THE UNIVERSITY OF TEXAS
MD Anderson ~~Cancer~~ Center
Making Cancer History®

Artificial Intelligence in Radiology

Naveen Garg, MD

Associate Professor

Abdominal Imaging Section

Department of Diagnostic Radiology

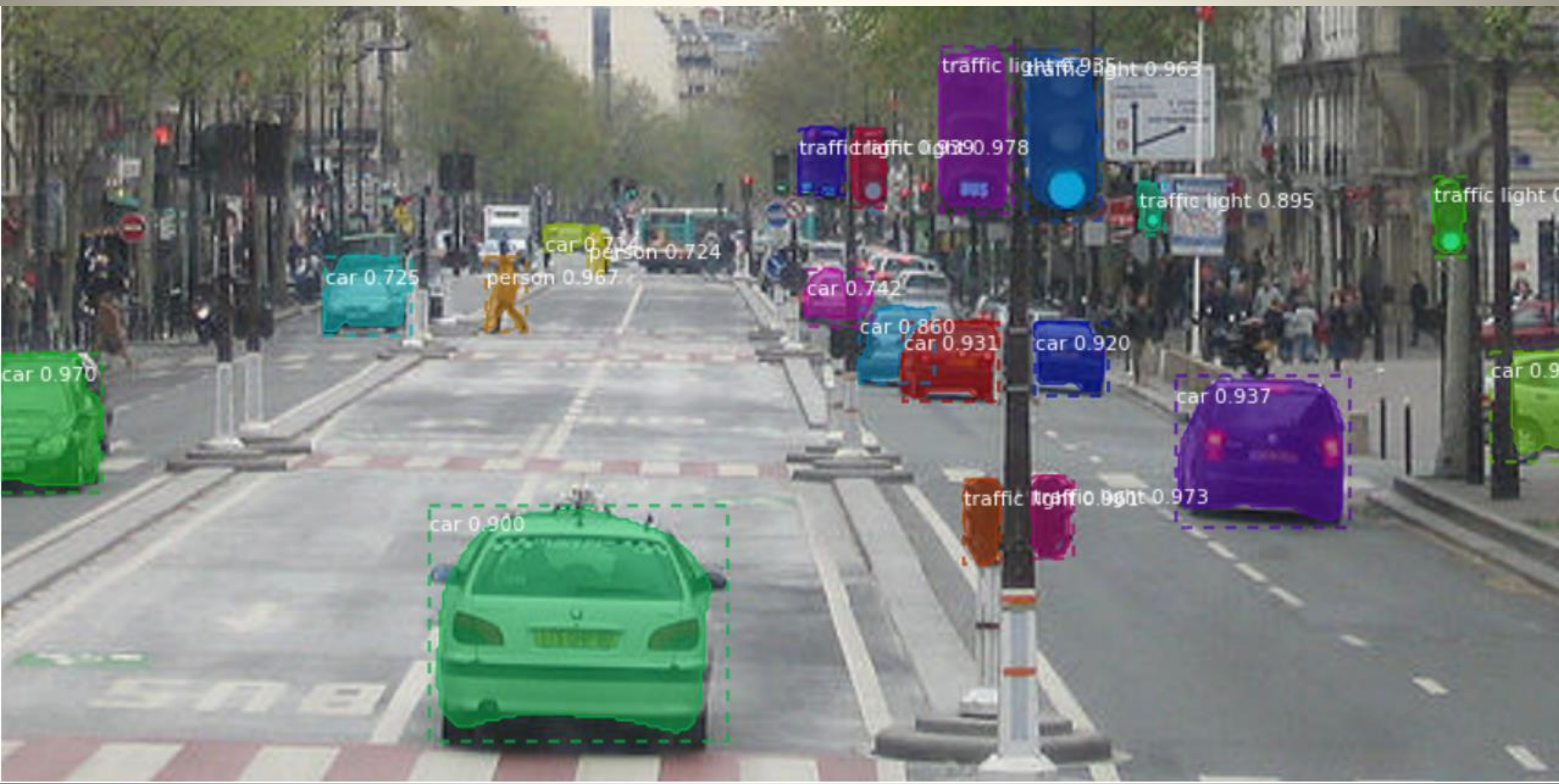
Disclosures

- Owner of Garglet LLC
- Advisor for Enlitic Inc.

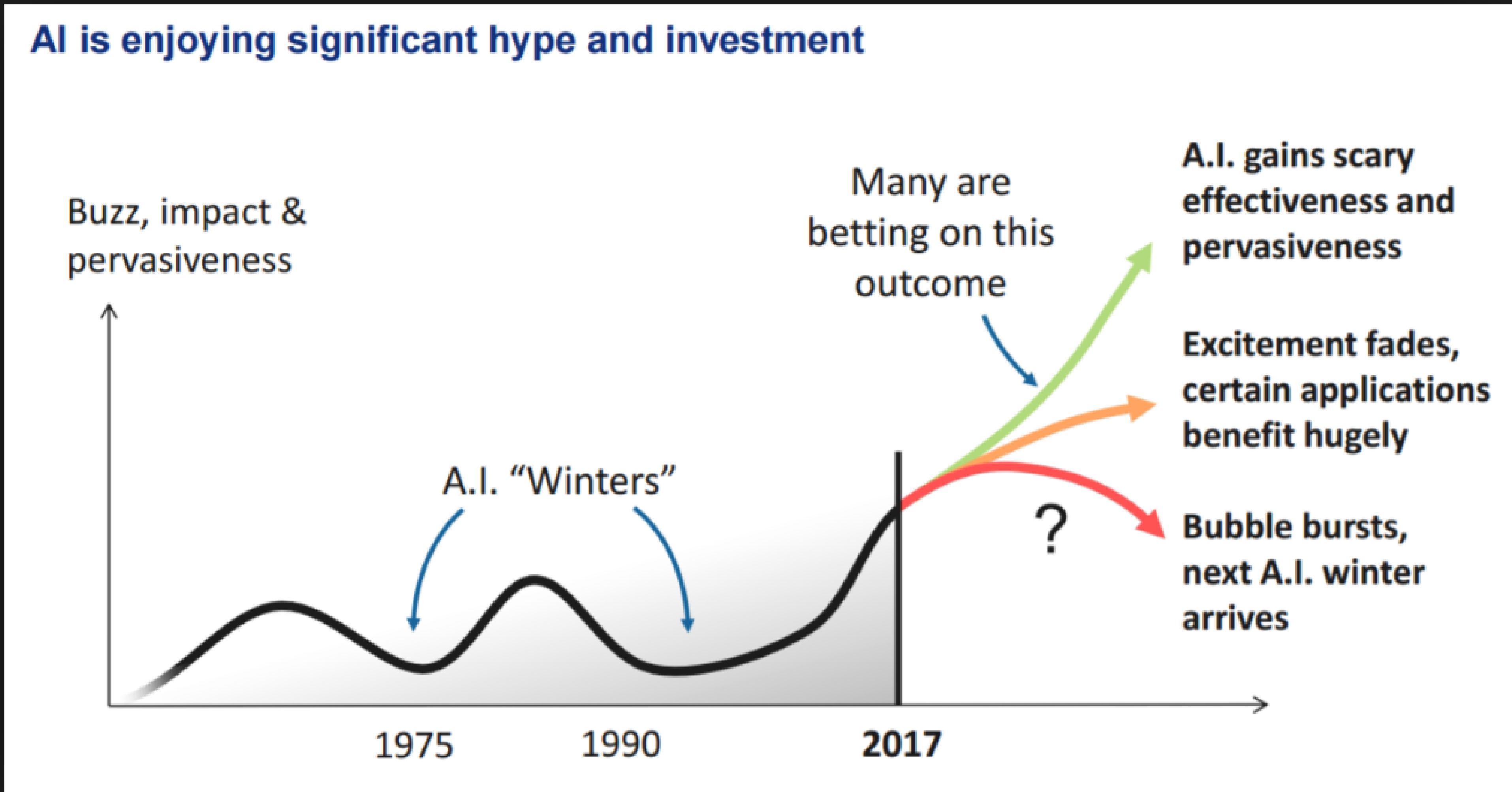
The possibility of AI doing radiology

Tesla autopilot

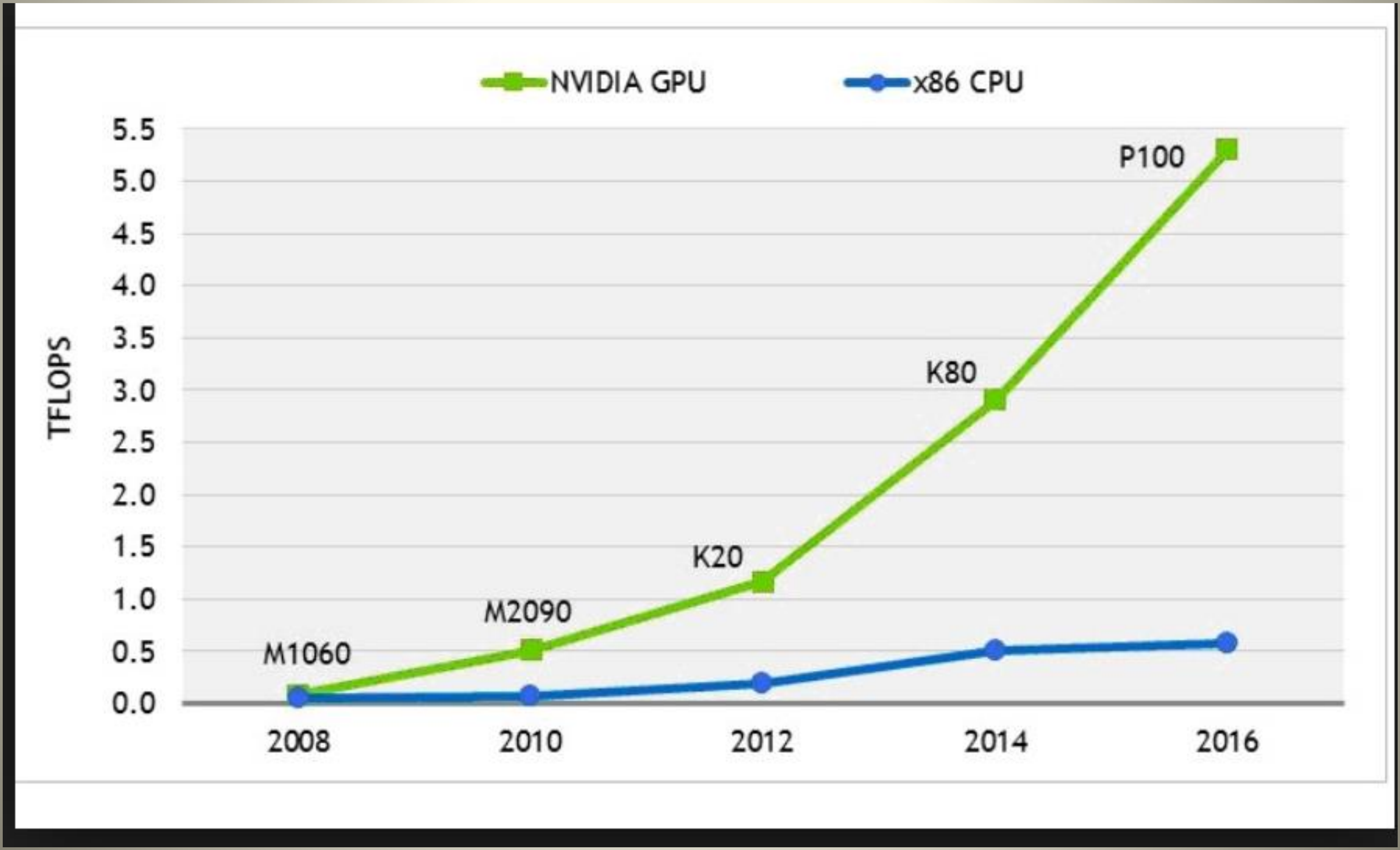
maskrcnn


car 0.970
car 0.725
person 0.967
person 0.724
car 0.742
car 0.860
car 0.931
car 0.920
car 0.937
car 0.900
traffic light 0.963
traffic light 0.895

AI Winters

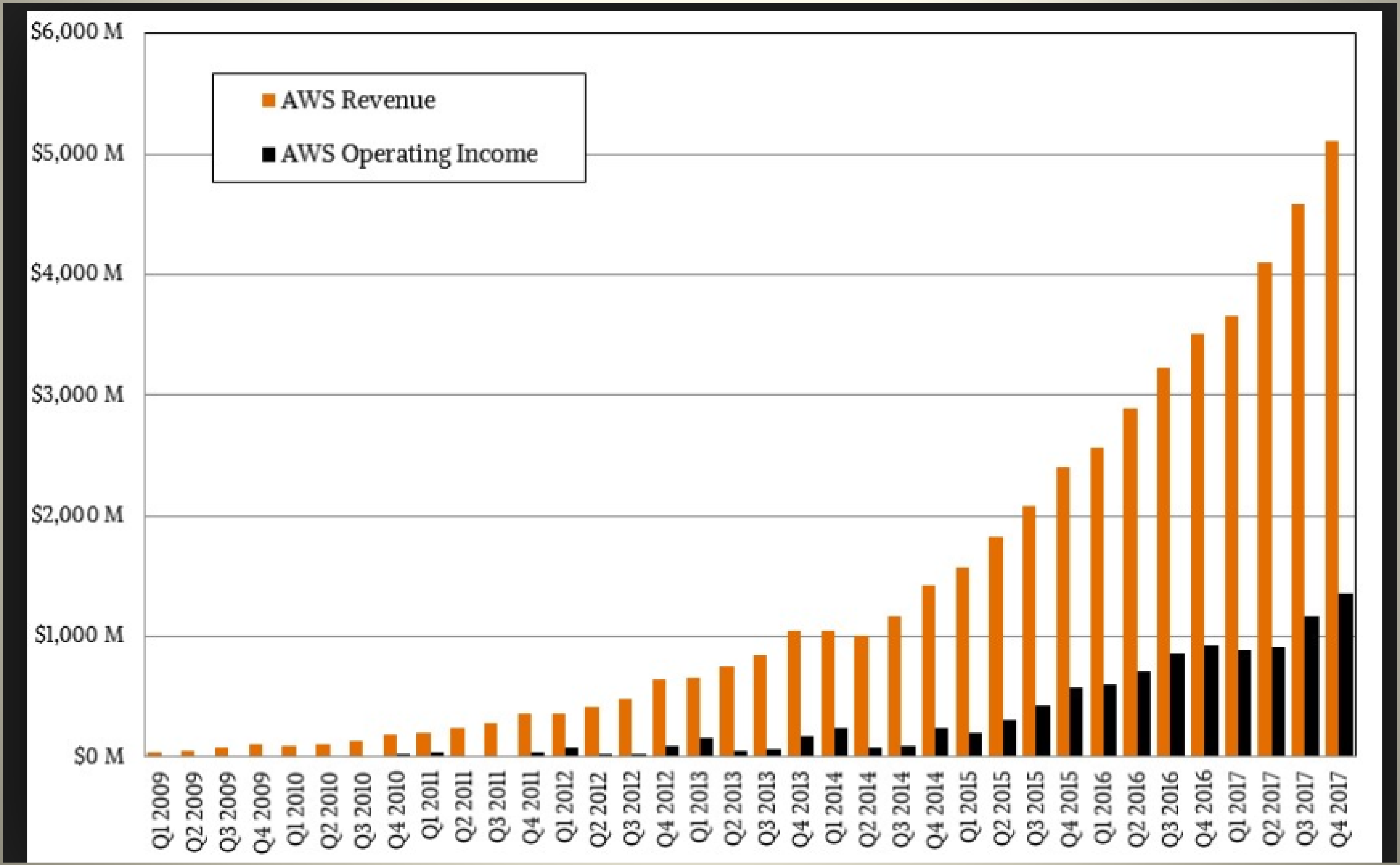

AI is enjoying significant hype and investment
Buzz, impact &
pervasiveness
A.I. "Winters"
Many are
betting on this
outcome
A.I. gains scary
effectiveness and
pervasiveness
Excitement fades,
certain applications
benefit hugely
?
Bubble bursts,
next A.I. winter
arrives
1975
1990
2017

Increases in computing power

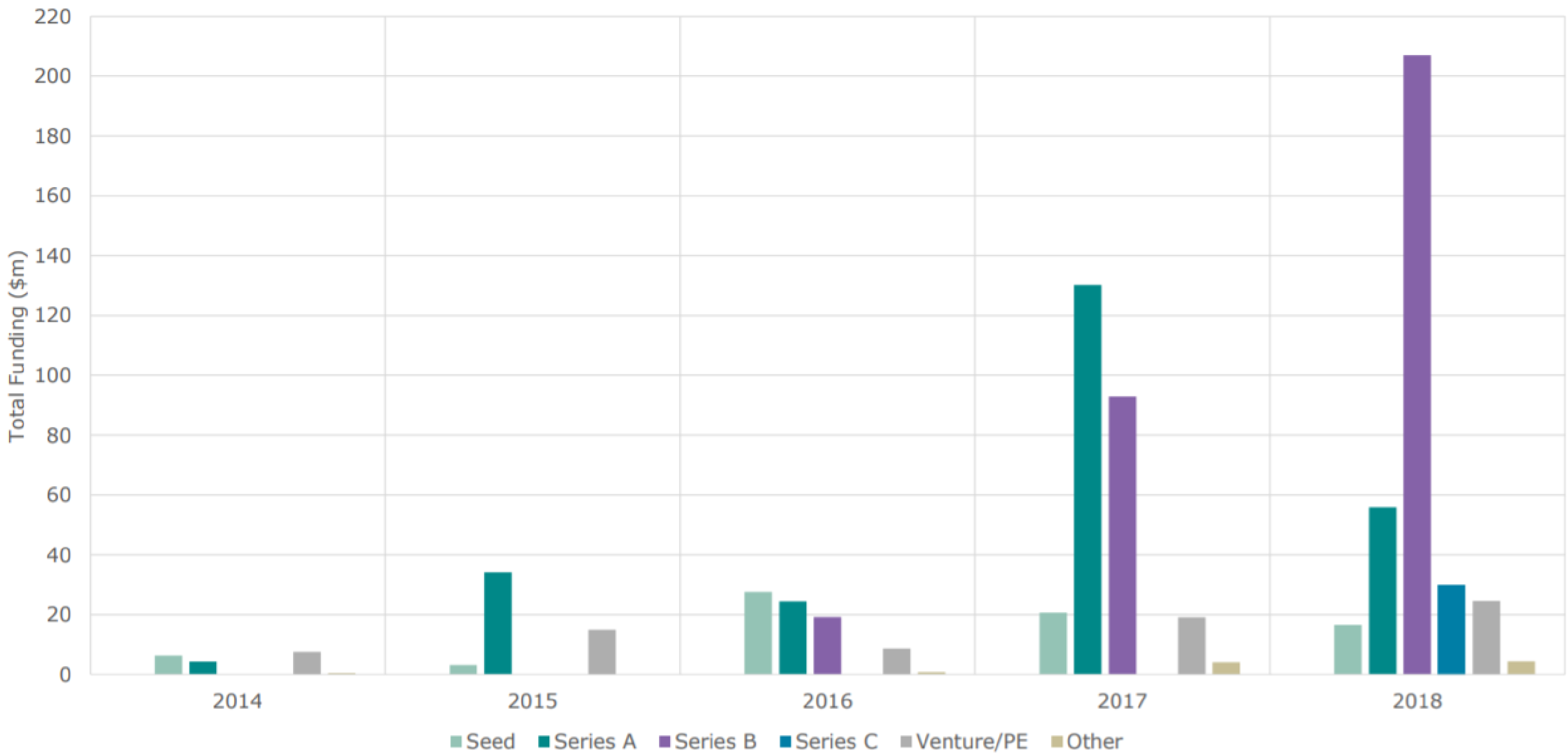


NVIDIA Stock price

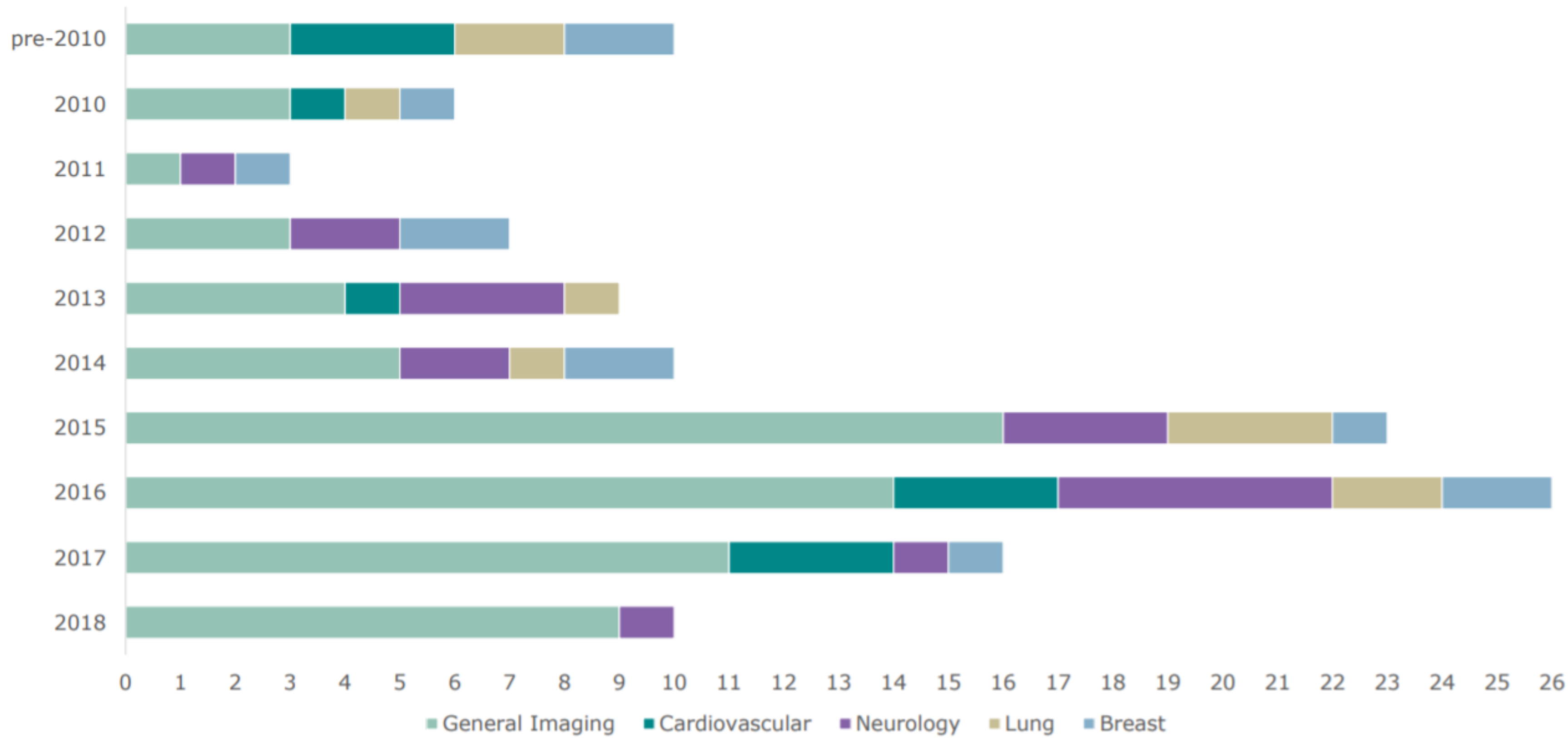
Rise of Cloud computing



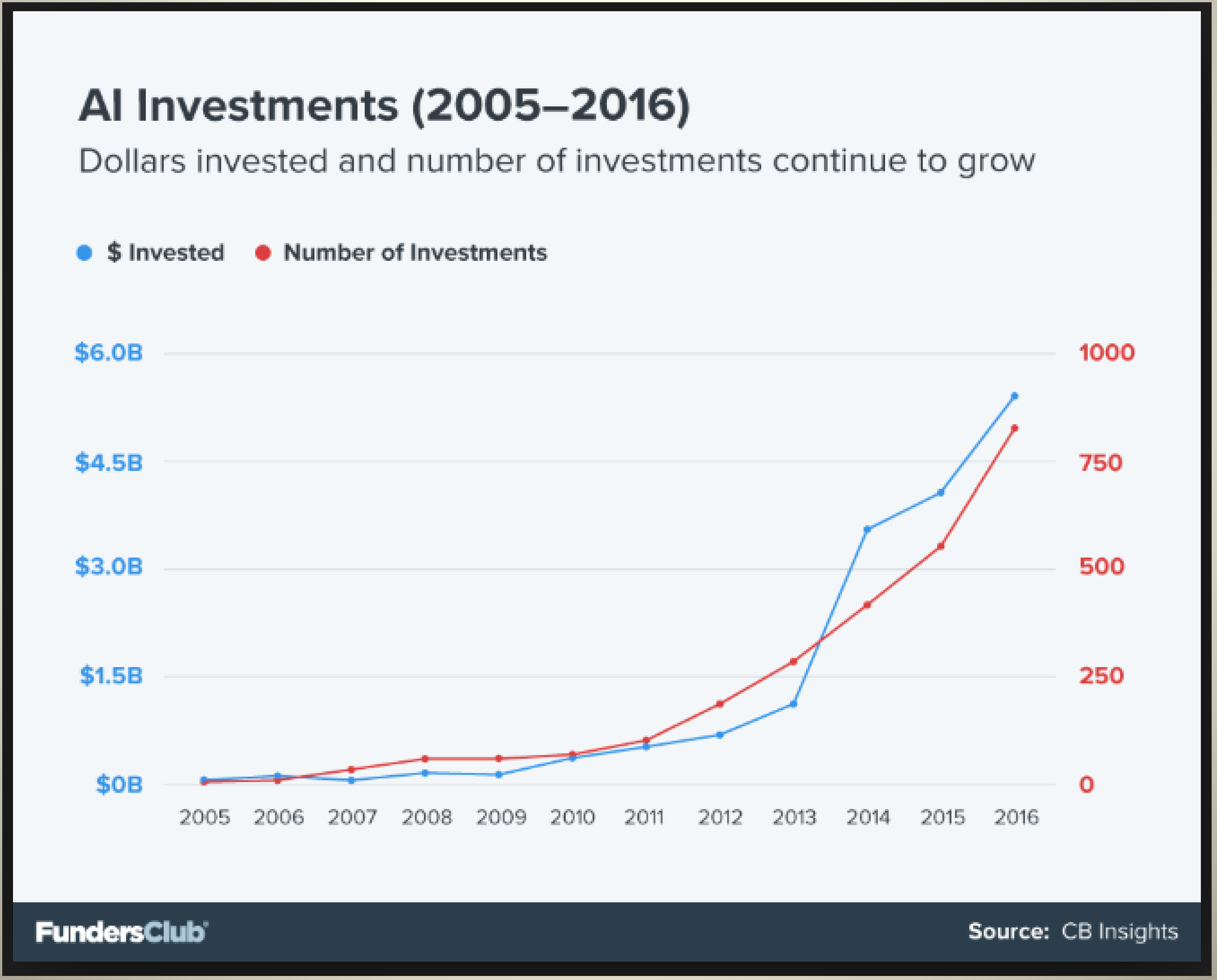
Venture Capital Funding of medical imaging AI companies over time



Number of Medical Imaging AI Companies Founded by Year



AI Investments over time



Venture Capital funding of AI in radiology

- 2017 - $270 million
- 2018 - $580 million

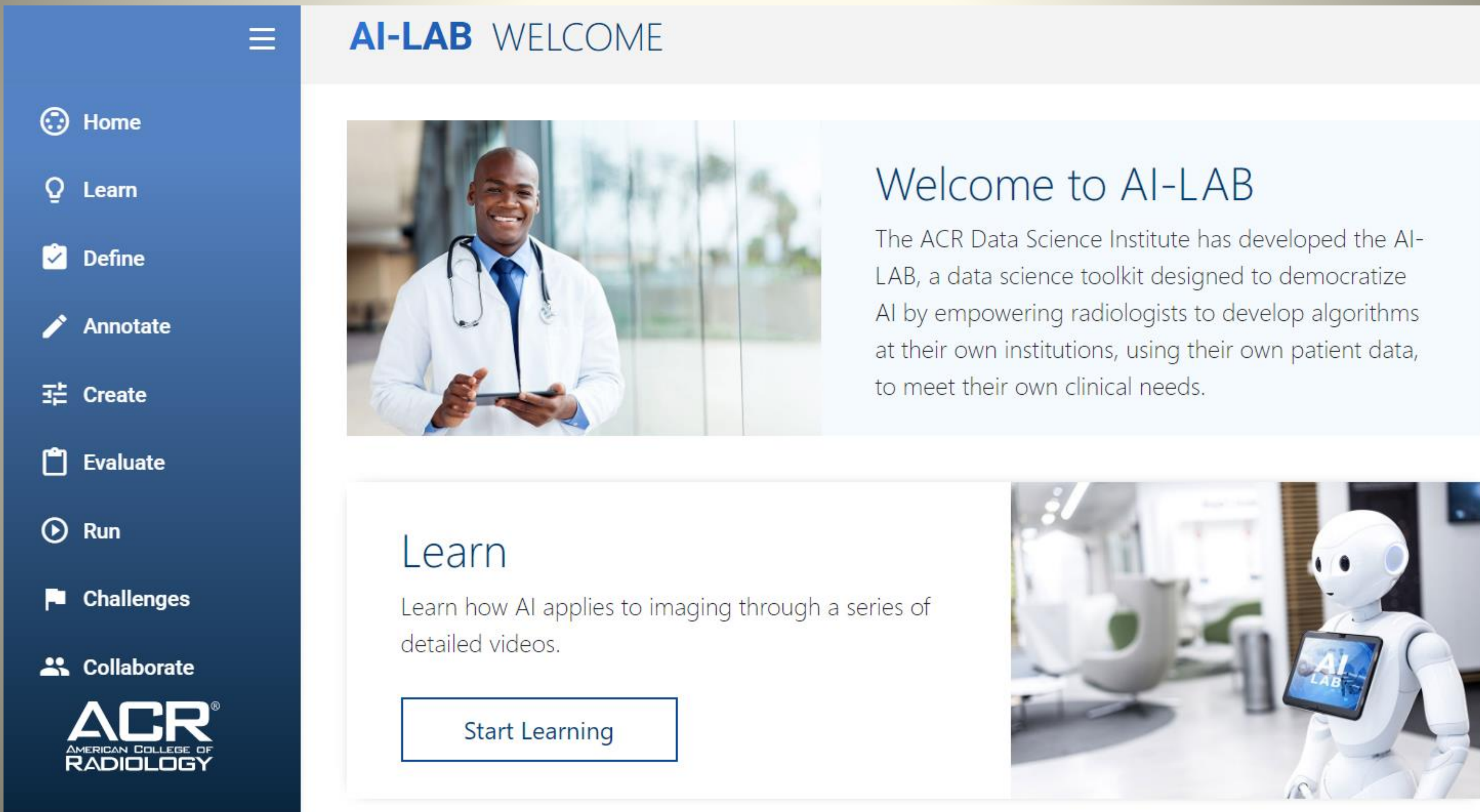
Commercial software for Radiology AI

Clara AI Lets Every Radiologist Teach Their Own AI

A system for radiologists to deliver AI-assisted annotation, adapt AI for their patients, and deploy it in the hospital.

March 18, 2019 by ABDUL HAMID HALABI

ACR AI Lab


AI-LAB WELCOME
Home
Learn
Define
Annotate
Create
Evaluate
Run
Challenges
Collaborate
ACR
AMERICAN COLLEGE OF RADIOLOGY
Welcome to AI-LAB
The ACR Data Science Institute has developed the AI-LAB, a data science toolkit designed to democratize AI by empowering radiologists to develop algorithms at their own institutions, using their own patient data, to meet their own clinical needs.
Learn
Learn how AI applies to imaging through a series of detailed videos.
Start Learning
AI LAB

Bone age from hand x-rays

RSNA®
Radiological Society
of North America

Search Competitions

Competition

Pediatric Bone Age Challenge

Organized by RSNA.organizing.committee - Current server time: June 19, 2019, 3:12 p.m.

▶ Current	Next
Test	Leaderboard
Oct. 7, 2017, midnight UTC	Sept. 1, 2017, midnight UTC

Detecting Pneumonia

RSNA®
Radiological Society of North America

Membership | Annual Meeting | Journals | Education | Research | Practice Tools

- Continuing medical education (CME)
- **AI resources and training**
  - **AI challenge**
    - RSNA Pneumonia Detection Challenge (2018)
    - RSNA Pediatric Bone Age Challenge (2017)
  - AI Webinars
- Professionalism and quality care
- Education awards
- Trainee resources
- Educator resources

RSNA Pneumonia Detection Challenge (2018)

As part of its efforts to help develop artificial intelligence (AI) tools for radiology, in 2018 RSNA organized an AI challenge to detect pneumonia, one of the leading causes of mortality worldwide.

About the challenge

We worked with colleagues at the Society for Thoracic Radiology and MD.ai to label pneumonia cases found in the database of chest x-rays made public by the National Institutes of Health (NIH).

Starting to see FDA Approvals

CISION PR Newswire

News Products Contact

FDA Clears the World's First AI Solution for Flagging Pulmonary Embolism

English

aidoc

Aidoc now leads the way in FDA approved AI solutions for radiologists

NEWS PROVIDED BY
Aidoc →
15 May, 2019, 16:00 IDT

SHARE THIS ARTICLE

Pneumothorax triage

HEALTH NEWS MAY 13, 2019 / 2:45 AM / A MONTH AGO

Israel's Zebra Medical gets FDA ok for AI chest X-ray product

1 MIN READ

TEL AVIV (Reuters) - Israel's Zebra Medical Vision said on Monday it received approval from the U.S. Food and Drug Administration for its artificial intelligence-based chest X-ray triage product.

The FDA approval focuses on an alert for urgent findings of pneumothorax,

Open Images Dataset

storage.googleapis.com/openimages/web/download.html

A dataset of ~9 million images that have been annotated with image-level labels and object bounding boxes

The images are very diverse and often contain complex scenes with several objects (8.4 per image on average) and the dataset is annotated with image-level labels spanning thousands of classes.

Youtube dataset

YouTube-8M Dataset

YouTube-8M is a large-scale labeled video dataset that consists of millions of YouTube video IDs, with high-quality machine-generated annotations from a diverse vocabulary of 3,800+ visual entities. It comes with precomputed audio-visual features from billions of frames and audio segments, designed to fit on a single hard disk. This makes it possible to train a strong baseline model on this dataset in less than a day on a single GPU! At the same time, the dataset's scale and diversity can enable deep exploration of complex audio-visual models that can take weeks to train even in a distributed fashion.

Our goal is to accelerate research on large-scale video understanding, representation learning, noisy data modeling, transfer learning, and domain adaptation approaches for video. More details about the dataset and initial experiments can be found in our technical report and in last year's workshop. Some statistics from the latest version of the dataset are included below.

6.1 Million Video IDs	**350,000** Hours of Video	**2.6 Billion** Audio/Visual Features	**3862** Classes	**3.0** Avg. Labels / Video

Patient data across hospitals in epic

Organizations on the Care Everywhere Network

Australia
- The Royal Children's Hospital

Canada
- Group Health Centre
- Mackenzie Health
- Ottawa Hospital

England
- Cambridge University Hospitals
- Great Ormond Street Hospital for Children NHS Foundation Trust
- University College London Hospitals

Lebanon
- American University of Beirut Medical Center

Netherlands
- Amphia Ziekenhuis
- Amsterdam UMC
- Elisabeth TweeSteden Ziekenhuis
- Medisch Centrum Leeuwarden

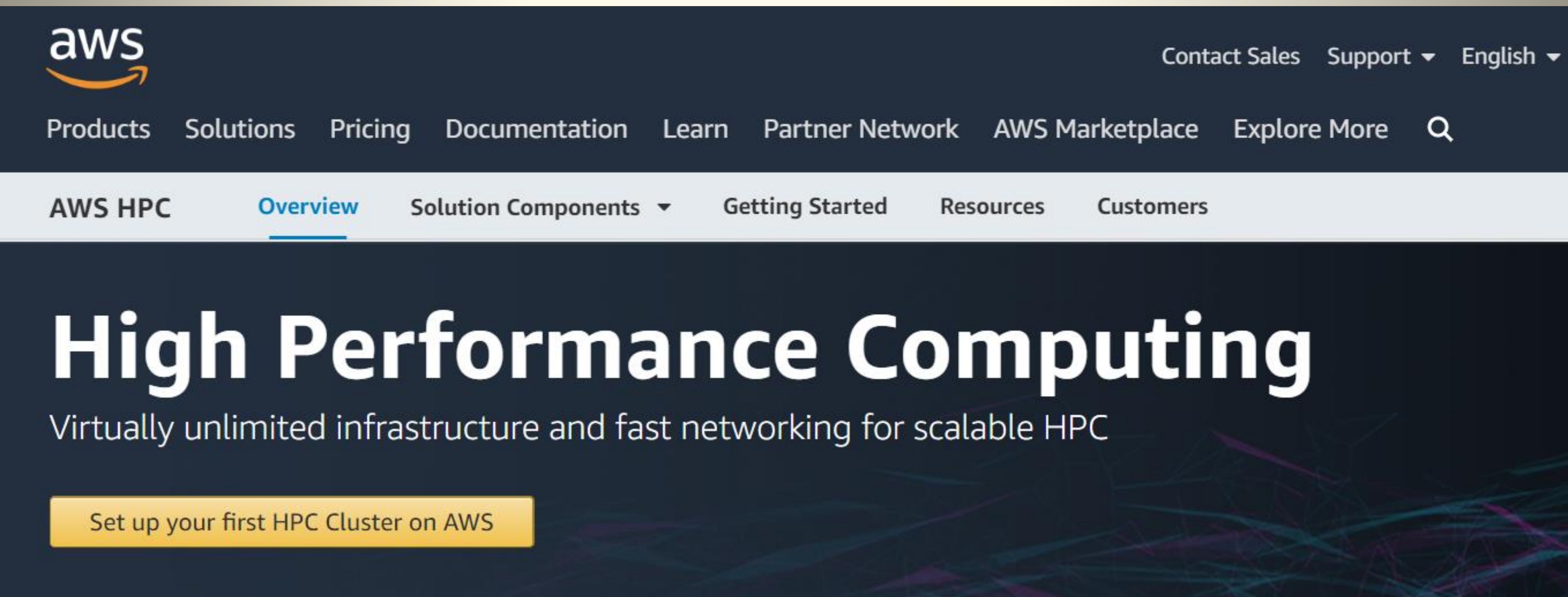
United Arab Emirates
- Cleveland Clinic Abu Dhabi
- Dubai Health Authority

Facebook Libra stablecoin

Problem Statement

The advent of the internet and mobile broadband has empowered billions of people globally to have access to the world's knowledge and information, high-fidelity communications, and a wide range of lower-cost, more convenient services. These services are now accessible using a $40 smartphone from almost anywhere in the world.[1] This connectivity has driven economic empowerment by enabling more people to access the financial ecosystem. Working together, technology companies and financial institutions have also found solutions to help increase economic empowerment around the world. Despite this progress, large swaths of the world's population are still left behind — 1.7 billion adults globally remain outside of the financial system with no access to a traditional bank, even though one billion have a mobile phone and nearly half a billion have internet access.[2]

HPC: inhouse vs. cloud



XSEDE national supercomputing center

PSC Bridges GPU-AI (Bridges GPU Artificial Intelligence)

PSC

Production Dates: 2019-01-01 - 2020-11-30

Startup Allocation Limit: 1500 GPU Hours (GPU-AI Nodes: 1 GPU-hour = 1 SU)

Description: The Bridges GPU-AI resource is integrated with Bridges and allocated as a distinct XSEDE resource, analogous to Bridges GPU, RM, LM, and Pylon. Bridges GPU-AI contains an NVIDIA DGX-2 enterprise research AI system, which tightly couples 16 NVIDIA Tesla V100 (Volta) GPUs each with 32 GB of GPU memory and connected by NVLink and NVSwitch, and 9 HPE Apollo 6500 servers, each with 8 NVIDIA Tesla V100 GPUs each with 16 GB of GPU memory connected by NVLink 2.0. Together, the DGX-2 and Apollo 6500 nodes provide maximum capability the most demanding of AI challenges and great capacity. Bridges-AI builds on Bridges' strength in converged HPC, AI, and Big Data to provide our nation's research community with an extraordinary platform for AI and AI-enabled simulation.

Fast.ai

Now anyone can train Imagenet in 18 minutes

Written: 10 Aug 2018 by *Jeremy Howard*

This post extends the work described in a previous post, Training Imagenet in 3 hours for $25; and CIFAR10 for $0.26.

A team of fast.ai alum Andrew Shaw, DIU researcher Yaroslav Bulatov, and I have managed to train Imagenet to 93% accuracy in just 18 minutes, using 16 public AWS cloud instances, each with 8 NVIDIA V100 GPUs, running the fastai and PyTorch libraries. This is a new speed record for training Imagenet to this accuracy on publicly available infrastructure, and is 40% faster than Google's DAWNBench record on their proprietary TPU Pod cluster. Our

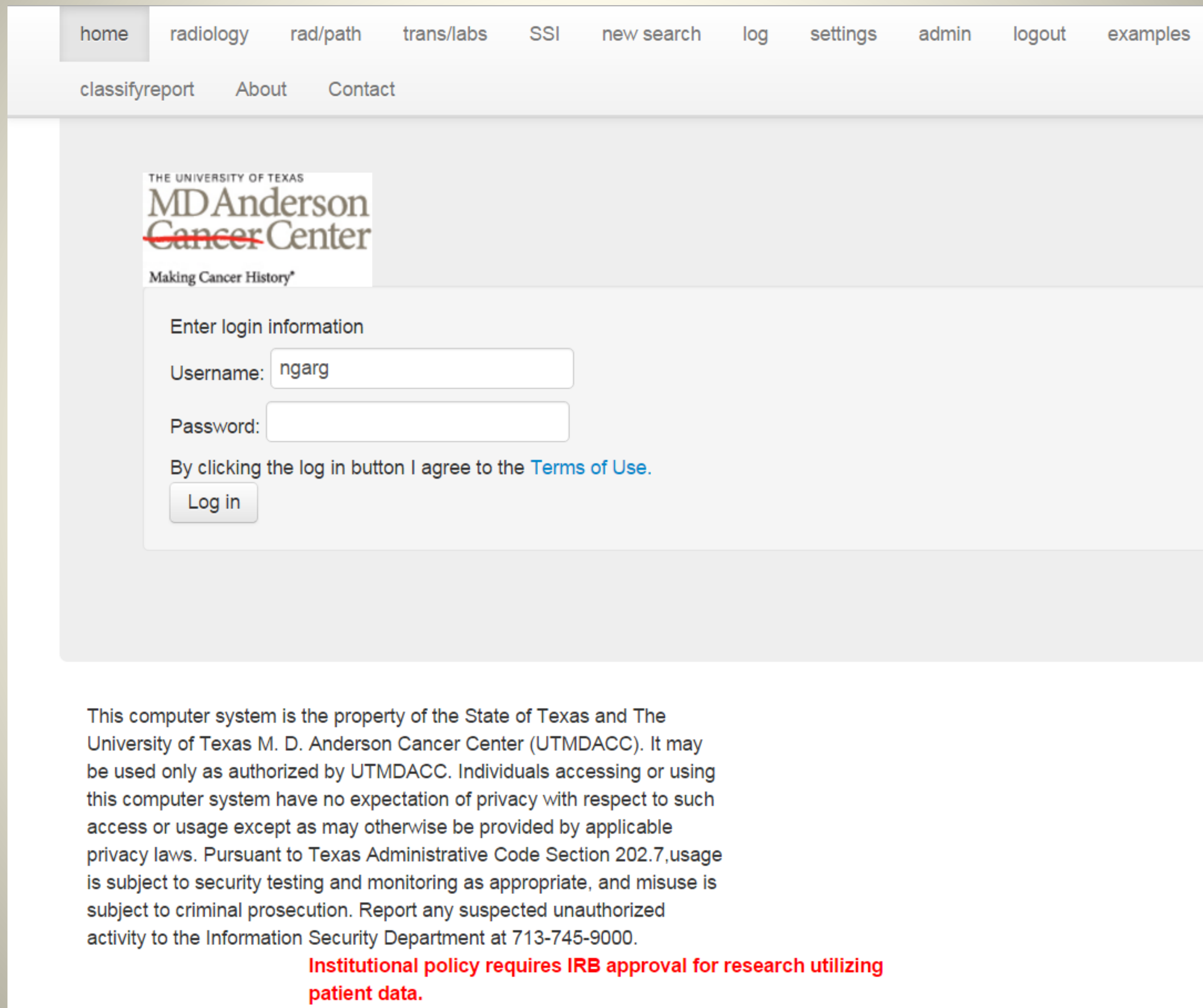
Milestones in Artificial Intelligence

- Computers beat humans at chess (1997)
  - IBM Deep Blue vs. Kasparov
- Computers beat humans at jeopardy (2011)
  - IBM Watson vs. Rutter and Jennings
- Computers beat humans at go ! (2016)
  - Google DeepMind's AlphaGo (version: Lee)[56] defeated Lee Sedol 4–1.
- AlphaZero beats StockFish at chess (2017)

Milestones in AI datasets

- MNIST: 70,000 handwritten digits (1998)
- Imagenet: 14 million hand annotated images (2009)
- DeepLesion: 40,000 recist tumor measurements (2018)
- MIMIC-CXR: 371,000 chest xrays (2019)

Garglab Search Tools



Data

Source	Since	millions of reports	gigabytes
Radiology	2000	6.3	9
Pathology	2000	7.8	7
Transcribed Documents	1993	16.4	121
Labs	2000	269	90

Data warehouse

- Electronic health records database
  - Radiology reports
  - Pathology reports
  - H/P’s, op notes, discharge summaries etc.
  - Labs
- Dicom
  - headers
  - presentation states / annotations
  - Images

Example Data lake for Radiology AI

– 1.5 million patients
– 10 million radiology reports
– 2 million presentation states
– 5 million annotated key images

Server

Server	virtual processors	RAM	Terabytes of storage	video ram	cuda cores
Garglab	32	768 gb	10	12 gb	2500

Features

- Cohort Identification
- Natural Language processing
- Sentence analysis
- Numeric range search
- Lesions size
- site of radiation
- relative time points

Rad Path Cross Search

home | radiology | rad/path | trans/labs | new search | log | settings | admin | logout | locations | examples | classifyreport | Abou

filtercache:

database: rad, path

radiology

radiologist:

diagnosis:

procedure:

modality: CT

start date: (yyyy-mm-dd)

end date: (yyyy-mm-dd)

query:

query2:

distance: (between query 1 and 2)

radpathdistance: 7 days between radiology and pathology matches

vectortype: mrns

list of medical record numbers

pathology

pathologist:

start date: (yyyy-mm-dd)

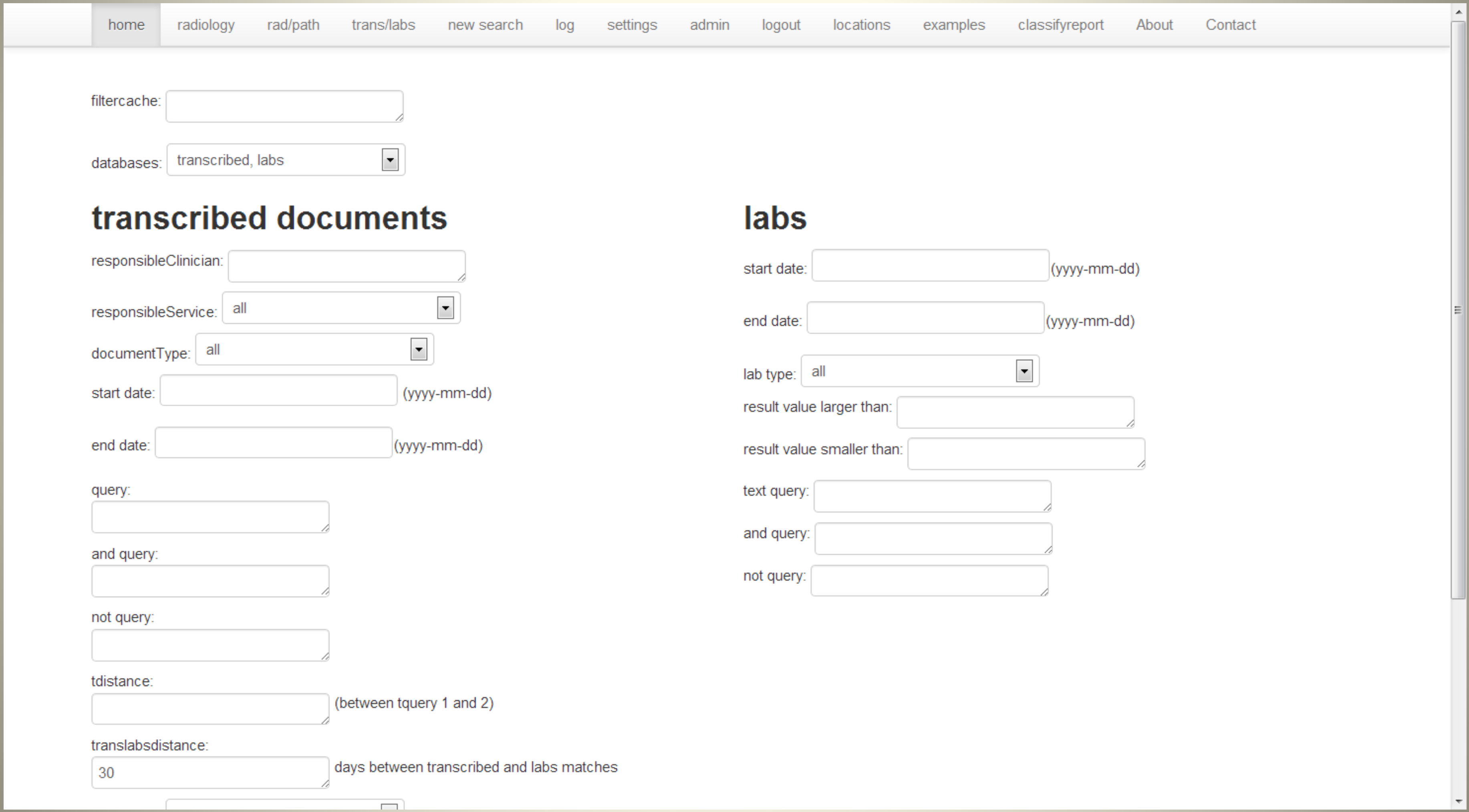
end date: (yyyy-mm-dd)

pquery:

pquery2:

pdistance: (between pquery 1 and 2)

Clinical Documents and Labs



SUV values on PET

```
if instr(r.modality, "pet")
if instr(r.diagnosis, "leukemia")
if regexmatch(r.report, "SUV\D+([\d.]+)", suv)
if (suvl > 10)
return {diagnosis: r.diagnosis, suv: suvl}
```

Garglet

filtername: diprodfilter1345121111306

previous page next page anon xml accessions mrns stats

diagnosis	suv
LYMPHOBLASTIC LYMPHOMA/LEUKEMIA,	10.6,
MDS/ LEUKEMIA,	18,
T CELL LYMPHOBLASTIC LEUKEMIA/LYMPHOMA,	12.1,
B-CELL ACUTE LYMPHOBLASTIC LEUKEMIA,	18,
CHRONIC LYMPHOCYTIC LEUKEMIA,	18,
MULT MYELOMA/ PLASMA CELL LEUKEMIA,	10.3,
CHRONIC LYMPHOCYTIC LEUKEMIA,	18,
RELAPSED REFEACTORY CHROMIC LYMPHOCYTIC LEUKEMIA CLL,	7.4.,
CLL/ CHRONIC LYMPHBLASTIC LEUKEMIA,	18,
BURKITTS LEUKEMIA,	8.9.,
NEW LEUKEMIA/LYMPHOMA/ALL/T-CELL,	4.5.,
LYMPHOMA/ NEW LEUKEMIA SUSPECT MUD,	18,
BURKITT LEUKEMIA,	15.9,
T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA,	18,
MDS (LEUKEMIA),	18,

tablename: pathology reports

filterscript:

```
r := report
if !instr(r.report, "prostate")
return 0
pos := regexmatch(r, "(\d+)(\.\d)?\s*[xX]\s*(\d+)(\.\d)?\s*[xX]\s*(\d+)(\.\d)?\s*[cC][Mm]", m)
if !pos
return 0
if m2
d1 := m1 + m2
else
d1 := m1
if m4
d2 := m3 + m4
else
d2 := m3
if m6
d3 := m5 + m6
else
d3 := m5
v := d1 * d2 * d3
return {d1: d1, d2: d2, d3: d3,
```

filtername: pathfilter1345121424291

previous page next page anon xml accessions mrns stats

d1	d2	d3	m	v
6.5,	3,	1,	6.5 x 3 x 1 cm,	19.500000 cm3,
0.6,	0.3,	0.3,	0.6 x 0.3 x 0.3 cm,	0.054000 cm3,
7.5,	6,	3.5,	7.5 x 6 x 3.5 cm,	157.500000 cm3,
3.5,	2.3,	0.8,	3.5 x 2.3 x 0.8 cm,	6.440000 cm3,
0.5,	0.4,	0.2,	0.5 x 0.4 x 0.2 cm,	0.040000 cm3,
1.2,	0.4,	0.3,	1.2 x 0.4 x 0.3 cm,	0.144000 cm3,
0.5,	0.5,	0.2,	0.5 x 0.5 x 0.2 cm,	0.050000 cm3,
10,	7.5,	5,	10 x 7.5 x 5 cm,	375.000000 cm3,
0.7,	0.6,	0.1,	0.7 x 0.6 x 0.1 cm,	0.042000 cm3,
4.8,	2.5,	3.2,	4.8 x 2.5 x 3.2 cm,	38.400000 cm3,
5.5,	4.5,	2.2,	5.5 x 4.5 x 2.2 cm,	54.450000 cm3,
1.5,	0.1,	0.1,	1.5 x 0.1 x 0.1 cm,	0.015000 cm3,
1.5,	0.5,	0.4,	1.5 x 0.5 x 0.4 cm,	0.300000 cm3,
4,	2.5,	0.7,	4.0 x 2.5 x 0.7 cm,	7.000000 cm3,
0.6,	0.4,	0.2,	0.6 x 0.4 x 0.2 cm,	0.048000 cm3,
5,	4,	2,	5.0 x 4.0 x 2.0 cm,	40 cm3,
7,	3,	1,	7 x 3 x 1 cm,	21 cm3,

Sample Queries

- melanoma
  - Path proven
  - treated with ipilimumab
  - had an FDG PET-CT.
- Liver masses
  - Ct abdomen with contrast
  - liver lesions 1 to 2 cm in size
- breast cancer
  - Recurrent
  - Invasive ductal
  - Triple negative
  - Family history of colon cancer
- Cholangiocarcinoma
  - Had radiation to the liver
  - Mri before and after radiation

GargLab EMR Search Request Form

http://garglab.com/projects/data-mining/

Requestor Name : ____________________ **Department:** ___________________________

Date of Request: ______________________ **Date Needed By:** ________________________

Purpose of information requested: IRB approved Protocol / Quality Improvement

/ Review preparatory to research / Educational /Other:______________________________

IRB Protocol #: _____________ _____________ (please also attach protocol summary)

QIAB project # _____________ _____________ (please also attach protocol summary)

Funding Source: ____

Patient consents : have IRB waiver of consent / have patient consent / QIAB / NA

Collaborators: from data source custodian department (optional, but recommended)

Radiologist: __________________________**Pathologist:** ________________________________

Clinician: ____

Type of Data: Number of matches only / anonymized data (with pseudo-mrns) / data with PHI (mrn)

Expected number of cases: ____

Inclusion Criterias:____

____

Exclusion Criterias: ____

____

Date Range Begin: ________________________ **Date Range End:** ______________________________

Keywords: ____

Existing Cohort: If you have an existing cohort (list of mrns) that you would like to filter, please attach.

Summary of search request: ____

I affirm that this Electronic Medical Record search is being requested in compliance with garglab terms of use available at: https://garglab:8090/TermsOfUse.html as well as institutional and statutory obligations, including 45 CFR 46 (IRB), HIPAA, and the policies of Office of Protocol Research.

Name of PI (print)	**Signature of PI**	**Date**

Submission Instructions:

Please send an email in the following format:

FROM: Principal Investigor's md anderson email account (and or sign form above and scan to pdf)

TO: 4info (4info@mdanderson.org)

CC: ngarg@mdanderson.org, Manuel.Guillen@mdanderson.org

CC: IAICustomerService@mdanderson.org (IAI Customer Service)

Attachment: (this form completed)

Subject: GargLab Search Request