2023 Spring Provider Symposium

Artificial Intelligence in Healthcare: Improving Health Outcomes and Advancing Equity with New Tools

Friday, May 19, 2023

Virtual Conference

healthfirst®
Health Insurance for New Yorkers

AI in Medical Imaging – From Image Quantitation to the Monitoring of Disease

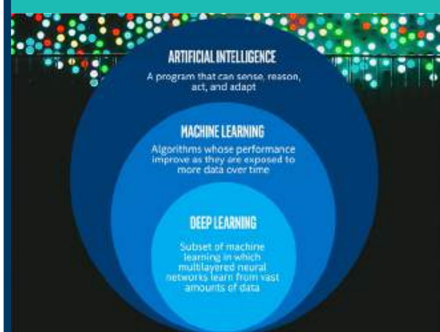
Hayit Greenspan, PhD.

Professor in the Department of Diagnostic, Molecular, and Interventional Radiology & Director of the AI Core at the BioMedical Engineering and Imaging Institute (BMEII)
Icahn School of Medicine at Mount Sinai

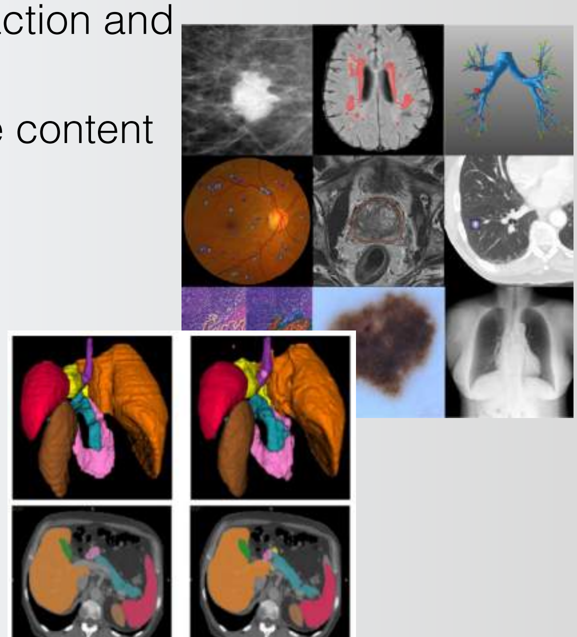
Affiliated with the AI and Human Health Dept at MS

Co-Director of AIET PhD Program at MS

AI in Medical Imaging



- Automate the extraction and analysis
- Quantify the Image content
  - Detection
  - Segmentation
  - Quantification
  - Categorization



The Data & Modelling Challenges

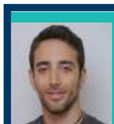
Data Challenge

- Collecting medical images is difficult; done by experts
- Labeling of medical images is difficult; done by experts
- Manual localization of lesions is challenging
- Ownership & privacy issues
- Very limited national-level open datasets

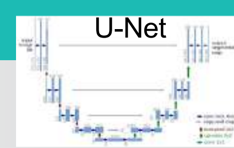
Modeling Challenge

- Robustness
  - Critical forgetting
  - Domain shift – Site adaptation
- Generalization
- Explainability
- Fuse across domains (Image+Clinical+Omics)

Our Challenge is to train networks with limited amount of data, and to build strong and robust systems



Transfer Learning: New pre-training using varying data & tasks



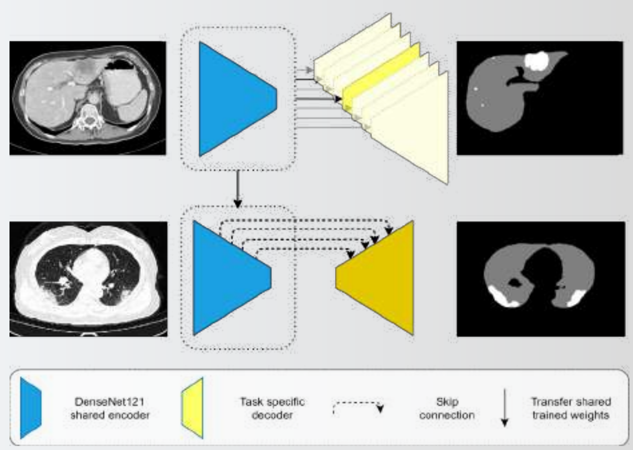
- Most common TL: ImageNet pretrained networks to initialize networks for medical tasks.
- The HydraNet architecture:

  Pretraining on several additional medical data sources and tasks: 7 segmentation datasets

  U-net architecture with 7 decoders for joint training

  Fine-tuning on a target task with single decoding head.

Nº	Dataset	Modality
1	Source Dataset - LiTS[12]	CT
2	Source Dataset - Pancreas[15]	CT
3	Source Dataset - Hepatic Vessel[15]	CT
4	Source Dataset - Spleen[15]	CT
5	Source Dataset - Prostate[15]	MRI
6	Source Dataset - BraTS[13]	MRI
7	Source Dataset - Left Atrial[14]	MRI
8	Target Dataset - MedSeg COVID-19[16]	CT



HydraNet - A Joint Learning Multi-Decoder Network

N. Sagie, S. Almog, A. Talby, and H. Greenspan, "**COVID-19 Opacity Segmentation in Chest CT via HydraNet - A Joint Learning Multi-Decoder Network**," Medical Imaging 2021: Computer-Aided Diagnosis. SPIE, 2021, vol. 11597.

HydraNet - Results

Data split, 20 non-contrast CT scans (3,071 slices):

Training: 16 scans (2296 slices)
Validation: 2 scans (452 slices)
Test: 2 scans (323 slices)

} 3-fold cross-validation

Dice: $\frac{2\cdot|TP|}{2\cdot|TP|+|FP|+|FN|}$

Precision: $\frac{|TP|}{|TP|+|FP|}$

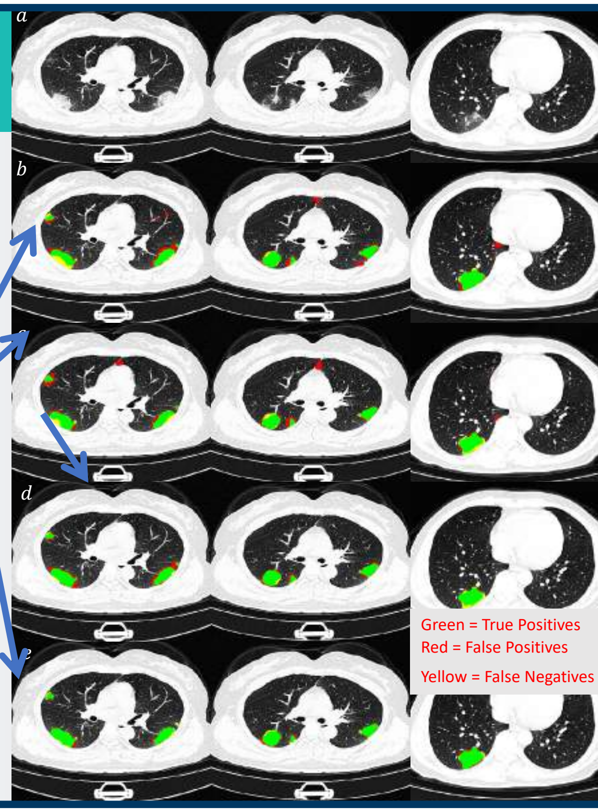
Sensitivity: $\frac{|TP|}{|TP|+|FN|}$

Experiment	Dice	Sensitivity	Precision
Random	0.65 ± 0.03	0.72 ± 0.14	0.64 ± 0.15
ImageNet	0.69 ± 0.06	0.72 ± 0.09	0.67 ± 0.07
HydraNet	0.71 ± 0.02	**0.75 ± 0.14**	0.72 ± 0.11
ImageNet+HydraNet	**0.73 ± 0.12**	0.68 ± 0.18	**0.81 ± 0.02**

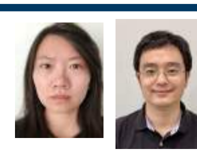
11% increase (P-value<0.0001)

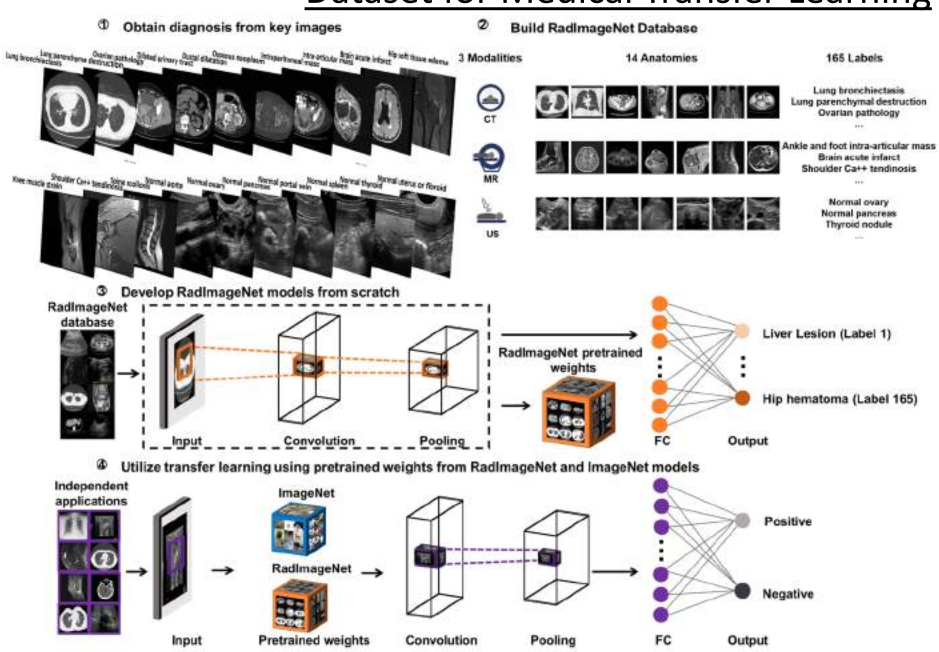
Transfer with Regularization – modify loss function to prevent overfitting:

Experiment	Dice	Sensitivity	Precision
MedSeg Dataset			
1. From scratch	0.65	0.726	0.645
2. ImageNet initialization	0.687	0.719	0.673
3. HydraNet from scratch	0.718	0.759	0.807
4. HydraNet w/ ImageNet	**0.764**	**0.789**	**0.825**

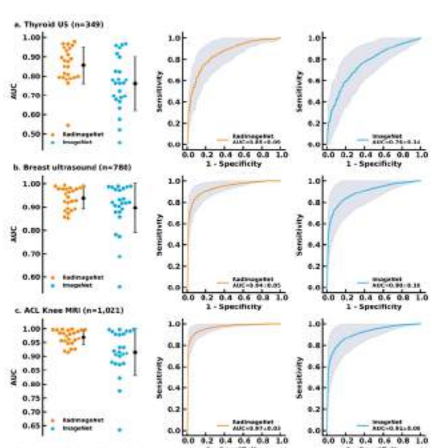


RadImageNet: An Open Radiological Deep Learning Research Dataset for Medical Transfer Learning





The RadImageNet Database: 1.35 million annotated radiologic images; Data request:radimagenet.com

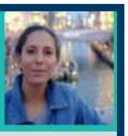


Xueyan Mei, Zelong Liu, Philip Robson, Brett Marinelli, Mingqian Huang, Amish Doshi, Adam Jacobi, Katherine Link, Thomas Yang, Chendi Cao, Ying Wang, Hayit Greenspan, Timothy Deyer, Zahi Fayad, Yang Yang

(Mei et al.) Radiology: Artificial Intelligence 2022

Learning using Very Weak Supervision: RV Strain Classification from 3D CTPA

Training with Single Label taken directly from the Report



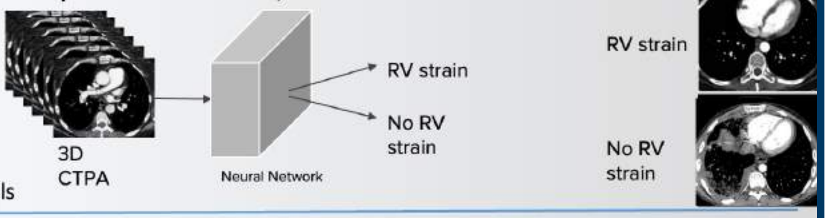
Pulmonary embolism (PE) is a life-threatening condition and early PE risk stratification can expedite treatment. *High-risk PE is caused by RV dysfunction* from acute pressure overload, which can be detected via CTPA.

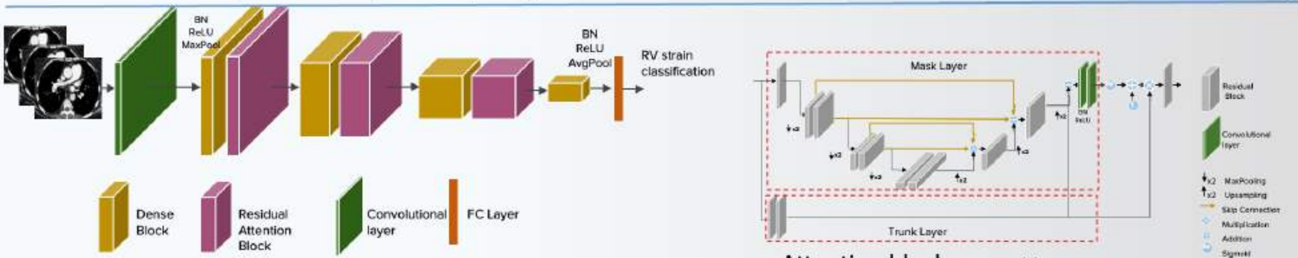
Goal:

Input=3D CTPA volume;
Output=Categorize yes/no RV strain;

Challenges:

Limited data: 363 CT scans;
Small number of positive examples: ~24%; Noisy labels





3D classification network (DenseNet3D-121)
With self attention blocks

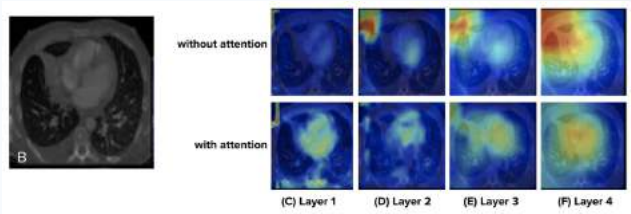
Attention blocks: resemble a U-net; provide weights for extracted features $A(x) = T(x) \cdot (1 + M(x))$

N. Cahan, E. M. Marom, S. Soffer, Y. Barash, E. Konen, E. Klang, H. Greenspan. "Weakly Supervised Attention Model for RV Strain Classification from volumetric CTPA Scans," Computer Methods and Programs in Biomedicine, 2022

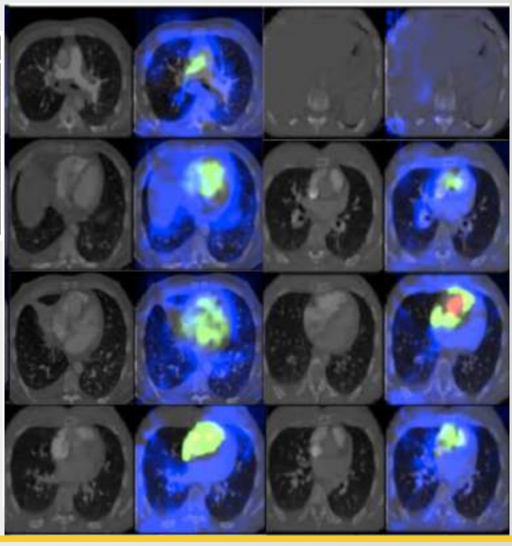
RV Strain Classification from 3D CTPA using Weak Supervision

Comparison with state-of-the-art 3D CNN models:
7% gain in AUC

Model Architecture	AUC	Accuracy	Specificity	Sensitivity	PPV	NPV	P-Value
ResNet3D-50 (basic block)	0.765 [0.66-0.87]	0.722	0.694	0.783	0.545	0.872	$p < 0.005$
ResNet3D-50 (residual block)	0.809 [0.7-0.89]	0.778	0.837	0.652	0.652	0.837	$p < 0.05$
ResNeXt3D-101	0.713 [0.61-0.81]	0.681	0.673	0.696	0.5	0.825	$p < 0.001$
DenseNet3D-121	0.805 [0.7-0.89]	0.792	0.796	0.783	0.643	0.886	$p < 0.05$
Models Genesis **	0.796 [0.69-0.89]	0.681	0.592	0.87	0.5	0.906	$p < 0.03$
Med3D *	0.801 [0.7-0.89]	0.75	0.714	0.826	0.576	0.897	$p < 0.05$
SANet	**0.877 [0.81-0.95]**	**0.847**	**0.837**	**0.87**	**0.714**	**0.932**	–



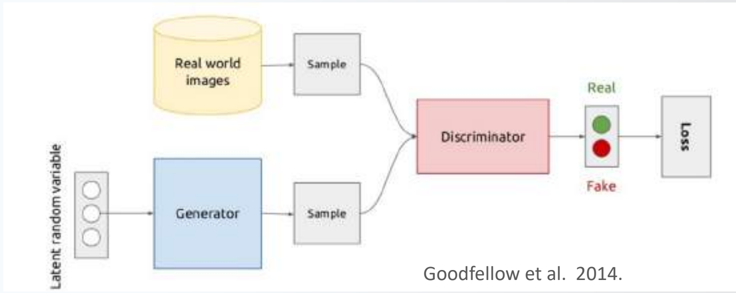
Heatmap visualization results



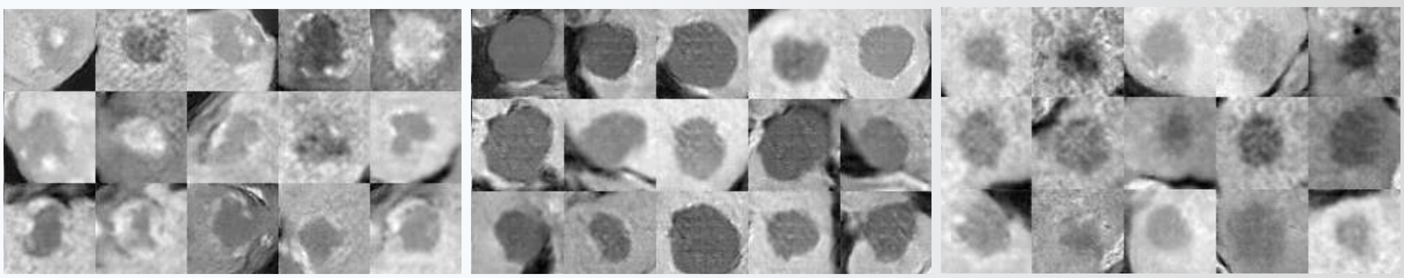
* Chen, S., Ma, K., and Zheng, Y., "**Med3d:**Transfer learning for 3d medical image analysis,"CoRRabs/1904.00625(2019).
** Zhou, Z., et al., "**Models genesis**: Generic autodidactic models for 3d medical image analysis," (2019).

Data Augmentation: Synthetic Lesions!

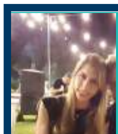
Generative Adverserial Networks (GANs)



Goodfellow et al. 2014.



Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H.: **Synthetic Data Augmentation using GAN for Improved Liver Lesion Classification**
IEEE-ISBI (2018);*Neurocomputing (2018)*

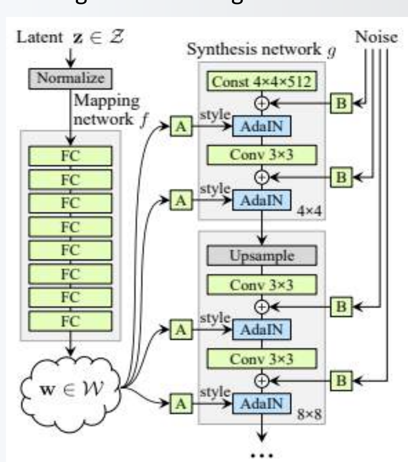


Data Augmentation using StyleGAN

StyleGAN (via Nvidia) proposes a unique generator architecture that creates high quality and high-resolution images: Maps the input latent vector into an intermediate disentangled latent space which we can use to manipulate the generated images.

Generated Faces

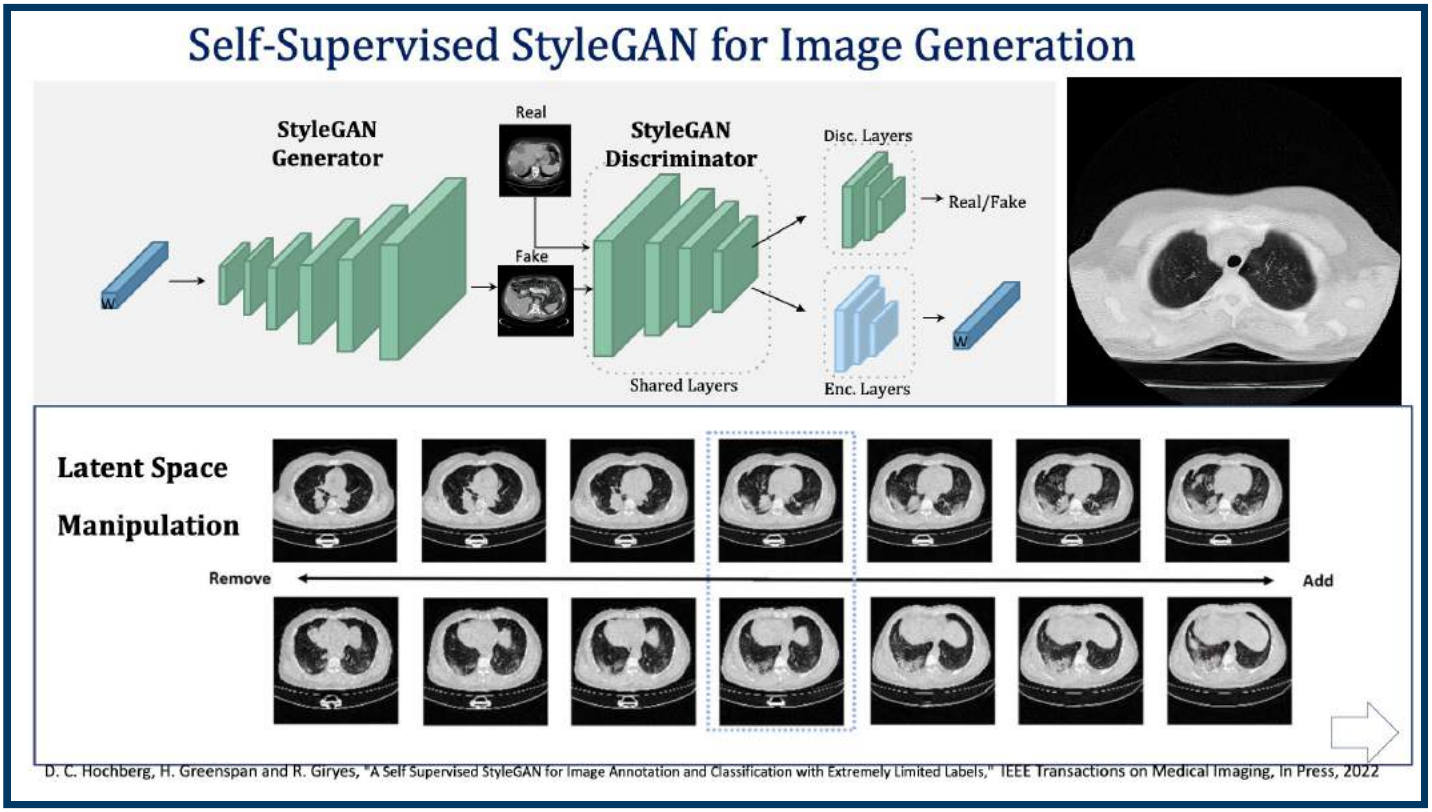
Karras et al. CVPR, 2019



Generated CT slices

Fine-tuning a pre-trained StyleGAN: Generate high-resolution realistic-looking CT images, using 11 CT scans from LiTS dataset

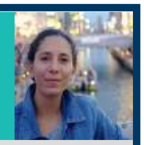
D. Cohen, R. Giryes, H. Greenspan, Style Encoding for Class-Specific Image Generation, SPIE, 2021
D. Cohen, R. Giryes, H. Greenspan, Encoding in Style Class-Specific Generation and Translation, MICCAI LABELS workshop, 2020


Self-Supervised StyleGAN for Image Generation
StyleGAN Generator
Real
Fake
StyleGAN Discriminator
Disc. Layers
Real/Fake
Shared Layers
Enc. Layers
Latent Space Manipulation
Remove
Add
D. C. Hochberg, H. Greenspan and R. Giryes, "A Self Supervised StyleGAN for Image Annotation and Classification with Extremely Limited Labels," IEEE Transactions on Medical Imaging, In Press, 2022


RadImageGAN: Synthetic Images
radimagegan_4k
12

From the Image to
the Patient
&
Disease Monitoring

Weakly Supervised Multimodal 30-day Mortality Prediction for Patients who were Diagnosed with Pulmonary Embolism

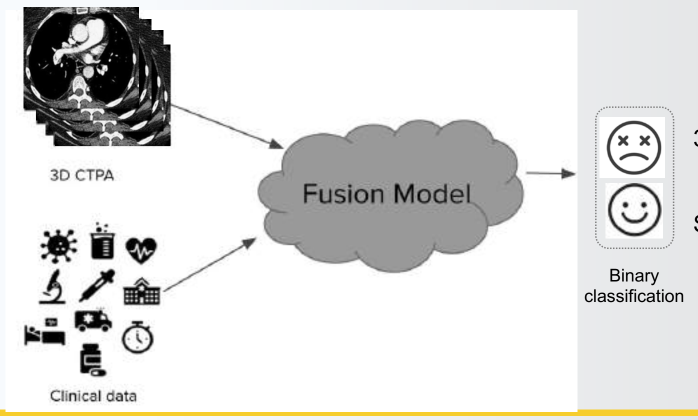


Goal

Develop a fully automated solution for 30-day mortality predictive model using both imaging data and clinical data from PE diagnosed patients

Challenges

- Limited data.
- Noisy labels with no annotations.
- Small number of positive examples
- Fusion is difficult!



N. Cahan, Y. Barash, S. Soffer, E. M. Marom, E. Konen and E. Klang and H. Greenspan, IEEE-ISBI 2022, Scientific Reports 2023

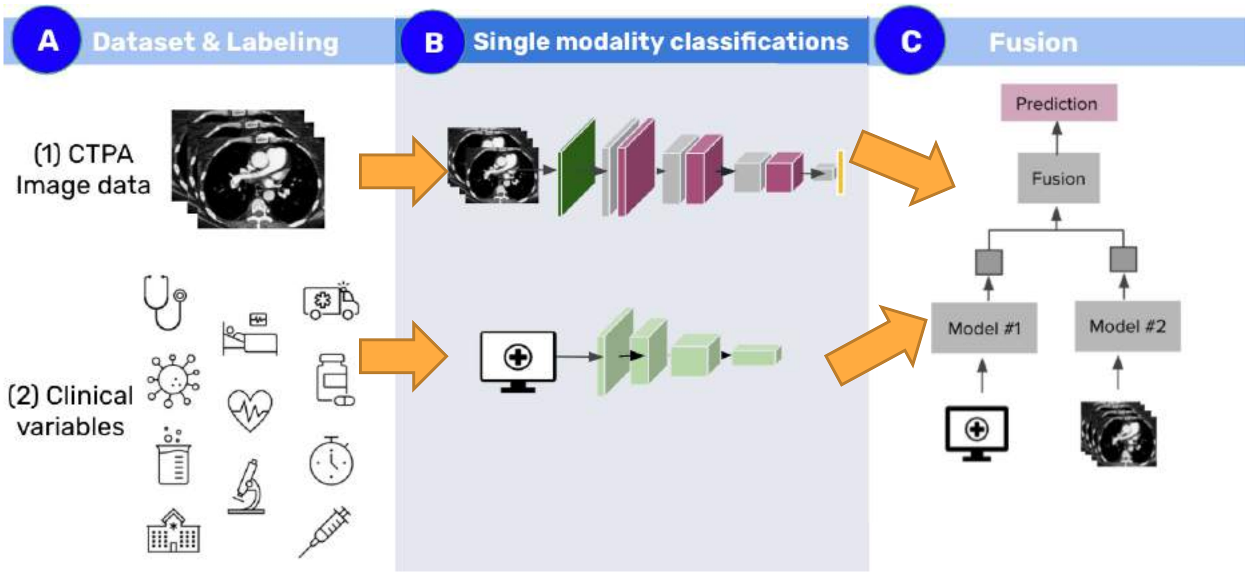
PE severity assessment

The key to an effective treatment of PE in the acute phase lies in the assessment of the patient's <u>early death risk</u>.

Existing prognostic tools include:

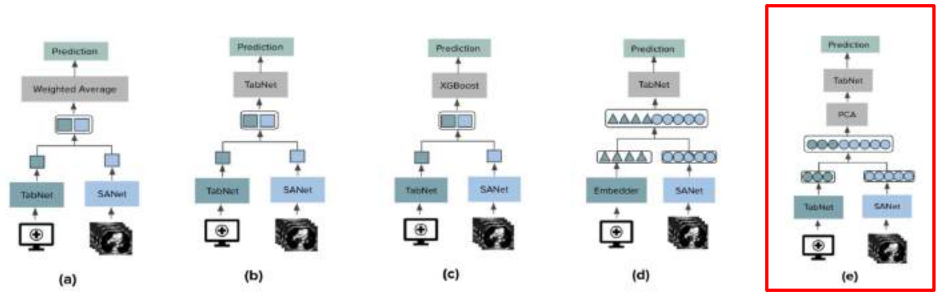
- Imaging:
  - CTPA
  - Echo
- Clinical predictive rules:
  - PESI
  - sPESI
- Laboratory markers:
  - Cardiac troponin I,T
  - Natriuretic peptides: BNP, NT-proBNP

Workflow

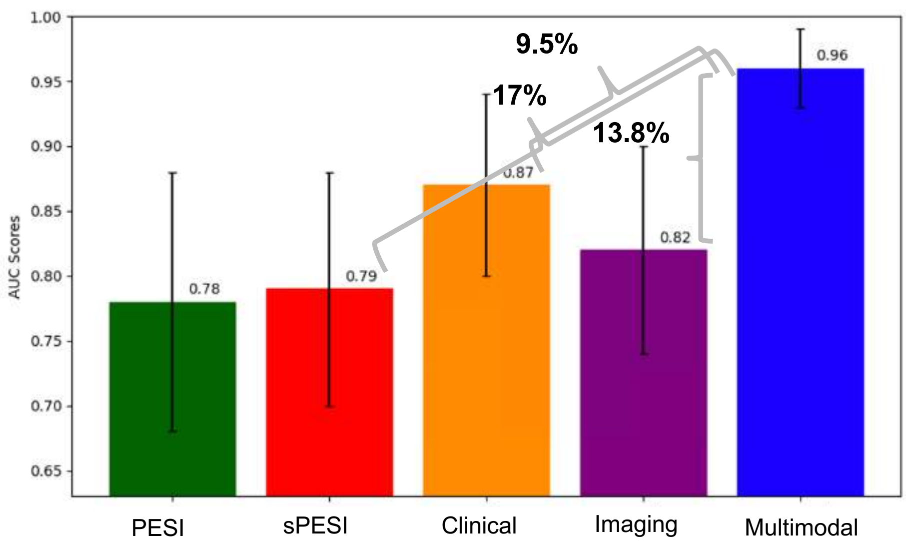


Multimodality comparison

Fusion Strategy	AUC	Accuracy	Specificity	Sensitivity
Late - average (a)	0.85 [0.75-0.93]	0.68	0.65	0.9
Late - TabNet (b)	0.92 [0.87-0.98]	0.89	0.89	0.9
Late - XGBoost (c)	0.88 [0.8-0.95]	0.76	0.74	0.9
Early (d)	0.9 [0.83-0.96]	0.81	0.79	0.9
Intermediate (e)	**0.96 [0.93-1.0]**	**0.93**	**0.94**	**0.9**



Summary of Results: Single vs Multimodal



Multimodality model outperforms the Imaging model by **13.8%** , the Clinical model by **9.5%** And the clinical rules by **17%**!

Weakly supervised end-to-end solution. Models are applied on **weakly labeled data, in variable and noisy** environments and for a highly **imbalanced** dataset -An **alternative prognostic score?**

PROJECTS/ COLLABORATIONS

Radiology/Pulmonology: Chest

Early detection of lung nodules and lung cancer;
Lung tissue analysis: fibrosis, ILD
Tracking progression over time

Gastro: Abdomen – Liver / PE

Automated segmentation of liver and liver lesions/tumors
Using multiple MRI sequences
Detection of PE; Characterization; mortality prediction

Neuro: MS Lesions in the brain; Stroke

Early detection
Identifying different population groups: forming depression/not

Dermatology

Surgery

Icahn School of Medicine at Mount Sinai
BioMedical Engineering and Imaging Institute

BMEII AI Core

Hayit Greenspan, PhD
Xueyan Mei, PhD
Valentin Fauveau, MS
Zahi A Fayad, PhD
May 9, 2023

Mount Sinai Hospital

Artificial Intelligence in Imaging

In today's world of rapid advancements in medical technologies, the need for efficient collaboration between healthcare professionals and researchers has never been greater. Artificial Intelligence (AI) opens new opportunities for exciting multidisciplinary research, combining the medical and engineering fields for improved diagnosis and treatment of a broad range of health conditions.

What We Do

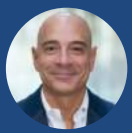
At the BMEII AI Core, we work collaboratively with researchers from various departments who aim to advance their ideas for artificial intelligence models related to medical imaging. We offer our **expertise in various aspects of AI development, including ideation, data collection, data annotation, AI model architecture, and deployment**. Our team is dedicated to provide guidance throughout the entire model creation journey to ensure that the highest quality standards are met. Additionally, we provide a platform to deploy, test and dynamically re-train deployed models to ensure they meet the demands of real clinical settings applications.

OUR RESOURCES

RadImageNet (RIN): Collection of pre-trained Deep Learning Networks for effective Transfer Learning

Discovery Viewer (DV): Medical AI Model Development Platform and Deployment Hub

THE TEAM



Zahi Fayad, PhD

Director, BMEII
AI Core Consultant

Xueyan Mei, PhD

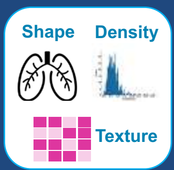
AI Research Lead

Hayit Greenspan, PhD

Director of AI Core;
Co-Director of AI-ET
PhD Program

Valentin Fauveau, MSc

AI Deployment Lead



RADIOMICS

CLASSIFICATION

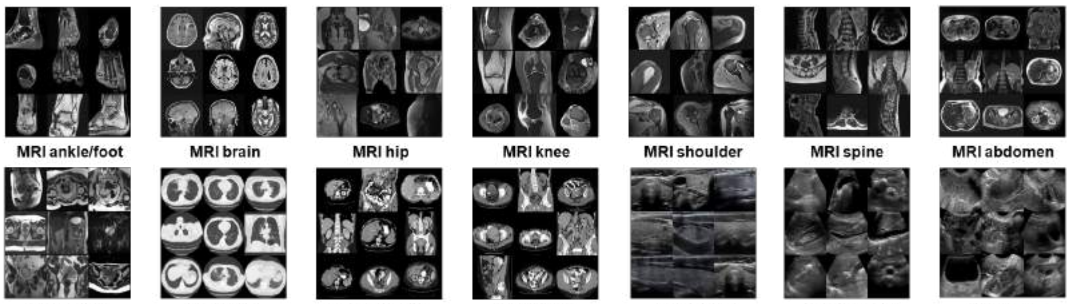
REGRESSION

SEGMENTATION

RadImageNet:
1.35 million annotated radiologic images

a

Data Structure of the RadImageNet Database



b

CT (292,353 total images)

Anatomy	# Labels	Average label size	Total # image
Lung	6	25,421	152,528
Abdomen/pelvis	28	4,994	139,825

c

Ultrasound (389,885 total images)

Anatomy	# Labels	Average Label size	Total # image
Thyroid	2	46,300	92,599
Abdomen/pelvis	13	22,868	297,286

d

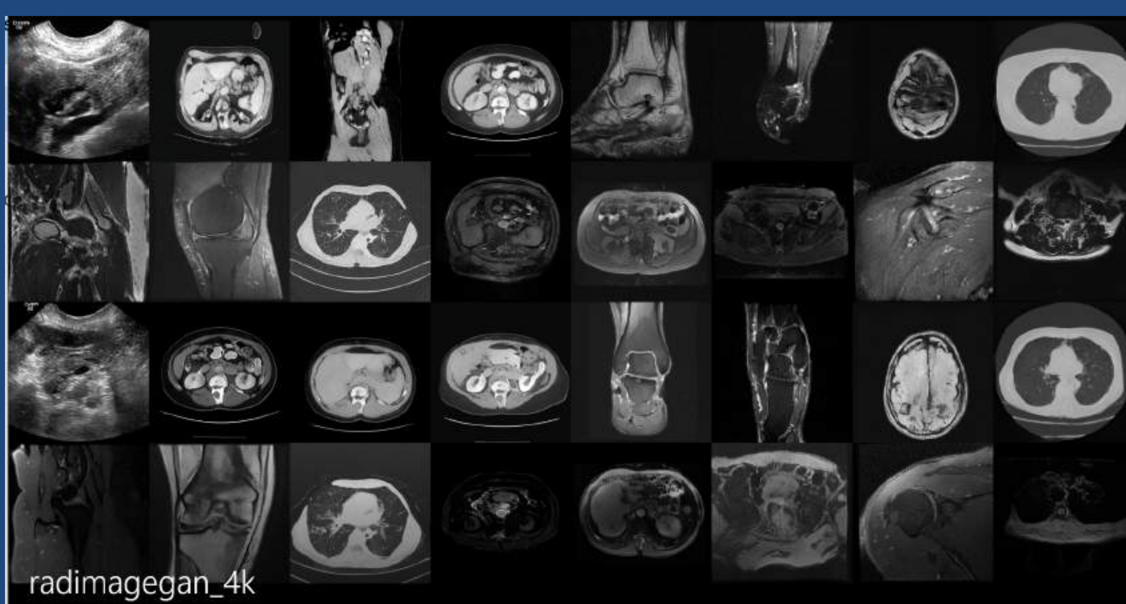
MRI (672,675 total images)

Anatomy	# Labels	Average Label size	Total # image
Knee	18	9,975	179,555
Shoulder	14	3,743	52,407
Spine	9	7,964	71,674
Ankle/foot	25	7,264	181,603
Abdomen/pelvis	26	3,513	91,348
Brain	10	4,467	44,671
Hip	14	3,673	51,417

(Mei et al.) Radiology: Artificial Intelligence 2022

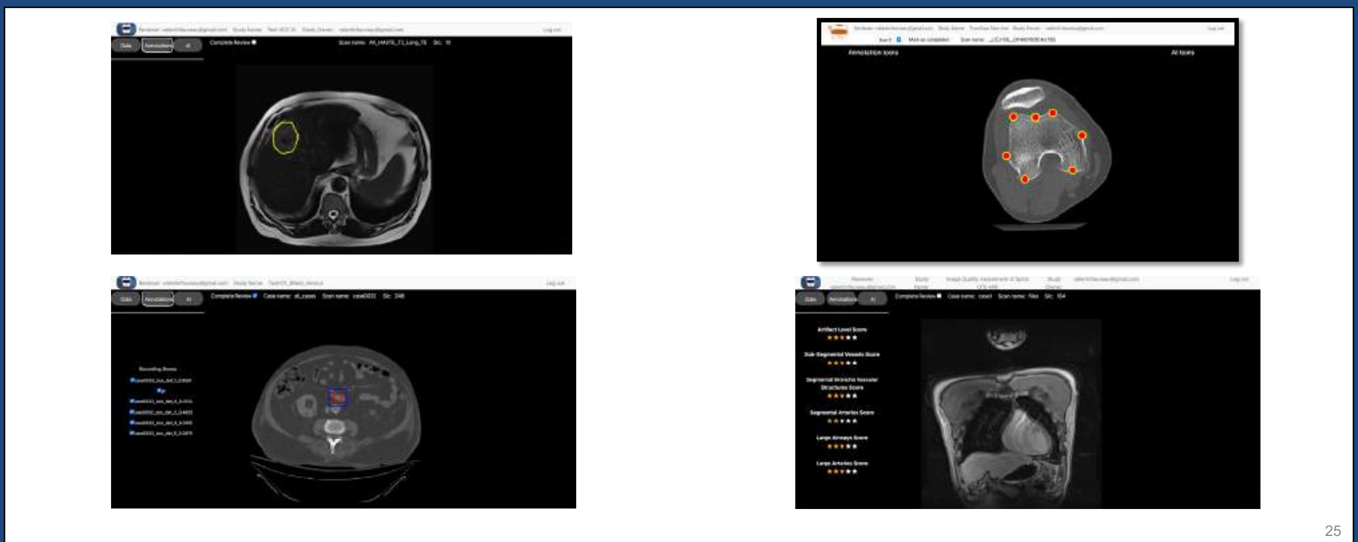
23

RadImageGAN: Synthetic Images

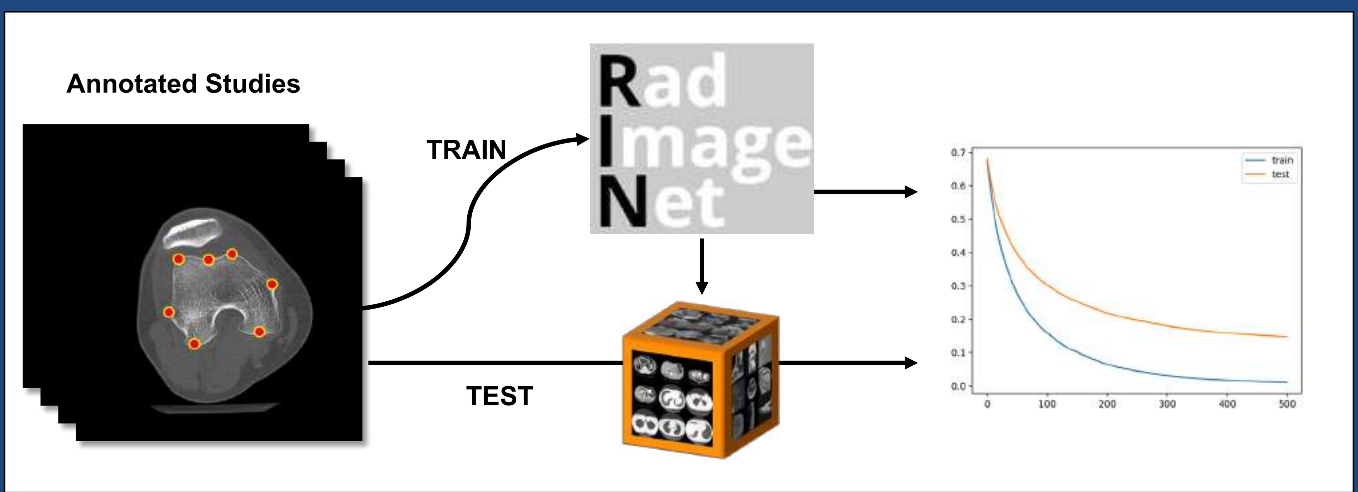


24

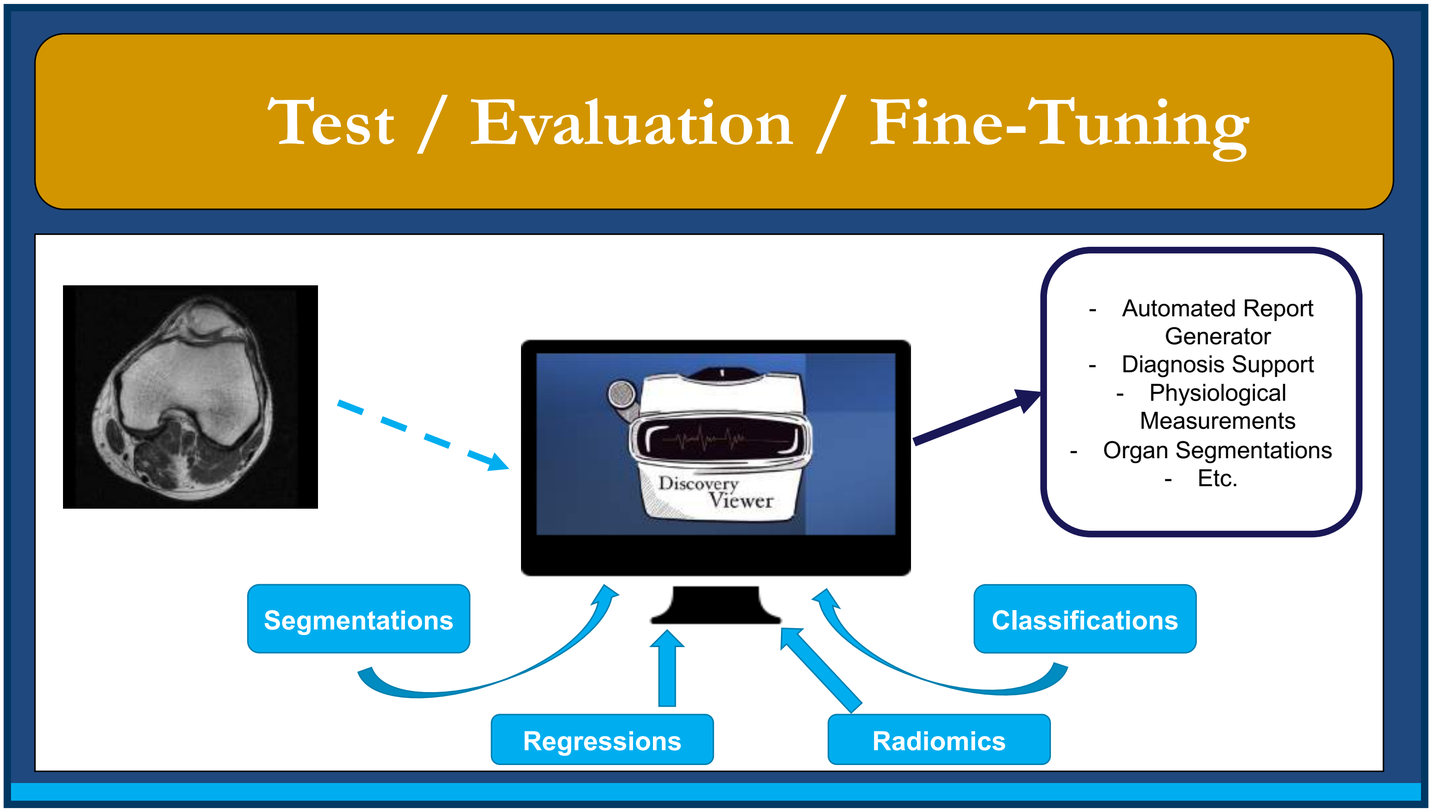
Data Annotation



Model Training



DV is bundled with **RadImageNet (RIN)**, making it easy to train and test models through a Transfer Learning approach



AIET

AI & Emerging Technologies

New PhD Program at Mt Sinai

Co-Directors:

Hayit Greenspan, PhD

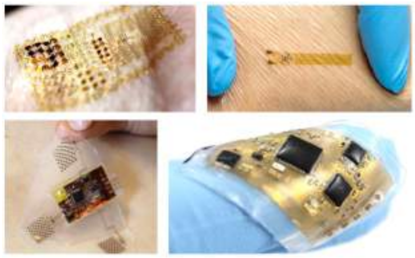
Alan Seifert, PhD

Icahn School of Medicine at Mount Sinai

BioMedical Engineering and Imaging Institute

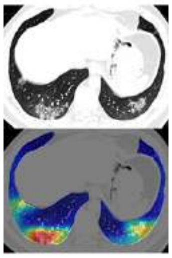
AIET Goals

- To offer students with solid quantitative and technical backgrounds educational and research opportunities in:
  - AI/machine learning
  - Medical imaging
  - Next generation medical technologies (devices, sensors, robotics, wearables)
  - Virtual/augmented reality for clinical applications or drug discovery

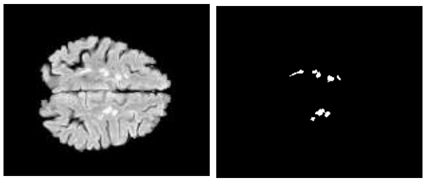


Stretchable skin-mounted electronic sensors for measurement of electrical activity
Yun Soung Kim, Ph.D., AIET Faculty Member

COVID automated image analysis in CT and Xray images



MS Lesion detection and segmentation in BRAIN MRI
Ranked 1st in DICE on ISBI-2015 challenge



Machine learning analysis of robotic surgery
Lee, Rapoport, et al.

Heart rate variability tracking for early diagnosis of COVID-19
Hirten et al., J Med Internet Res. 2021

https://icahn.mssm.edu/education/phd/biomedical-sciences/artificial-intelligence-technologies

AIET PhD

Our Disciplines

Artificial Intelligence and a variety of other powerful technologies (e.g., imaging, biotechnology, nanotechnology, information technology, cognitive science, and robotics) are paving the way for a new era of biomedical research, offering unparalleled opportunities to improve human health.

Collaborations

Relationships with several well-regarded higher education institutions:

--State University of New York at Stony Brook,
--Rensselaer Polytechnic Institute (RPI),
--the Grove School of Engineering at the City College of New York,
--the Cooper Union - Albert Nerken School of Engineering,
--the Hasso Plattner Institute of the University of Potsdam, in Germany
offer complementary technical expertise to expand collaborative research and enrichment opportunities for trainees and faculty.

https://icahn.mssm.edu/education/phd/biomedical-sciences/artificial-intelligence-technologies

Mortality Risk AI: A use-case involving ‘off-label’ use and considerations for equity

Vincent J. Major, PhD

May 19, 2023
HealthFirst Provider Symposium. Artificial Intelligence & Health: Addressing Health Outcomes and Equity with New Tools.

@vincentjmajor
vincent.major@nyulangone.org
vincentmajor.com

NYU Langone Health

Who Am I?

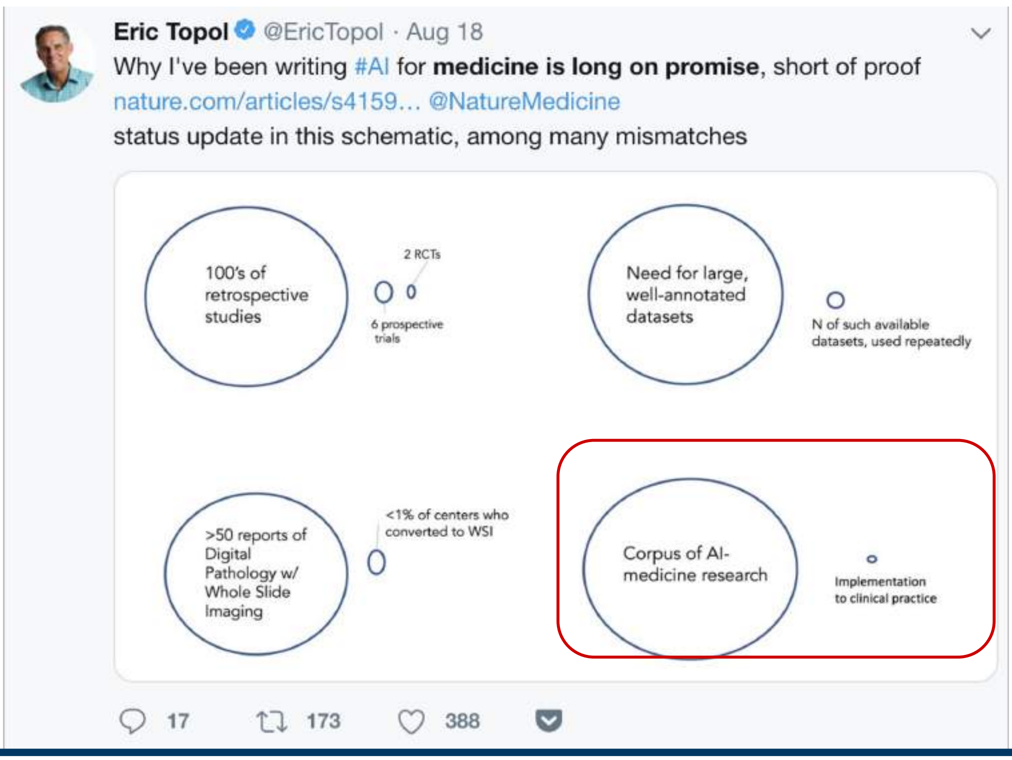
- Trained engineer, working in healthcare applications for nine years.
- PhD in biomedical informatics from New York University where I:
  - developed machine learning models to identify hospitalized patients at very high-risk of mortality and
  - deployed these to notify physicians within minutes of admission

Two hats:

- Academic (NYU Grossman School of Medicine)
  - Assistant Professor, Center for Health Innovations and Delivery Science, Department of Population Health
- Operations (NYU Langone Health)
  - Lead, Operational Machine Learning EHR Integration, Medical Center IT

NYU Langone Health

The Predictive Analytics Unit at NYU



NYU Langone Health

Purpose of the Predictive Analytics Unit

one

- Provide a predictive modeling capacity to:
  - Increase operational efficiency.
  - Drive value.
  - Translate AI to the point of care.

two

- Lead specific projects from inception through:
  - Development,
  - Deployment, and
  - Evaluation

three

- Provide education and professional development to support data-driven decision making.

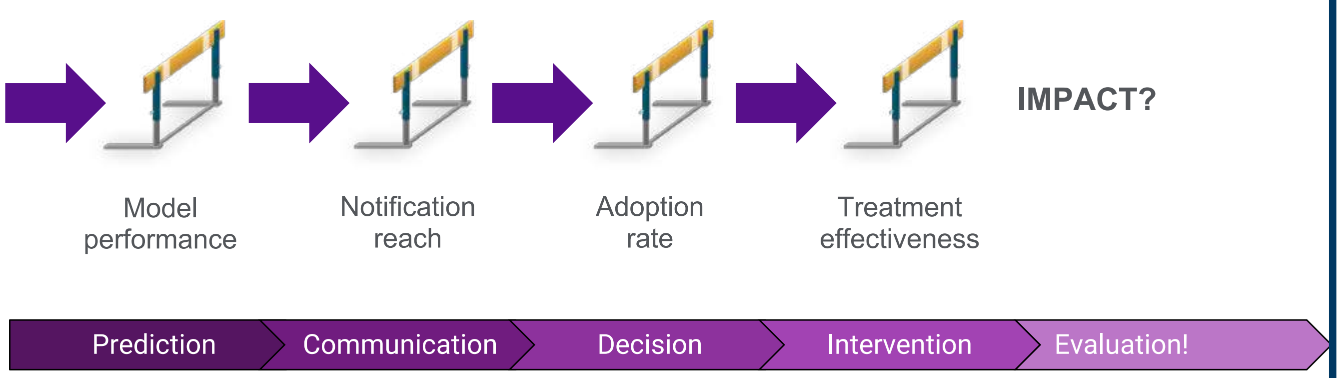
four

- Provide central umbrella organization for clinical data science at NYU and for external collaboration

NYU Langone Health

Big Picture of Clinical Predictive Models

A good model doesn't guarantee success. We need to consider the entire clinical workflow.



NYU Langone Health

Background: AI for End-of-Life

Background – *Decline at the end-of-life*

- Patients with life-limiting disease can maintain their health status for months or years.

    - A 'sentinel hospitalization' may mark start of a rapid decline before death. (Lin et al., 2014)

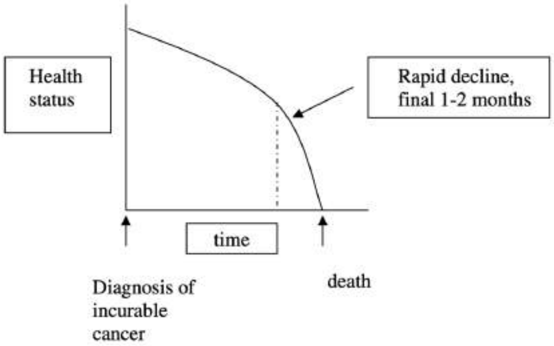


Fig. 1 – Conceptual model of cancer death trajectory.[122,123]

(Glare et al., 2008)

Background – *Prognosis in practice*

- Estimating when a patient will die is difficult.
  - Physicians are known to overestimate survival. (Christakis and Lamont, 2000; Glare et al., 2003)
  - Physicians are ⅓ accurate at estimating survival within +/- ⅓ (Taniyama et al., 2014)
- The estimate and confidence are the driver for downstream interventions.
  - Optimism and uncertainty can lead to continued treatment.
- Many different scores, indices and tools exist.
  - Performance varies widely by study. (Siontis, Tzoulaki and Ioannidis, 2011)
  - Few are widely used in practice. (Yourman et al., 2012)
  - None have demonstrated consistent improvements in ACP.

Can we use machine learning?

NYU Langone Health

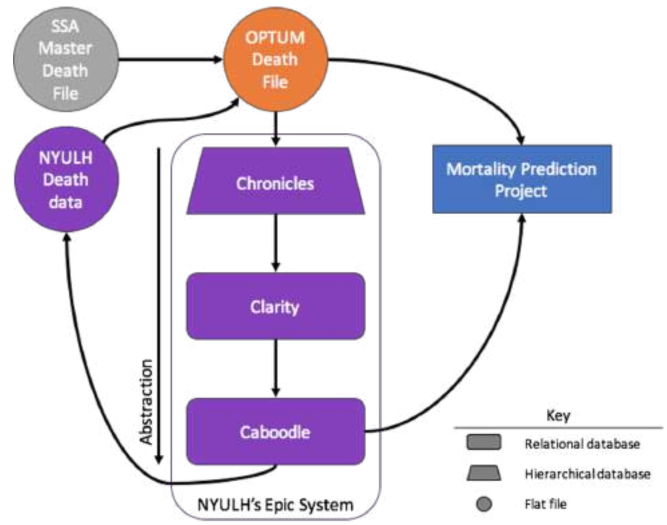
Objective

- To identify hospitalized patients at risk of dying who may benefit from advance care planning and prompt their providers to act
- To develop, validate, and implement machine learning models to predict patients at very high risk of dying
  - Adapt prior work to:
    - Predict on (almost) all adult admissions,
    - Predict at (or near) the time of admission,
    - Predict with high precision (75% positive predictive value), and
    - Predict death within 60 days of admission.

NYU Langone Health

Development – *Data sources*

- Data pulled from Epic Electronic Health Record.
  - All adult admissions from 2015–2017 across 3 sites.
    - 128k inpatient admissions.



NYU Langone Health

Development – *Experimental Design*

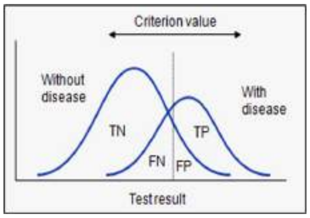
- Three years of data, split it into groups:
  - Training, and
  - Temporal testing set
- Ensure one patient is only in one group.

NYU Langone Health

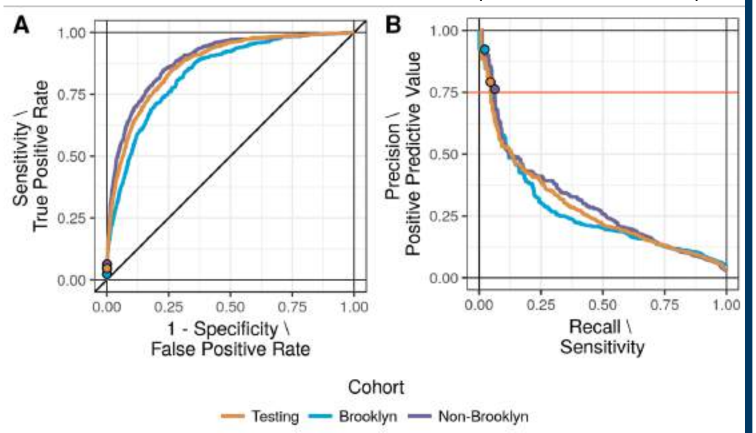
Development – *Modeling*

- Random forest
  - Parameters selected within cross-validation
  - Retrained and applied to testing

- Results
  - Good discrimination and moderate average precision
    - Maintained across sites
  - Operate at 75% PPV
    - 4.6% recall
    - Overly conservative at Brooklyn



(MEDCALC, 2019)



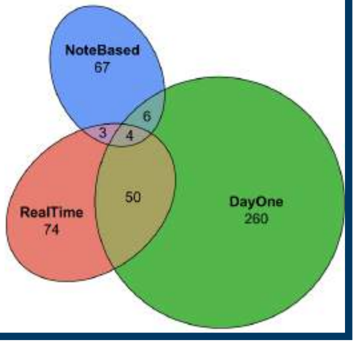
[Major, 2020. BMC Medical Informatics and Decision Making.](#)

Implementation – *Live models*

We have 3 models that are live:

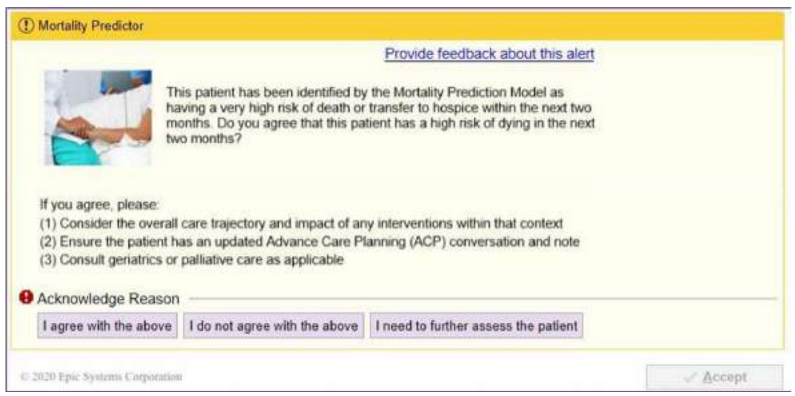
1. [RealTime (Major, 2020)](#),
2. DayOne,
3. [NoteBased (Bilaloglu, 2021)](#).

The models tend to find similar patients but each adds a different group.



Implementation – *Clinical Alert*

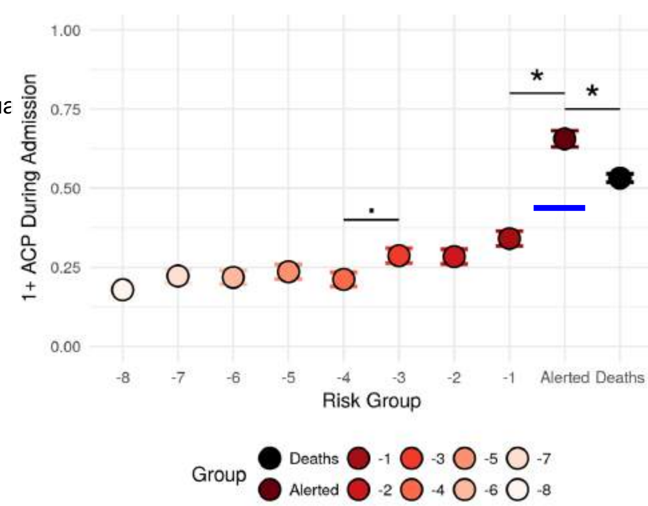
- Presented to patient's attending physician,
- Interruptive (cannot click away),
  - One chance to defer
- Asks if the user agrees with the patient's risk.
  - Agree
  - Disagree
  - Defer



[Wang, 2021. NEJM Catalyst.](#)

Evaluation – *Adoption of ACP above baseline?*

- No control group, use other groups
  - Lesser risk
    - Using the models to find the next group of equa
  - Deaths
    - All deaths within 60 days of admission



- Results
  - Alerted > Deaths
  - Alerted > Lesser risk
  - ACP = f(Risk)

Physicians have adopted the alert and increased ACP rates

NYU Langone Health

Evaluation – *Key Observations*

1. Disagrees are common but few patients with All-Disagree (16%)
   a. All-Disagrees have better survival outcomes
      i. Providers can identify lower risk patients.
2. Providers are adopting the ACP recommendations (66%)
   a. Less when they disagree (34% vs 72%)
   b. Higher than the next group of lesser risk and all deaths
3. Overall, the system is performing well and identifying patients at risk

NYU Langone Health

Off-Label Use of AI for Care Management

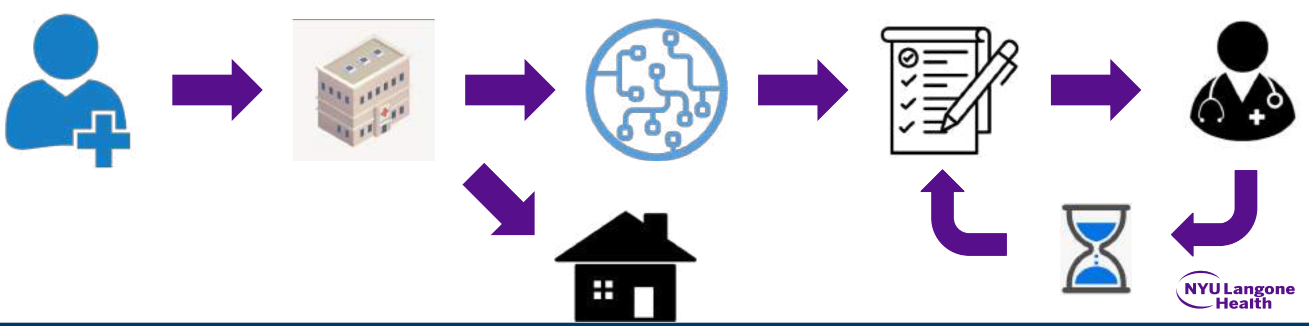
Adaptation of the Inpatient AI

- The AI was designed with a goal: to develop, validate, and implement machine learning models to predict patients at very high risk of dying.
  - Predict on (almost) all adult admissions,
  - Predict death within 60 days of admission
  - Predict with high precision (75% positive predictive value), and
  - Predict at (or near) the time of admission, to
  - Notify Attending physicians within minutes of admission.

- Pivoted this to help screen patients in the community:
  - Focusing on attributed patients in the community, while
  - Recycling predictions from a recent hospitalization, to
  - Prioritize screening and outreach by risk (the top ~8% of patients)

Human-in-the-loop screening for Supportive Care

1. Patient comes to the hospital,
2. Models predict their risk, which is added to a list,
3. Patient (survives,) is discharged and goes home,
4. A Supportive Care RN reviews the list daily, reviews their chart, and contacts patients directly. Follow-up with Community Health Workers and Clinical Care Coordinators.



Intervention and Patient Engagement

Over 3,600 patients have been identified since the program's inception in late 2019. Cases are reviewed by our interdisciplinary team for eligibility and outreached if appropriate.

Supportive Care RN	• Goals of care discussions • ACP Notes • Hospice Enrollment	• 352 Patients eligible for SC intervention • 277 Patients engaged, ***61% Engagement Rate*** • Outreach in progress for 60 additional patients • ***29 (10%) Patients enrolled in hospice***
Community Health Workers	• SDOH Support	• 221 Patients engaged, ***36% Engagement Rate*** • Top 3 goals of care: acquiring medical care, acquiring food support and acquiring transportation
Clinical Care Coordinators	• High Risk Case Review and Referral to SC RN	• 291 Patients engaged, ***30% Engagement Rate*** • Top 3 goals of care: Healthy Behaviors, Chronic Disease Management, Diabetes Management

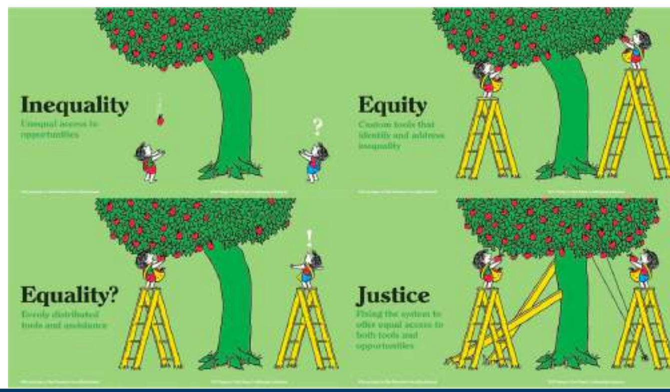
Algorithmic Fairness and Equity

The impossibility of being fair

- Many different concepts and definitions for an algorithm being ‘fair’
  - Model decision unchanged by protected/sensitive attributes (such as race/ethnicity or gender)
  - Equal likelihood across attributes, etc.
- Statistically, it is impossible to satisfy these two without either:
  - No disparity in outcomes, or
  - A trivially useless AI.

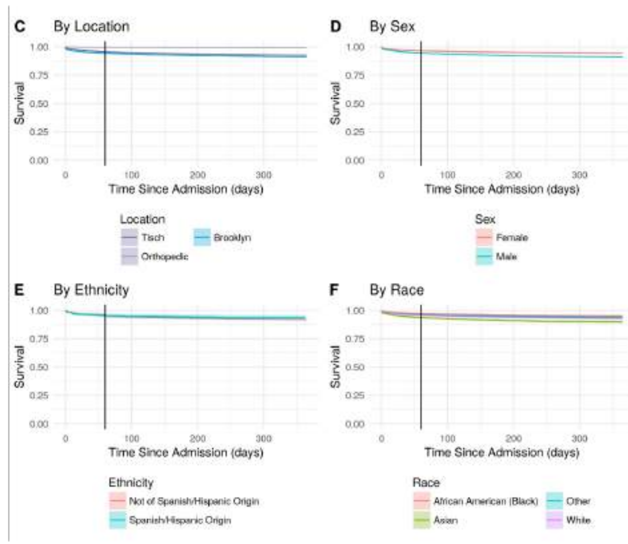
Said another way, given a known disparity, it is impossible to generate the same AI-based decisions for different groups while treating all patients equally.

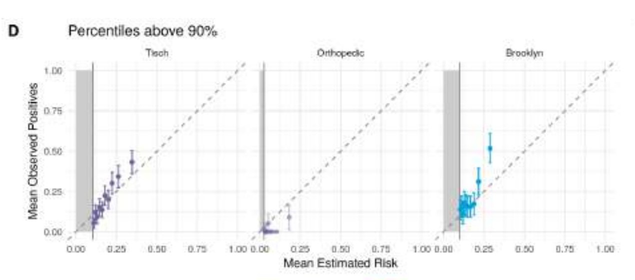
Which would you pick? Treat patients the same or give more resource to patients who need it?



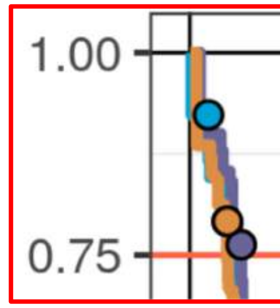
Our model

1. Disparities exist
   a. By sex, by race, and by location
   b. We underestimate risk at the highest 1%

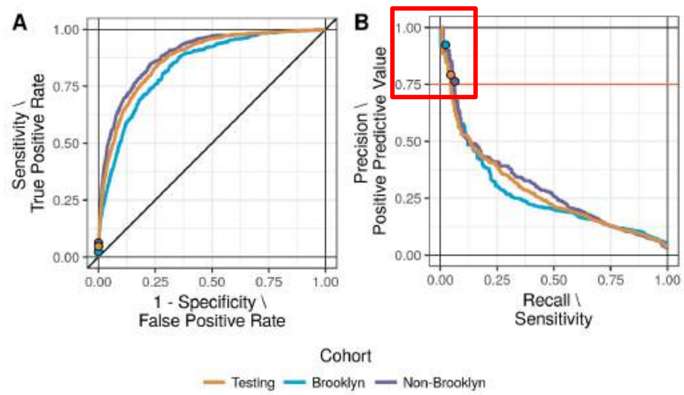


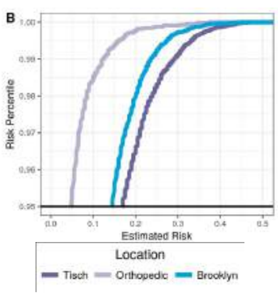


Our model



1. Disparities exist
2. One threshold → disparity in treatment
    a. One threshold for all patients (orange), we achieve wildly more conservative actions at Brooklyn (blue vs. purple)
    b. Because the distributions are different ↓



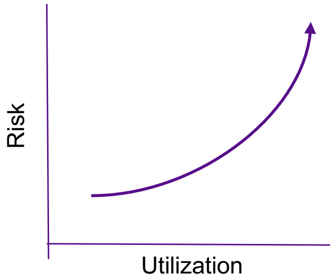


Our model

1. Disparities exist
2. One threshold → disparity in treatment
3. Our inputs are encounters and coded EHR data
    a. Reasonable to infer that a model may learn:

    utilization $\propto$ risk

    b. Patients with lower access to affordable care may be disadvantaged



With all of this said, what are we trying to do?

Help patients by allocating a limited resource where it may be needed.

Either we can forget about a model, or, we can override it and help the people who need it most.

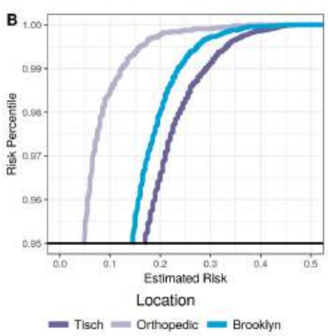
NYU Langone Health

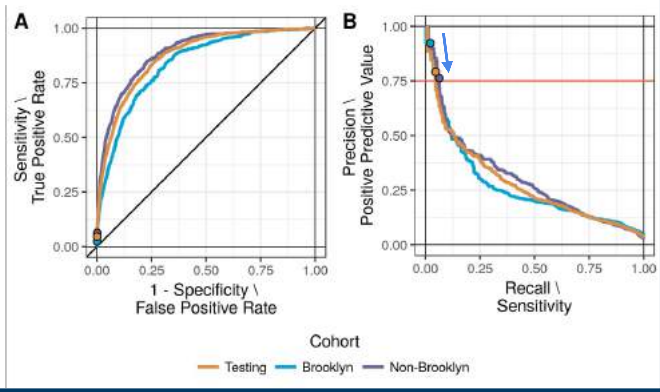
One solution: use an imperfect model and choose to treat patients *un*equally with the goal of equity.

For our inpatient use-case, we cared about PPV—how confident we were in the identified patients (positives).

We could choose to lower the threshold at Brooklyn to equalize PPV.

Or, we could choose to equalize a different metric, e.g. recall/sensitivity—the proportion of patients who die who are outreached.





Parting thoughts

1. AI for end-of-life planning and supportive care is possible (for us, since 2018)
2. An accurate model will learn the existing disparities in data and outcomes.
   a. A model can either:
      i. pretend all patients are the same (ignore their sensitive attributes), or
      ii. Act in a way to achieve equal actions (equal performance across groups), but **not both**.
3. Don't dismiss bias as unhelpful.
   a. It is a limitation that can be wrangled.
4. Awareness of existing disparities and potential paths for correction are more important than an accurate AI.
   a. A simpler model may be easier to understand and correct than a 'black-box'.

NYU Langone Health

Thank you!

Questions?

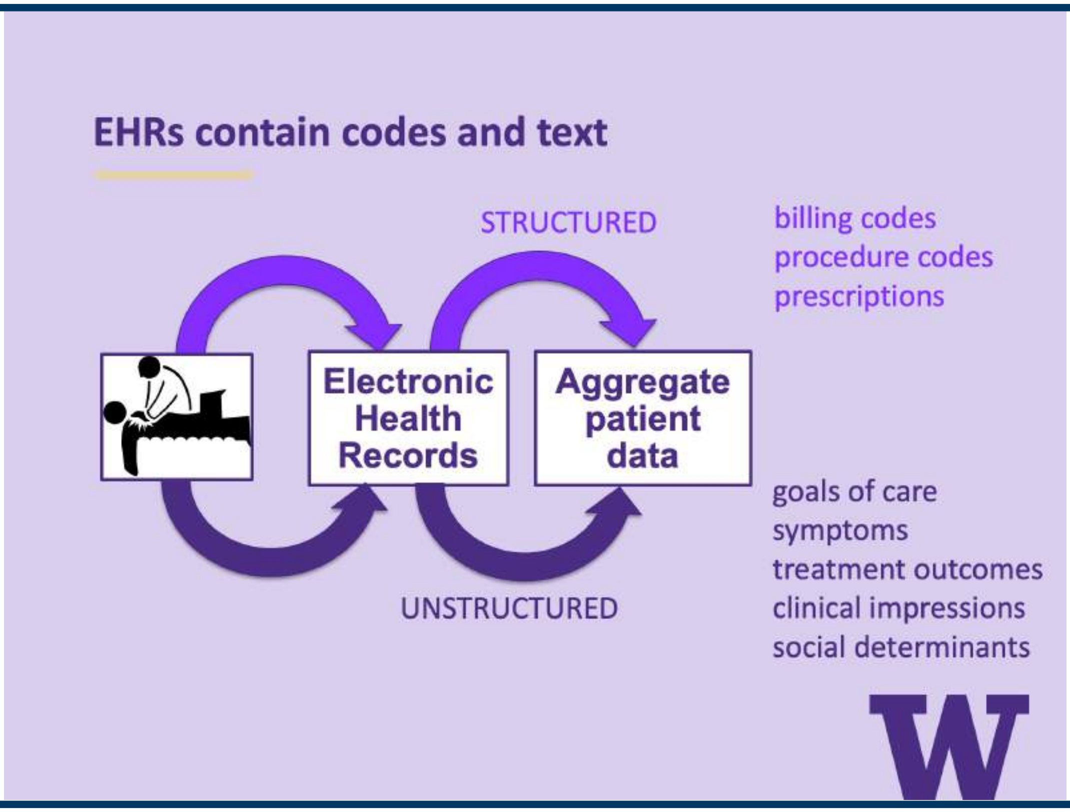
@vincentjmajor
vincent.major@nyulangone.org
vincentmajor.com

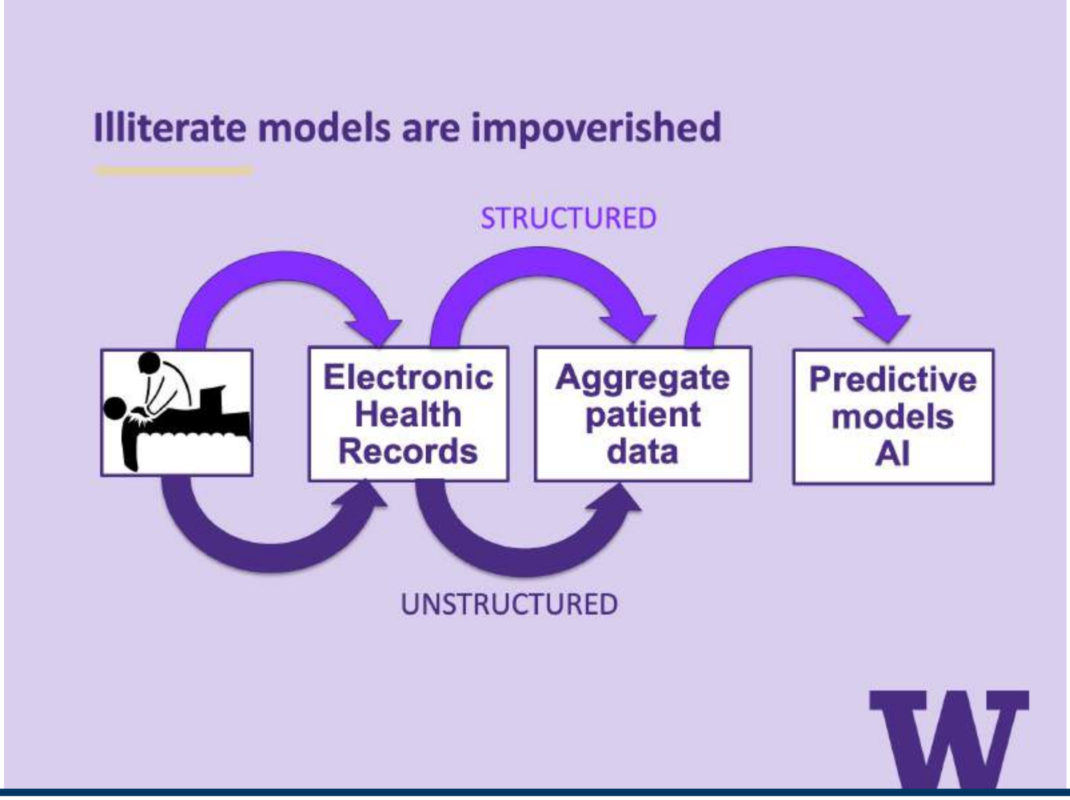
NYU Langone Health

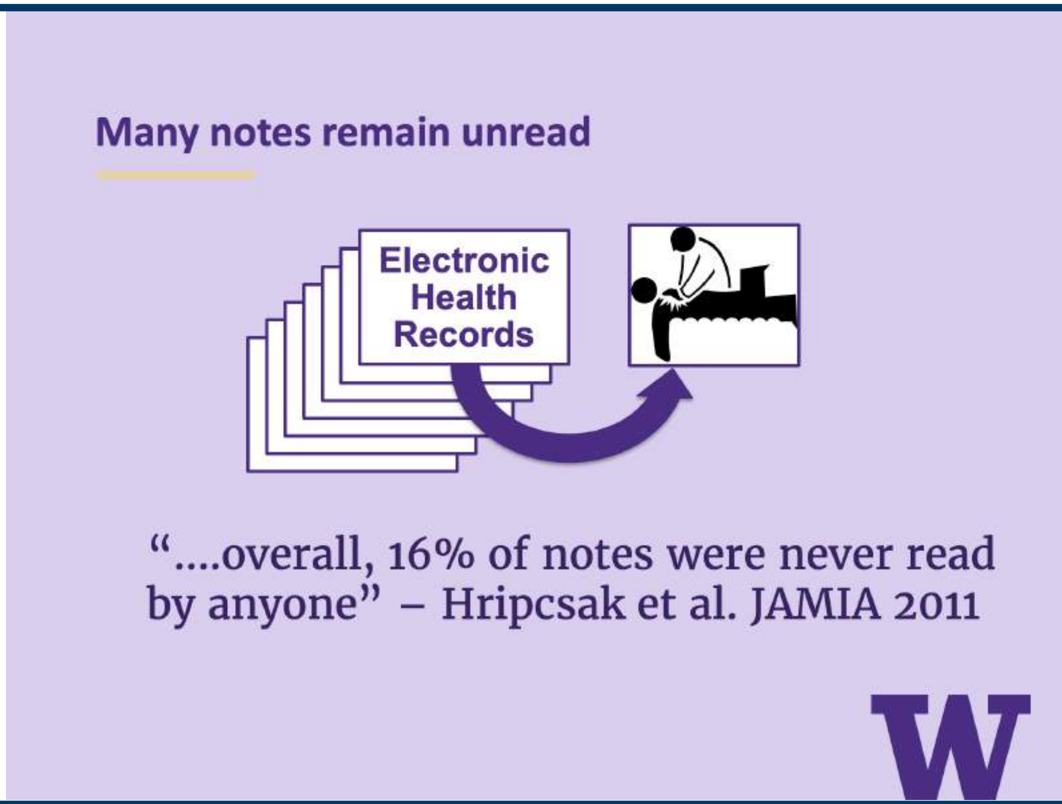
Identifying Improbable Expressions with Neural Language Models to Measure Mental State and Status.

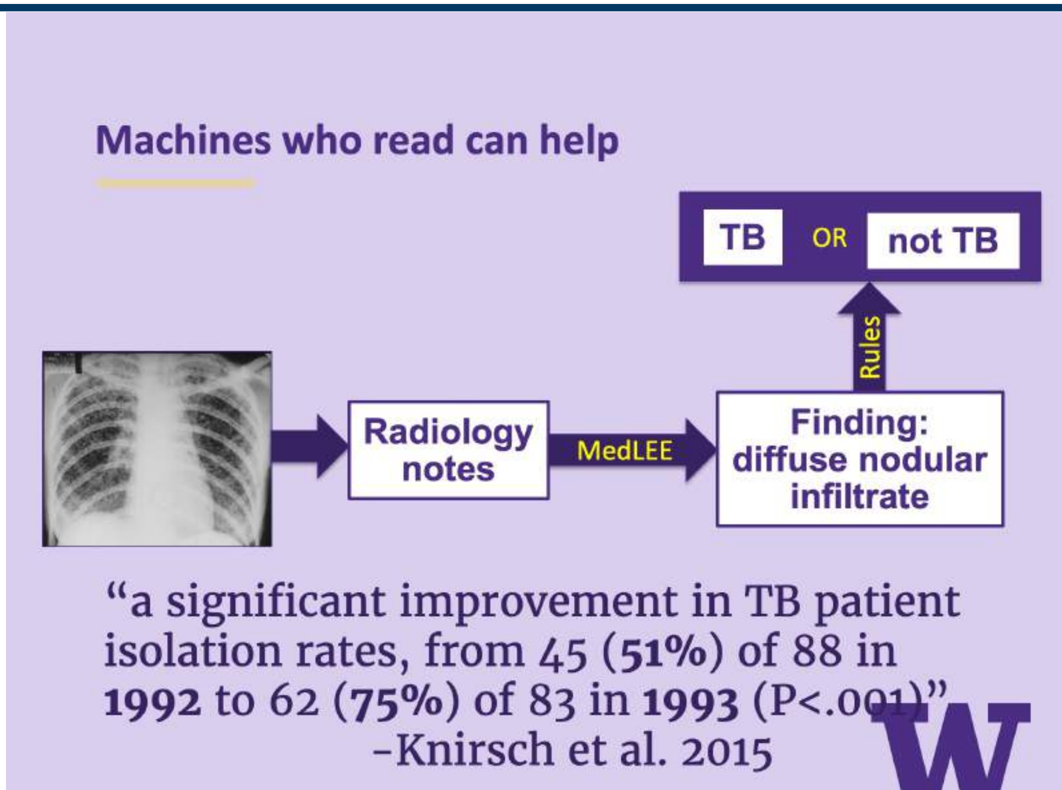
Trevor Cohen
Biomedical Informatics and Medical Education
UNIVERSITY of WASHINGTON

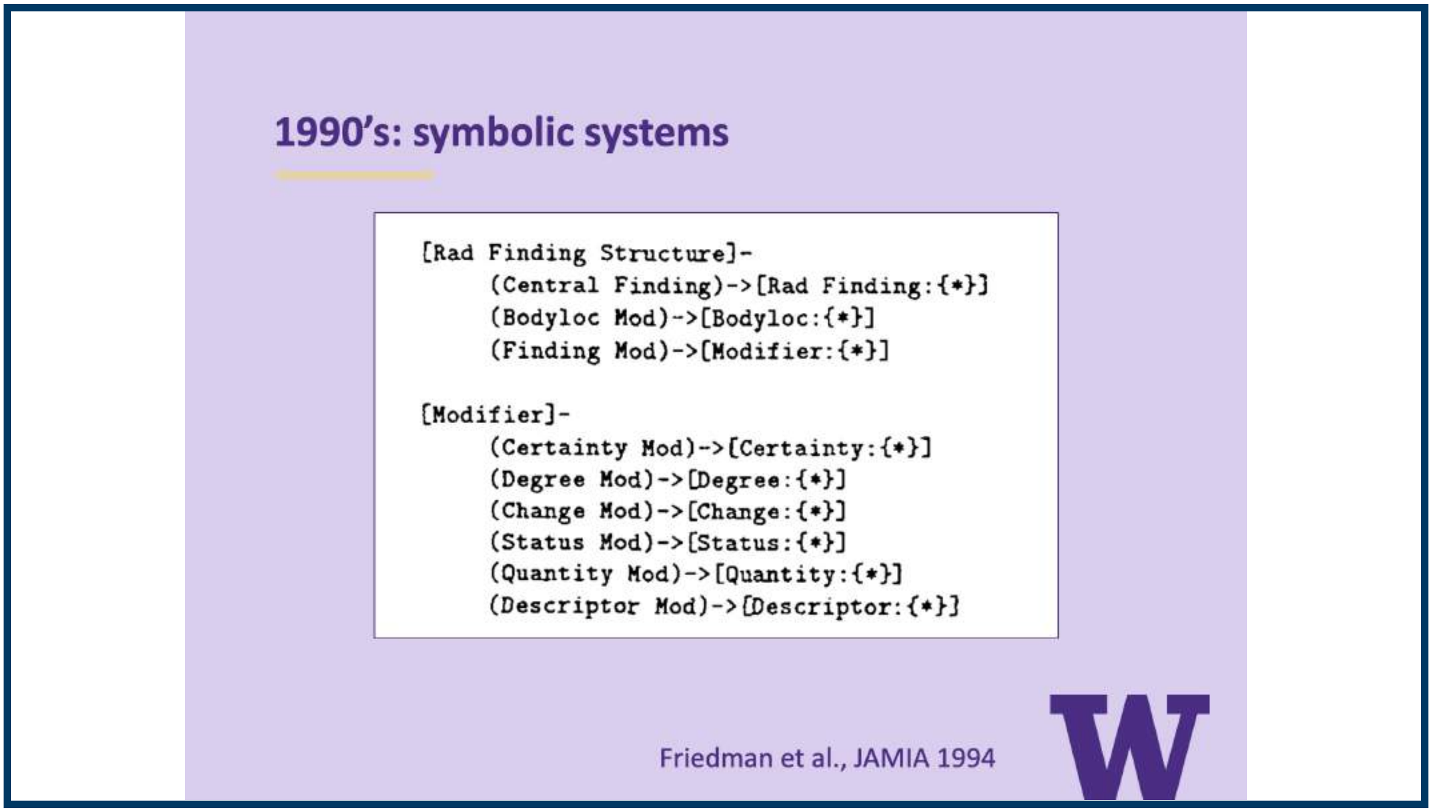
Why natural language processing?

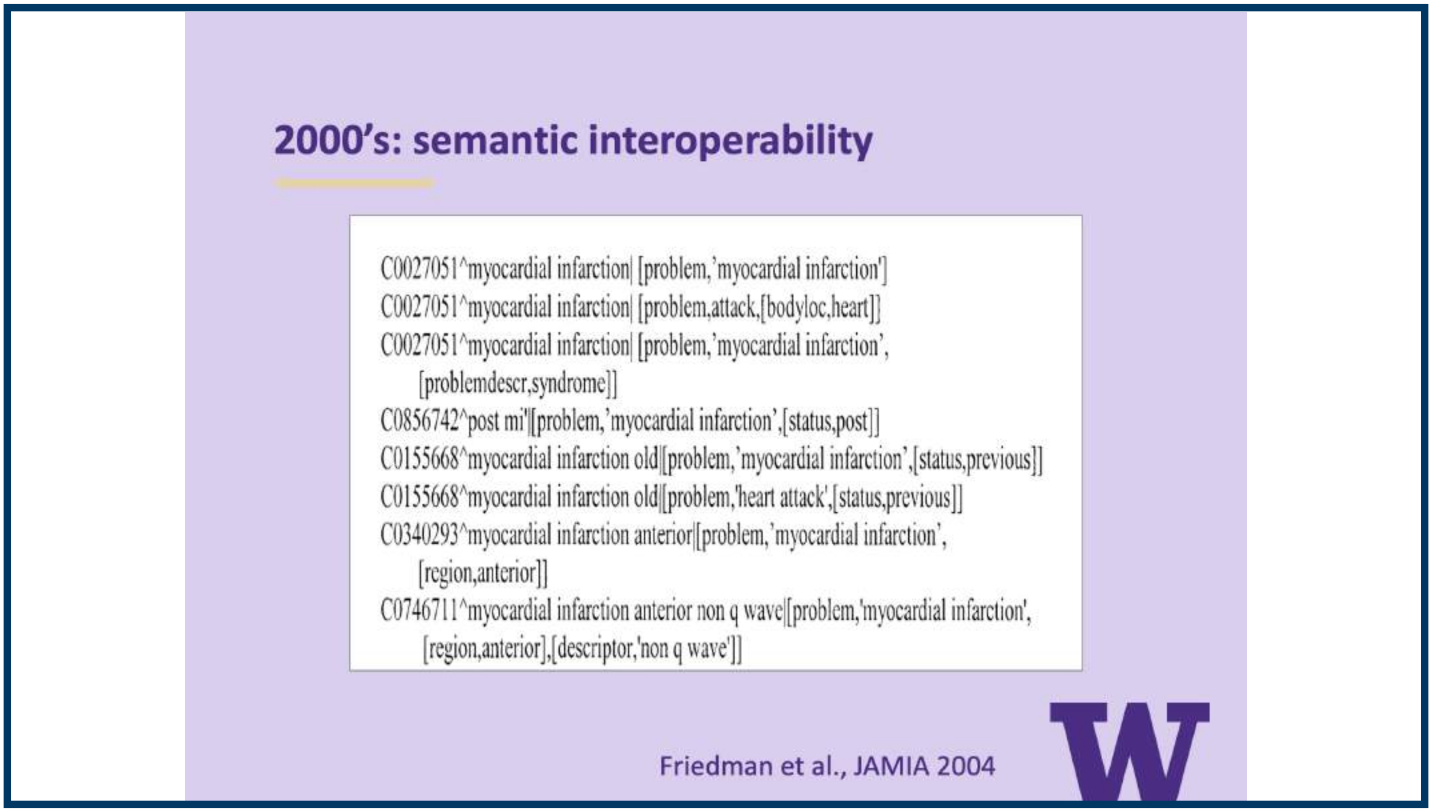

EHRs contain codes and text
STRUCTURED
billing codes
procedure codes
prescriptions
Electronic Health Records
Aggregate patient data
goals of care
symptoms
treatment outcomes
clinical impressions
social determinants
UNSTRUCTURED

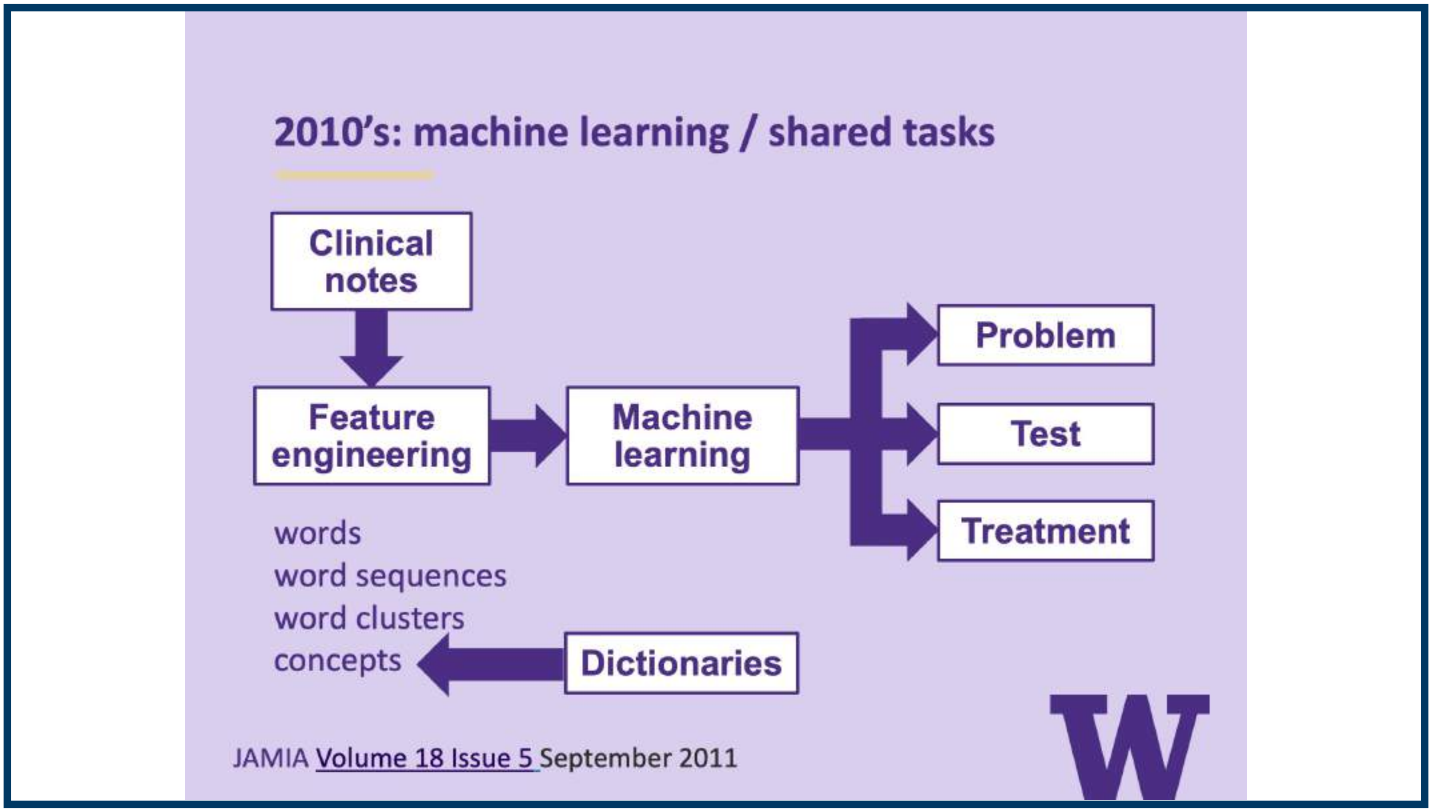

Illiterate models are impoverished
STRUCTURED
Electronic Health Records
Aggregate patient data
Predictive models AI
UNSTRUCTURED

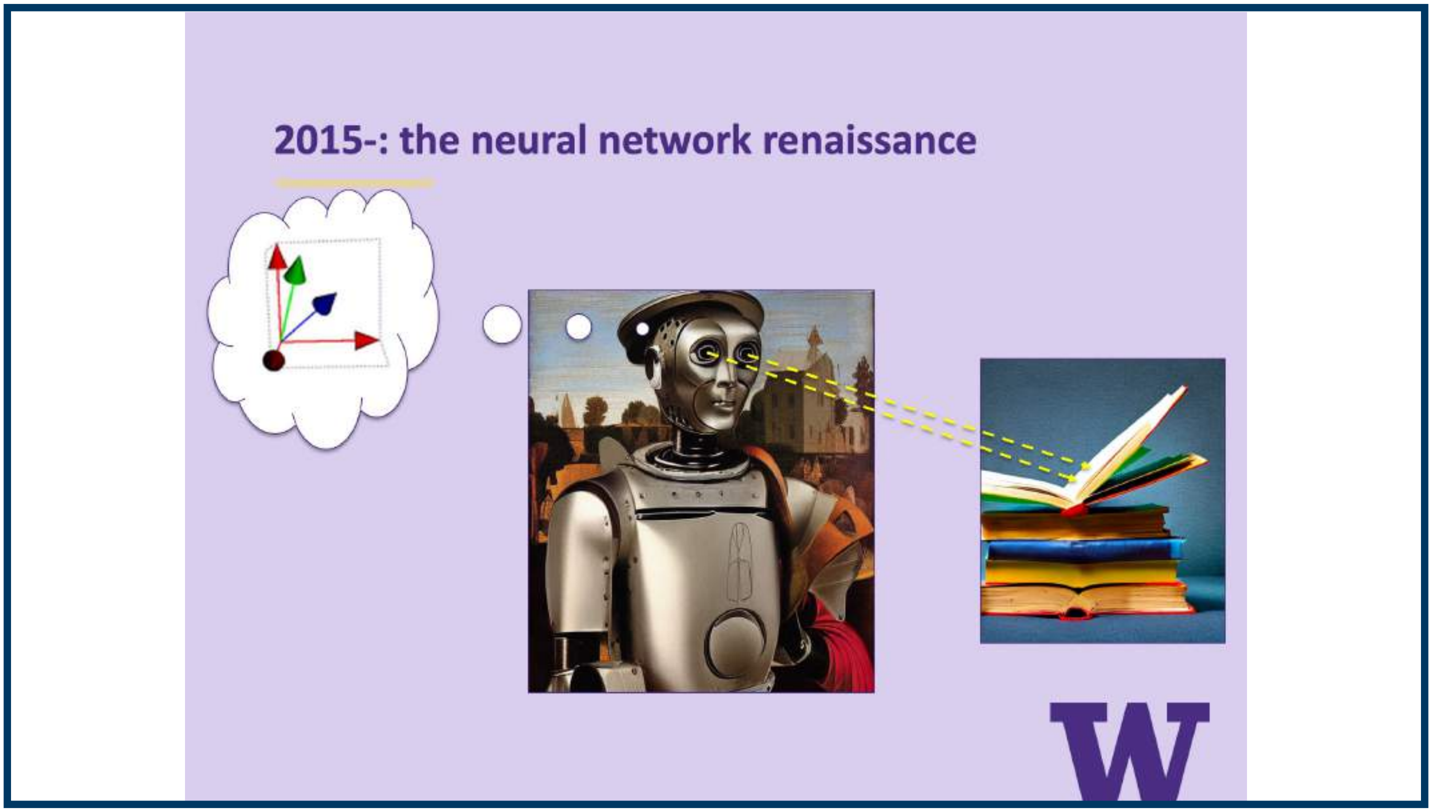

Many notes remain unread
Electronic Health Records
"....overall, 16% of notes were never read by anyone" – Hripcsak et al. JAMIA 2011
W

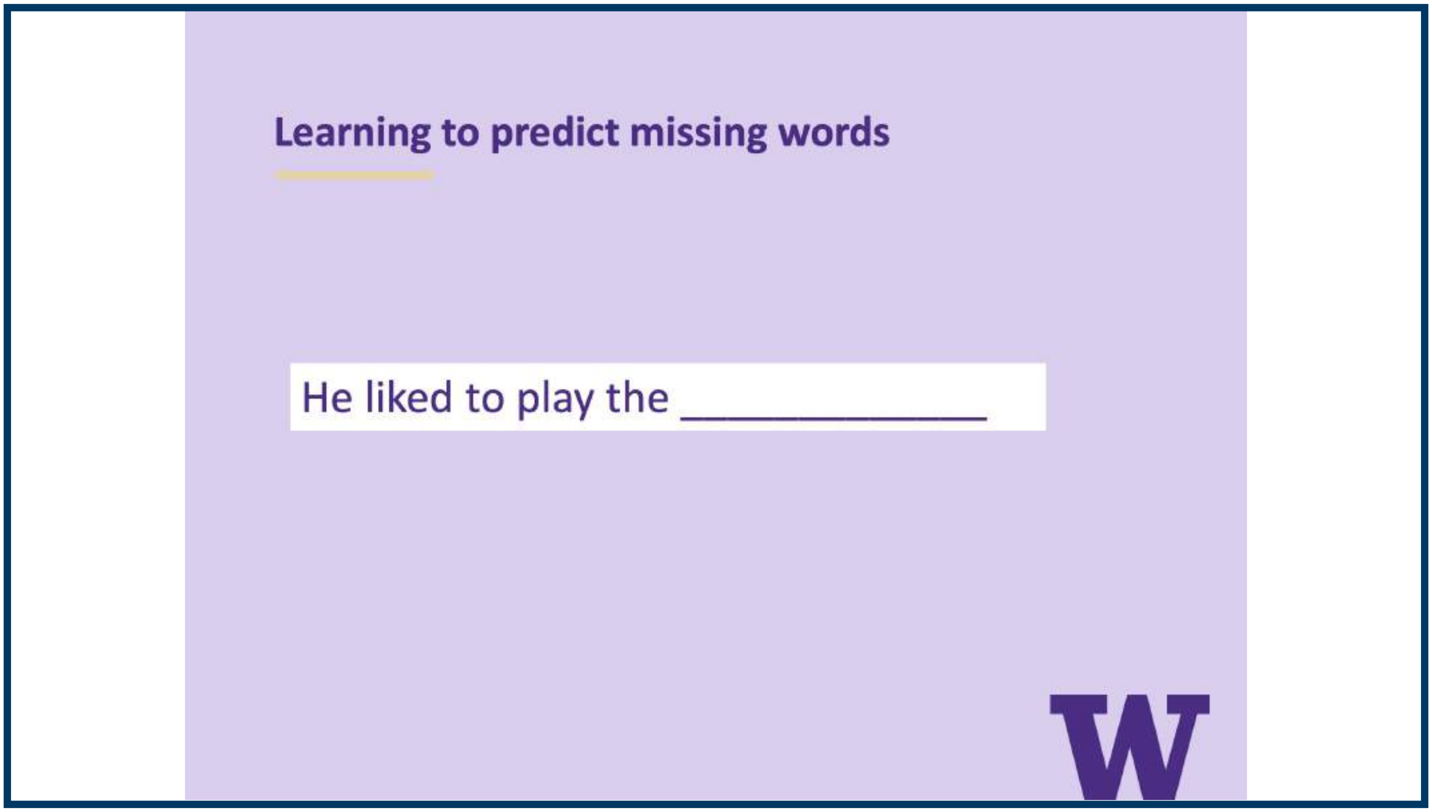

Machines who read can help
TB
OR
not TB
Rules
Radiology notes
MedLEE
Finding: diffuse nodular infiltrate
"a significant improvement in TB patient isolation rates, from 45 (51%) of 88 in 1992 to 62 (75%) of 83 in 1993 (P<.001)"
–Knirsch et al. 2015
W

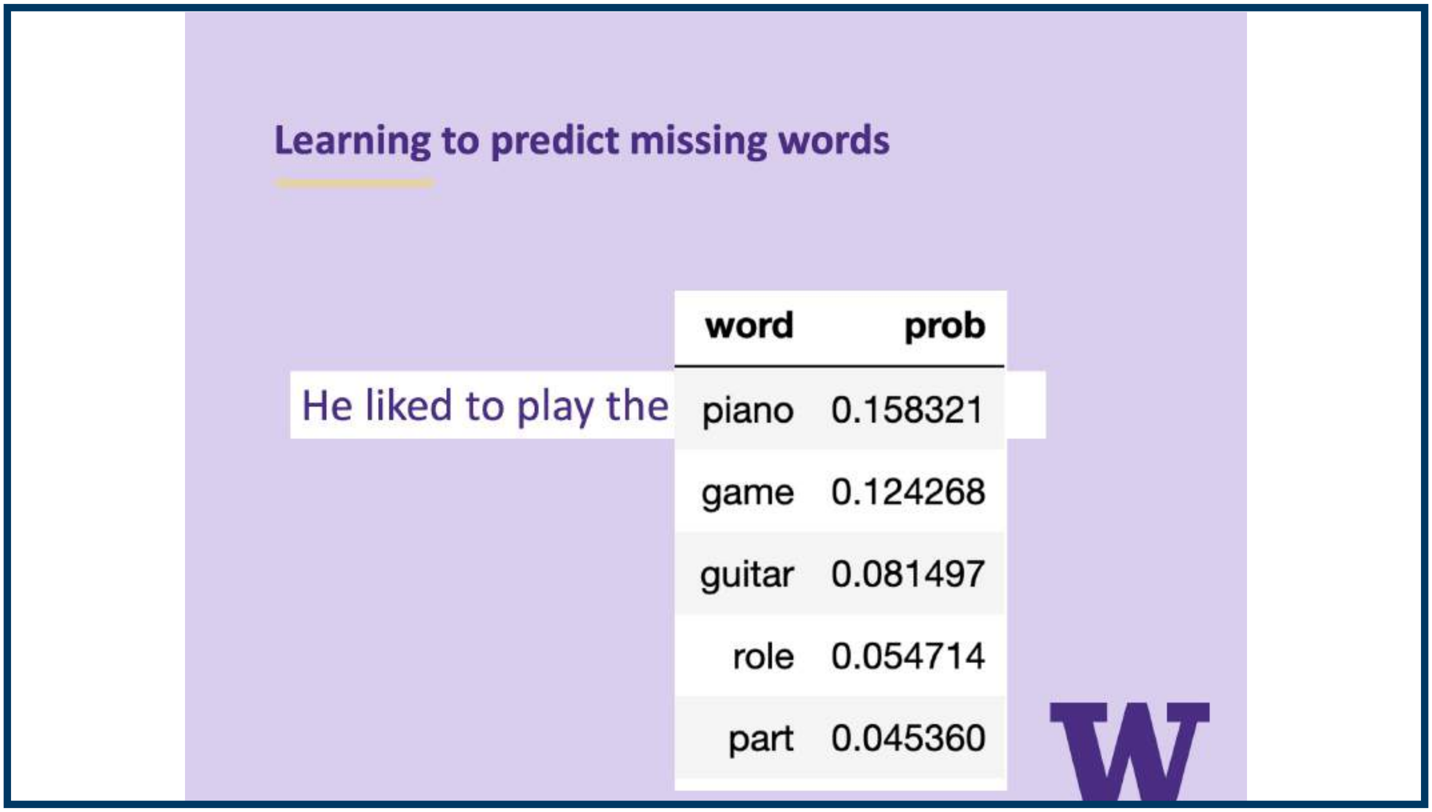

1990's: symbolic systems
[Rad Finding Structure]-
(Central Finding)->[Rad Finding:{*}]
(Bodyloc Mod)->[Bodyloc:{*}]
(Finding Mod)->[Modifier:{*}]
[Modifier]-
(Certainty Mod)->[Certainty:{*}]
(Degree Mod)->[Degree:{*}]
(Change Mod)->[Change:{*}]
(Status Mod)->[Status:{*}]
(Quantity Mod)->[Quantity:{*}]
(Descriptor Mod)->[Descriptor:{*}]
Friedman et al., JAMIA 1994
W

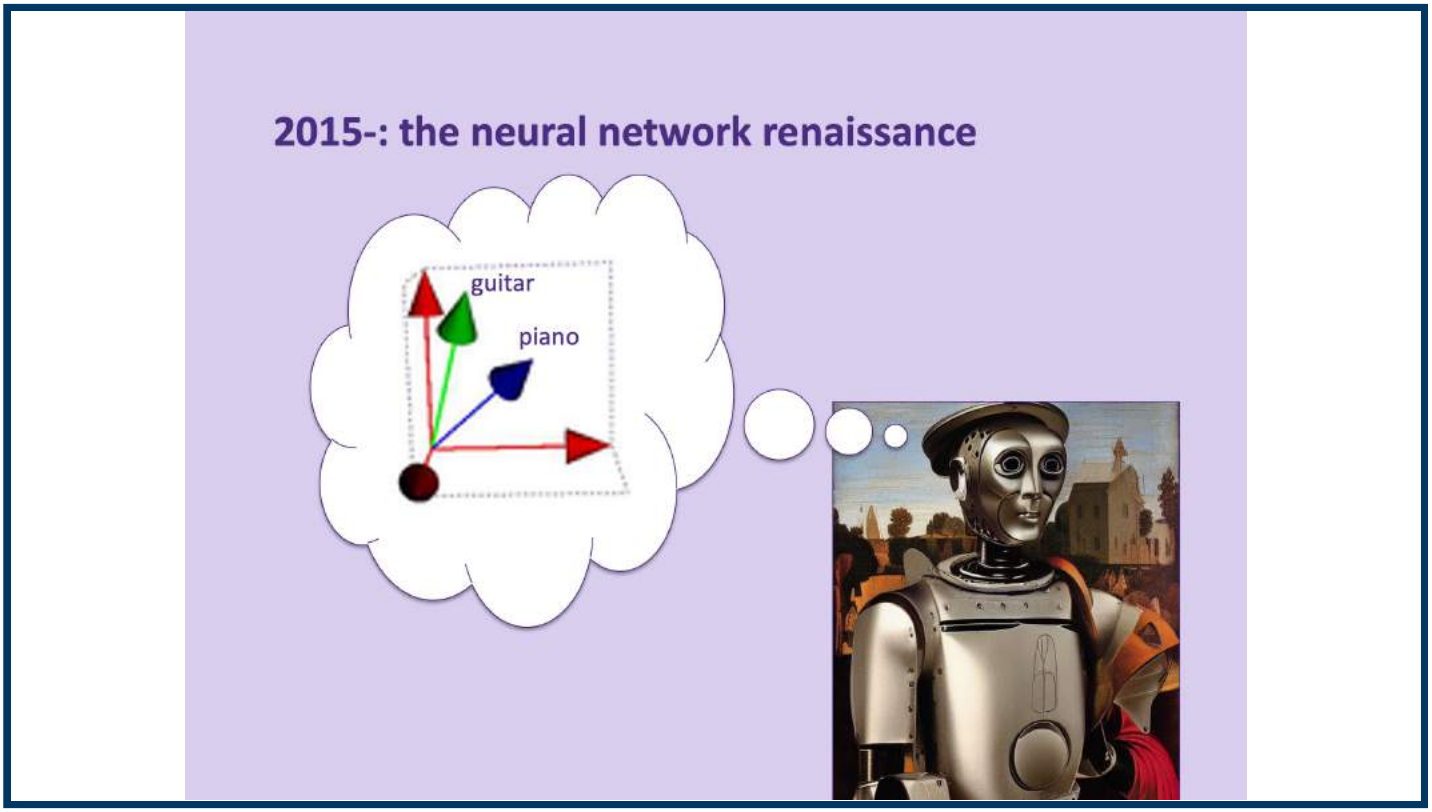

2000's: semantic interoperability
C0027051^myocardial infarction| [problem,'myocardial infarction']
C0027051^myocardial infarction| [problem,attack,[bodyloc,heart]]
C0027051^myocardial infarction| [problem,'myocardial infarction',
[problemdescr,syndrome]]
C0856742^post mi|[problem,'myocardial infarction',[status,post]]
C0155668^myocardial infarction old|[problem,'myocardial infarction',[status,previous]]
C0155668^myocardial infarction old|[problem,'heart attack',[status,previous]]
C0340293^myocardial infarction anterior|[problem,'myocardial infarction',
[region,anterior]]
C0746711^myocardial infarction anterior non q wave|[problem,'myocardial infarction',
[region,anterior],[descriptor,'non q wave']]
Friedman et al., JAMIA 2004
W

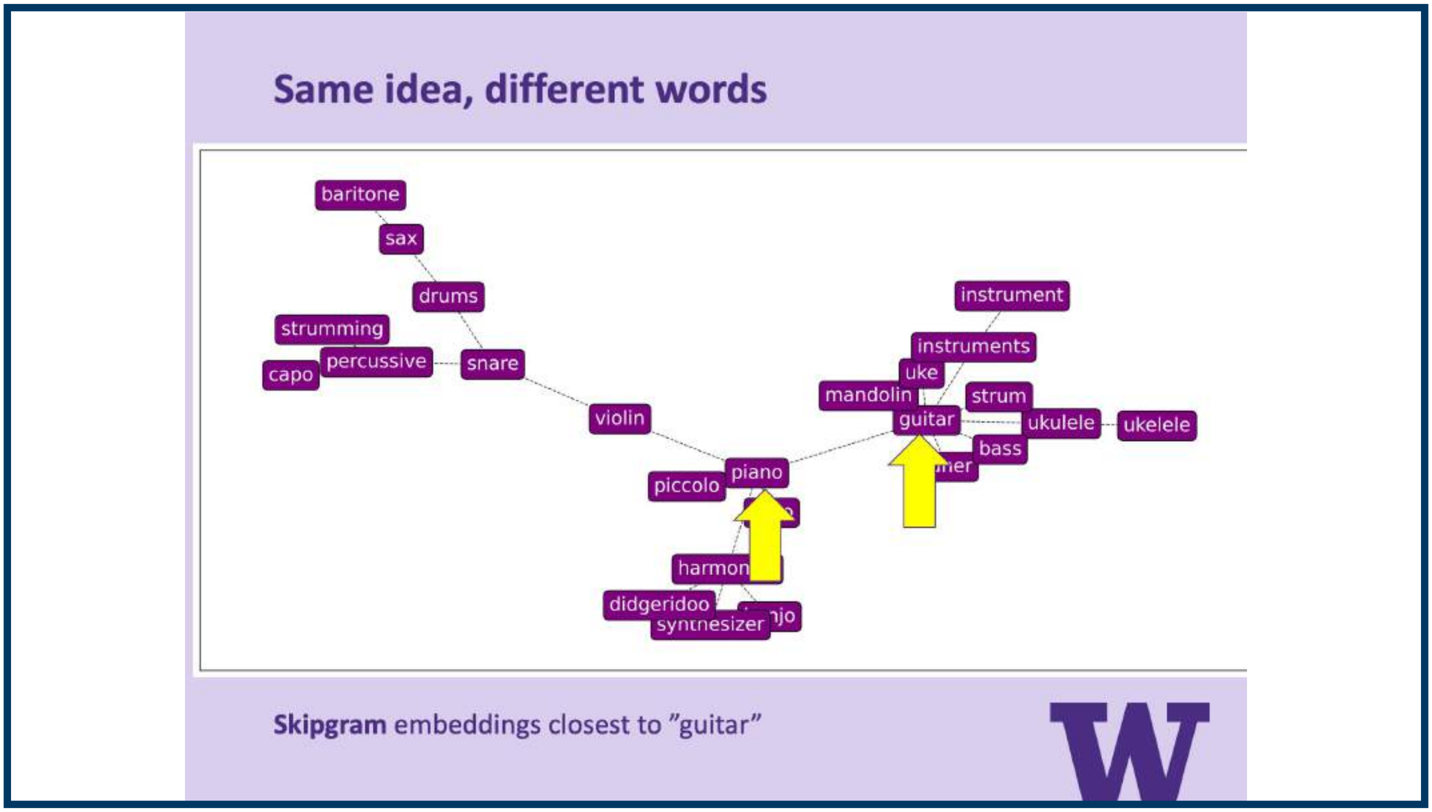

2010's: machine learning / shared tasks
Clinical notes
Feature engineering
Machine learning
Problem
Test
Treatment
words
word sequences
word clusters
concepts
Dictionaries
JAMIA Volume 18 Issue 5 September 2011
W

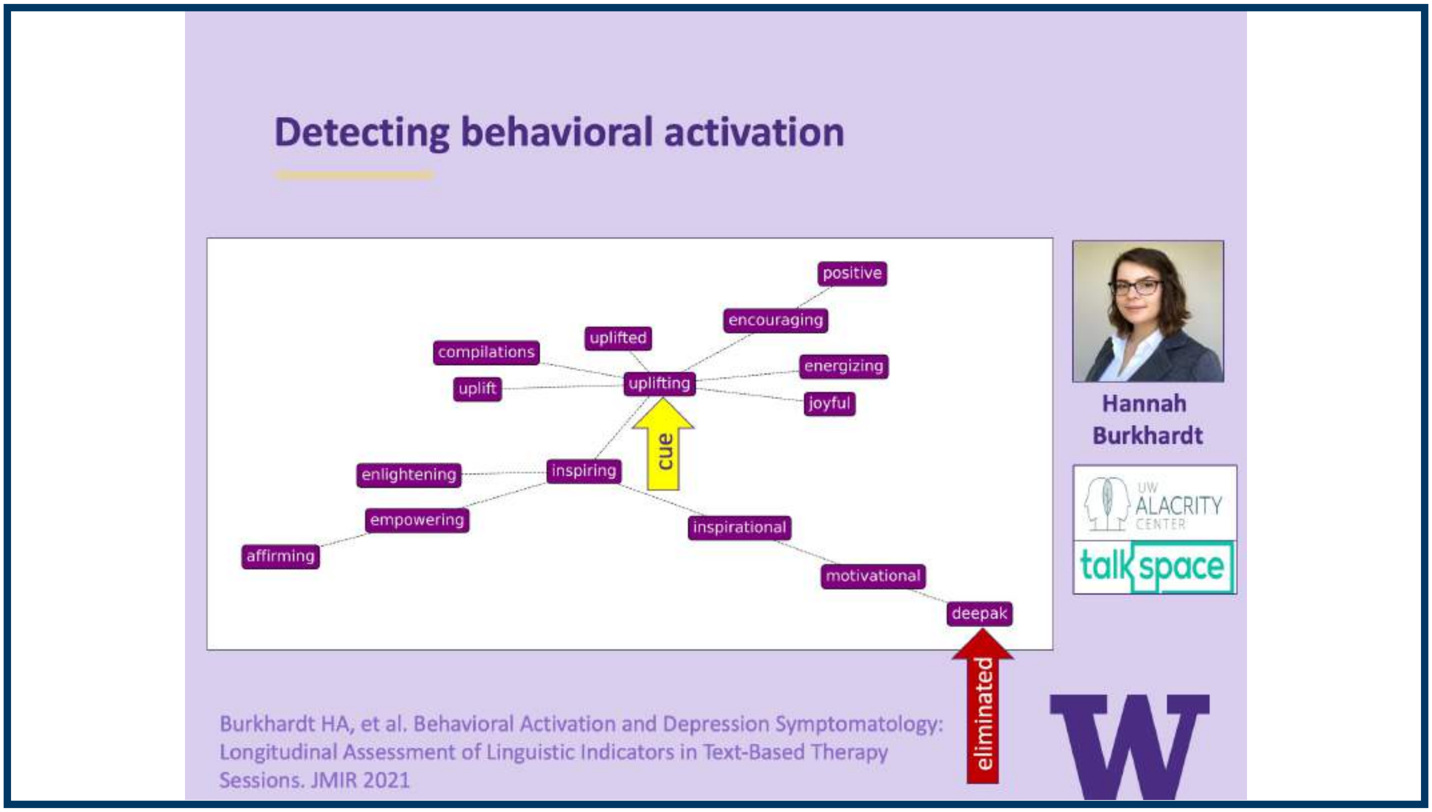

2015-: the neural network renaissance
W

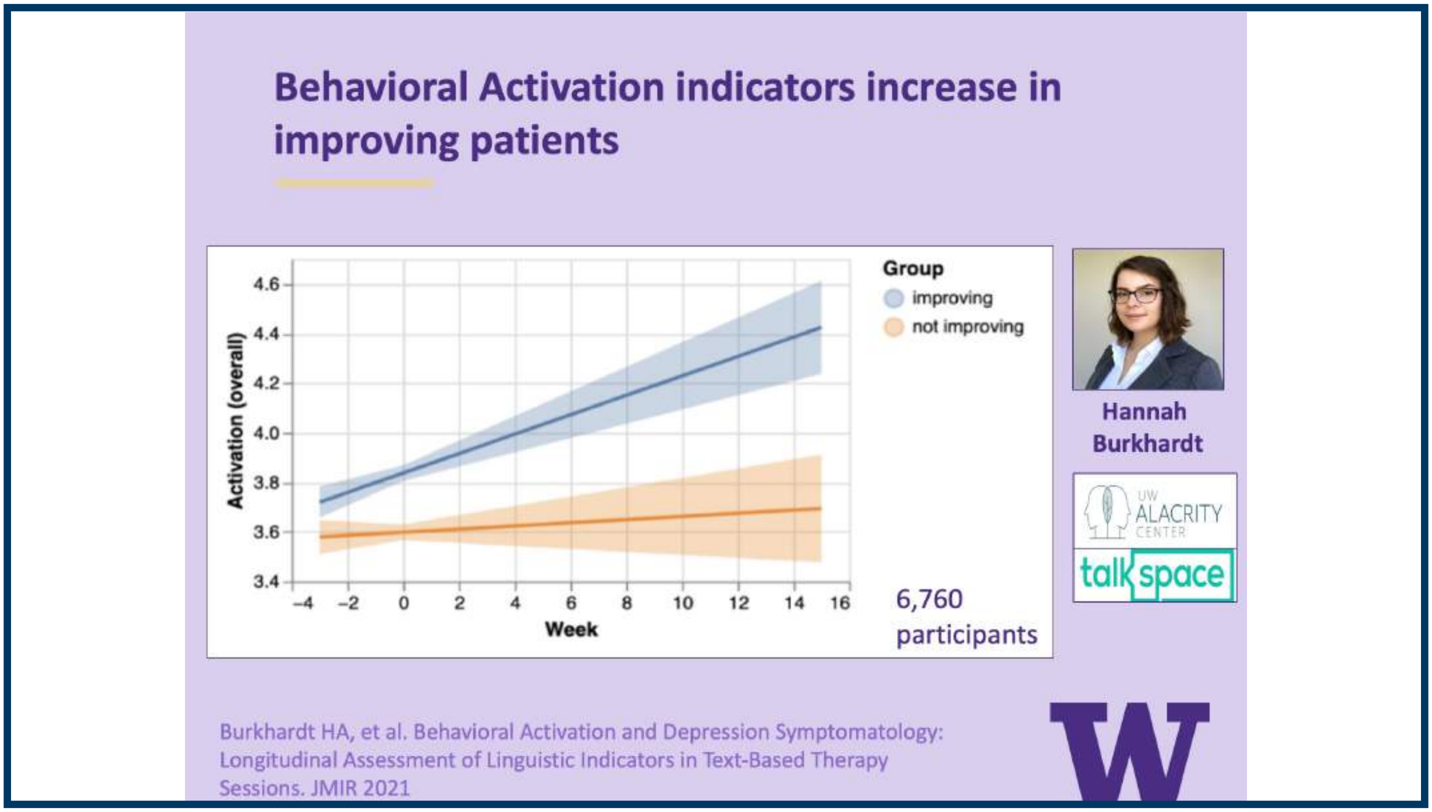

Learning to predict missing words
He liked to play the ______________

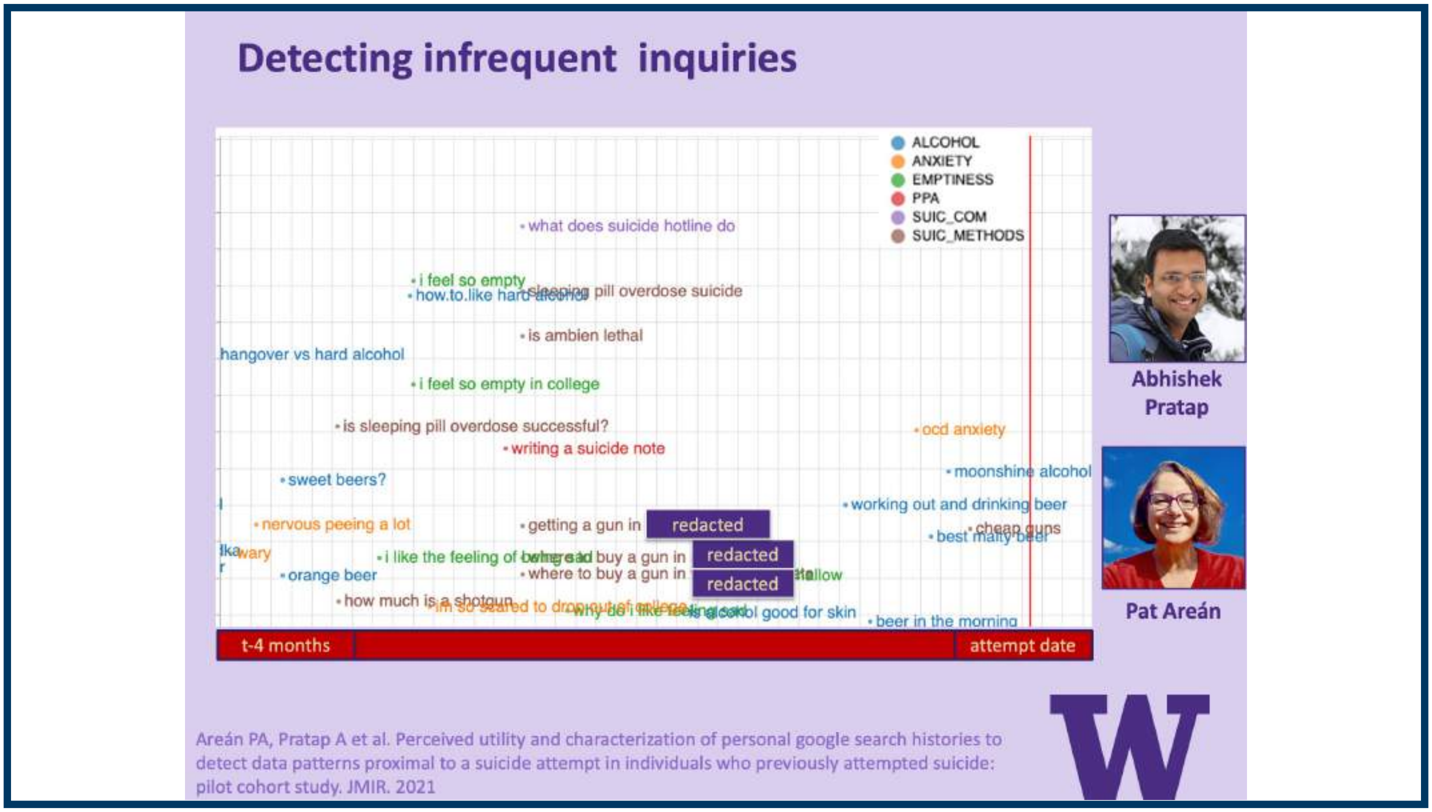

Learning to predict missing words
He liked to play the
word prob
piano 0.158321
game 0.124268
guitar 0.081497
role 0.054714
part 0.045360

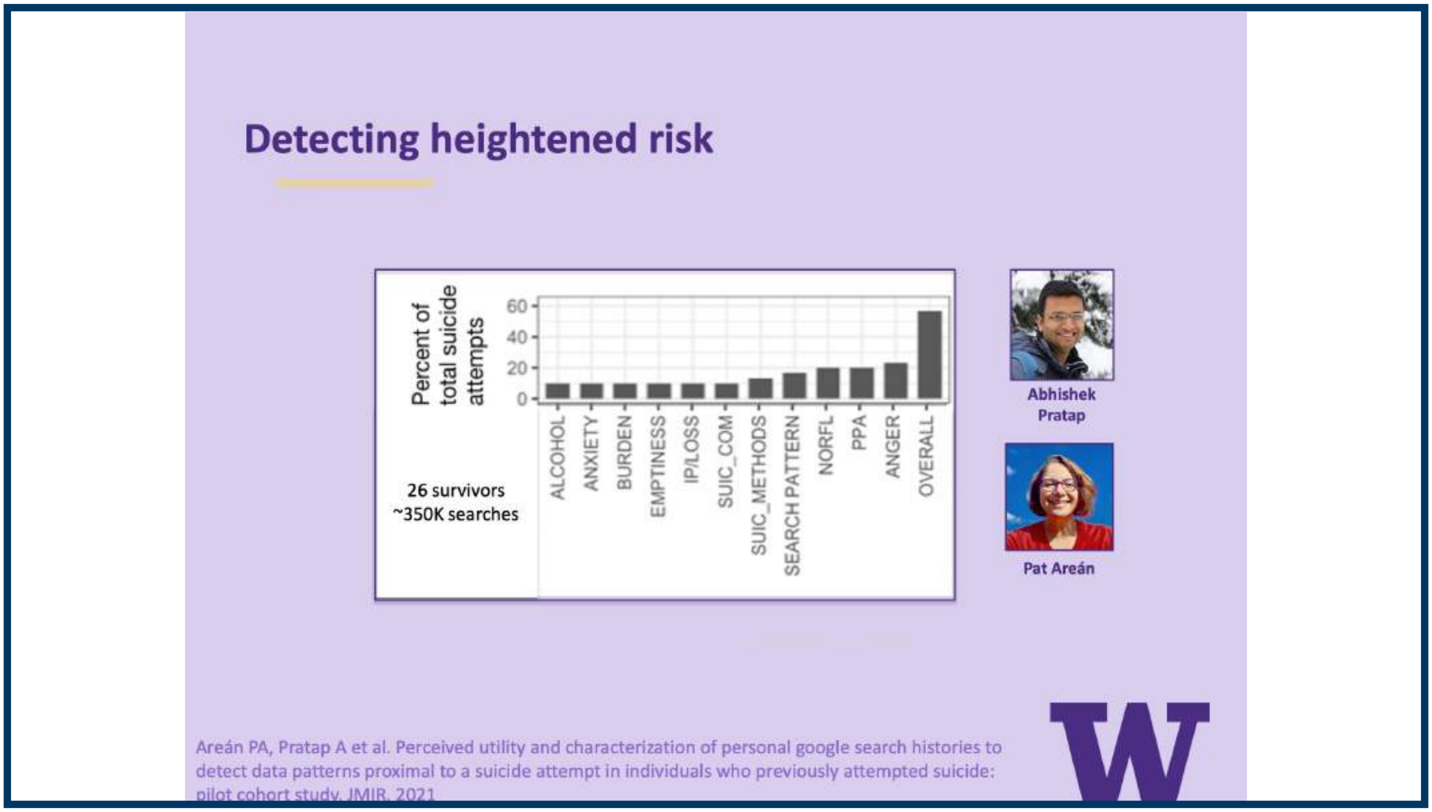

2015-: the neural network renaissance
guitar
piano

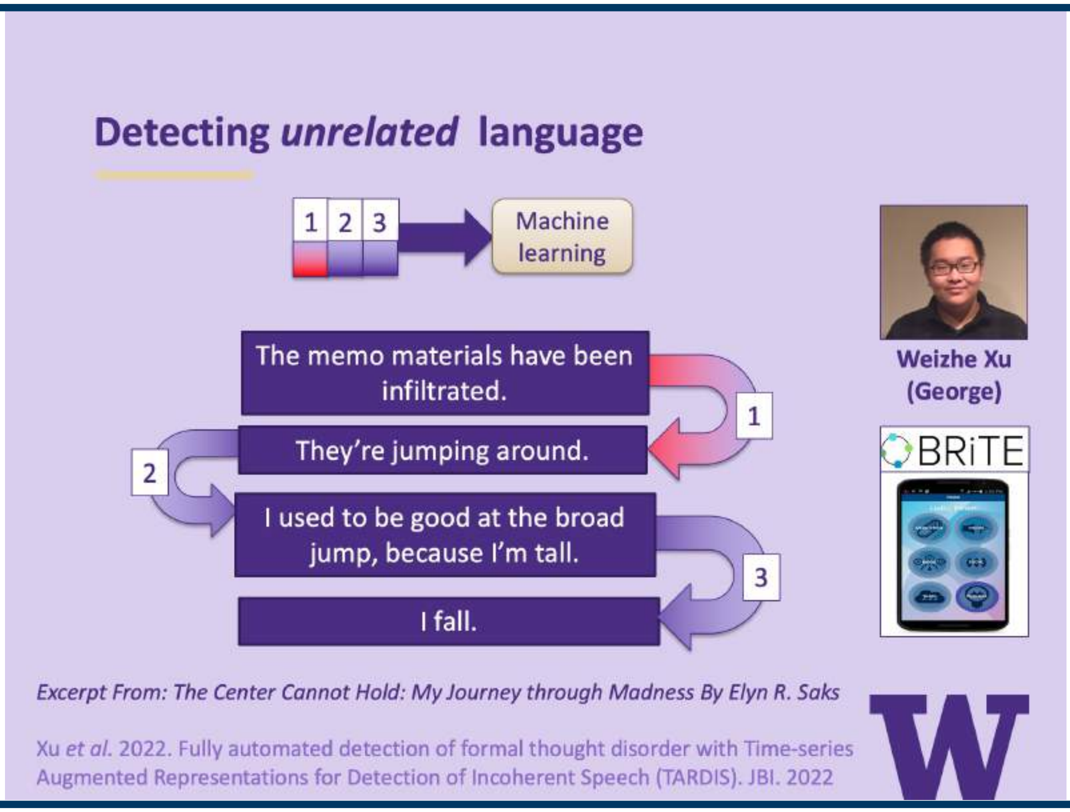

Same idea, different words
baritone
sax
drums
strumming
percussive
capo
snare
violin
instrument
instruments
uke
mandolin
strum
guitar
ukulele
ukelele
bass
piano
piccolo
didgeridoo
synthesizer
Skipgram embeddings closest to "guitar"

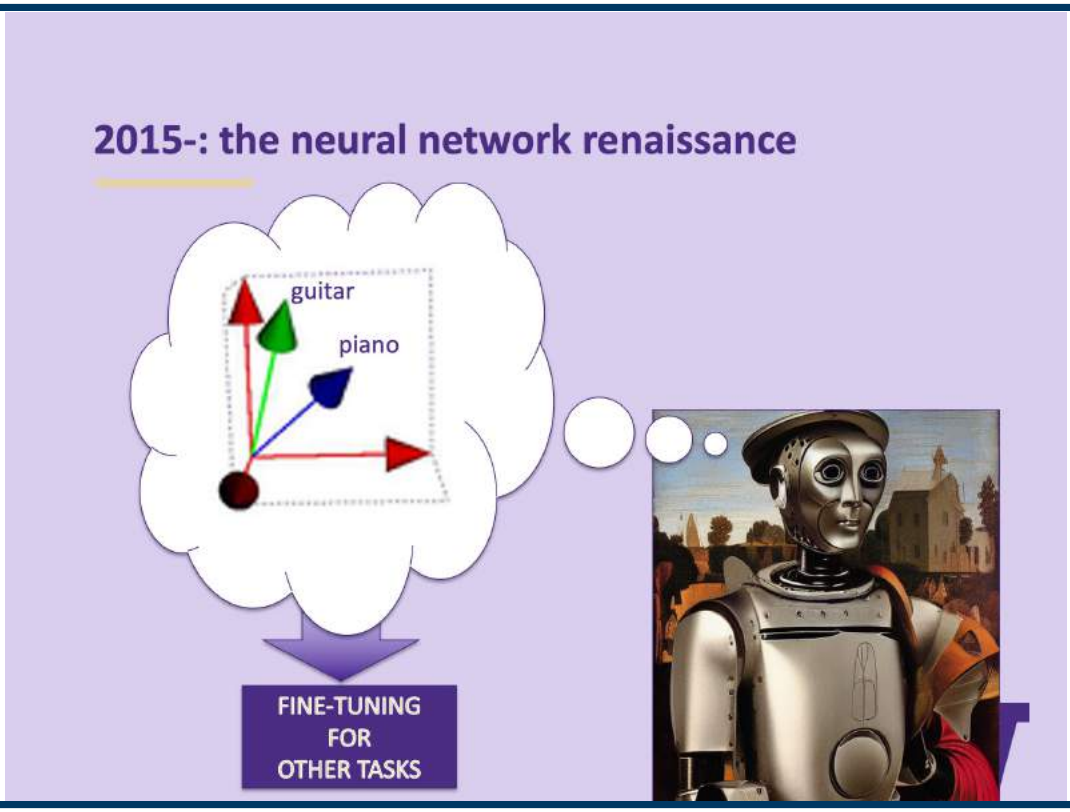

Detecting behavioral activation
positive
encouraging
uplifted
compilations
energizing
uplifting
uplift
joyful
cue
enlightening
inspiring
empowering
inspirational
affirming
motivational
deepak
eliminated
Hannah Burkhardt
UW ALACRITY CENTER
talkspace
Burkhardt HA, et al. Behavioral Activation and Depression Symptomatology: Longitudinal Assessment of Linguistic Indicators in Text-Based Therapy Sessions. JMIR 2021

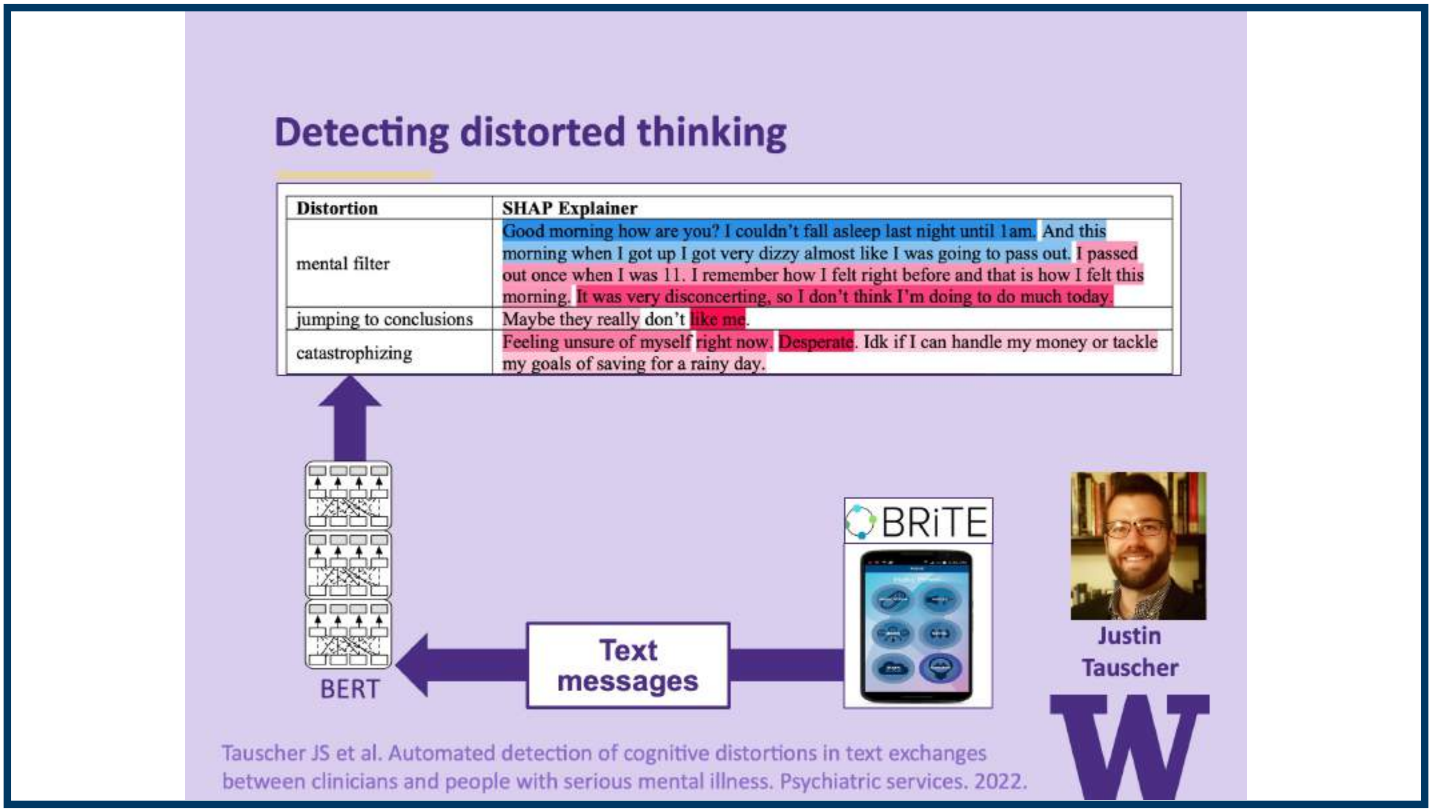

Behavioral Activation indicators increase in improving patients
Group
improving
not improving
Activation (overall)
4.6
4.4
4.2
4.0
3.8
3.6
3.4
−4 −2 0 2 4 6 8 10 12 14 16
Week
6,760 participants
Hannah Burkhardt
UW ALACRITY CENTER
talkspace
Burkhardt HA, et al. Behavioral Activation and Depression Symptomatology: Longitudinal Assessment of Linguistic Indicators in Text-Based Therapy Sessions. JMIR 2021


Detecting infrequent inquiries
ALCOHOL
ANXIETY
EMPTINESS
PPA
SUIC_COM
SUIC_METHODS
what does suicide hotline do
i feel so empty
sleeping pill overdose suicide
is ambien lethal
hangover vs hard alcohol
i feel so empty in college
is sleeping pill overdose successful?
writing a suicide note
ocd anxiety
sweet beers?
moonshine alcohol
working out and drinking beer
nervous peeing a lot
getting a gun in
redacted
cheap guns
where to buy a gun in
redacted
redacted
orange beer
how much is a shotgun
beer in the morning
t-4 months
attempt date
Abhishek Pratap
Pat Areán
Areán PA, Pratap A et al. Perceived utility and characterization of personal google search histories to detect data patterns proximal to a suicide attempt in individuals who previously attempted suicide: pilot cohort study. JMIR. 2021


Detecting heightened risk
Percent of total suicide attempts
60
40
20
0
ALCOHOL
ANXIETY
BURDEN
EMPTINESS
IP/LOSS
SUIC_COM
SUIC_METHODS
SEARCH PATTERN
NORFL
PPA
ANGER
OVERALL
26 survivors
~350K searches
Abhishek Pratap
Pat Areán
Areán PA, Pratap A et al. Perceived utility and characterization of personal google search histories to detect data patterns proximal to a suicide attempt in individuals who previously attempted suicide: pilot cohort study. JMIR. 2021

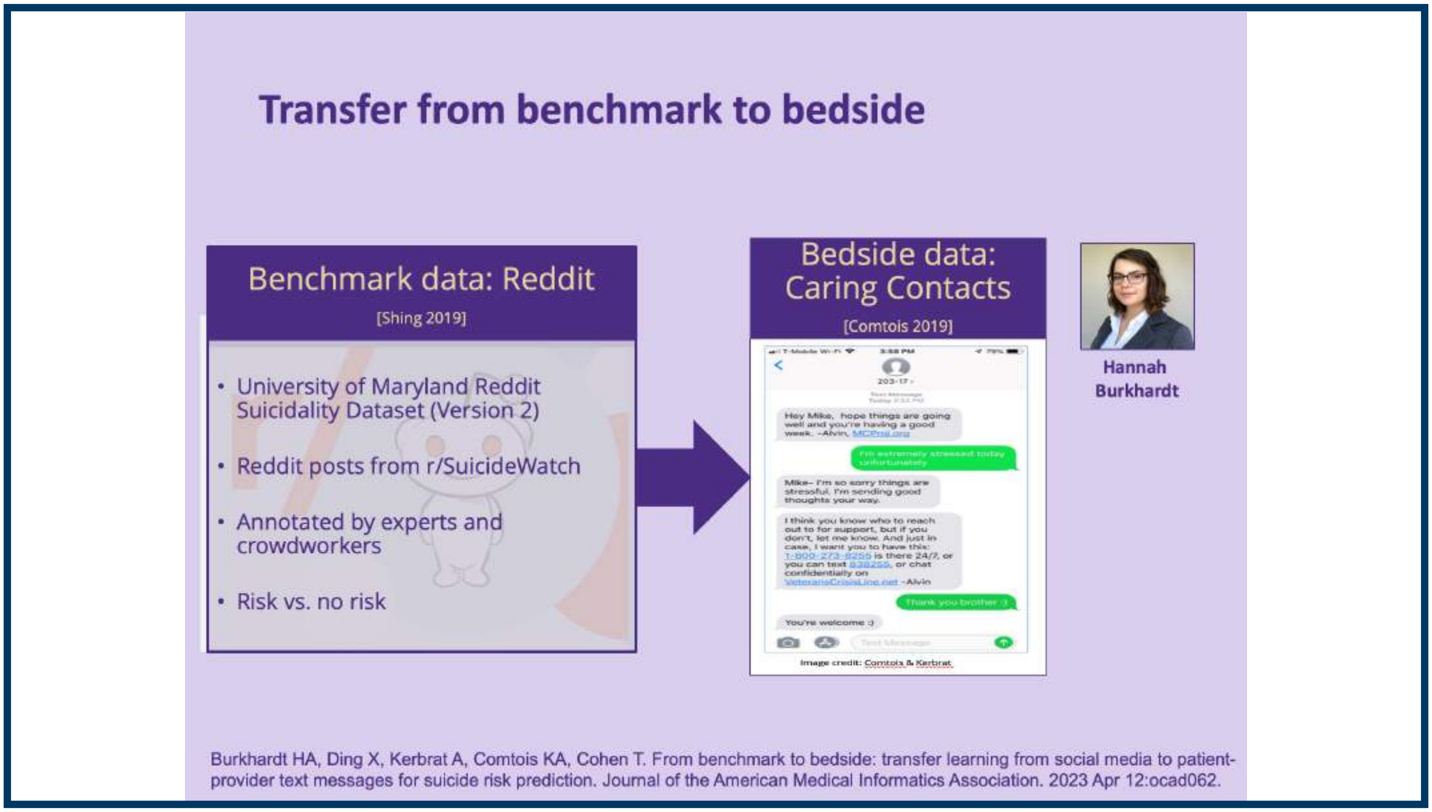

Detecting *unrelated* language
1 2 3
Machine learning
The memo materials have been infiltrated.
They're jumping around.
I used to be good at the broad jump, because I'm tall.
I fall.
1
2
3
Weizhe Xu (George)
BRiTE
Excerpt From: The Center Cannot Hold: My Journey through Madness By Elyn R. Saks
Xu et al. 2022. Fully automated detection of formal thought disorder with Time-series Augmented Representations for Detection of Incoherent Speech (TARDIS). JBI. 2022


2015-: the neural network renaissance
guitar
piano
FINE-TUNING FOR OTHER TASKS

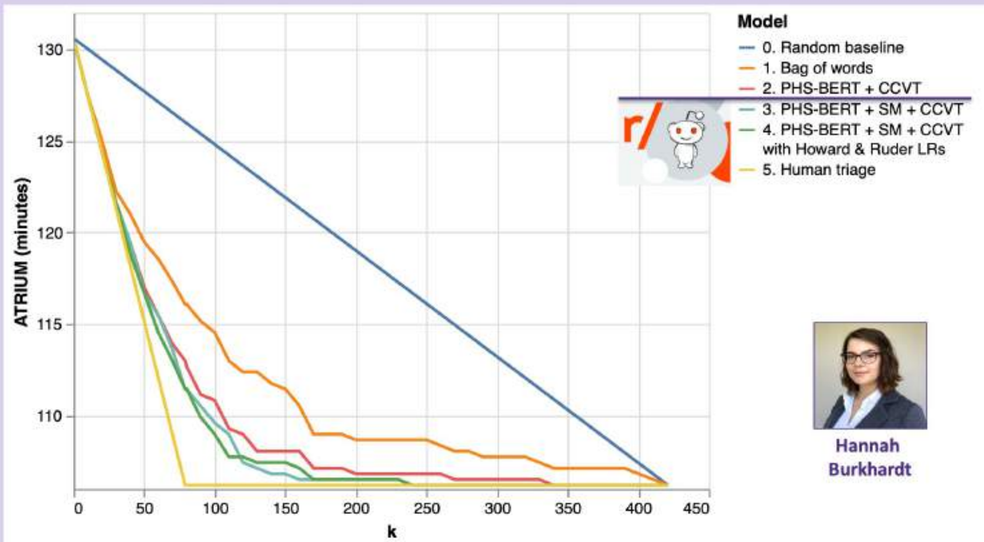


Detecting distorted thinking

Distortion	SHAP Explainer
mental filter	Good morning how are you? I couldn't fall asleep last night until 1am. And this morning when I got up I got very dizzy almost like I was going to pass out. I passed out once when I was 11. I remember how I felt right before and that is how I felt this morning. It was very disconcerting, so I don't think I'm doing to do much today.
jumping to conclusions	Maybe they really don't like me.
catastrophizing	Feeling unsure of myself right now. Desperate. Idk if I can handle my money or tackle my goals of saving for a rainy day.



From benchmark to bedside



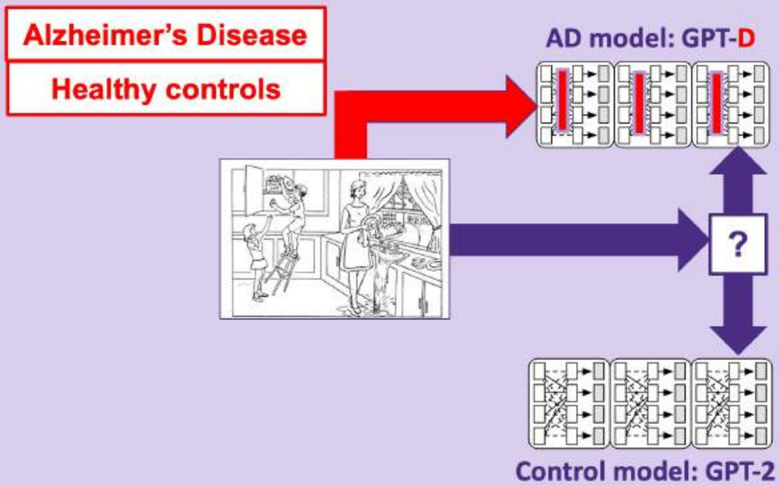
Burkhardt HA, Ding X, Kerbrat A, Comtois KA, Cohen T. From benchmark to bedside: transfer learning from social media to patient-provider text messages for suicide risk prediction. Journal of the American Medical Informatics Association. 2023 Apr 12:ocad062.

Learning to predict missing words

He liked to play the

word	prob
piano	0.158321
game	0.124268
guitar	0.081497
role	0.054714
part	0.045360

Detecting *improbable* language



Serguei Pakhomov

Changye Li

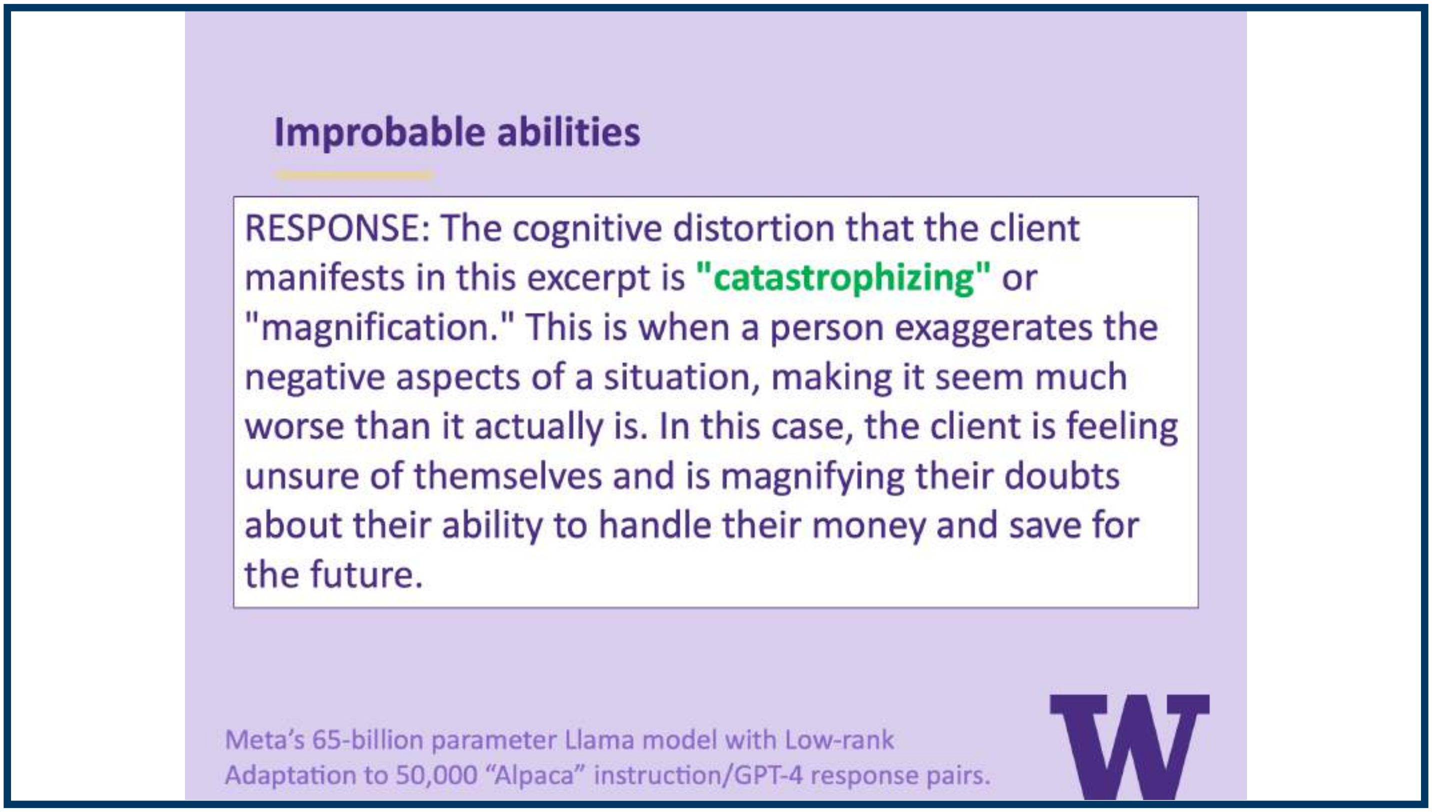
Li, Changye, et al. GPT-D: Inducing dementia-related linguistic anomalies by deliberate degradation of artificial neural language models. ACL 2022

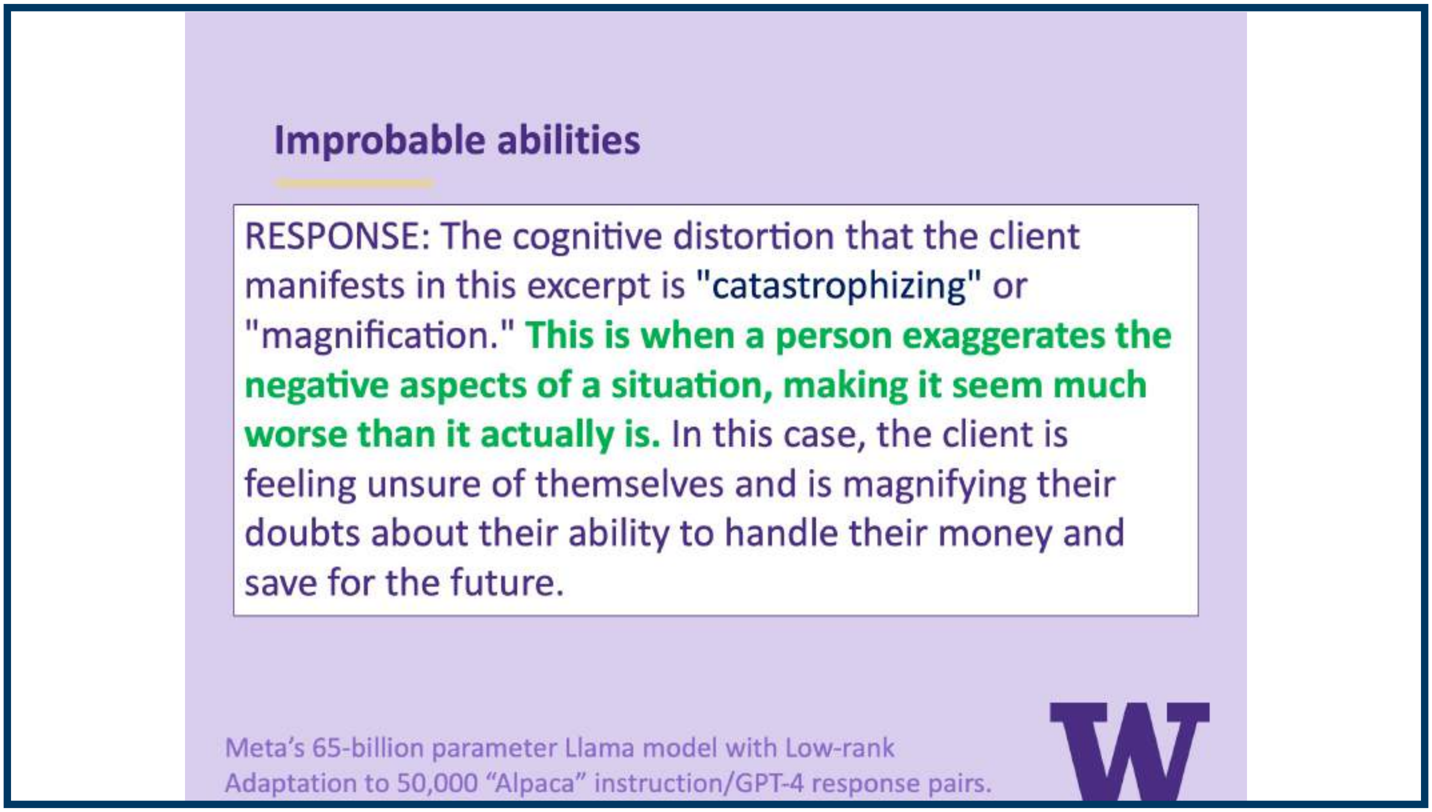
Improbable abilities

INSTRUCTION: "What cognitive distortion does the client manifest in the following excerpt?"

INPUT: "Feeling unsure of myself right now. Desperate. Idk if I can handle my money or tackle my goals of saving for a rainy day."

Meta's 65-billion parameter Llama model with Low-rank Adaptation to 50,000 "Alpaca" instruction/GPT-4 response


Improbable abilities
RESPONSE: The cognitive distortion that the client manifests in this excerpt is "catastrophizing" or "magnification." This is when a person exaggerates the negative aspects of a situation, making it seem much worse than it actually is. In this case, the client is feeling unsure of themselves and is magnifying their doubts about their ability to handle their money and save for the future.
Meta's 65-billion parameter Llama model with Low-rank Adaptation to 50,000 "Alpaca" instruction/GPT-4 response pairs.


Improbable abilities
RESPONSE: The cognitive distortion that the client manifests in this excerpt is "catastrophizing" or "magnification." This is when a person exaggerates the negative aspects of a situation, making it seem much worse than it actually is. In this case, the client is feeling unsure of themselves and is magnifying their doubts about their ability to handle their money and save for the future.
Meta's 65-billion parameter Llama model with Low-rank Adaptation to 50,000 "Alpaca" instruction/GPT-4 response pairs.

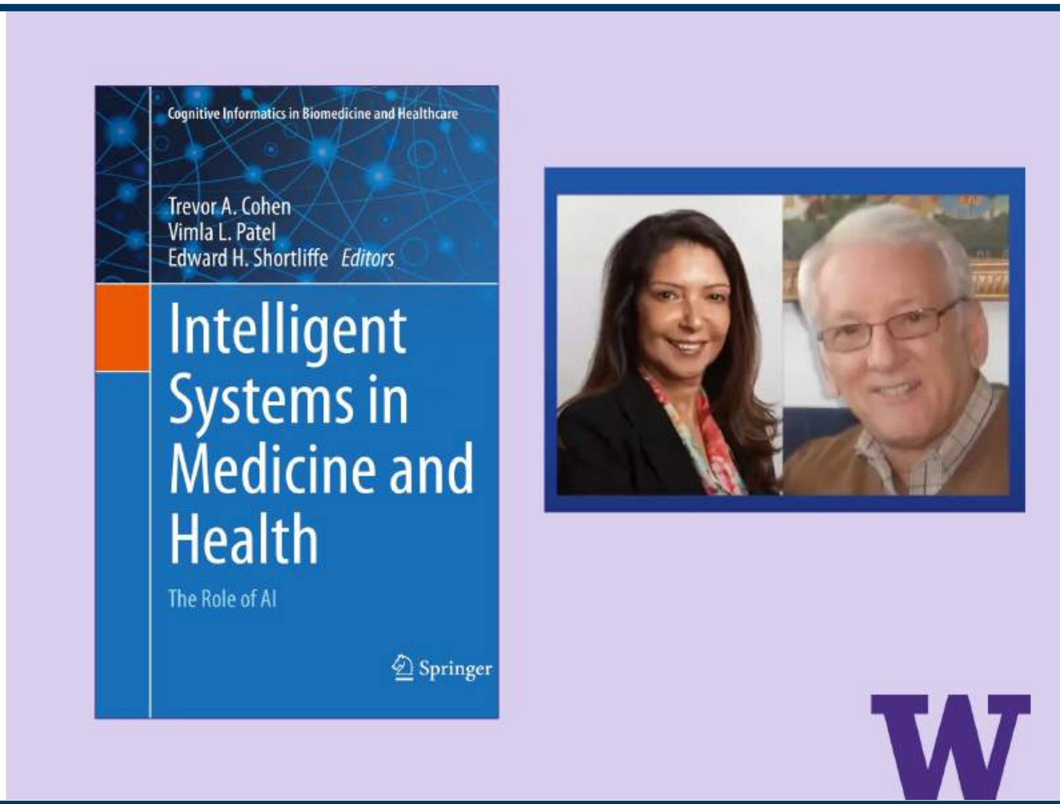
Improbable abilities

RESPONSE: The cognitive distortion that the client manifests in this excerpt is "catastrophizing" or "magnification." This is when a person exaggerates the negative aspects of a situation, making it seem much worse than it actually is. **In this case, the client is feeling unsure of themselves and is magnifying their doubts about their ability to handle their money and save for the future.**

Meta's 65-billion parameter Llama model with Low-rank Adaptation to 50,000 "Alpaca" instruction/GPT-4 response pairs.

Acknowledgments

- **Additional collaborators (with apologies)**
- **Support**
  - **Activation:** NLM 67-3780, R01LM011563, NIMH R01 MH102252, P50 MH113838, P50 MH 115837
  - **Coherence:** NIMH 3R01MH112641, 3R01MH112641-S2;
  - **Dementia:** NLM R01LM011563-S1, NIA AG069792
  - **Distortions:** Innovation Grant from the UW Medicine Garvey Institute for Brain Health Solutions
  - **Suicide risk:** R01MH123484-01A1, P50 MH115837, R01 MH 10230f4, R33 MH110509


Cognitive Informatics in Biomedicine and Healthcare
Trevor A. Cohen
Vimla L. Patel
Edward H. Shortliffe Editors
Intelligent Systems in Medicine and Health
The Role of AI
Springer
W

About Healthfirst

Healthfirst is New York's largest not-for-profit health insurer, earning the trust of 1.8 million members by offering access to affordable healthcare. Sponsored by New York City's leading hospitals, Healthfirst's unique advantage is rooted in its mission to put members first by working closely with its broad network of providers on shared goals. Healthfirst takes pride in being pioneers of the value-based care model, recognized as a national best practice. For nearly 30 years, Healthfirst has built its reputation in the community for top-quality products and services New Yorkers can depend on. It has grown significantly to serve the needs of members, offering market-leading products to fit every life stage, including Medicaid plans, Medicare Advantage plans, long-term care plans, qualified health plans, and individual and small group plans. Healthfirst serves members in New York City and on Long Island, as well as in Westchester, Sullivan, and Orange counties. For more information about Healthfirst, visit healthfirst.org.

About Healthfirst Advance

Healthfirst ADVANCE represents a culmination of Healthfirst's 30-year commitment to value-based care and the belief that health equity can only be achieved when organizations work together to tackle systemic barriers. Healthfirst's health equity programs are only possible because of the hard work and collaboration with its invaluable provider and community partners. Read case studies, first-person health equity views, and more at ADVANCE | Healthfirst.

Thank you for attending
Artificial Intelligence in Healthcare: Improving Health Outcomes and Advancing Equity with New Tools

healthfirst®
Health Insurance for New Yorkers

0363-23