UCF Center for Research in Computer Vision

# CAP6412
# Advanced Computer Vision

Mubarak Shah

shah@crcv.ucf.edu

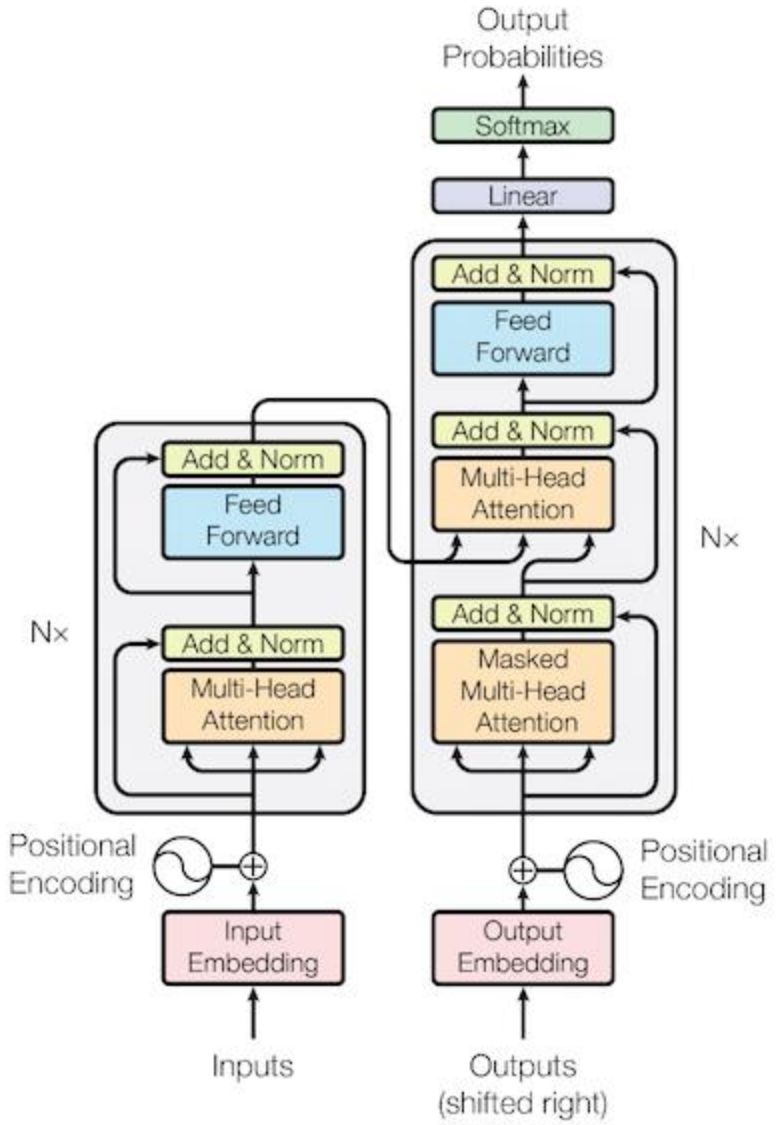
HEC-245

# CAP6412, Spring 2022

- **Instructor**: Dr. Mubarak Shah
- **Email:** shah@crcv.ucf.edu
- **Office**: HEC 245
- **Phone**: 4078235077
- **Co-Instructor**: Dr. Navid Kardan, kardan@knights.ucf.edu
- **Time**: Mondays and Wednesdays 3:00 to 4:15PM
- **Location:** Zoom
- **Office Hours:** Mondays 2:00 to 3:00PM; Wednesdays 4:15 to 5:00PM; Fridays 1:00 to 2:00; and by appointment
- **Pre-requisite**: CAP5415
- Course webpage: https://www.crcv.ucf.edu/courses/cap6412-spring-2022/

# Course Objective

- To expose graduate students to the cutting-edge research in Computer Vision.
  - We will discuss research papers employing transformers



# Students Learning Outcomes

- Read and understand a research paper.
- Write a comprehensive review of the paper.
- To identify strong and weak points of the paper.
- To generate own ideas to solve the same problem
- To work on research project and write a research paper

# Grading Policy

| | |
|---|---|
| Reports (you have to do only 50% of the papers) | 10% |
| Presentations (roughly two) | 20% |
| Lead Discussion Session (roughly two) | 5% |
| Attendance | 5% |
| Projects | 50% |

No Mid Term, Final, Homework

- **Late Policy**
  - 0 for late reports
  - Projects
    - 20% off per day
    - up to 4 days

# Project Timeline

- Three in-class presentations (see class schedule)
  - Project ideas / proposal [5%]
  - Update 1 [5%]
  - Update 2 [10%]
  - Final presentation/Paper [30%]

# Programming/ GPU Cluster

- Pytorch, TensorFlow
- Each student will get an account on UCF HPC Newton
- Watch video: https://www.youtube.com/watch?v=niQ5hvABvKg&list=PLd3hlSJsX_Ikm5il1HgmDB_z62BeoikFX&index=19

# Papers/ Reports

- We will discuss one paper in each class
- All students will read the assigned paper before the class and write a report
- One group of students will make presentation and another group will lead discussion session for each paper
- Presentation should not be more than 30 minutes
- You can select the paper you want to present from the list on course webpage
- Reports will be due just before the class meeting through Web Courses

- Schedule Table: https://www.crcv.ucf.edu/courses/cap6412-spring-2022/schedule/

# Reports (one page)

- Parts of Report
  - Summary
  - Good points
  - Weak points
  - Questions
  - Ideas

- Reports will not be graded
- We will follow the honor system, by submitting report you will pledge that you have read the paper, have written report yourself, have not copied word by word from the paper and from any other student or resource

# Statement of Academic Integrity

- The UCF Golden Rule (http://goldenrule.sdes.ucf.edu/ ) will be observed in the class. Plagiarism and Cheating of any kind on an examination, quiz, or assignment will result at least in an "F" for that assignment (and may, depending on the severity of the case, lead to an "F" for the entire course) and may be subject to appropriate referral to the Office of Student Conduct for further action. I will assume for this course that you will adhere to the academic creed of this University and will maintain the highest standards of academic integrity. In other words, don't cheat by giving answers to others or taking them from anyone else. I will also adhere to the highest standards of academic integrity, so please do not ask me to change (or expect me to change) your grade illegitimately or to bend or break rules for one person that will not apply to everyone.

# Paper presentation review/Rehearsal Schedule

- For Monday presentation
  - Friday a week before the presentation during 1:00PM office hour
- For Wednesday presentation
  - Monday the week of presentation during 2:00PM office hour

# Other Resources

- CAp5415 Computer  Vision https://www.youtube.com/playlist?list=PLd3hlSJsX_IkXSinyREhlMjFvpNfpazfN
- Survey paper
  - Transformers in vision: A survey S Khan, M Naseer, M Hayat, SW Zamir, FS Khan, M Shah arXiv preprint arXiv:2101.01169

# Textbook

There is no textbook for this class. We will discuss recent research papers.

A good deep learning textbook:

- Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning. (free electronic version)

# Research Paper

- Each paper has following parts
  - Title
  - Abstract
  - Introduction
  - Rest of the paper
    - Related Work
    - Method
    - Results
    - Conclusion

- Each part is equally important (25% each!)

# How to read a research paper?

- You must read the paper several times to understand it.
  - When you read the paper first time, if you do not understand something do not get stuck, keep reading assuming you will figure out that later.
  - When you read it the second time, you will understand much more, and the third time even more …
- Read the abstract first then look at the figures with captions and then conclusion

# How to read a research paper?

- Try first to get a general idea of the paper
  - What problem is being solved?
  - What are the main steps?
  - How can I implement the method?, Even though I do not understand why each step is performed the way it is performed
- Try to relate the method to other methods you know, and conceptually find similarities and differences.

# How to read a research paper?

- In the first reading it may be a good idea to skip the related work.
- Do not use dictionary to just look up the meaning of technical terms like
- Try to understand each concept in isolation, and then integrate them to understand the whole paper.

# Useful Blogs about how to read a paper?

- https://web.stanford.edu/class/ee384m/Handouts/HowtoReadPaper.pdf

- https://blogs.lse.ac.uk/impactofsocialsciences/2016/05/09/how-to-read-and-understand-a-scientific-paper-a-guide-for-non-scientists/

# Artificial intelligence revolution

- Computer Vision is a branch of  Artificial Intelligence

# Artificial intelligence (AI) revolution

AI is new **Electricity!**

Andrew Ng, Stanford University

# Artificial intelligence revolution

- **AlphaGo beats Ke Jie again to wrap up three-part match**
- Google's Automated AI Tool Surpasses **Human** Experts
- **Machines 'beat humans' for a growing number of tasks**
- **Artificial intelligence** is the next **revolution** in customer service
- A Nobel Prize-Winning Economist Explains **Why Good AI Will Always Outsmart Humans.**
  - Economist Daniel Kahneman on why the human brain makes avoidable mistakes when solving problems

# Self-Driving Cars

- **Ford** intends to launch **self-driving car** service 'at scale' by 2021
- Will Self-Driving Cars End The **Big Automakers**?
- Every big tech firm in China is becoming a **self-driving** car company
- How safe should we expect **self-driving cars** to be?
- **Intel** acquired self-driving car platform company **Mobileye** early in 2017 for $15.3 billion.

# Artificial intelligence revolution

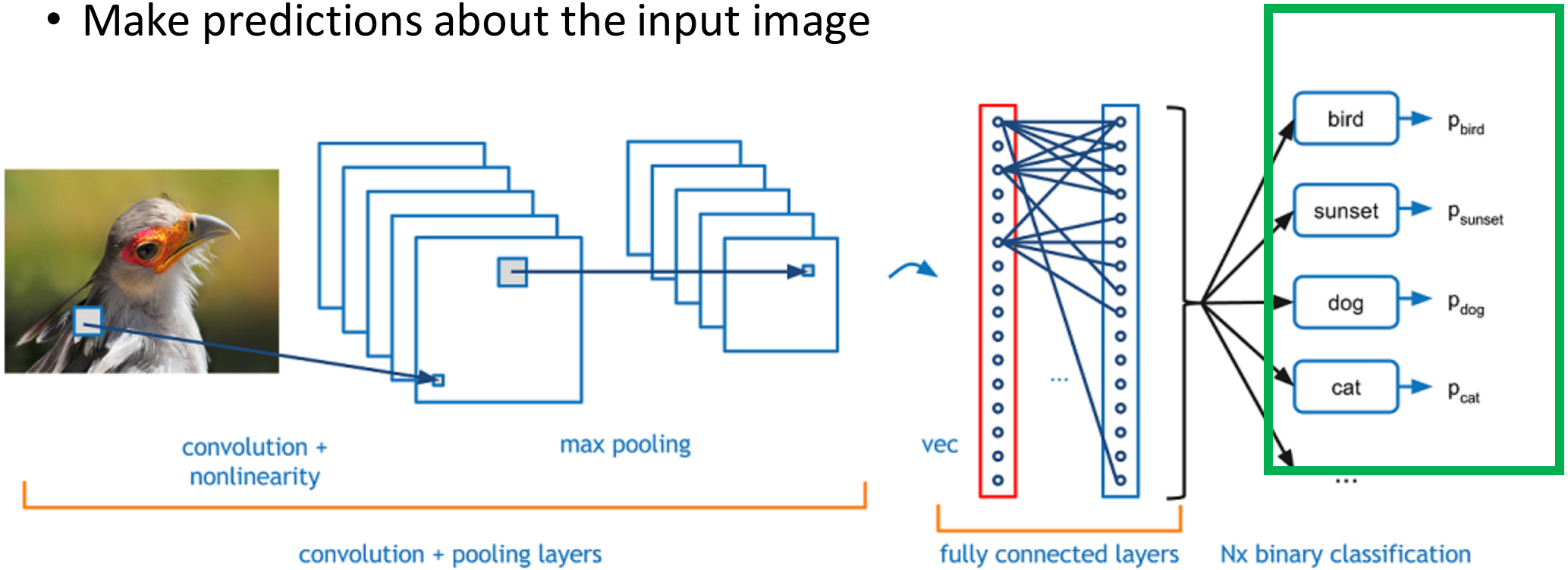
- DeepMind's new 280 billion-parameter language model kicks GPT-3's butt in accuracy
- Data Scientist Michael Jordan of Berkley Calls for a More Economics-Aware Approach to AI
- A New Link to an Old Model Could Crack the Mystery of Deep Learning | Quanta Magazine
- Deep North, which uses AI to track people from camera footage, raises $16.7M — VentureBeat
- Teaching AI to perceive the world through your eyes

# Artificial intelligence revolution

- The truth about artificial intelligence? It isn't that honest — The Guardian (US/CA) Tests of natural language processing models show that the bigger they are, the bigger liars they are. Should we be worried?
- IEEE Spectrum Deep Learning's Diminishing Returns The cost of improvement is becoming unsustainable Sep 24th, 2021 (193 kB)

# Convolutional Neural Network (CNN)

- A class of Neural Networks
  - Takes image as input
  - Make predictions about the input image




*Source : https://adeshpande3.github.io*

# Deep Learning Has Been Disruptive

- Learning
- Supervised
- Semi-Supervised
- Weakly-Supervised
- Unsupervised
- Self Supervised
- Reinforcement

- All data driven

# Deep Learning Has Been Disruptive

- Very Different Paradigm
- Real Learning
- Excellent results
- Rapid Progress
- Computer Vision is impacting other areas

# Deep Learning Has Been Disruptive

- Computer Vision is impacting other areas
  - Natural Language Understanding
  - Robotics
  - Computer Graphics
  - Speech Recognition
  - Games: Alpha Go
  - ….

# Deep Learning Has Been Disruptive

Dramatic Increase in Number of

- Publications
- Attendance in Conferences
- Datasets
- Startups
- Academics moving to Industry
- Software platforms/libraries

- GPUs
- Tutorials, videos, online courses

# Computer Vision Conferences and Journals

- Conferences
  - International Conference on Computer Vision (ICCV)
  - Computer Vision and Pattern Recognition (CVPR)
  - European Conference on Computer Vision (ECCV)
- Journals
  - IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)
  - International Journal of Computer Vision (IJCV)
  - Computer Vision and Image Understanding (CVIU)

# Machine Learning Conferences

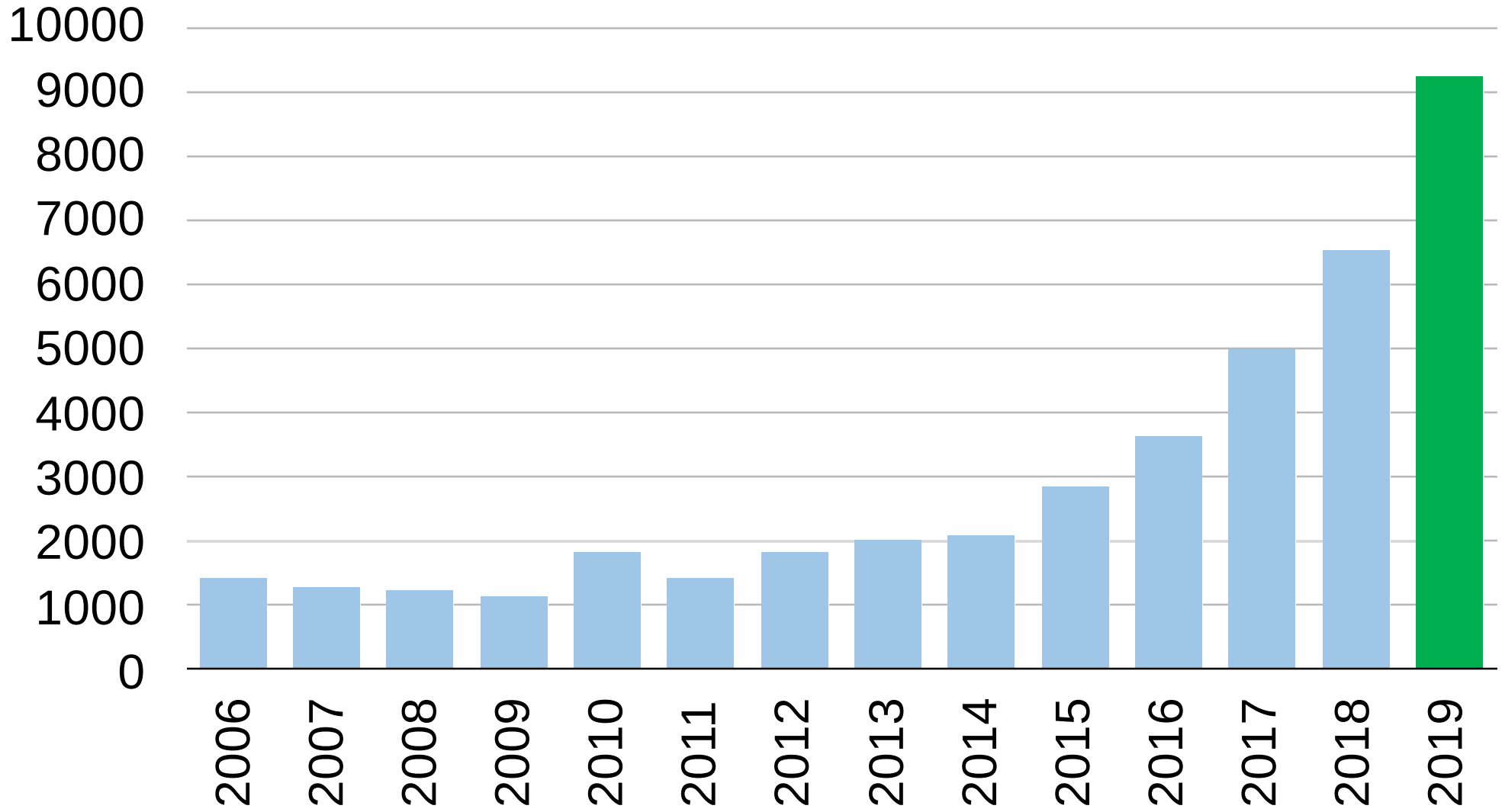
- Neural Information Processing Systems (**NIPS/NeurIPS**)
- International Conference on Machine Learning (**ICML**)
- International Conference on Learning Representations (**ICLR**)

# NIPS 2018 Tickets Sell Out in Less Than 12 Minutes



**Faster than this year's Burning Man, but still not as fast as a Beyonce concert! Maybe next year.**

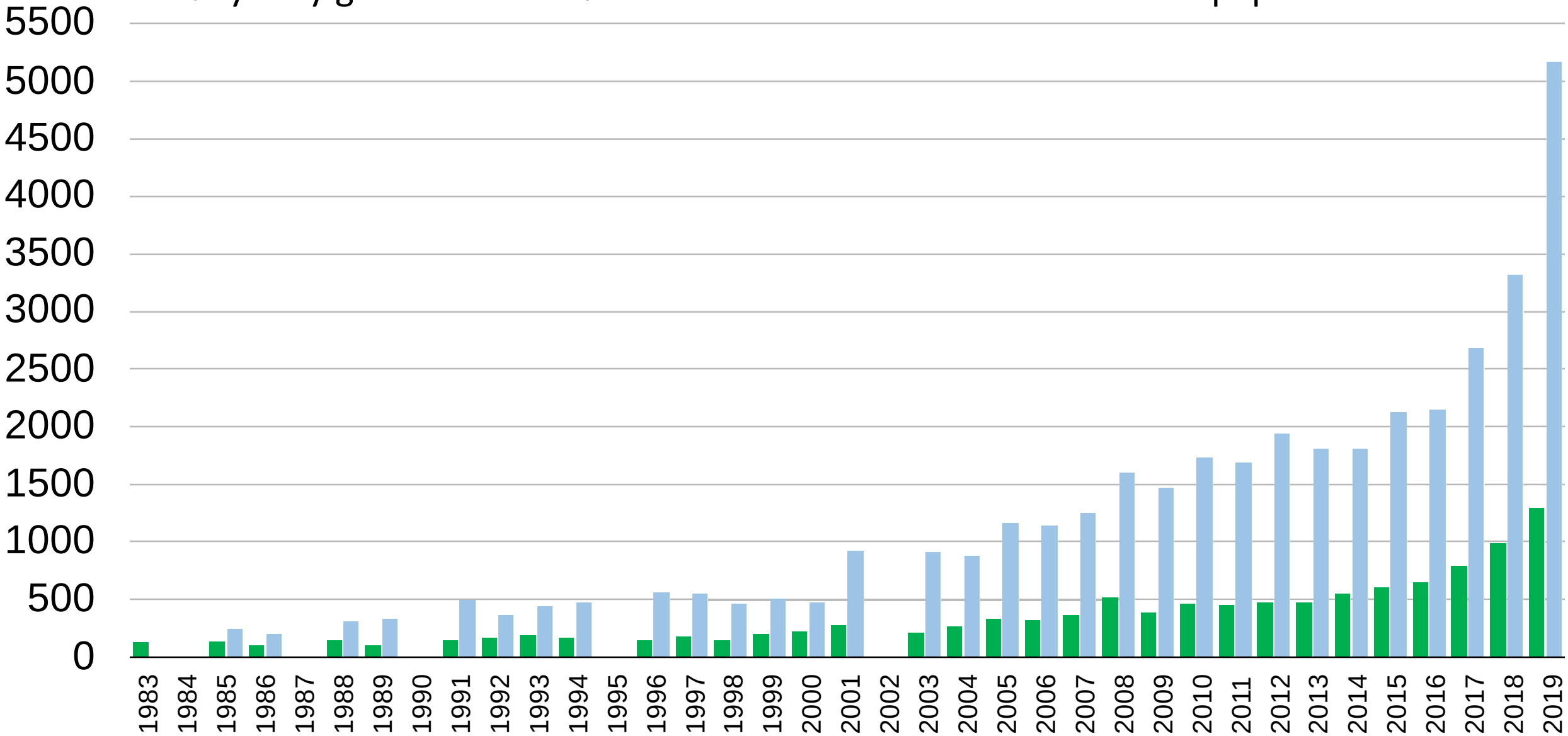
# CVPR Attendance Trend



UCF Center for Research in Computer Vision

# CVPR Submitted and Accepted Papers

56% yearly growth with 26% acceleration → 10.8B submitted papers in 2028



# CVPR is 4th best venue among all sciences!

- https://scholar.google.com/citations?view_op=top_venues&hl=en

# Top Computer Science Conferences

- https://research.com/conference-rankings/computer-science

Google Scholar



# Geoffrey Hinton

FOLLOW

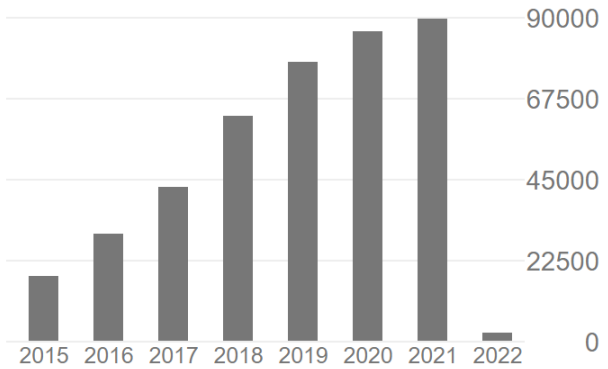
Emeritus Prof. Comp Sci, U.Toronto & Engineering Fellow, Google
Verified email at cs.toronto.edu - Homepage

machine learning psychology artificial intelligence cognitive science computer science

| TITLE | CITED BY | YEAR |
|---|---|---|
| Imagenet classification with deep convolutional neural networks<br>A Krizhevsky, I Sutskever, GE Hinton<br>Advances in neural information processing systems 25, 1097-1105 | 101030 | 2012 |
| Deep learning<br>Y LeCun, Y Bengio, G Hinton<br>Nature 521 (7553), 436-44 | 46543 | 2015 |
| Dropout: a simple way to prevent neural networks from overfitting<br>N Srivastava, G Hinton, A Krizhevsky, I Sutskever, R Salakhutdinov<br>The journal of machine learning research 15 (1), 1929-1958 | 33106 | 2014 |
| Learning internal representations by error-propagation<br>DE Rumelhart, GE Hinton, RJ Williams<br>Parallel Distributed Processing: Explorations in the Microstructure of … | 29718 | 1986 |
| Learning representations by back-propagating errors<br>DE Rumelhart, GE Hinton, RJ Williams<br>Nature 323 (6088), 533-536 | 27032 | 1986 |

Cited by VIEW ALL

| | All | Since 2017 |
|---|---|---|
| Citations | 534847 | 362461 |
| h-index | 166 | 119 |
| i10-index | 391 | 286 |




Public access VIEW ALL

0 articles 8 articles

not available available

Based on funding mandates

Google Scholar



# Kaiming He

FOLLOW

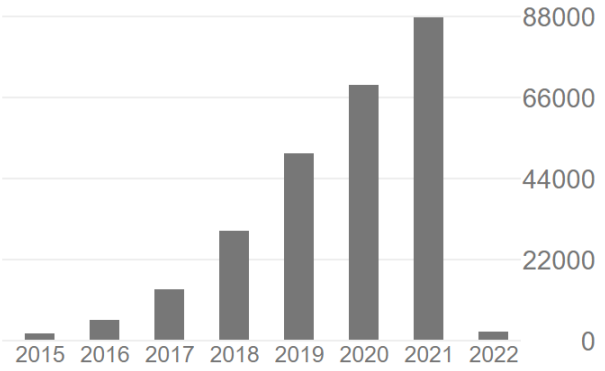
Research Scientist, Facebook AI Research (FAIR)
Verified email at fb.com - Homepage

Computer Vision  Machine Learning

| TITLE | CITED BY | YEAR |
|---|---|---|
| Deep Residual Learning for Image Recognition<br>K He, X Zhang, S Ren, J Sun<br>Computer Vision and Pattern Recognition (CVPR), 2016 | 102850 | 2016 |
| Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks<br>S Ren, K He, R Girshick, J Sun<br>Neural Information Processing Systems (NIPS), 2015 | 38183 | 2015 |
| Mask R-CNN<br>K He, G Gkioxari, P Dollár, R Girshick<br>International Conference on Computer Vision (ICCV), 2017 | 16506 | 2017 |
| Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification<br>K He, X Zhang, S Ren, J Sun<br>International Conference on Computer Vision (ICCV), 2015 | 14273 | 2015 |
| Focal Loss for Dense Object Detection | 11056 | 2017 |

Cited by

| | All | Since 2017 |
|---|---|---|
| Citations | 266859 | 254497 |
| h-index | 59 | 57 |
| i10-index | 68 | 68 |

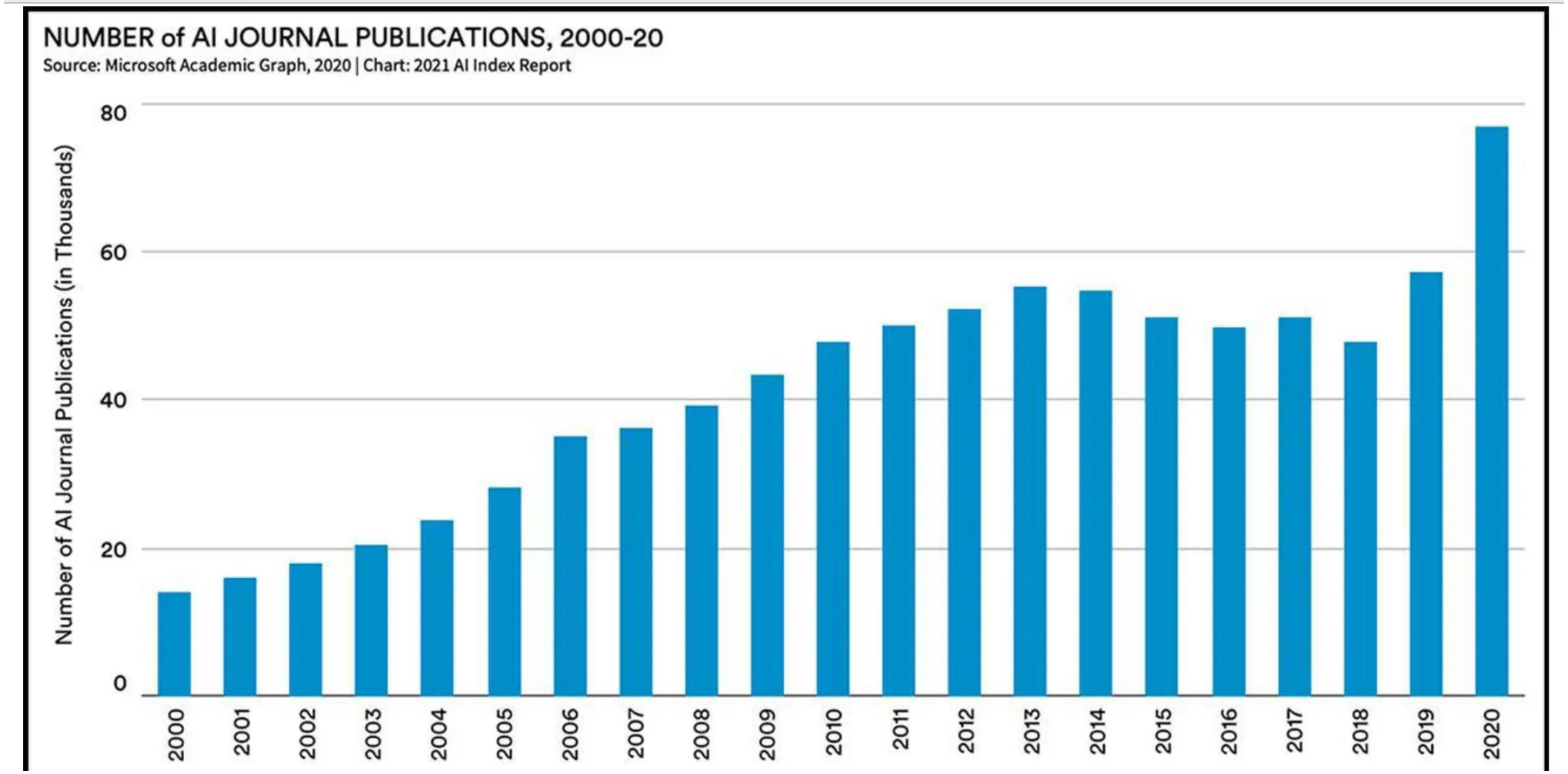



Public access  VIEW ALL

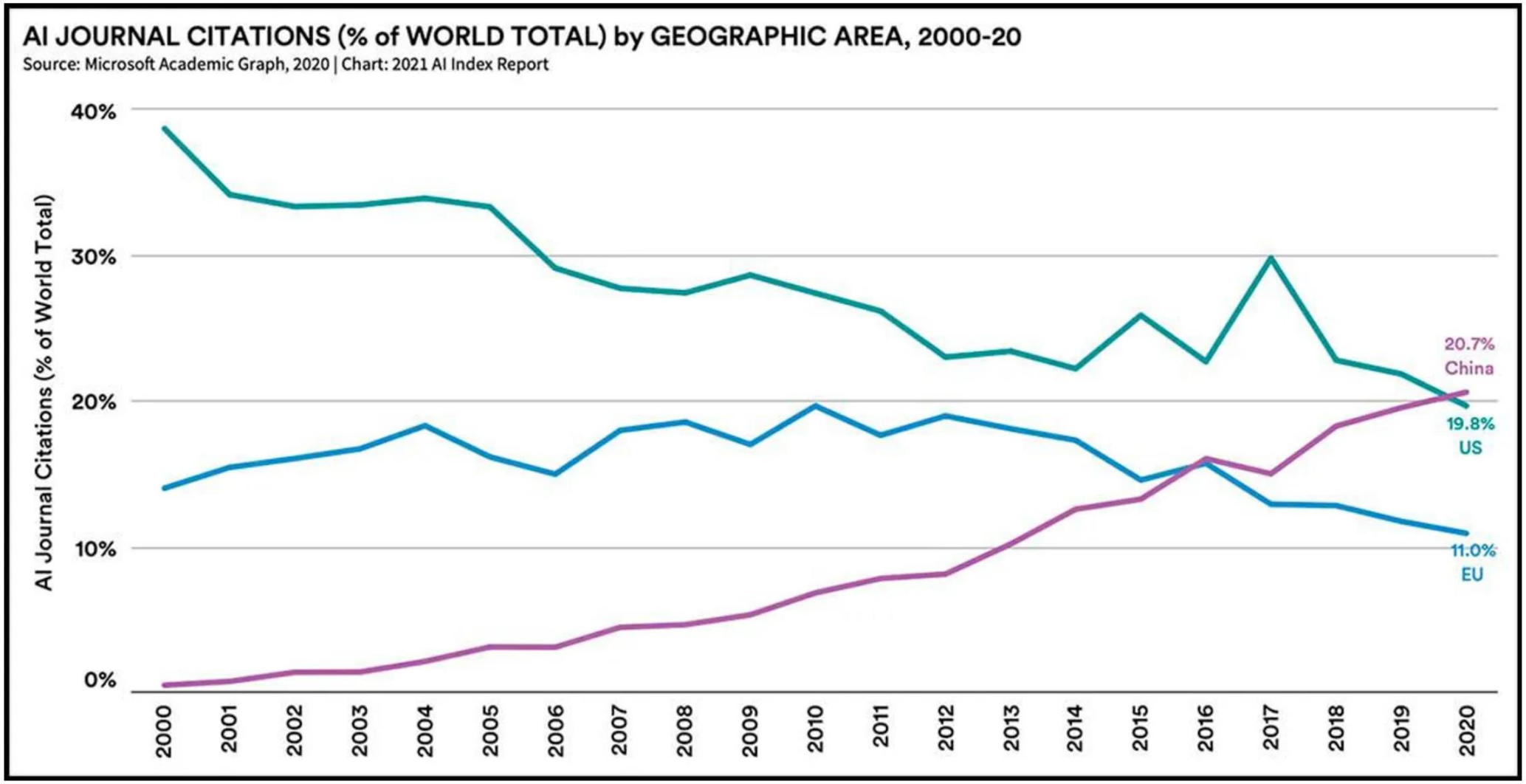
0 articles  10 articles

not available  available

Based on funding mandates

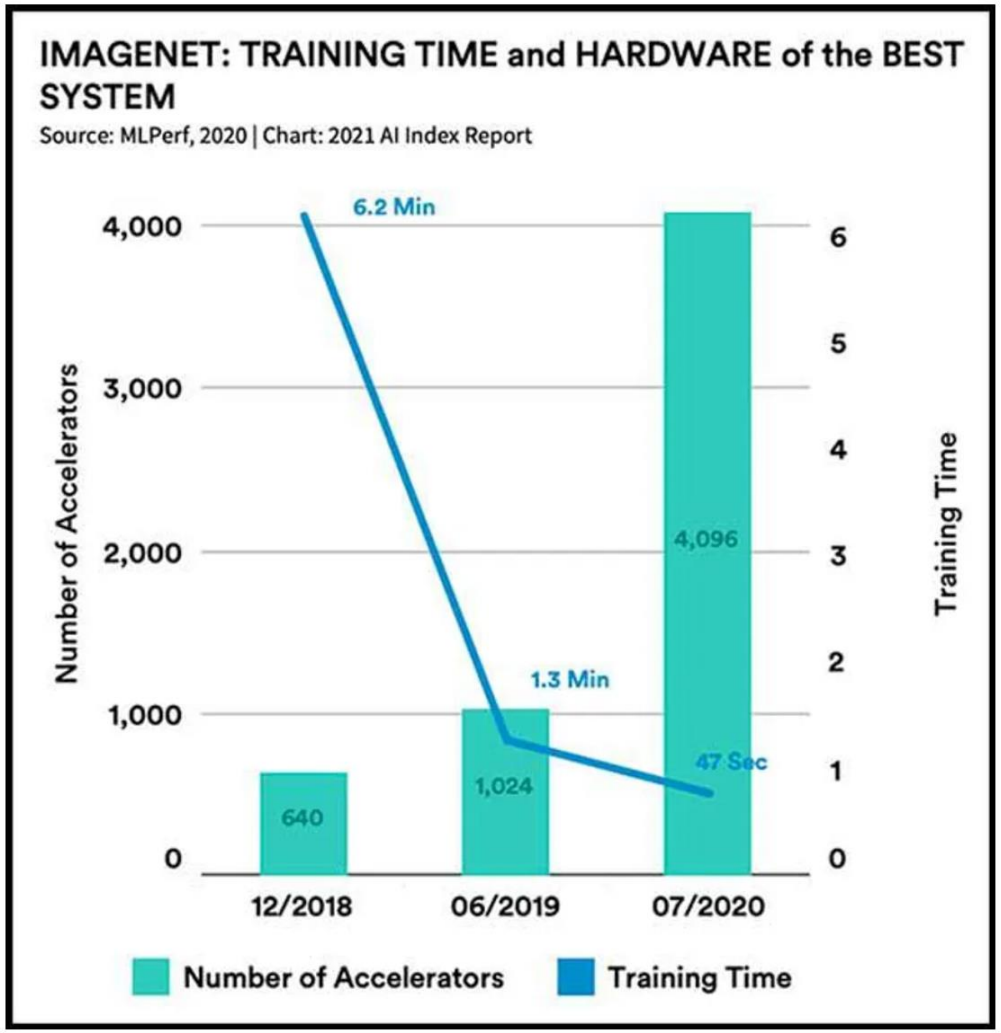



15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM



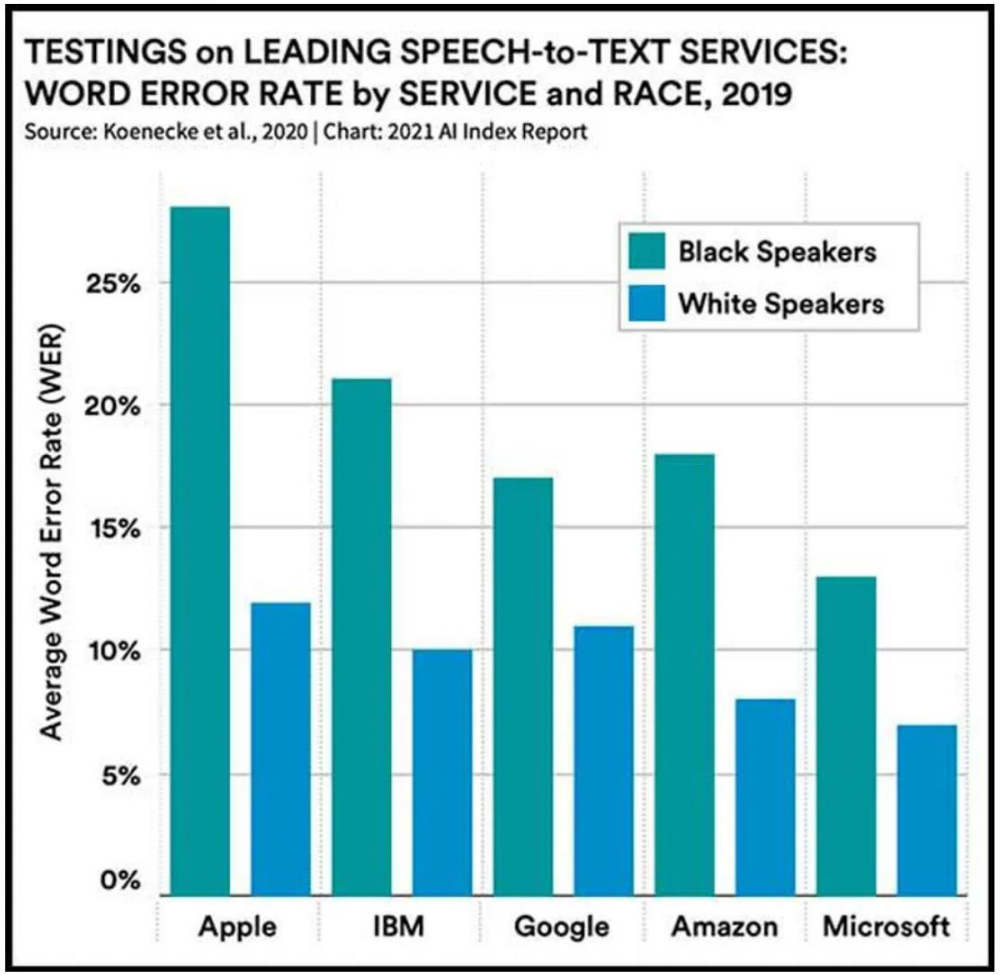

15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM

# 3. Faster Training = Better AI

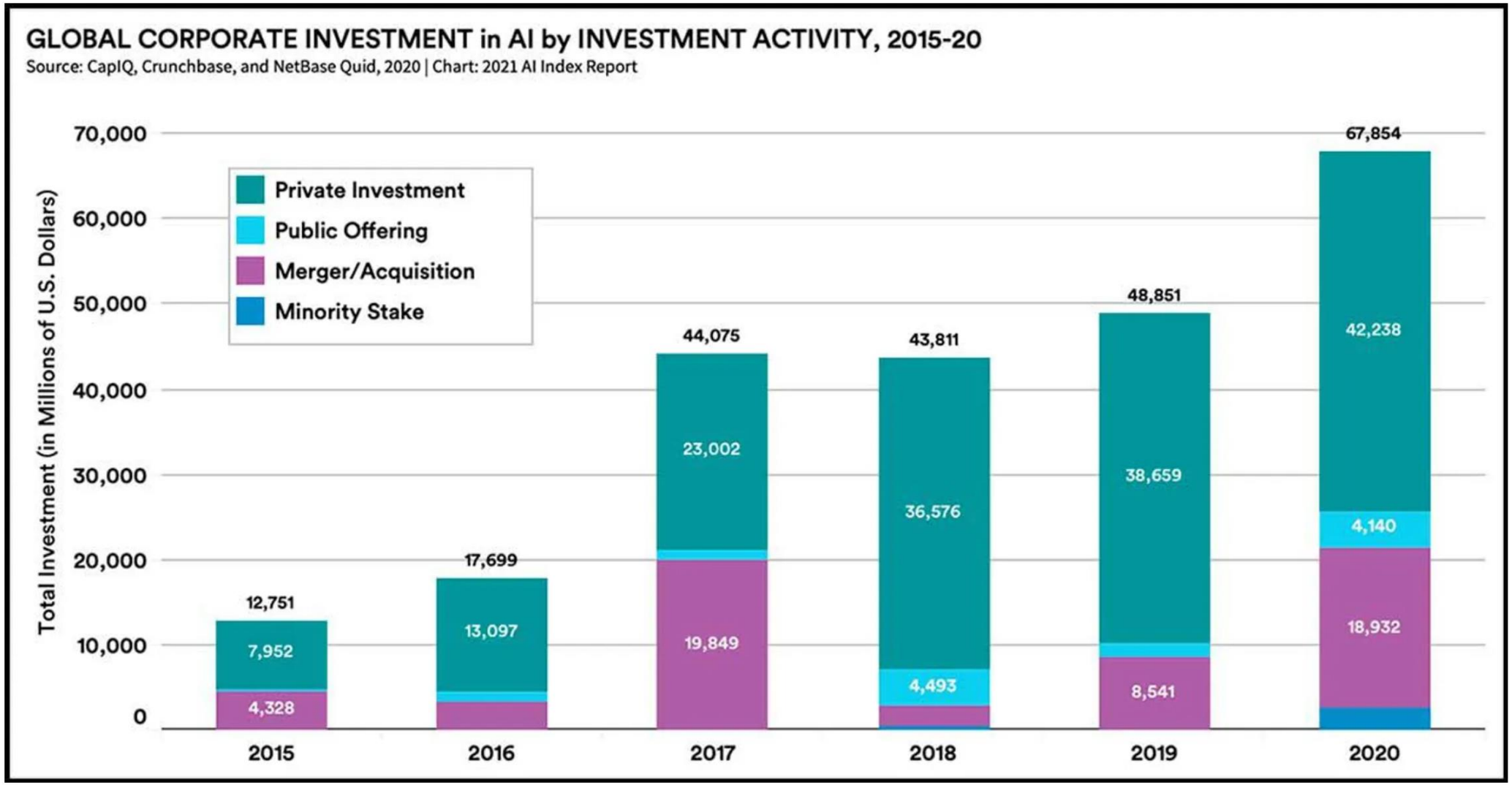



15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM

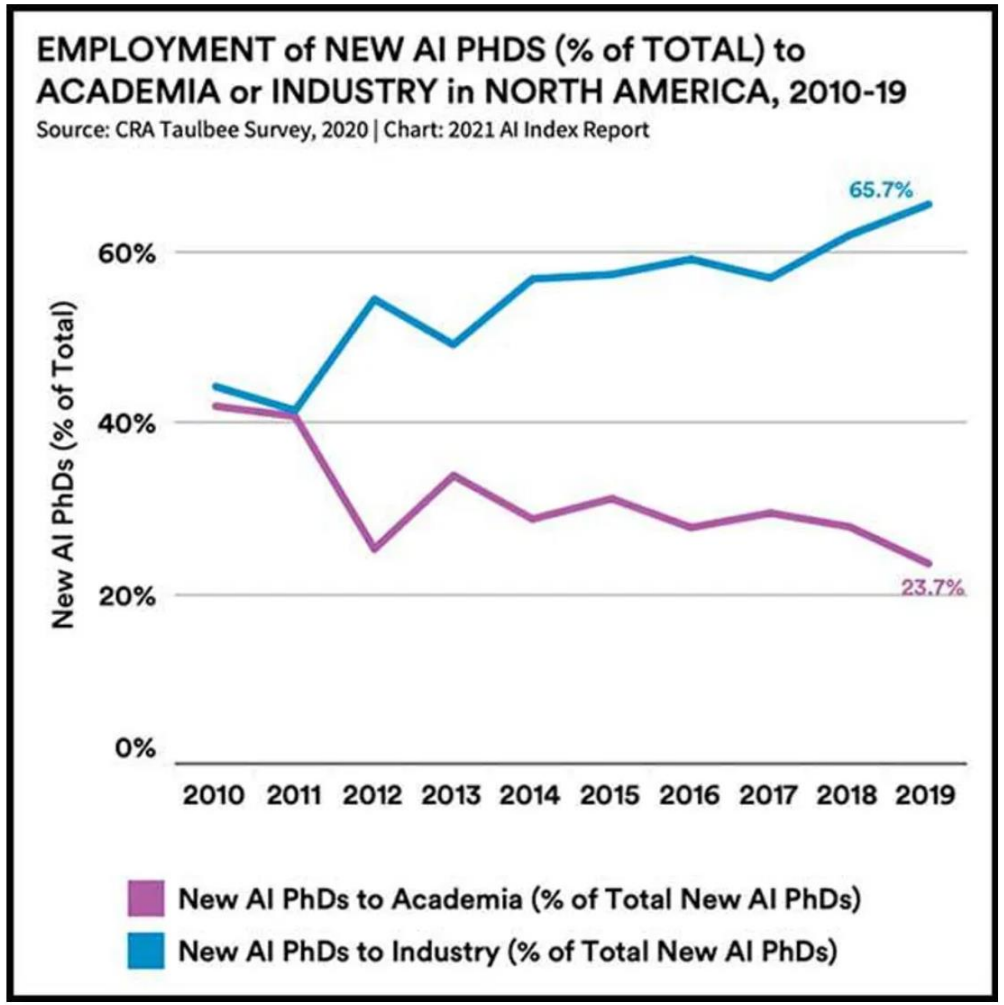
# 6. A Huge Caveat



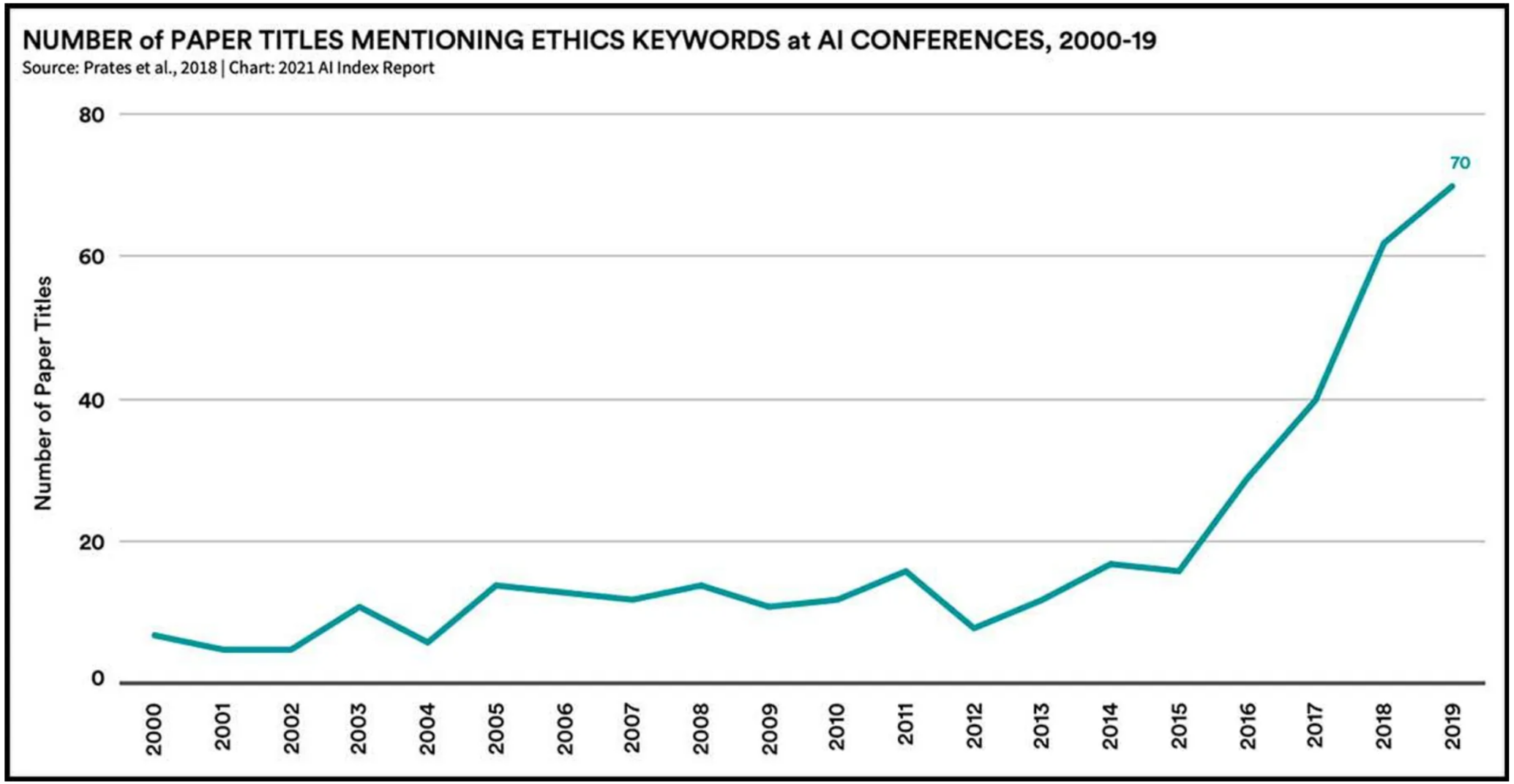

15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM


GLOBAL CORPORATE INVESTMENT in AI by INVESTMENT ACTIVITY, 2015-20
Source: CapIQ, Crunchbase, and NetBase Quid, 2020 | Chart: 2021 AI Index Report
Total Investment (in Millions of U.S. Dollars)
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
Private Investment
Public Offering
Merger/Acquisition
Minority Stake
12,751
7,952
4,328
17,699
13,097
44,075
23,002
19,849
43,811
36,576
4,493
48,851
38,659
8,541
67,854
42,238
4,140
18,932
2015
2016
2017
2018
2019
2020

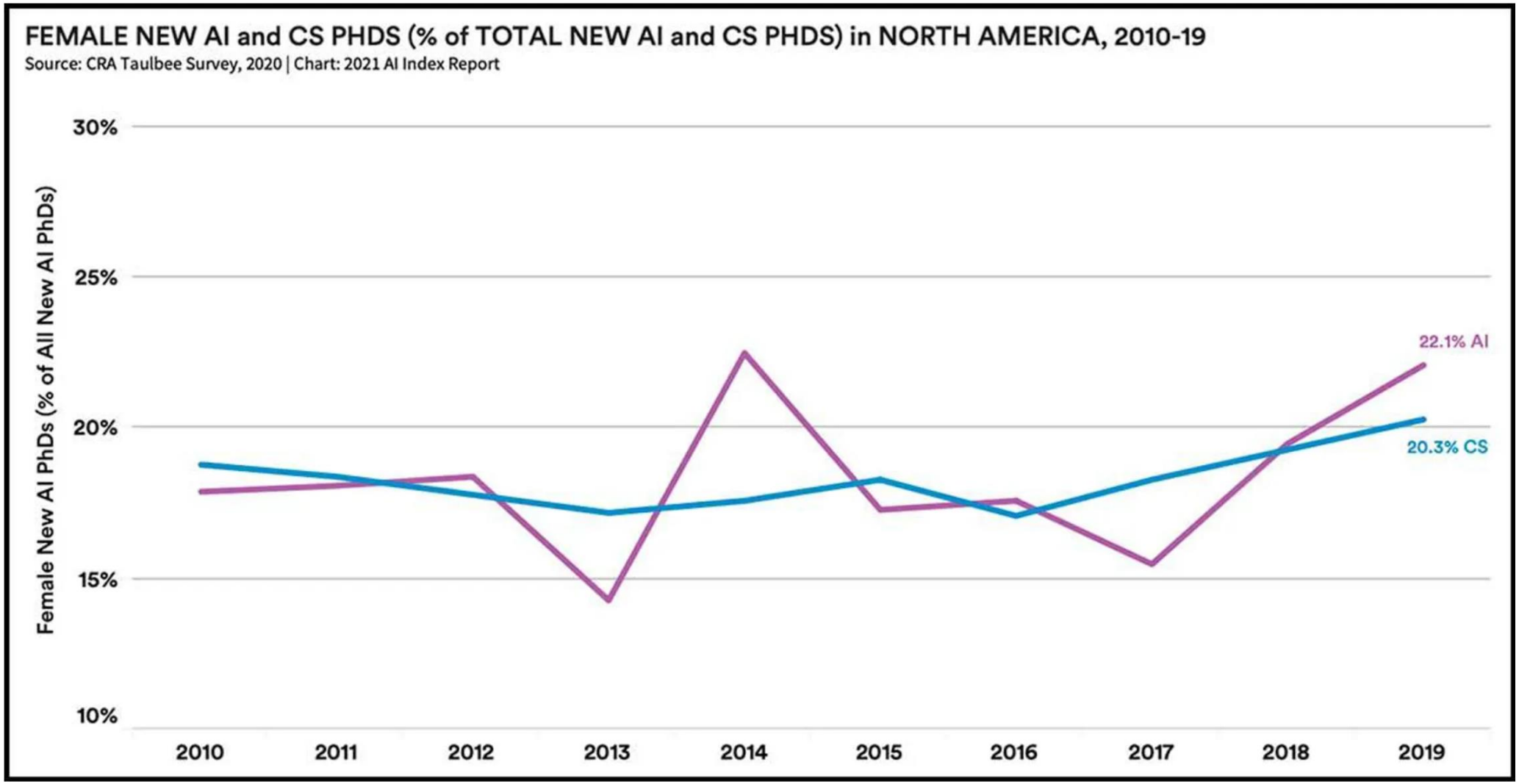
# 12. Ph.D.'s Hear the Siren Call of Industry



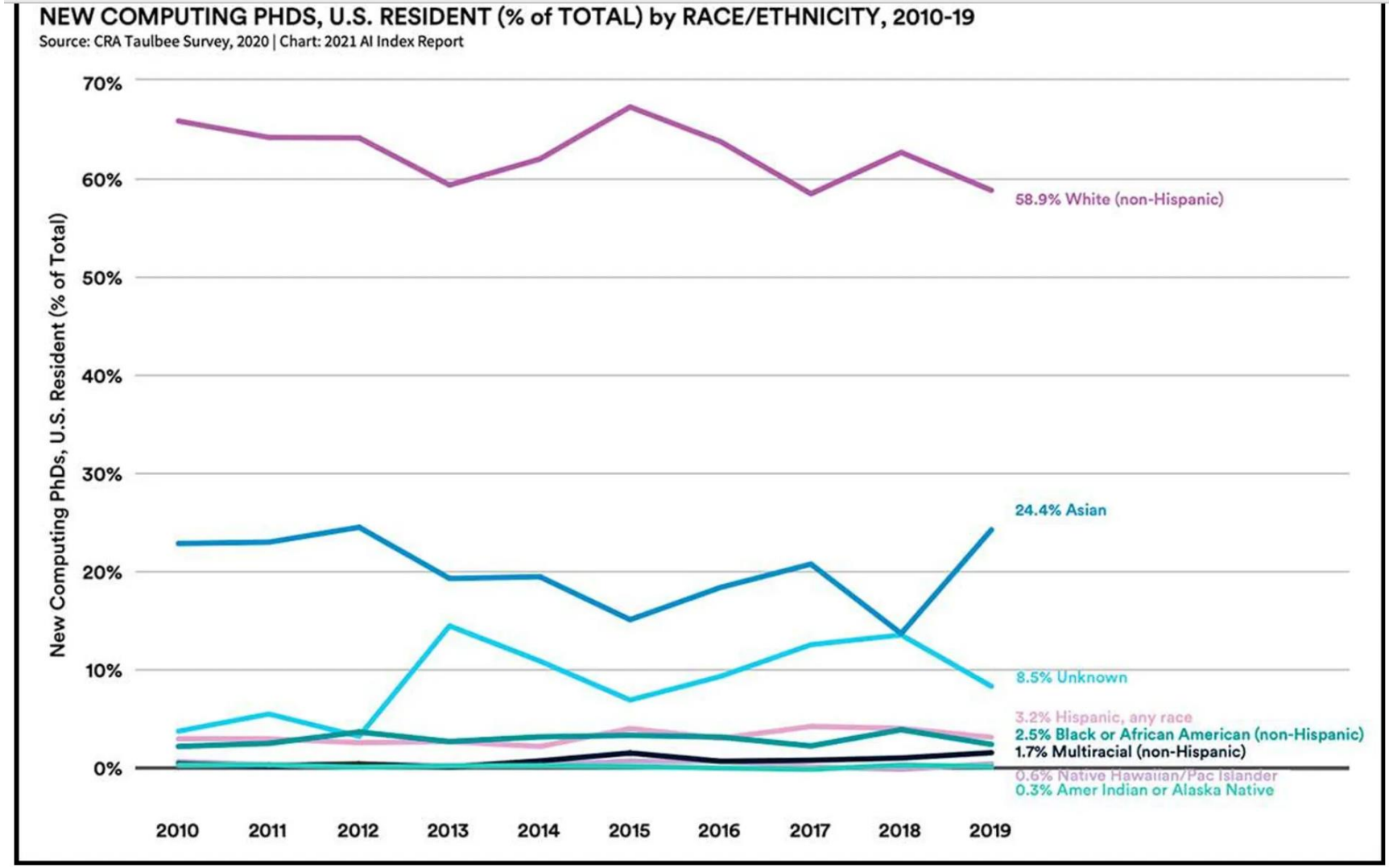

15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM




15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM




15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM




15 Graphs You Need to See to Understand AI in 2021 The 2021 AI Index provides insight into jobs, publications, diversity, and more ELIZA STRICKLAND 15 APR 2021, IEEE SPECTRUM

# Artificial intelligence revolution

Machine Learning Has become like Alchemy!

Ali Rahimi, Google

# Limitations

- Deep Learning systems are brittle
  - Small perturbations can cause state-of-the-art DNNs to misclassify an image

- Deep Learning systems are Blackbox
  - Their predictions are not explainable

# Adversarial Attack

$\epsilon$

$$\boldsymbol{\rho} = \epsilon \, \text{sign}\left(\nabla \mathcal{J}(\boldsymbol{\theta}, \mathbf{I}_c, \ell)\right)$$



$\mathbf{I}_c$

"panda"
57.7% confidence

$+ .007 \times$



$\text{sign}\left(\nabla \mathcal{J}(\boldsymbol{\theta}, \mathbf{I}_c, \ell)\right)$

"nematode"
8.2% confidence

$=$



$\mathbf{I}_c + \epsilon \, \text{sign}\left(\nabla \mathcal{J}(\boldsymbol{\theta}, \mathbf{I}_c, \ell)\right)$

"gibbon"
99.3 % confidence

[ Ribeiro et al. 2016 ]

# Predict Wolf vs Husky

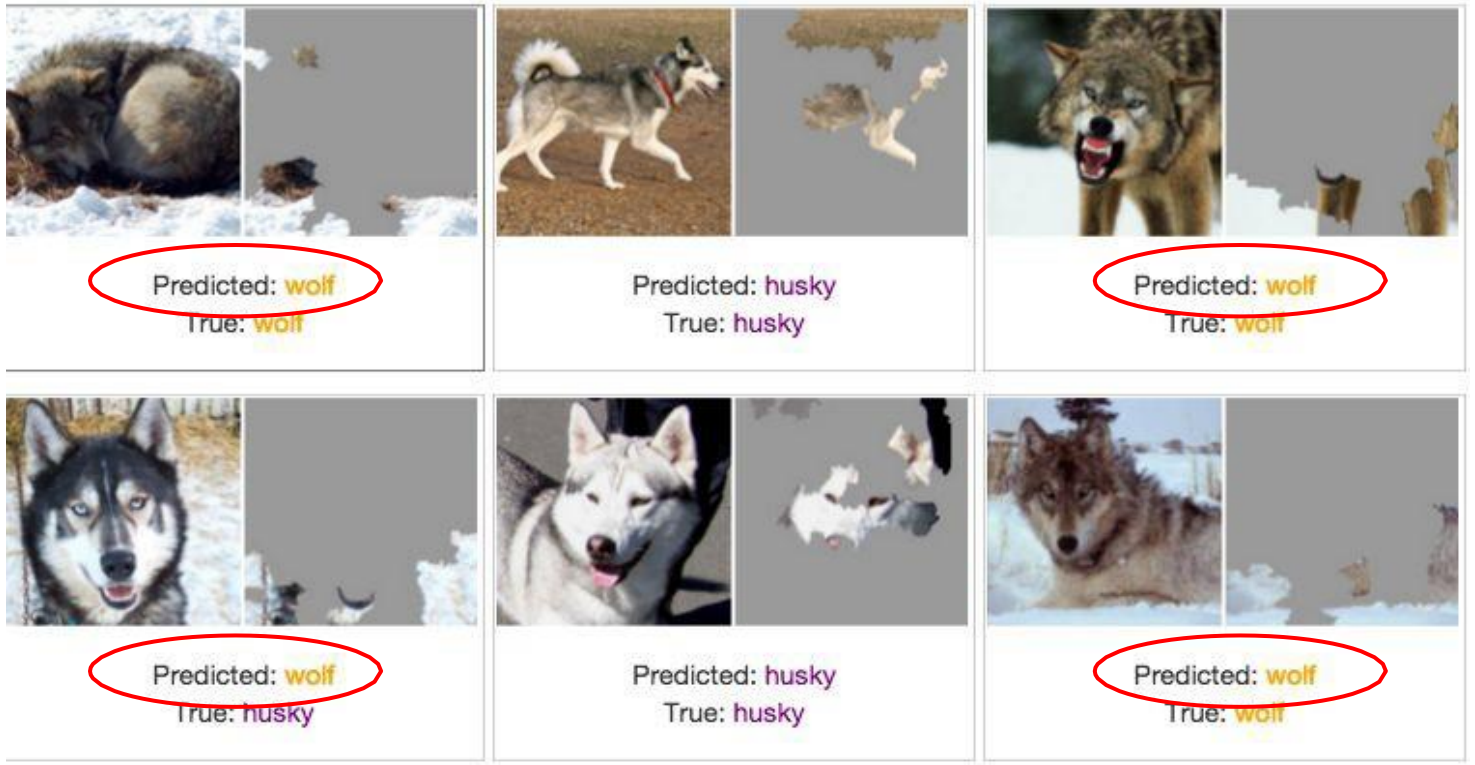
Only 1 mistake!




Slide credit: Hima, Lakkaraju Julius, Adebayo Sameer Singh

[ Ribeiro et al. 2016 ]

# Predict Wolf vs Husky




We've built a great snow detector…

Slide credit: Hima Lakkaraju, Julius Adebayo, Sameer Singh

**Slides and Video:** explainml-tutorial.github.io

# AI Experts Speak: Memorable Quotes from Spectrum's AI Coverage

- “In terms of how much progress we've made in this work over the last two decades: I don't think we're anywhere close today to the level of intelligence of a two-year-old child. But maybe we have algorithms that are equivalent to lower animals for perception.” YOSHUA BENGIO
- “Dig into every industry, and you'll find AI changing the nature of work.” DANIELA RUS,MIT AI Lab
- “We need to think about how to steer AI. How do we transform today's buggy and hackable AI systems into systems we can really trust and make sure they really do what we intend them to do? How do we teach machines to understand our goals, adopt our goals, and retain our goals?” MAX TEGMARK

# AI Experts Speak: Memorable Quotes from Spectrum's AI Coverage

- “The policy goal is to get the regulators to enforce their laws in the age of algorithms. You can't give up on enforcing antidiscrimination laws because you don't understand the technology.” **CATHY O'NEIL**
- “General intelligence doesn't mean you have to understand humans. There will be a lot of intelligent machines that don't need to know that stuff. [Take the] example of robotic construction workers on Mars: I don't think they need to understand human emotions and human desires to be able to construct things.” **JEFF HAWKINS**
- “AI will take many single-task, single-domain jobs away. You can argue that humans have abilities that AI does not: We can conceptualize, strategize, create. Whereas today's AI is just a really smart pattern recognizer that can take in data [and] optimize. But how many jobs in the world are simple repetitions of tasks that can be optimized?” **KAI-FU LEE**

# AI Experts Speak: Memorable Quotes from Spectrum's AI Coverage

- "AI is fundamentally an applied technology that's going to serve our society. Humanistic AI not only raises the awareness of the importance of the technology. It's a really important way to attract diverse students, technologists, and innovators to participate." **FEI-FEI LI**
- "Those of us in machine learning are really good at doing well on a test set, but unfortunately deploying a system takes more than doing well on a test set. All of AI…has a proof-of-concept-to-production gap." **ANDREW NG**
- "While the science-fiction discussions about AI and superintelligence are fun, they are a distraction. There's not been enough focus on the real problem, which is building planetary-scale machine learning–based systems that actually work, deliver value to humans, and do not amplify inequities." MICHAEL JORDAN,

# AI Experts Speak: Memorable Quotes from Spectrum's AI Coverage

- “What's missing is a principle that would allow our machine to learn how the world works by observation and by interaction with the world. A learning predictive world model is what we're missing today, and in my opinion is the biggest obstacle to significant progress in AI.” YANN LECUN
- “The beautiful thing about AI and robotics is that you're never done.” MANUELA VELOSO
- “Those who argue that the risk from AI is negligible have failed to explain why superintelligent AI systems will necessarily remain under human control; and they have not even tried to explain why superintelligent AI systems will never be developed.” STUART RUSSELL

# Jürgen Schmidhuber

- ACM explicitly mentions "astonishing" deep learning breakthroughs in 4 fields: (A) speech recognition, (B) natural language processing, (C) robotics, (D) computer vision, as well as "powerful" new deep learning tools in 3 fields: medicine, astronomy, materials science. Most of these breakthroughs and tools, however, were direct consequences of the breakthroughs **of my lab** and other labs in the past 3 decades.

# Introductions

- My Introduction

- Your Introduction

Slide credit: Judy Hoffman