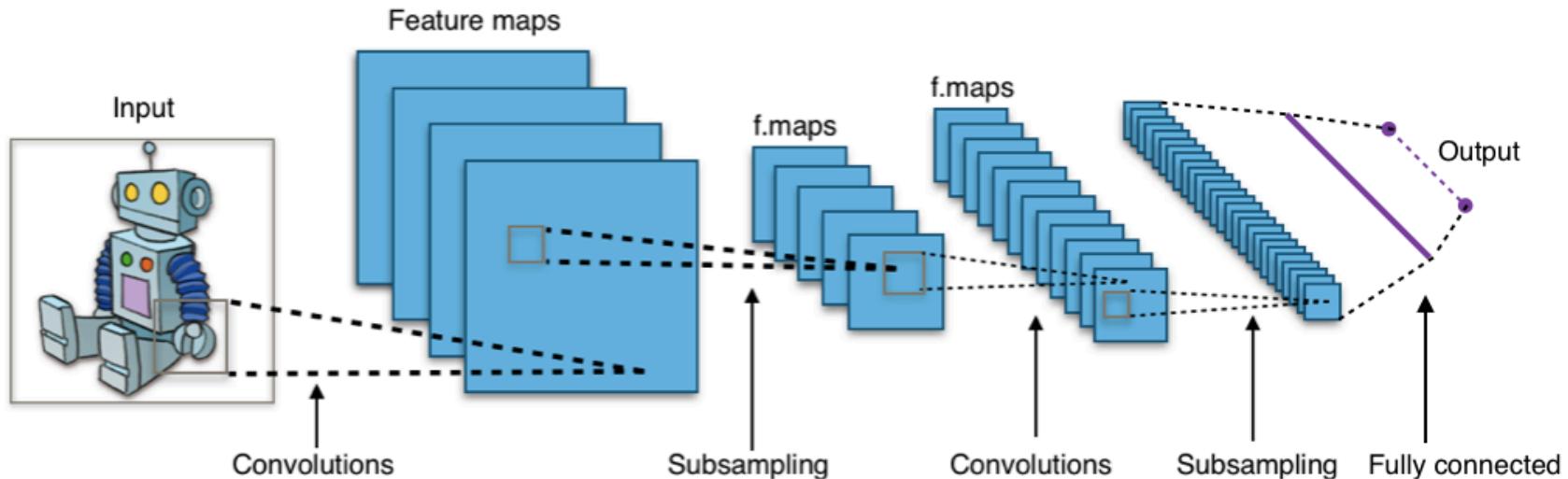
LEARNING WITH

LIMITED LABELED DATA

NIPS 2017

Today, training data is the  
biggest bottleneck in ML

# The Rise of Representation Learning



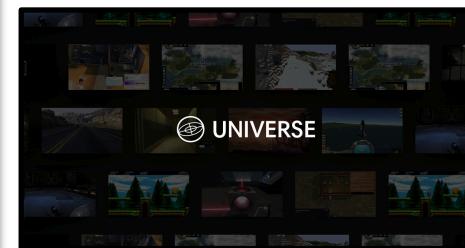
# Representation Learning is Data Hungry

- Those feature maps and transformation functions have lots of data-dependent parameters
- Good generalization requires *at least* tens of thousands of labeled training examples



# Now We Need a Lot of Training Data

Hard to move beyond a few benchmarks



# Now We Need a Lot of Training Data

Proprietary data,  
labeled at  
immense cost

W I R E D

GOOGLE'S HAND-FED AI NOW  
GIVES ANSWERS, NOT JUST  
SEARCH RESULTS

“Google trains these neural networks using data handcrafted by a massive team of PhD linguists...”

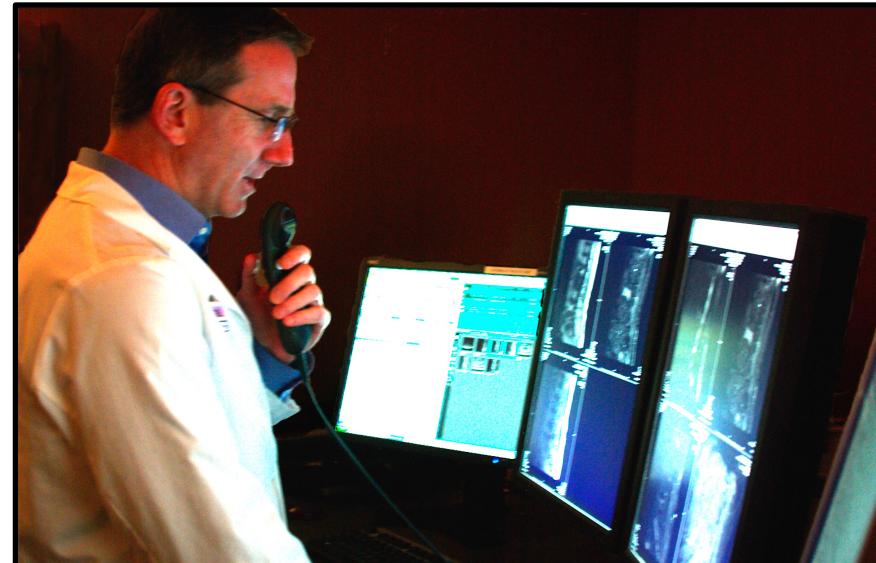
# Other Barriers to Curating Data

Expensive  
collection  
procedures



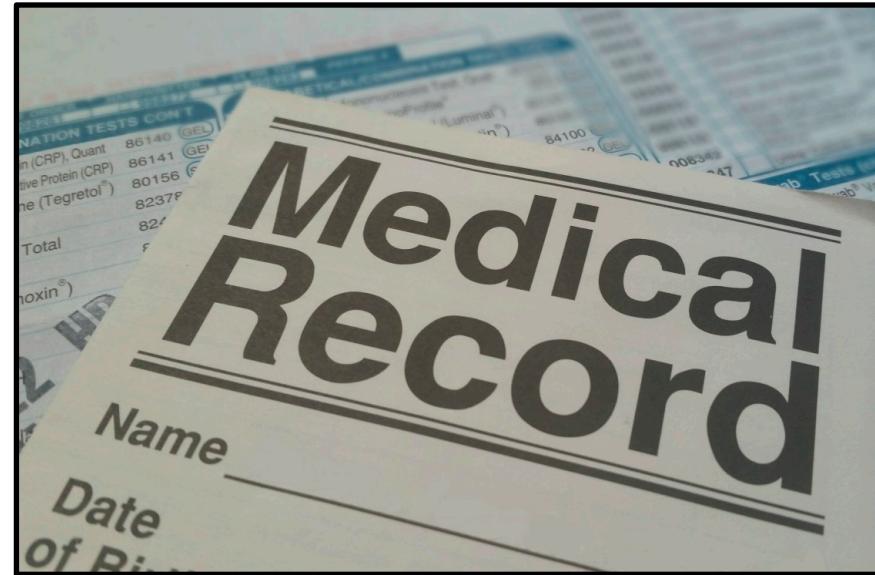
# Other Barriers to Curating Data

Need for  
domain  
expertise



# Other Barriers to Curating Data

Data  
regulation



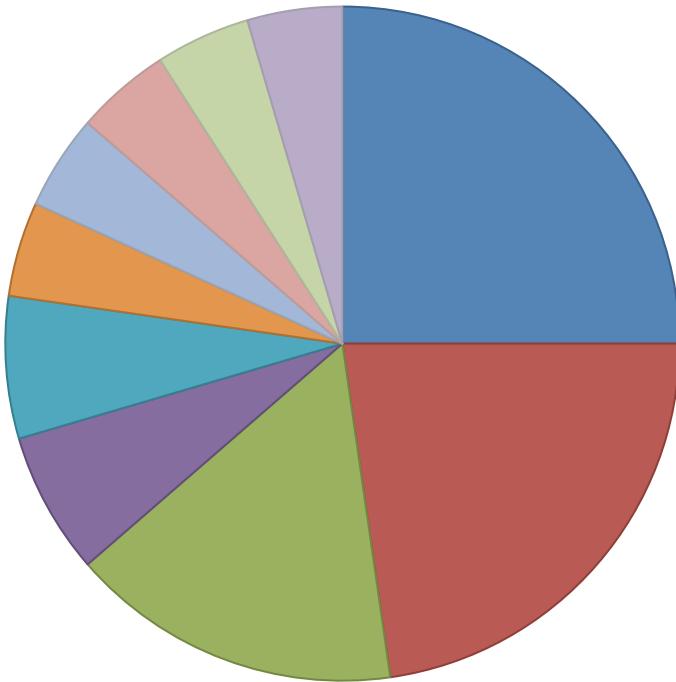
# How Will We Feed the Next Generation of Data-Hungry ML?

# Program Stats

- 65 submissions
- 56 reviewers
- 44 accepted papers

**Thank you!**

# Topics of Accepted Papers



- Semi-Supervised Learning
- Weak Supervision
- Transfer Learning
- Representation Learning
- Applications
- Multi-Task Learning
- Data Augmentation
- Active Learning
- Self-Training
- Knowledge Distillation

# Invited Speakers



**Gaël Varoquaux**

8:40 AM



**Nina Balcan**

3:30 PM



**Tom Mitchell**

9:10 AM



**Sameer Singh**

4:15 PM



**Andrew McCallum**

11:00 AM



**Ian Goodfellow**

4:45 PM



**Sebastian Riedel**

11:30 AM

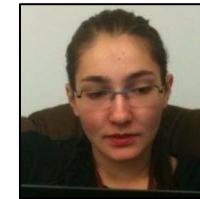


**Alan Ritter**

5:45 PM

# Other Program Highlights

- Panel on limited labeled data in medical imaging  
2:00 PM
- Award ceremony  
6:15 PM



# Information for Poster Presenters

- Two sets of one-minute spotlights: 9:55 AM, 2:30 PM
- Followed immediately by poster sessions
- Can still send a spotlight slide to [bach@cs.stanford.edu](mailto:bach@cs.stanford.edu)

More Information:

<http://lld-workshop.github.io>