The Gift From Yesterday

Matthew Lungren, MD MPH
Assistant Professor, Radiology, Stanford University School of Medicine
COMPUTER SCIENTIST TASKS ARE SIMILAR TO DOCTOR TASKS
a man on a skateboard. man riding a bicycle. orange cone on the ground. man riding a bicycle. two people riding a skateboard. red helmet on the man. skateboard on the ground. white shirt with red and white stripes. orange and white cone. trees are behind the people.

IMPRESSION:
Status post left upper lobectomy for lung cancer.

Multiple pulmonary nodules some which are solid and some of which are groundglass with little significant change since March 2013. However as compared with May 2012 the lingular opacification has definitely increased in size. There is also a new groundglass focal opacity measuring about a centimeter in the right upper lobe. The findings are concerning for recurrent adenocarcinoma.
“No acute airspace disease”

“No acute intracranial abnormality”

“Sagittal knee MRI with acute ACL tear”

“Heterogenously dense. BI-RADS 2”

“PET CT metabolically active GE junction mass”

“Mild cerebellar atrophy”
OPEN SOURCE DATA SETS ARE THE PRIMARY CATALYST FOR MACHINE LEARNING ADVANCEMENTS
Negascout Planning Algorithm Developed

1983

14 years

The Extended Book of Chess Games Dataset Released

1991

6 years

Deep Blue Defeats Kasparov

1997
Who is Stoker? ($1,000)

Who is Bram Stoker? ($17,973)

Who is Bram Stoker? ($5,600)
Mixture of Experts Algorithm Developed

8.6 Million documents from Wikipedia and Project Gutenberg Released to Public

Watson becomes world Jeopardy Champion
Convolutional Neural Network Created

1.5 Million classified images with 1,000 object categories released to public

Microsoft Research surpasses human recognition performance
MEDICAL DATA ARE A PUBLIC GOOD AND SHOULD BE OPEN SOURCED TO BENEFIT ALL
Big Data Market Size Revenue Forecast Worldwide From 2011 To 2027 (in billion U.S. dollars)
DATA are not Traditional Form of Property

- Not divided or consumed
- Easily replicated

"Ownership" of data is an imprecise concept:
- Rights to control access
- Rights to share of profit

Value of data and information is relative:
- Exclusive access to data may give relative advantage
- Advantage negated when others have access to the same data

Credit: David Larson MD MBA
People With Cancer By State Email/Postal/Phone Mailing List

The People With Cancer By State mailing list consists of individuals who suffer from different types of cancer, and would be great prospects for health products, treatments, medications, and much more. These active, opt-in consumers yield proven results for a wide variety of direct marketing campaigns. This list is the ideal tool to effectively build relationships with your unique audience.

<table>
<thead>
<tr>
<th>SEGMENTS</th>
<th>COUNTS THROUGH 09/06/2016</th>
<th>MARKET:</th>
<th>CHANNELS:</th>
</tr>
</thead>
<tbody>
<tr>
<td>562,694 TOTAL UNIVERSE / BASE RATE</td>
<td>$220.00/M</td>
<td>CONSUMER</td>
<td></td>
</tr>
<tr>
<td>562,694 POSTAL</td>
<td>+ $220.00/M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47,548 PHONE</td>
<td>+ $130.00/M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>279,096 FACEBOOK</td>
<td>$249.00/F</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DESCRIPTION

About Exact Data's Consumer Database:
but... Medical Data are PEOPLE
Our position:

Once clinical data have been used to provide care, the primary purpose for acquiring the data is fulfilled.

For secondary use:

Clinical data should be treated as a form of public good used for the benefit of all thorough open source research and education **NOT for sale or exclusive partner access**
Benchmark Performance with Consensus Expert Ground Truth

7,200+ teams worldwide competing in medical imaging challenges hosted at Stanford
Publications on AI in Radiology

Pesapane F. et al. Eur Radiol Exp. 2018
Process for Open Source Medical Imaging Data

1. Curate and label data
2. Manual PHI review of all images
3. Data user agreement and registration
ALTERNATIVE STRATEGIES TO OPEN SOURCE MEDICAL IMAGING DATA ARE COMPLIMENTARY NOT REPLACEMENTS
Detecting and Simulating Artifacts in GAN Fake Images (Extended Version) Xu Zhang, Svebor Karaman, and Shih-Fu Chang
1. Sending encrypted gradients
2. Secure aggregation
3. Sending back model updates
4. Updating models
Responsible ethically sound processes exist to make medical data available.

Open source medical data and benchmark competitions will accelerate innovation.

Publishing without sharing data is just marketing NOT science.

Alternative approaches to solve data availability are complimentary.
Thank You

Matthew P Lungren MD MPH
Assistant Professor of Radiology
Associate Director
Center for Artificial Intelligence in Medicine & Imaging
Stanford University School of Medicine
mlungren@stanford.edu

@mattlungrenMD

#FDAPDSsymp | #AlinHealth