Medical Imaging Informatics

William Hsu, PhD
Associate Professor of Radiology
Medical Imaging Informatics

David Geffen School of Medicine

http://willhsu.discoveryinformatics.org

@uclawillhsu

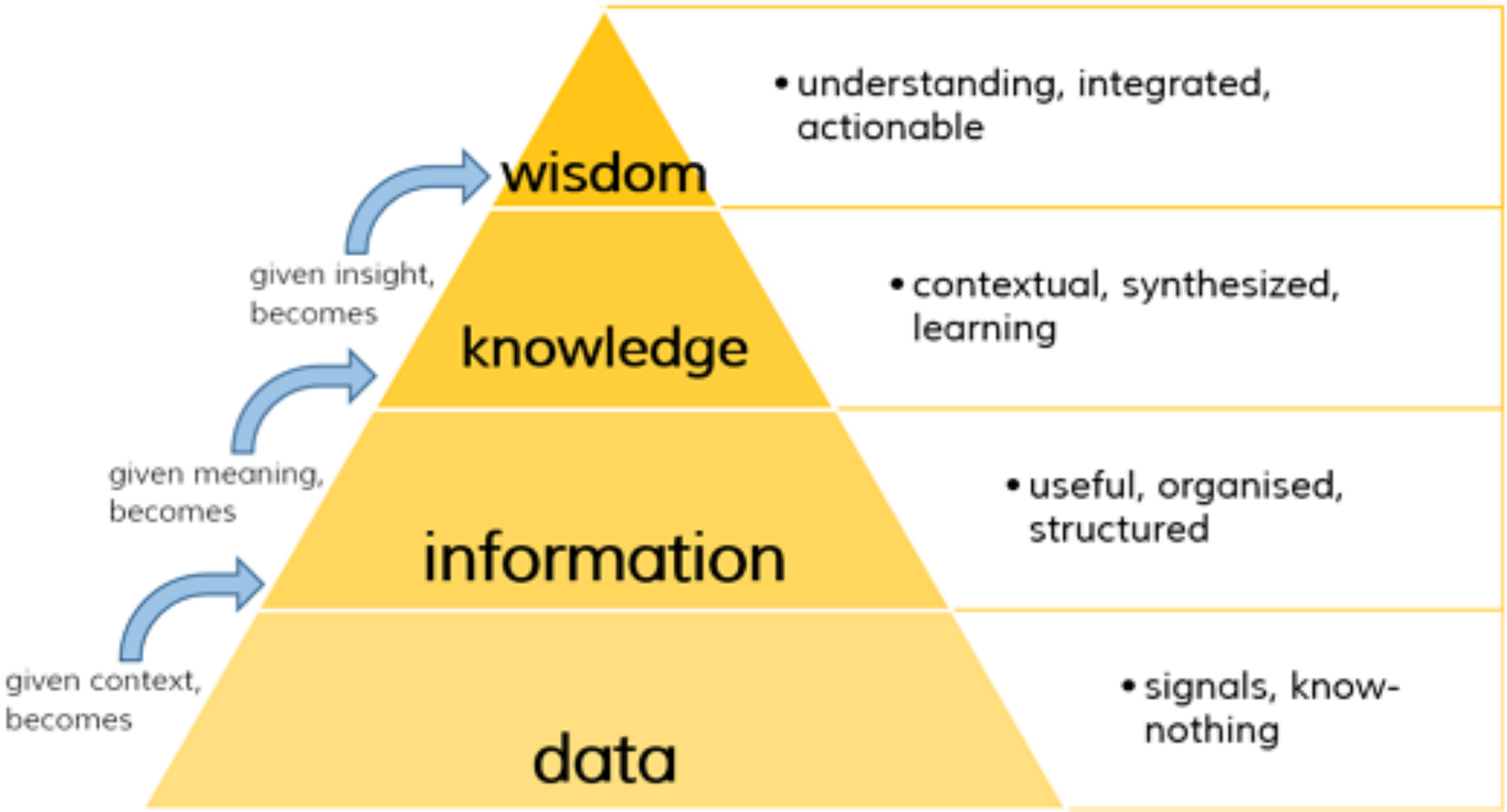
UCLA Health Radiology

What is Imaging Informatics?

Biomedical informatics is the scientific field that deals with biomedical *data*, *information*, and *knowledge* – their storage, retrieval, and optimal use for problem solving and decision making.

Figure 1. A "Fundamental Theorem" of informatics.

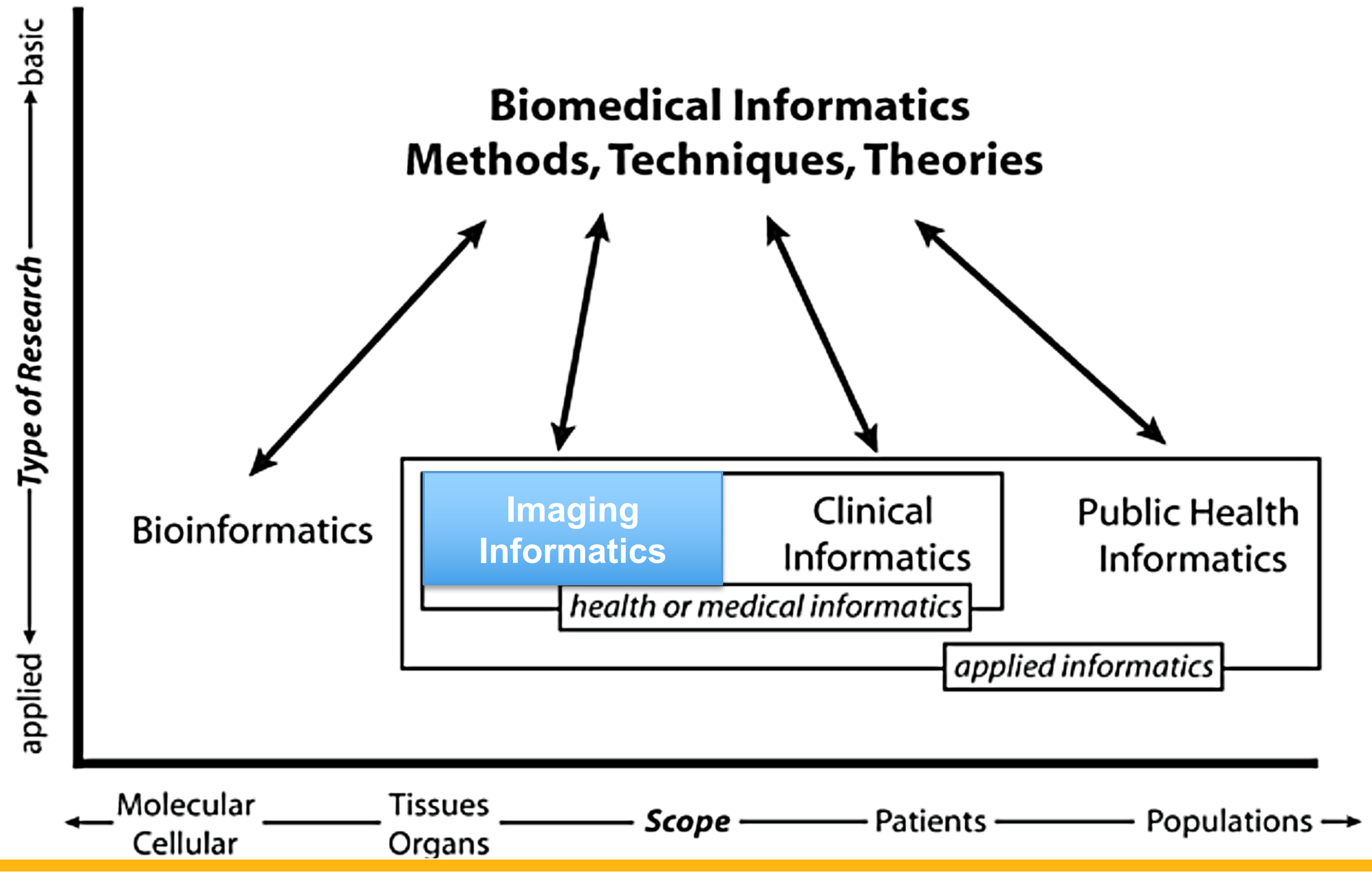
Ref: Friedman CP (2009). https://dx.doi.org/10.1197%2Fjamia.M3092


wisdom
knowledge
information
data
given insight, becomes
given meaning, becomes
given context, becomes
• understanding, integrated, actionable
• contextual, synthesized, learning
• useful, organised, structured
• signals, know-nothing

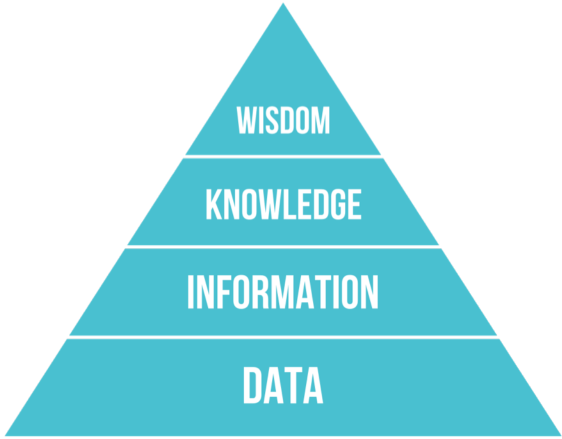
What is Imaging Informatics?

Biomedical informatics is the scientific field that deals with biomedical data, information, and knowledge – their storage, retrieval, and optimal use for problem solving and decision making.

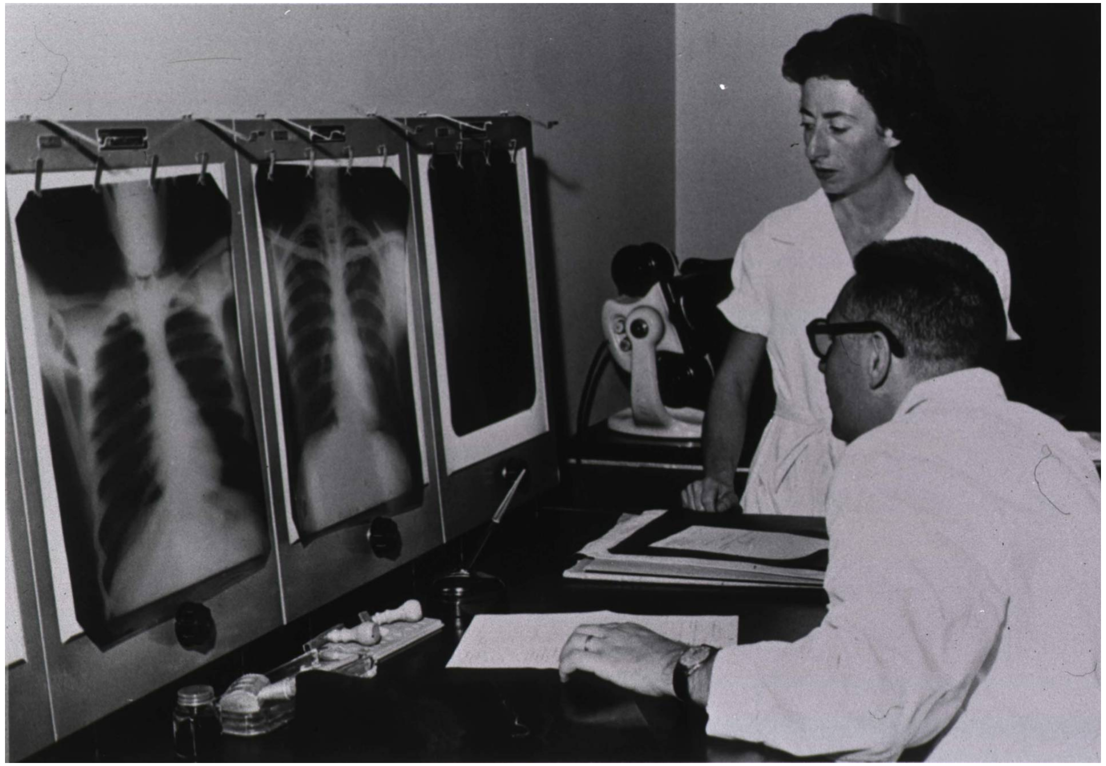
The primary aim of **imaging informatics** is to improve the efficiency, accuracy, usability, and reliability of medical imaging services within the healthcare enterprise.

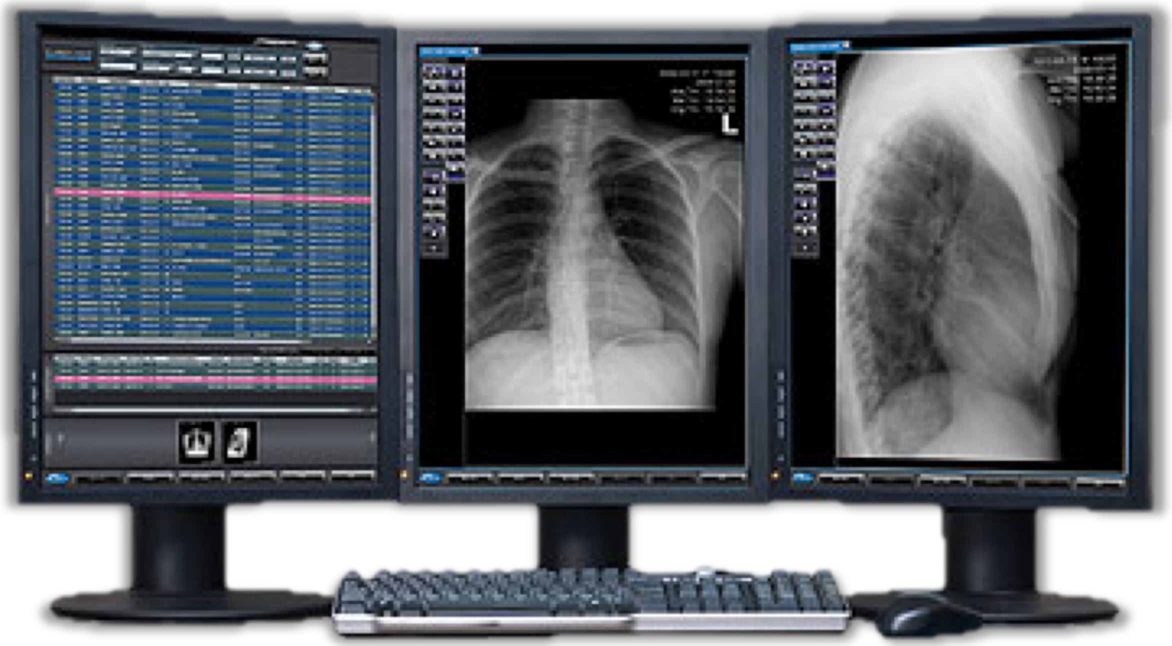

basic
Type of Research
applied
Biomedical Informatics
Methods, Techniques, Theories
Bioinformatics
Imaging Informatics
Clinical Informatics
Public Health Informatics
health or medical informatics
applied informatics
Molecular Cellular
Tissues Organs
Scope
Patients
Populations

Data

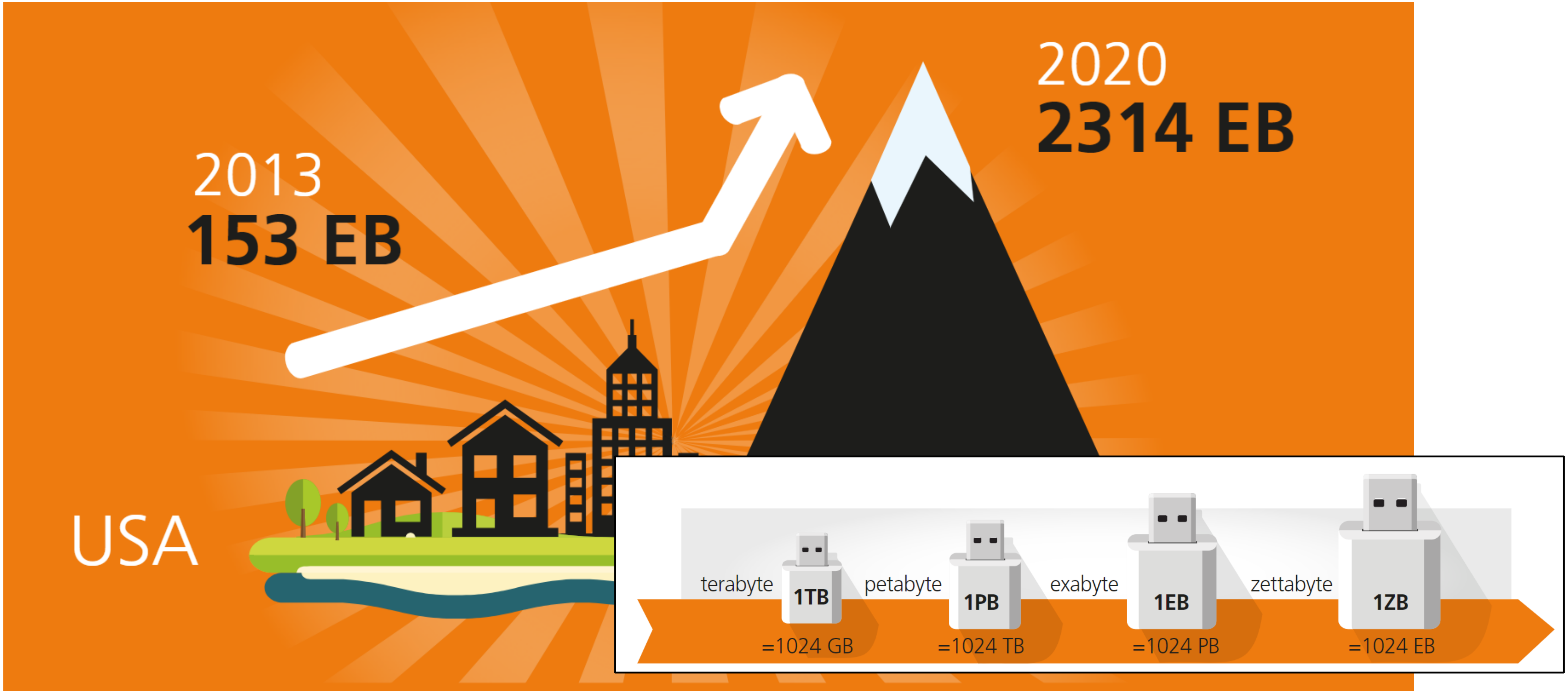


Radiology: Past and Present





Source: https://www.kqed.org/futureofyou/256816/how-technology-ruined-the-radiology-profession



2013
153 EB
2020
2314 EB
USA
terabyte
1TB
=1024 GB
petabyte
1PB
=1024 TB
exabyte
1EB
=1024 PB
zettabyte
1ZB
=1024 EB

Types of Information Systems

Hospital information system (HIS, CareConnect)

- Supports the comprehensive information requirements of hospitals an medical centers, including patient, clinical, ancillary, and financial management

Radiology information system (RIS, Radiant)

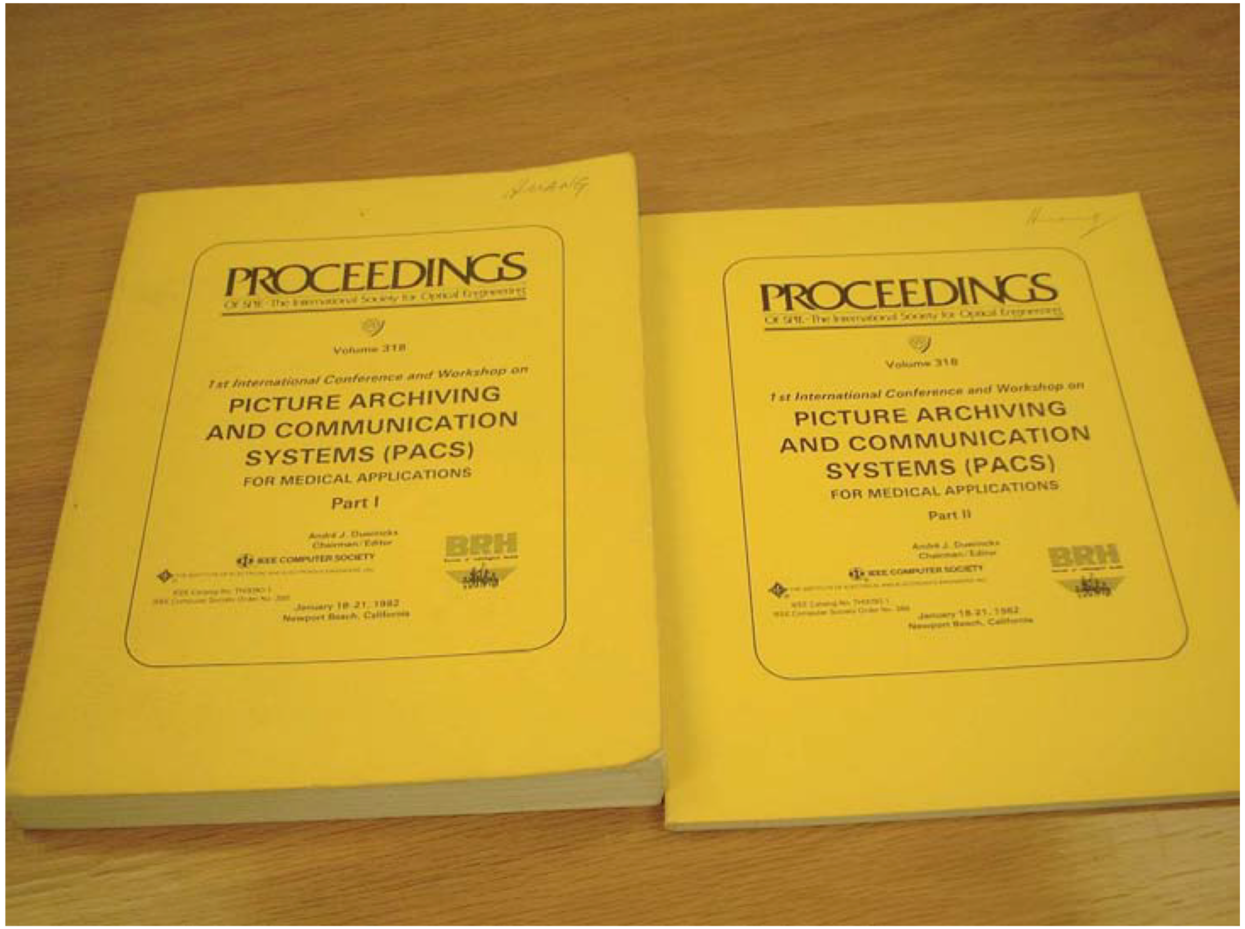
- Supports radiology department operations: scheduling exams, reporting results, billing

Picture archiving and communication system (PACS, Centricity)

- Acquires, stores, retrieves, and displays digital images

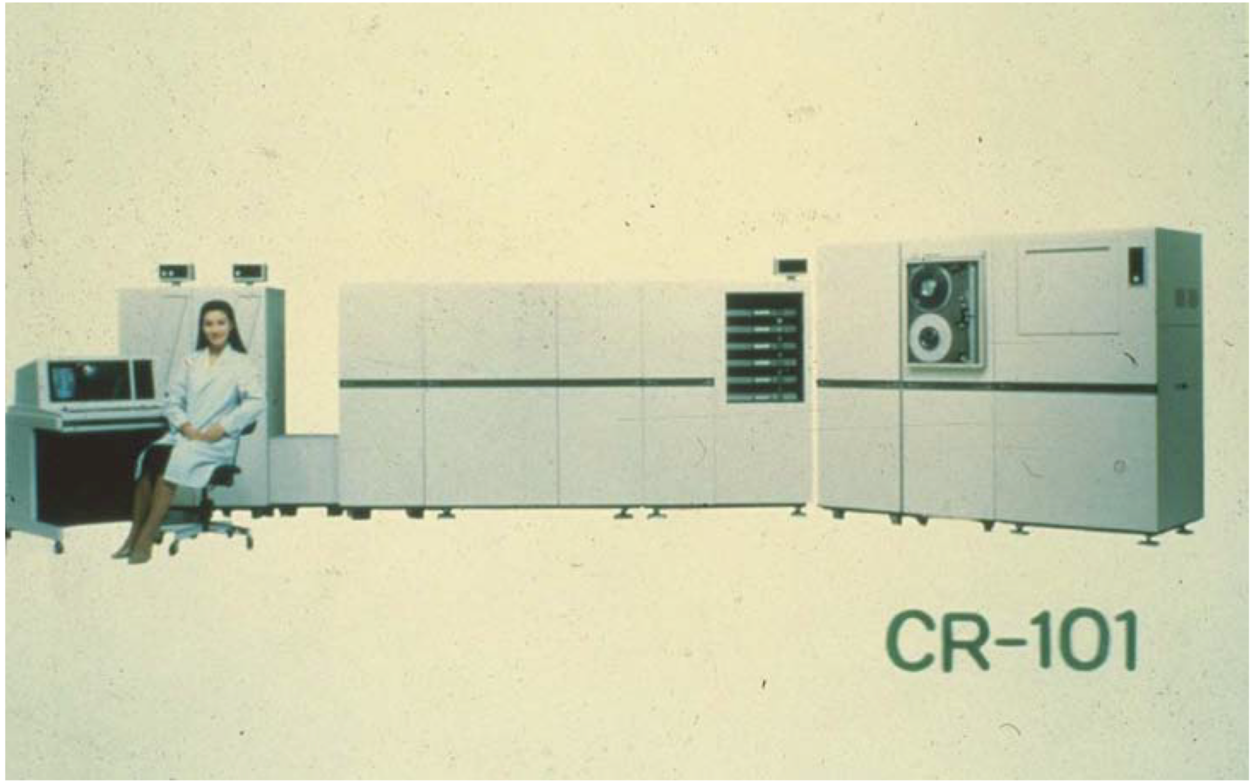
Clinical decision support system (CDSS, ACR Select…)

- Facilitates the integration and use of data in decision-making tasks

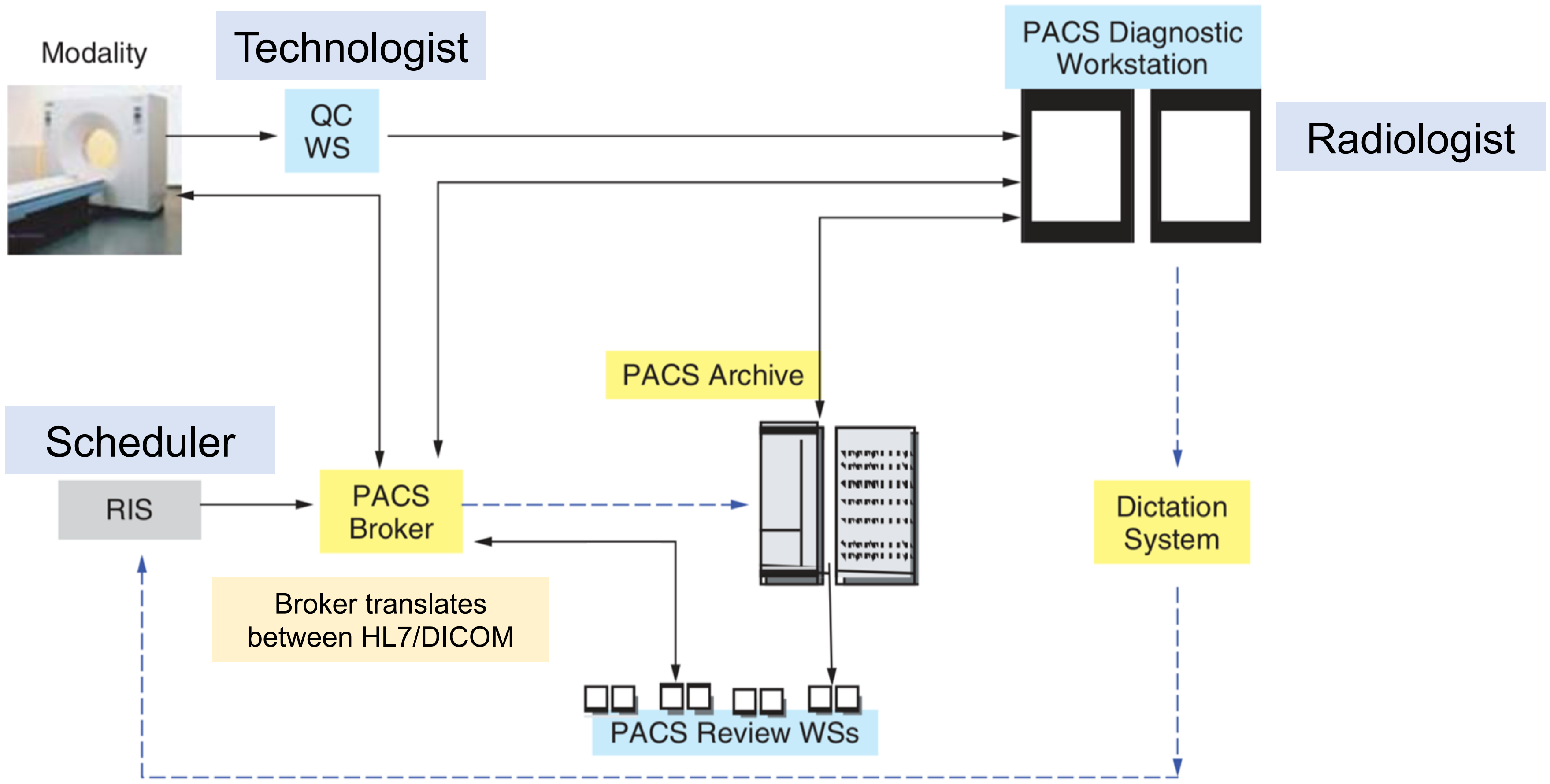


Proceedings of the first PACS specific meeting -- 1982

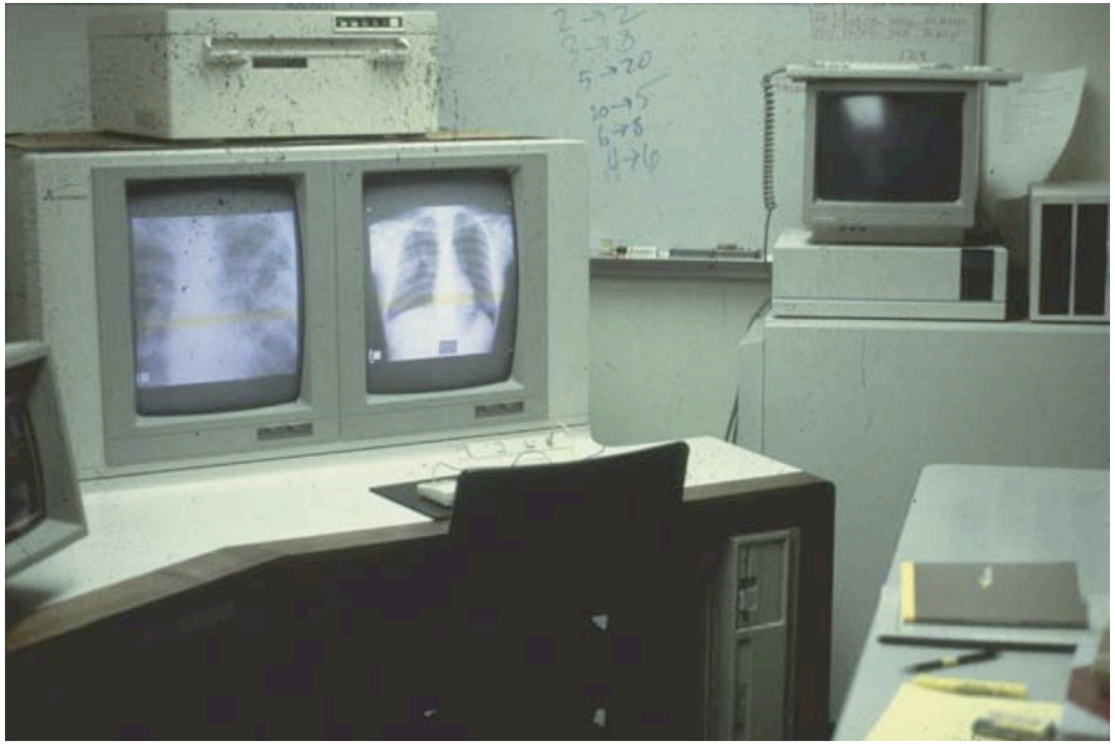
First commercial PACS in New Orleans (Fuji) -- 1985

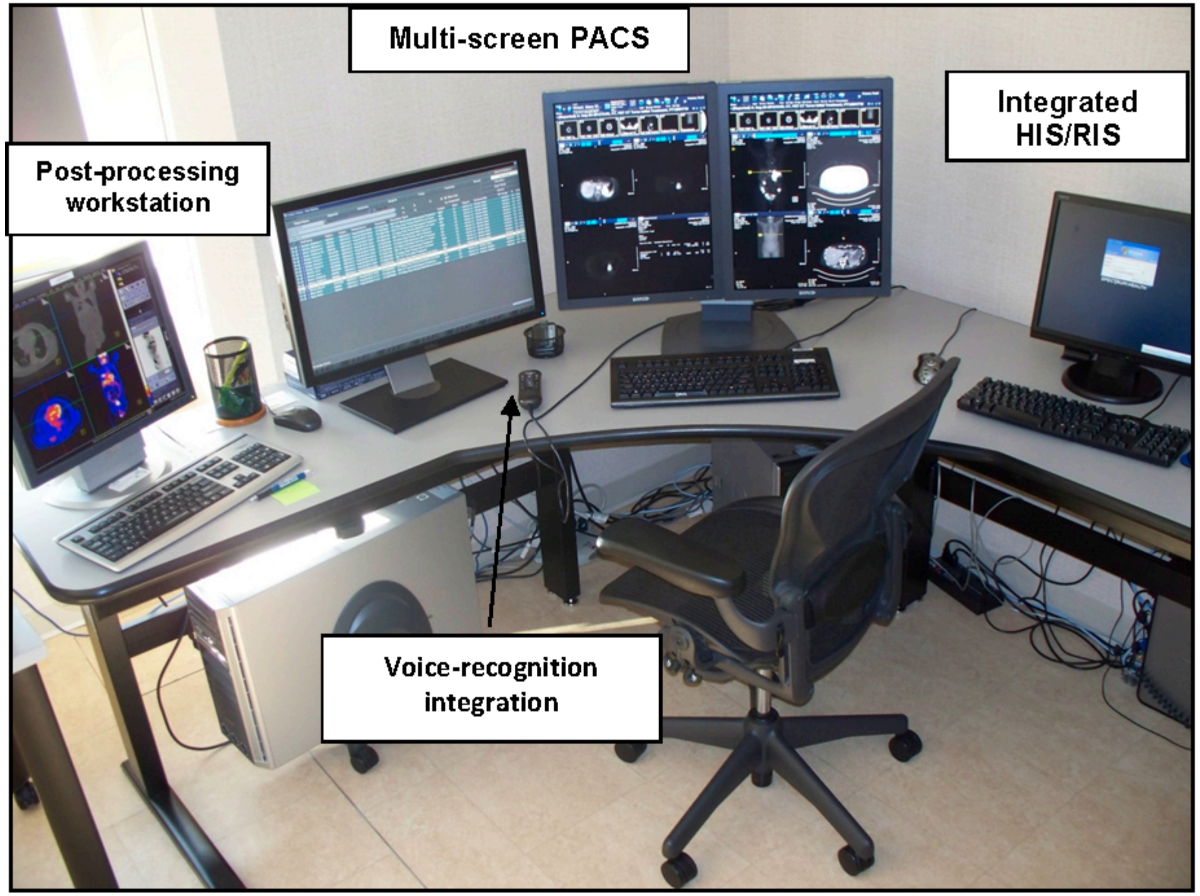


Basic PACS Workflow

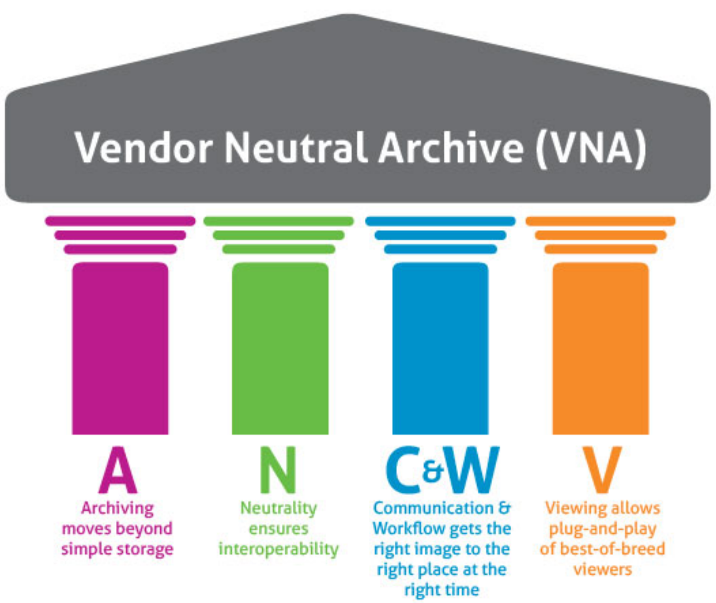


Evolution of PACS





Source: Faasse et al. Positron Emission Tomography-Computed Tomography Data Acquisition and Image Management



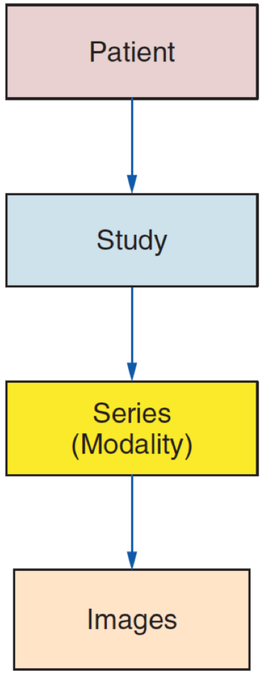
Messaging

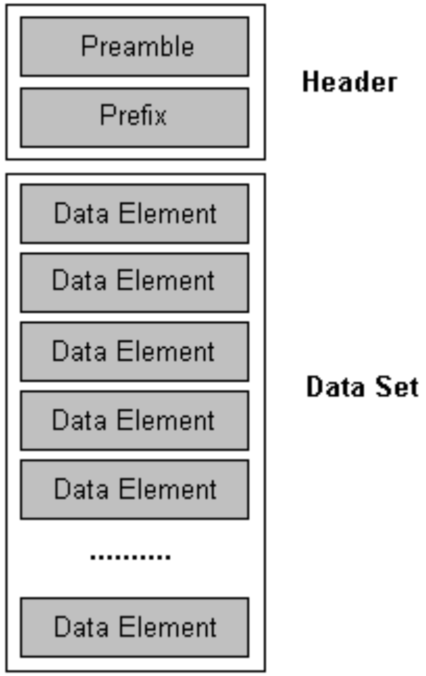
- HL7: Health Level 7
    - Standard used to send messages between the hospital information system (CareConnect) and ancillary systems (Radiant/PACS)

```
(1) Message header segment
MSH||STORE|HOLLYWOOD|MIME|VERMONT|200305181007|security|
ADT|MSG00201|||<CR>
(2) Event type segment
EVN|01|200305181005||<CR>
(3) Patient identification segment
PID|||PATID1234567||Doe^John^B^II||19470701|M||C|
```

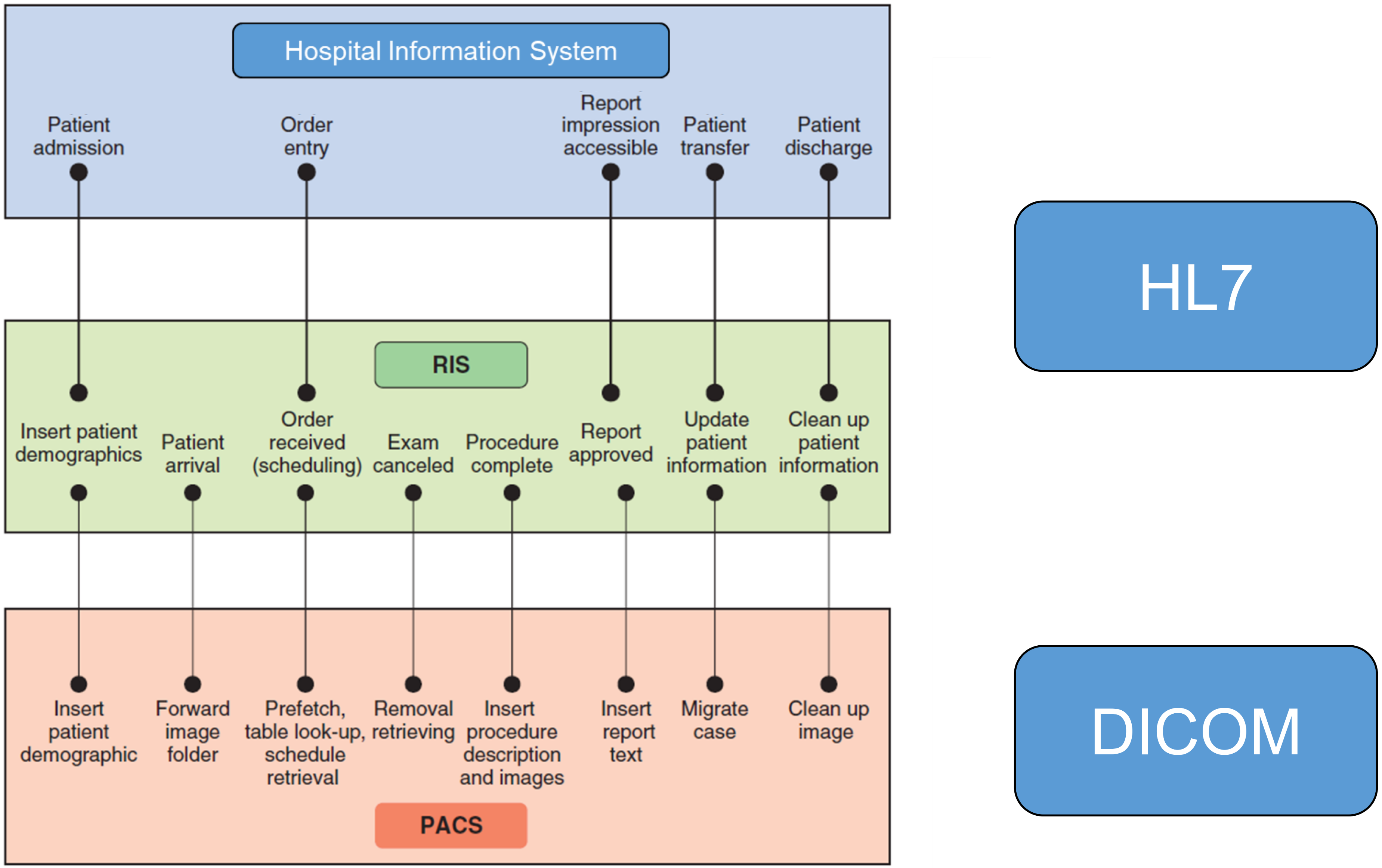
Messaging

- DICOM: Digital Imaging COmmunications in Medicine
  - Standard used to communicate among acquisition devices, PACS, and radiologist workstations





Group - Element	Description	Type	Length	Value
0002 0000	Group 0002 Length	UL	4	200
0002 0001	File Meta Information Version	OB	2	(binary data)
0002 0002	Media Storage SOP Class UID	UI	26	1.2.840.10008.5.1.4.1.1.7
0002 0003	Media Storage SOP Instance UID	UI	54	1.2.826.0.1.3680043.2.1208.34948545565201041914916296
0002 0010	Transfer Syntax UID	UI	18	1.2.840.10008.1.2
0002 0012	Implementation Class UID	UI	18	1.2.804.114118.3
0002 0013	Implementation Version Name	SH	6	
0002 0016	Source Application Entity Title	AE	18	
0002 0100	Private Information Creator UID	UI	0	(empty)
0002 0102	Private Information	OB	0	(binary data)
0008 0005	Specific Character Set	CS	12	ISO_IR 100
0008 0008	Image Type	CS	24	ORIGINAL\PRIMARY\OTHER
0008 0012	Instance Creation Date	DA	10	20100419
0008 0013	Instance Creation Time	TM	14	140916.000000
0008 0016	SOP Class UID	UI	26	1.2.840.10008.5.1.4.1.1.7
0008 0018	SOP Instance UID	UI	54	1.2.826.0.1.3680043.2.1208.34948545565201041914916296
0008 0020	Study Date	DA	10	20100419
0008 0021	Series Date	DA	10	20100419
0008 0030	Study Time	TM	14	140916.000000
0008 0031	Series Time	TM	14	140916.000000
0008 0032	Acquisition Time	TM	14	140916.000000
0008 0033	Content Time	TM	14	140916.000000

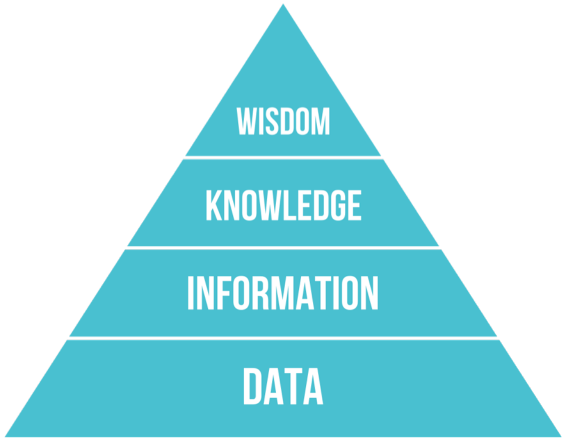


Adapted from: Huang, H. K. PACS and imaging informatics: basic principles and applications. John Wiley & Sons, 2011.

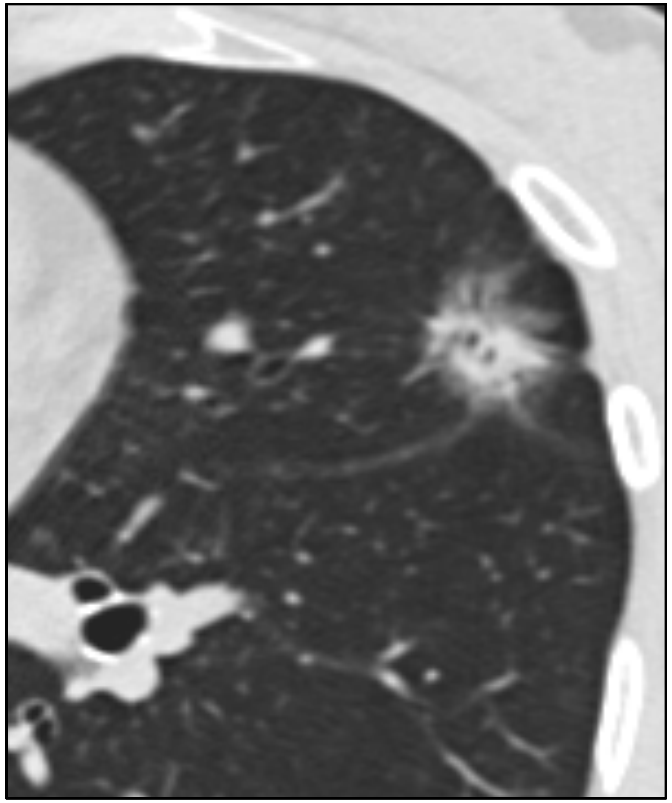
Takeaways

- Imaging informatics is important because
  - Volume of imaging data is rapidly increasing
  - Images contain a treasure trove of information
- Data is stored in large information systems (HIS/RIS/PACS)
- Coordination is required to prepare for and accept each imaging study → communication is possible through messaging standards (HL7, DICOM)
- Effective data management is the first step towards clinical decision support

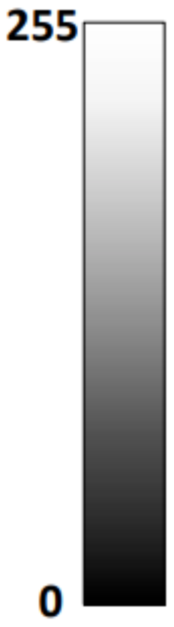
Information

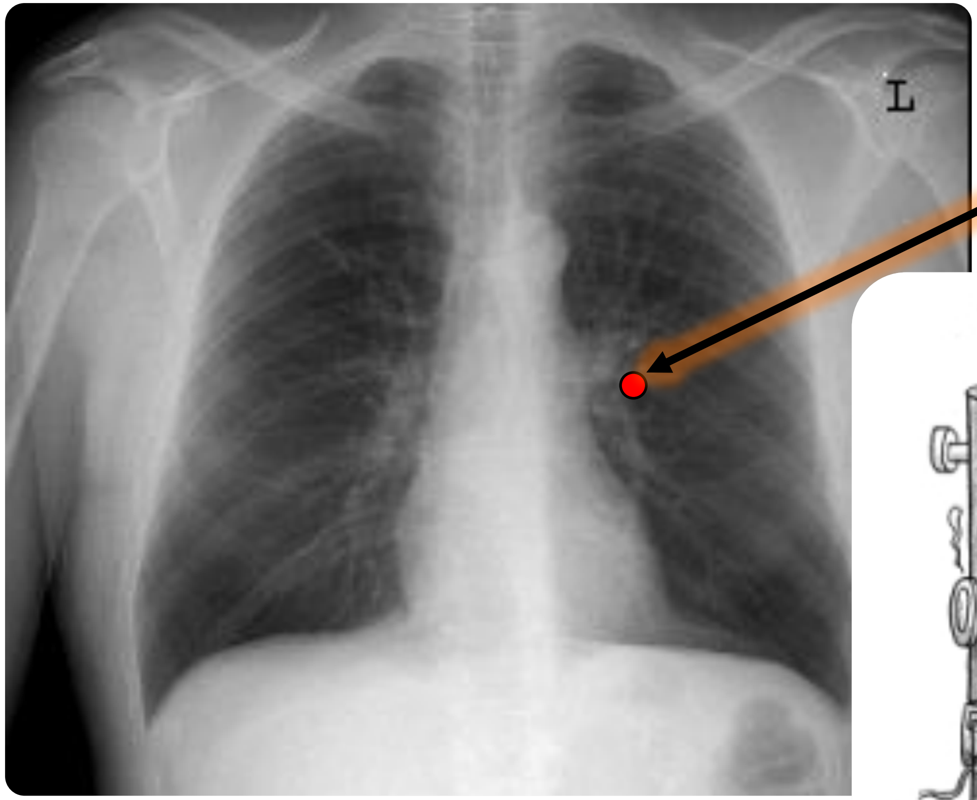


What is an Image?

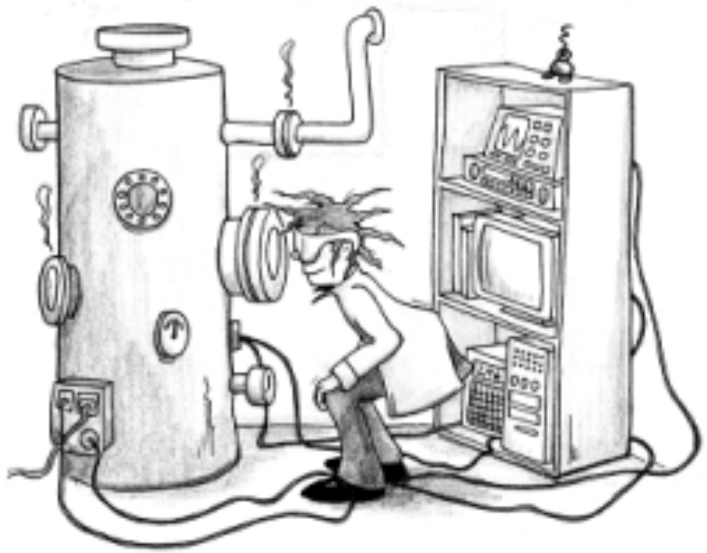


19	50	48	88	11	73	74	77	6	60
64	29	58	47	17	39	99	56	82	41
57	79	30	33	134	145	26	1	92	43
76	8	70	100	184	173	156	176	51	8
62	5	75	118	176	189	189	163	49	74
68	79	8	38	103	127	110	164	7	14
86	35	13	12	198	108	57	61	3	32
72	14	38	29	91	28	39	49	87	3
45	85	23	98	65	84	26	71	32	59
44	85	32	96	53	48	51	76	87	12

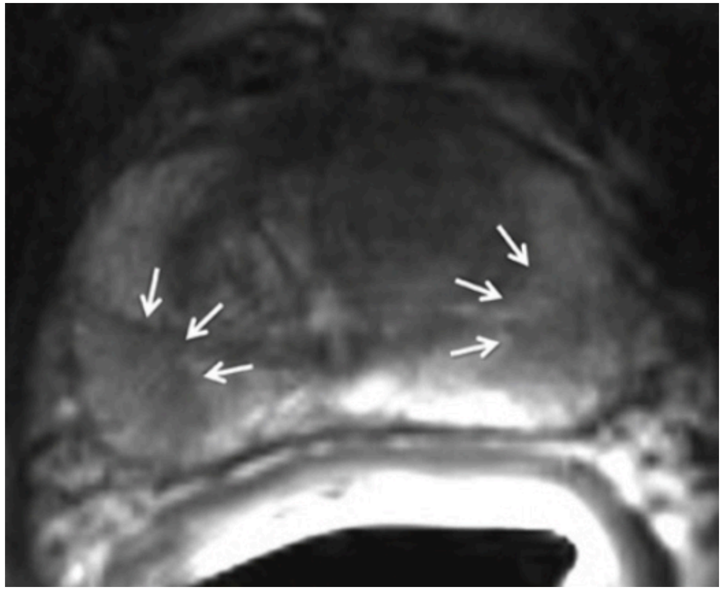




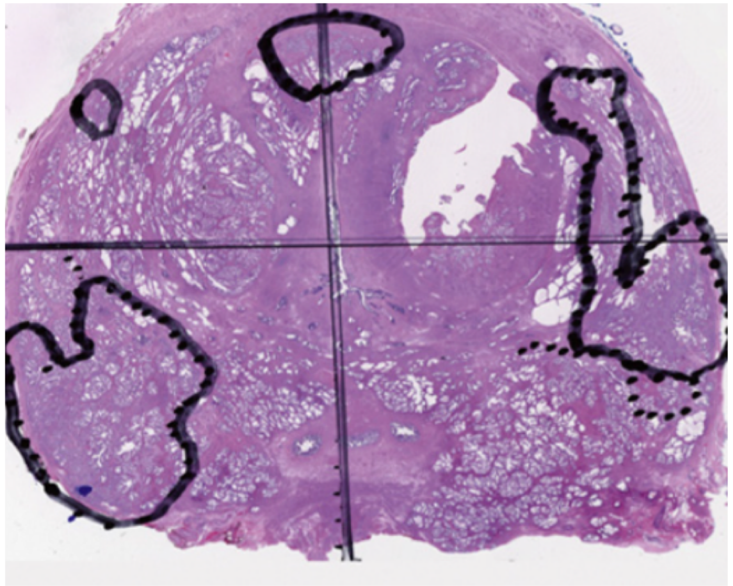
What does one pixel represent?



What effect are you observing?



T2-MRI



Whole-mount Path

62-yr-old male with intermediate-risk prostate cancer (biopsy Gleason score 4 + 3, PSA level 11 ng/ml)

- Pixel data
  - What does it look like?
- Information content
  - What is it?
- Knowledge → Wisdom
  - What does it mean?
  - What should we do about it?

Image Analysis Pipeline

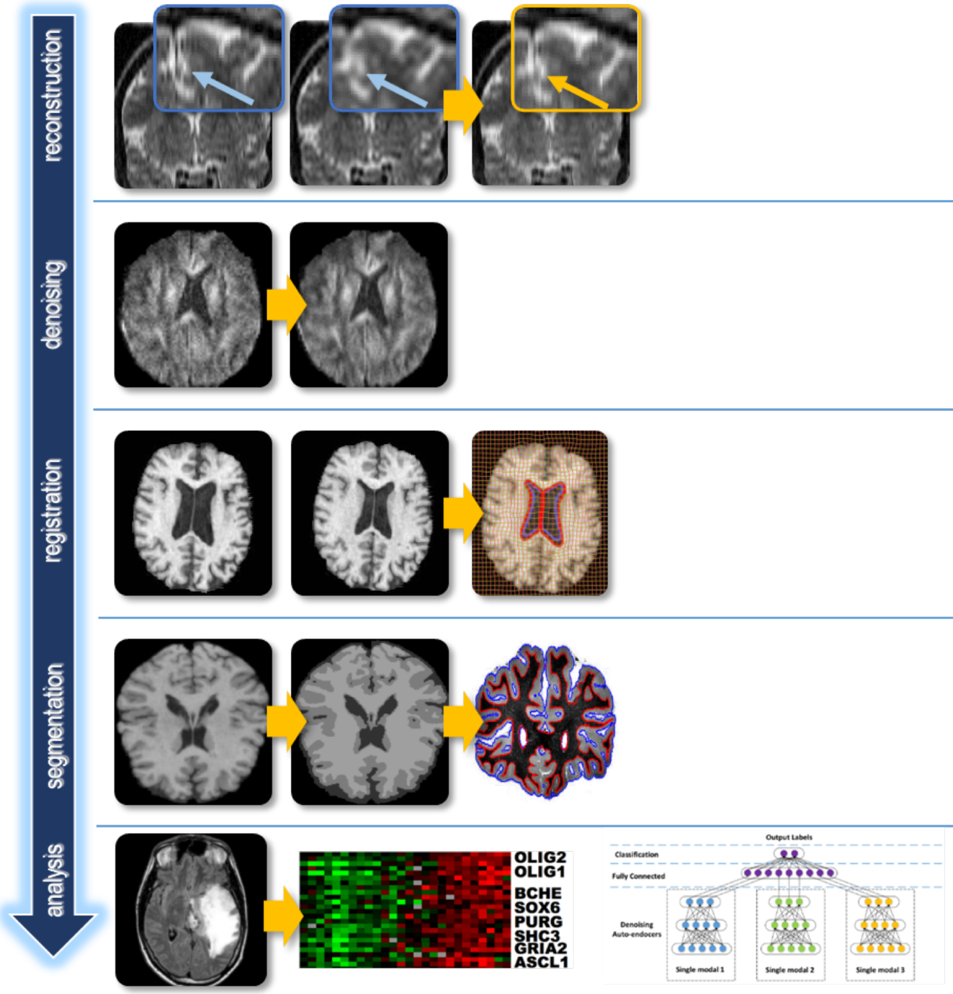


image reconstruction

pre-processing & normalization

spatial registration

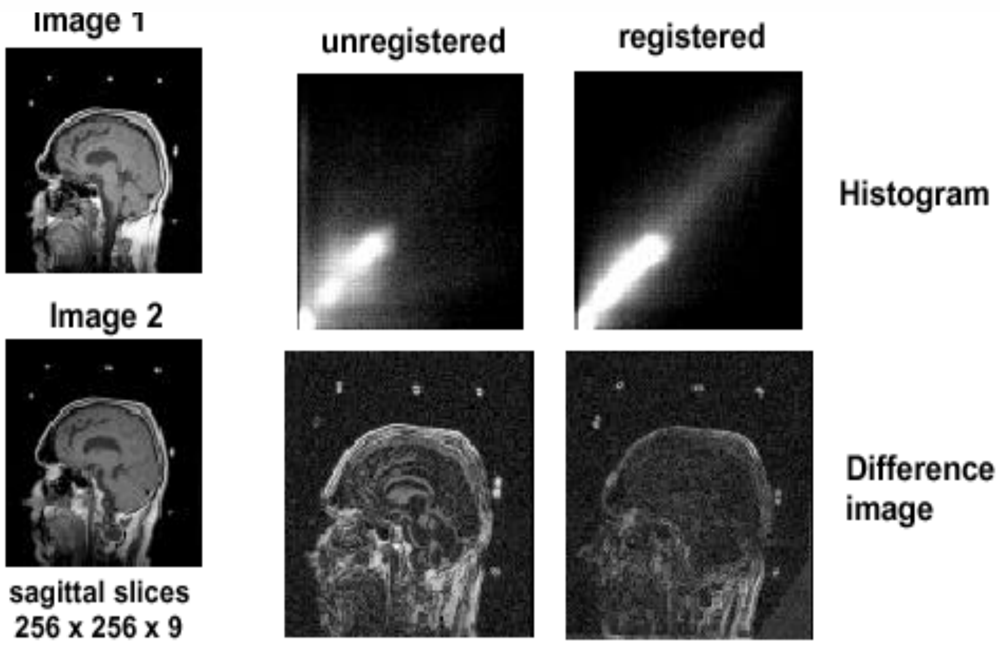
region of interest delineation

feature extraction machine learning

Components: Registration

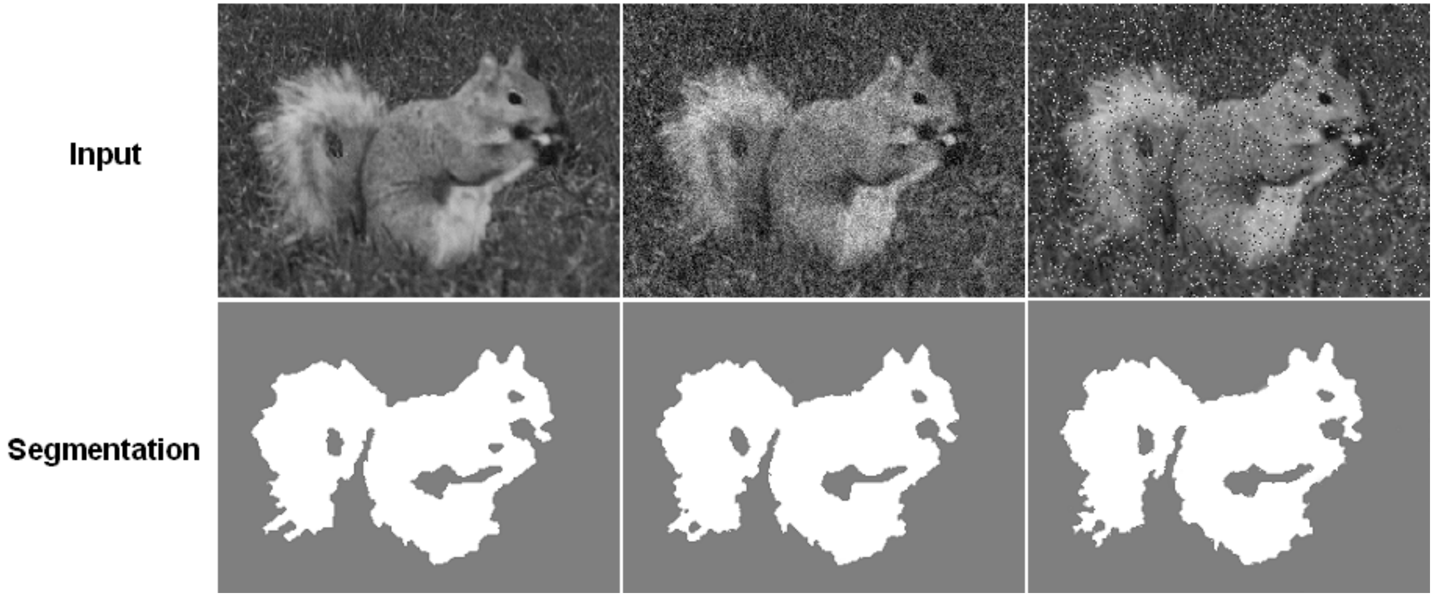
- Process of aligning images so that the correspondences between them can be seen more easily

$$\arg\max_{T}\left\{similarity(Image_1, T(Image_2))\right\}$$



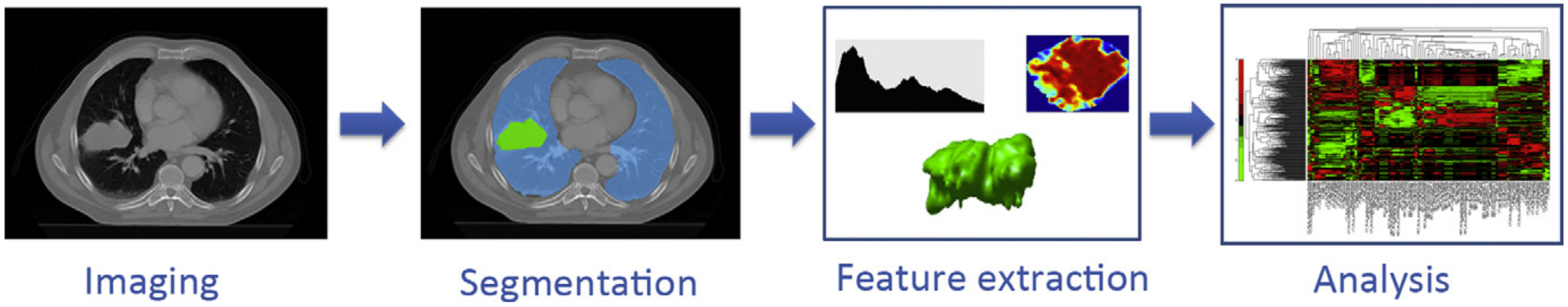
Components: Segmentation

- Partitioning image pixels into different classes (e.g., foreground, background)
- Used when we want to analyze just the pixels within a particular region



Components: Feature Extraction

- **Radiomics** is the extraction and analysis of large amounts of advanced quantitative imaging features with high throughput from medical images
- Feature types:
  - Color, shape, texture



Kumar, V et al (2012). Radiomics: The process and the challenges. *Magnetic Resonance Imaging*, *30*(9), 1234–1248. http://doi.org/10.1016/j.mri.2012.06.010

Radiology Reports are *Unstructured*

History: 57 year old man with 2 day history of thumb tingling and numbness and history of colon cancer 12 years prior.

Technique: An MRI of the brain was performed on a 1.5T scanner utilizing the following sequences: thin-cut axial SPGR and thin-cut axial T2W.

Findings: An enhancing mass is again seen in the anterioinferior aspect of the left parietal lobe along the primary sensory cortex abutting the central sulcus. The mass also abuts the sylvian fissure. There is a small region of central necrosis. The lesion measures 1.7 cm AP x 2.0 cm TR x 1.8 cm CC. There is moderate perilesional edema which extends into the posterior temporal and frontal lobes. There is no significant mass effect from the lesion. The high convexity cerebral sulci are normal and there is no significant midline shift or mass effect on the ventricles. There is no hydrocephalus. No other lesions or regions of abnormal enhancement are seen within the brain, dura or leptomeninges. The skull is normal in appearance. Minimal T2W/FLAIR high signal intensity surrounds the periventricular white matter, consistent with mild microvascular ischemic disease. The brainstem and cerebellum are unremarkable in appearance. The basal cisterns are clear.

Impression: Contrast-enhanced MRI of the brain again demonstrates an enhancing mass in the anterioinferior aspect of the left parietal lobe, located on the primary sensory cortex. There is moderate perilesional edema. No other lesions are seen.

Structured Reporting

CT lung cancer screening

Technical parameters
kVp: [ ]

mA: [ ]

DLP: [ ]

Clinical information
Screening visit: [Baseline | Year 1 | Year 2]

[Lung cancer screening.]

Comparison
[None.]

Findings

Exam parameters
Diagnostic quality: [Satisfactory | Limited, but interpretable | Non-diagnostic]

Comments: Comments [ ]

Lung nodules
[None. | Present, detailed below:]

[-- | Right lung | Left lung] [ ] mm [-- | solid | semi-solid | ground-glass] [-- | unchanged | increased | decreased | new] image # [ ]

[-- | Right lung | Left lung] [ ] mm [-- | solid | semi-solid | ground-glass] [-- | unchanged | increased | decreased | new] image # [ ]

[-- | Right lung | Left lung] [ ] mm [-- | solid | semi-solid | ground-glass] [-- | unchanged | increased | decreased | new] image # [ ]

RSNA Informatics Reporting

RSNA's Reporting Initiative is improving radiology reporting practices by building IT standards and a library of clear and consistent report templates.

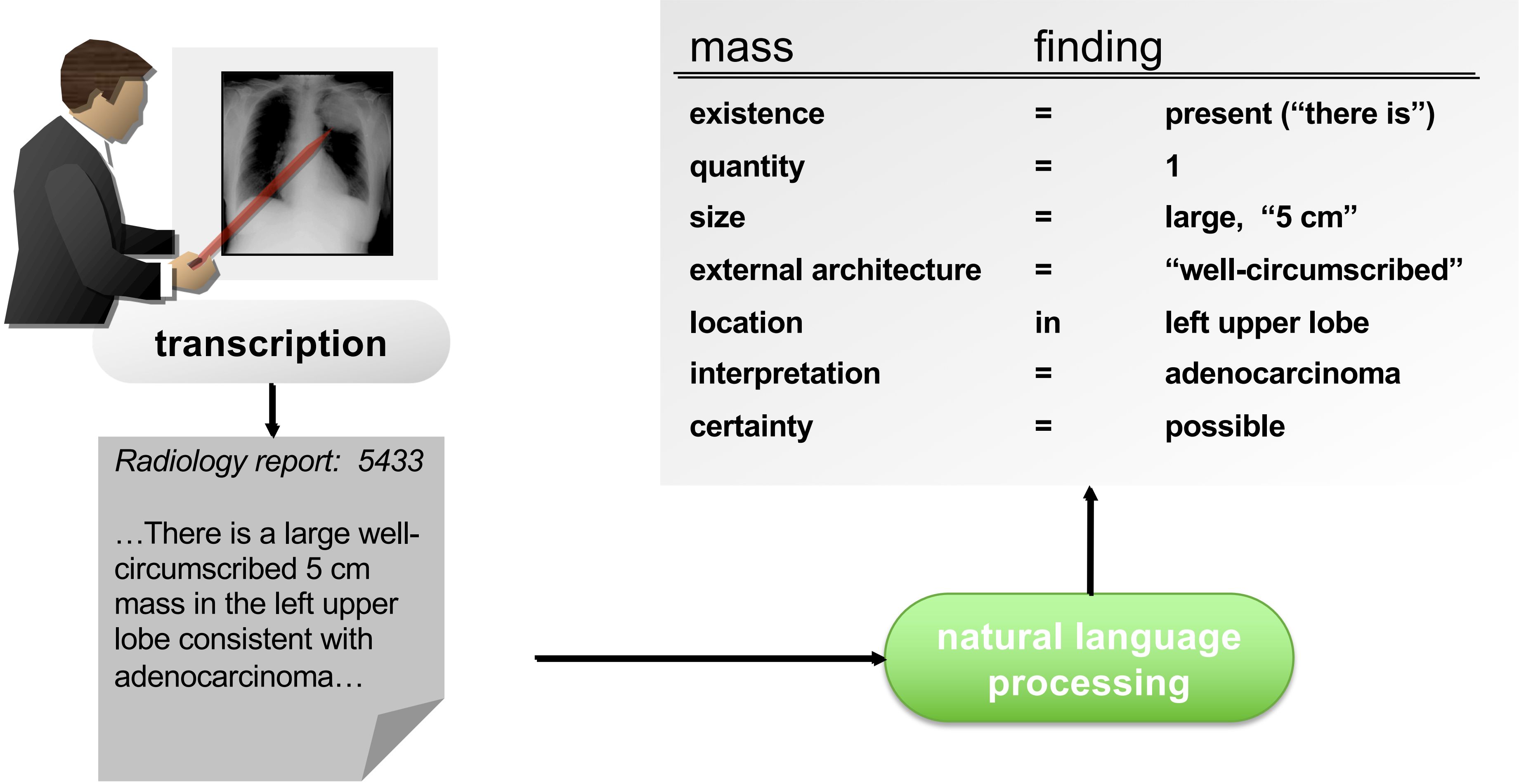
Supported in part by NIH / NIBIB.

Specialties | Organizations | Languages | Popular | New

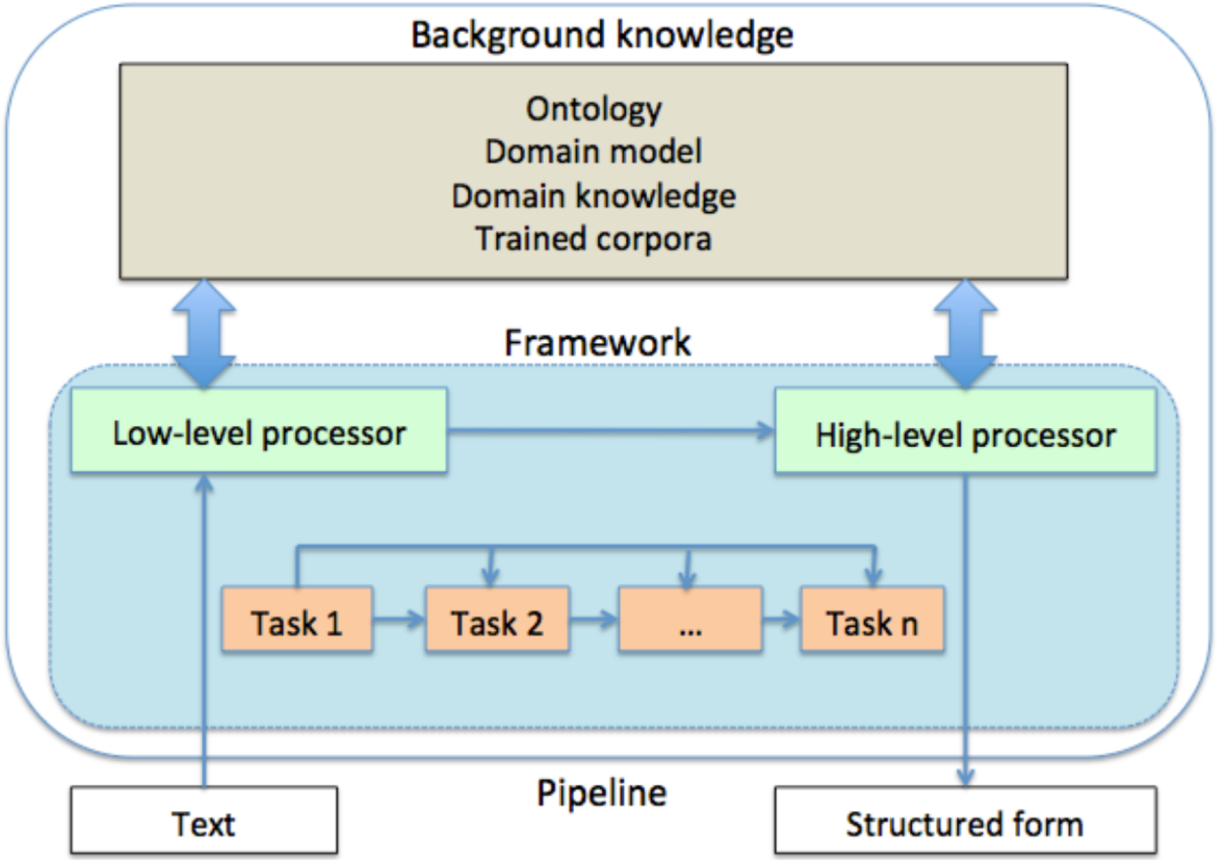
- CA **Cardiac Radiology** — CT Calcium Score · 7 more
- CH **Chest Radiology** — MR Brachial Plexus · 21 more
- CT **Computed Tomography** — CT Onco Renal Mass · 51 more
- DX **Diagnostic Radiology** — Skeletal Survey · 54 more
- ER **Emergency Radiology** — US Retroperitoneum · 23 more
- GI **Gastrointestinal Radiology** — Defecography · 46 more
- GU **Genitourinary Radiology** — CT Onco Renal Mass · 36 more
- HN **Head and Neck** — Parathyroid SPECT · 15 more
- IR **Interventional Radiology** — CT-guided Drainage Catheter Placement · 16 more
- MR **Magnetic Resonance Imaging** — MR Neck · 29 more
- MK **Musculoskeletal Radiology** — Skeletal Survey · 42 more
- NR **Neuroradiology** — MR Neck · 24 more
- NM **Nuclear Medicine** — White Blood Cell Scan · 28 more
- OB **Obstetric/Gynecologic Radiology** — US Pelvis · 6 more
- OI **Oncologic Imaging** — CT Onco Renal Mass · 21 more
- PD **Pediatric Radiology** — Skeletal Survey · 12 more
- QI **Quality Improvement** — Communication of Actionable Findings
- RS **Research** — US Carotid Arteries (with Stenosis Calculator) · 2 more
- US **Ultrasound** — US Retroperitoneum · 26 more
- VI **Vascular Imaging** — US Abdominal Aorta · 20 more

http://radreport.org/

Natural Language Processing (NLP)



Components of an NLP System

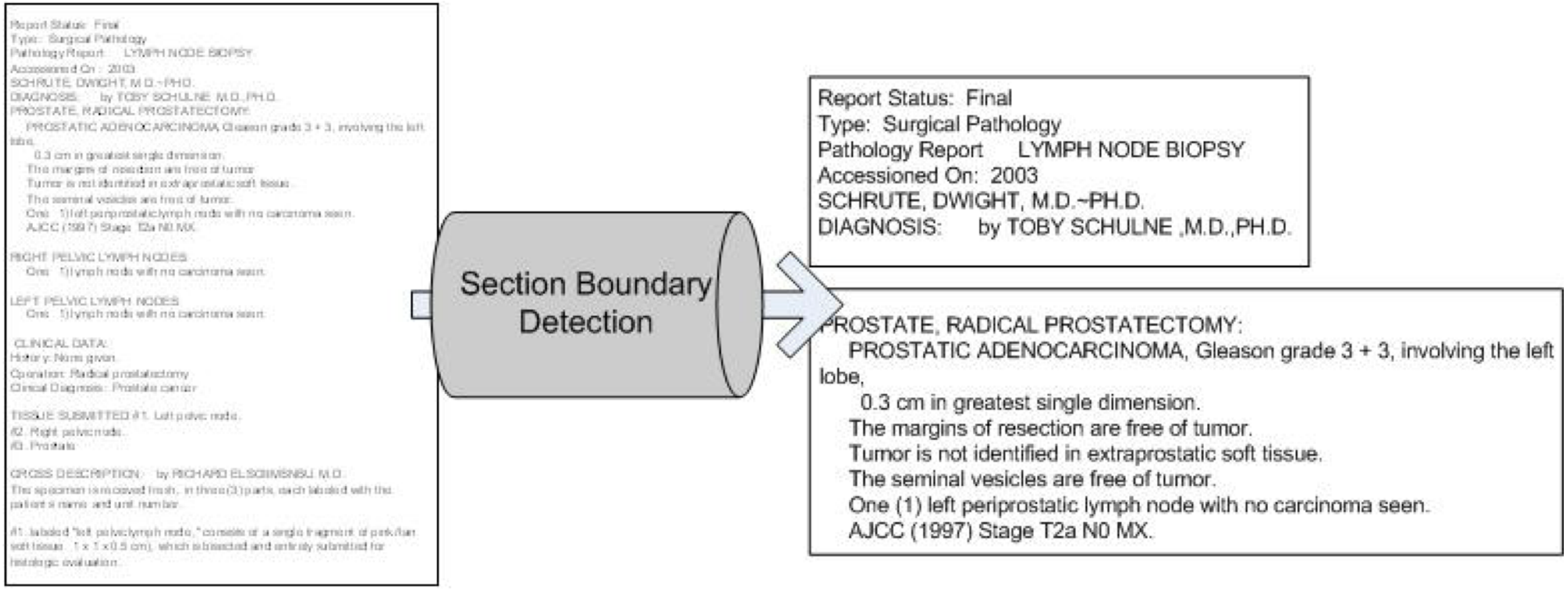


Doan, Son, et al. "Natural language processing in biomedicine: a unified system architecture overview." *Clinical Bioinformatics* (2014): 275-294.

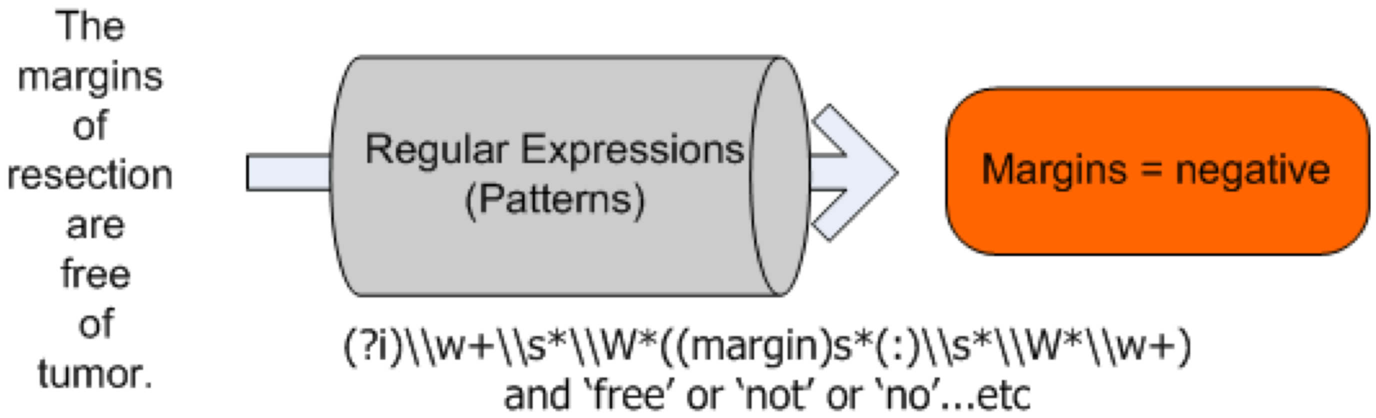
Components: Background Knowledge

- A **lexicon** is a collection of information about the words of a language about the lexical categories to which they belong
- e.g., RadLex lexicon
  - A radiology-specific lexicon with over 75,000 terms and synonyms developed by RSNA
  - http://www.rsna.org/RadLex.aspx
- e.g., Radiology Gamuts Ontology
  - A **gamut** is a set of conditions that can cause a specified imaging finding
  - Gamuts Ontology contains 16,912 entities (12,878 causes / 4,662 effects)
  - https://www.gamuts.net

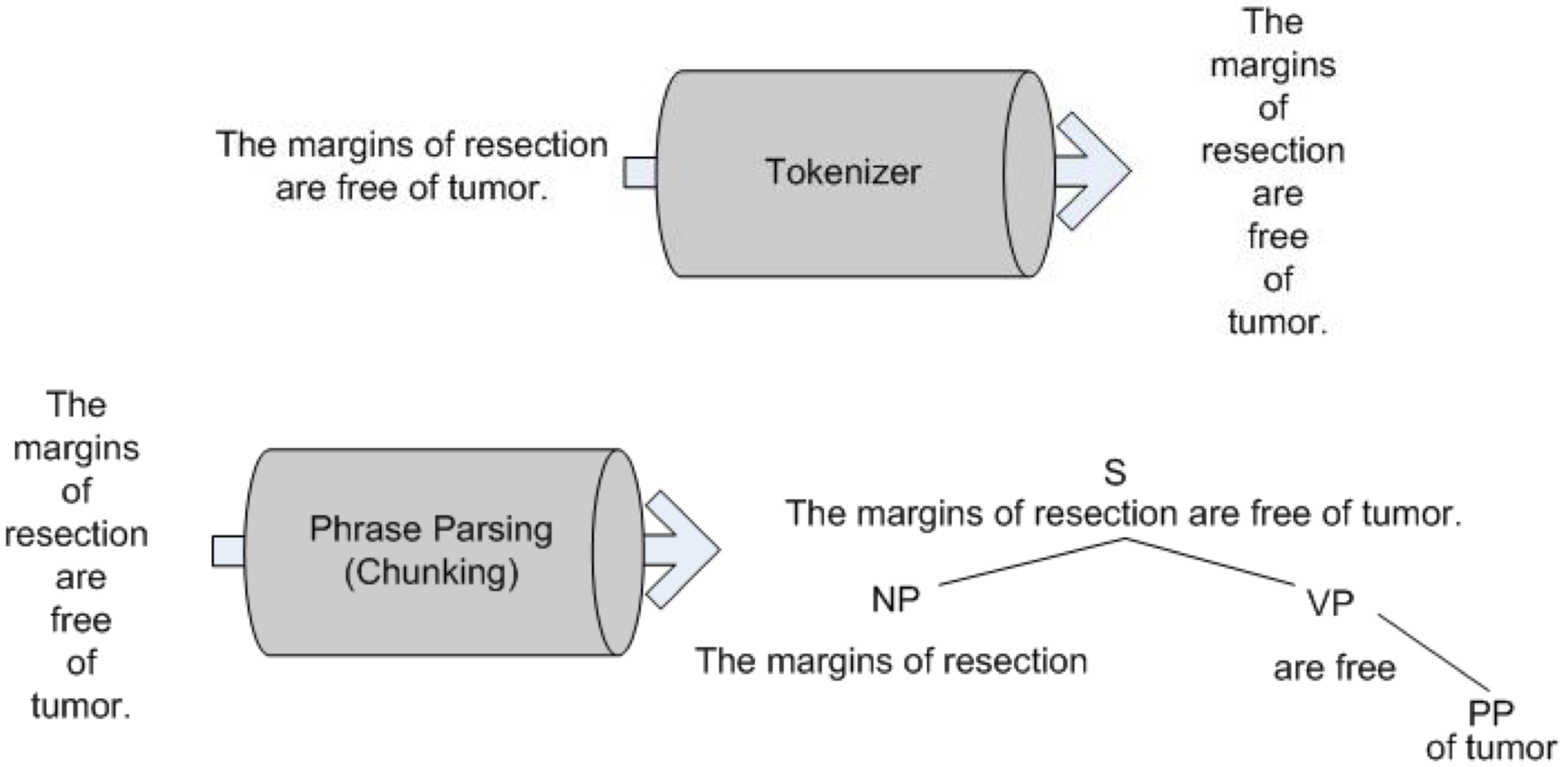
Components: Structural Analyzer



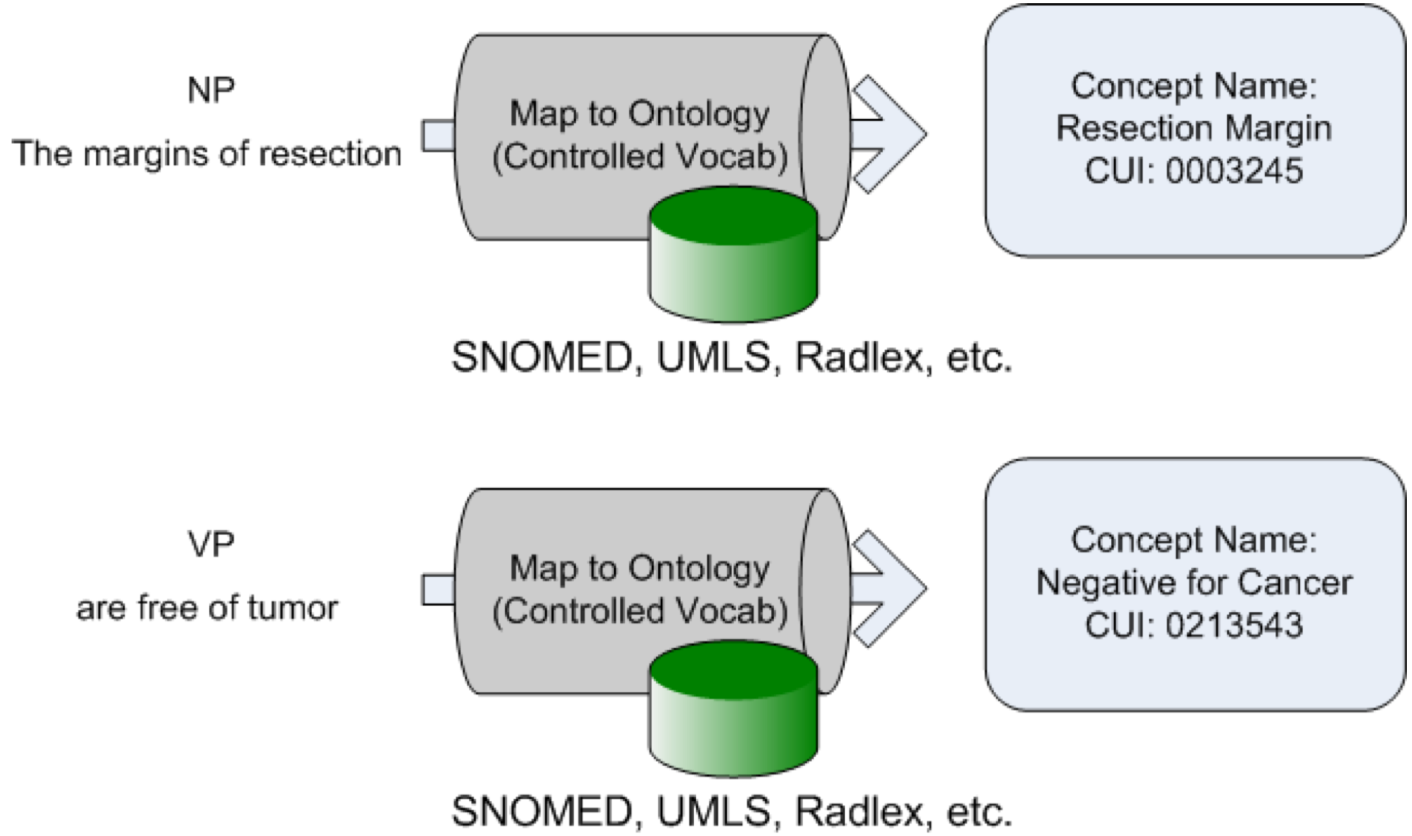
Components: Rule-based Approach



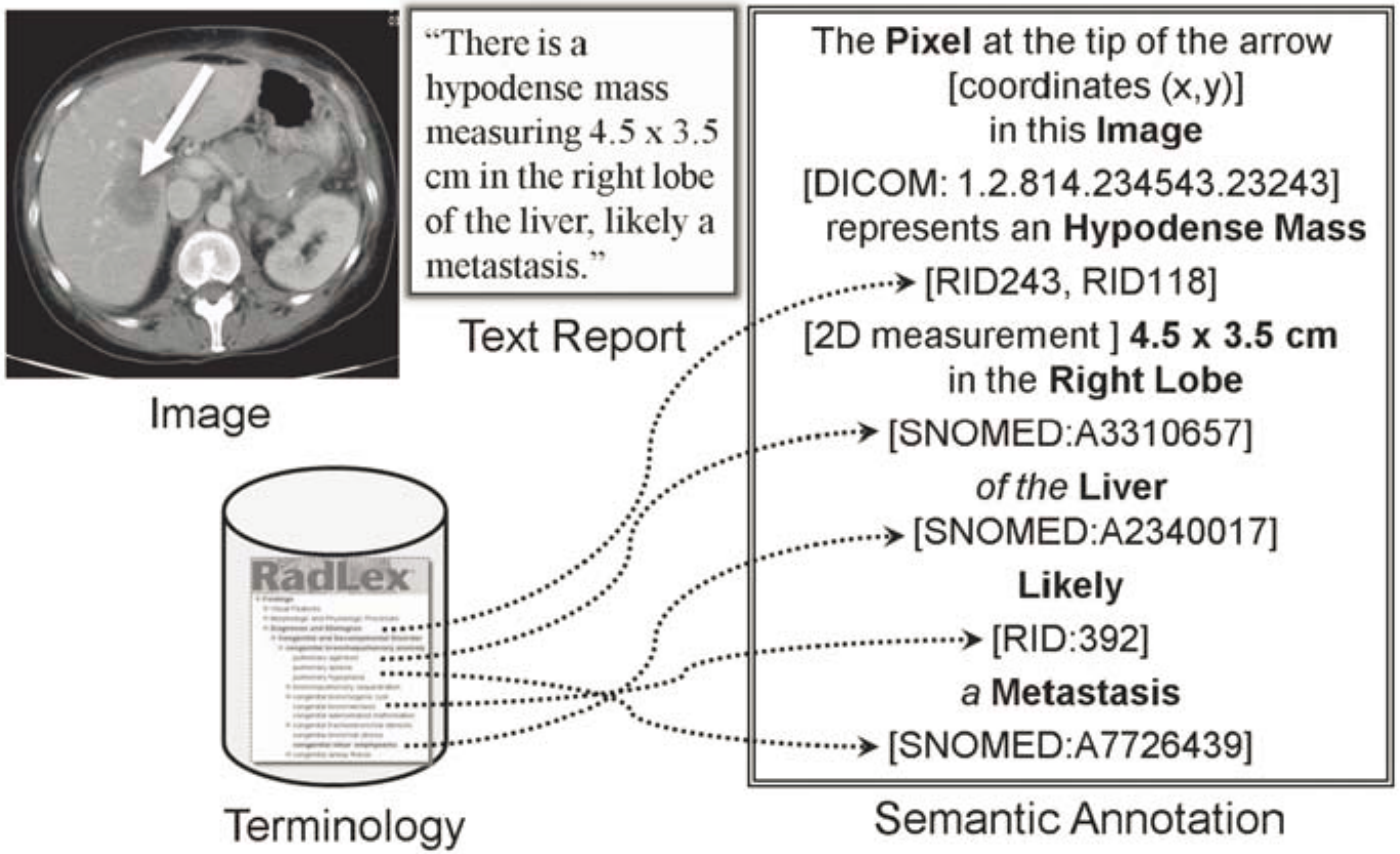
Components: Lexical Analyzer/Parser



Components: Semantic Interpreter



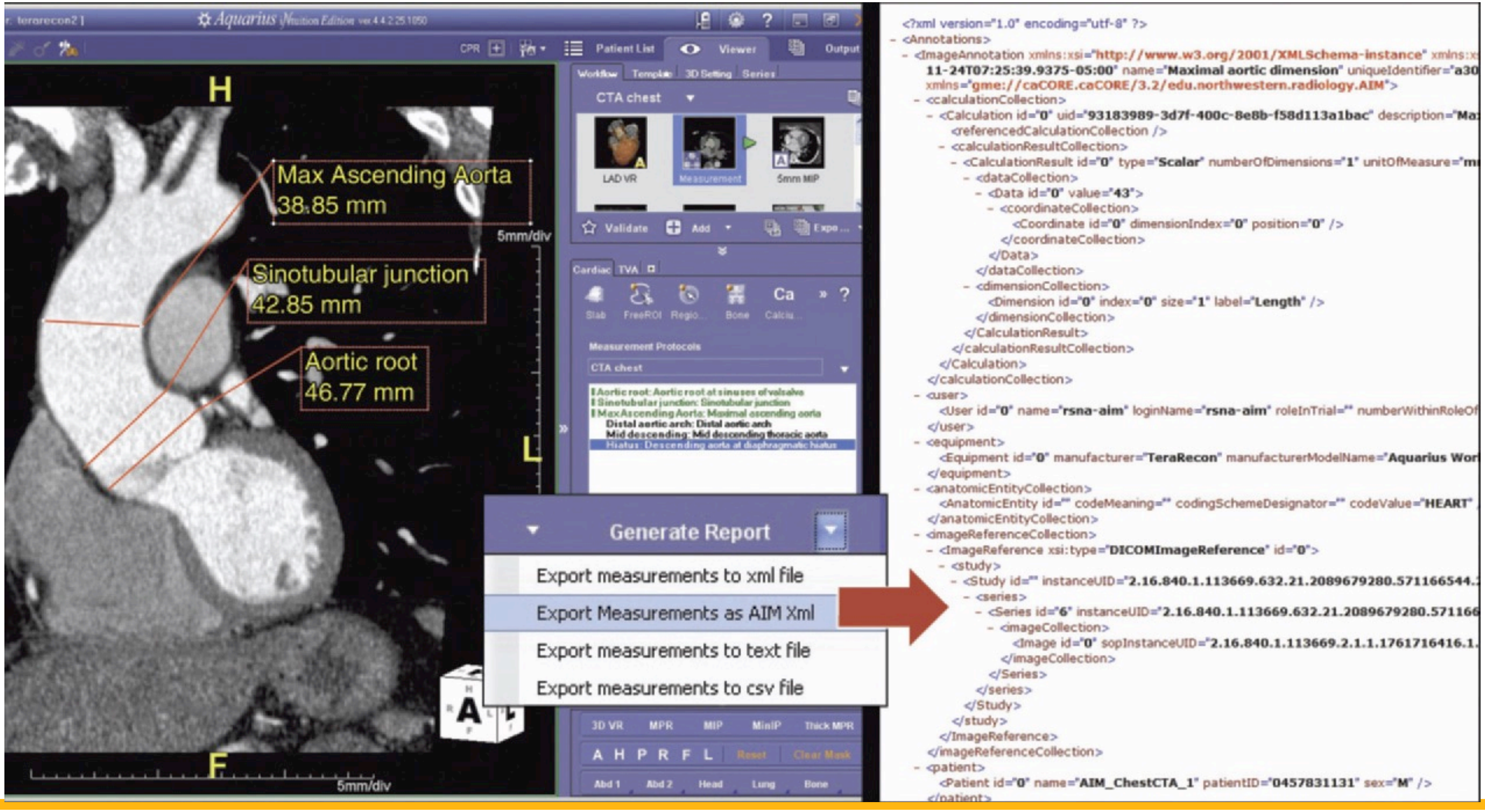
Combining Images & Text → Meaning



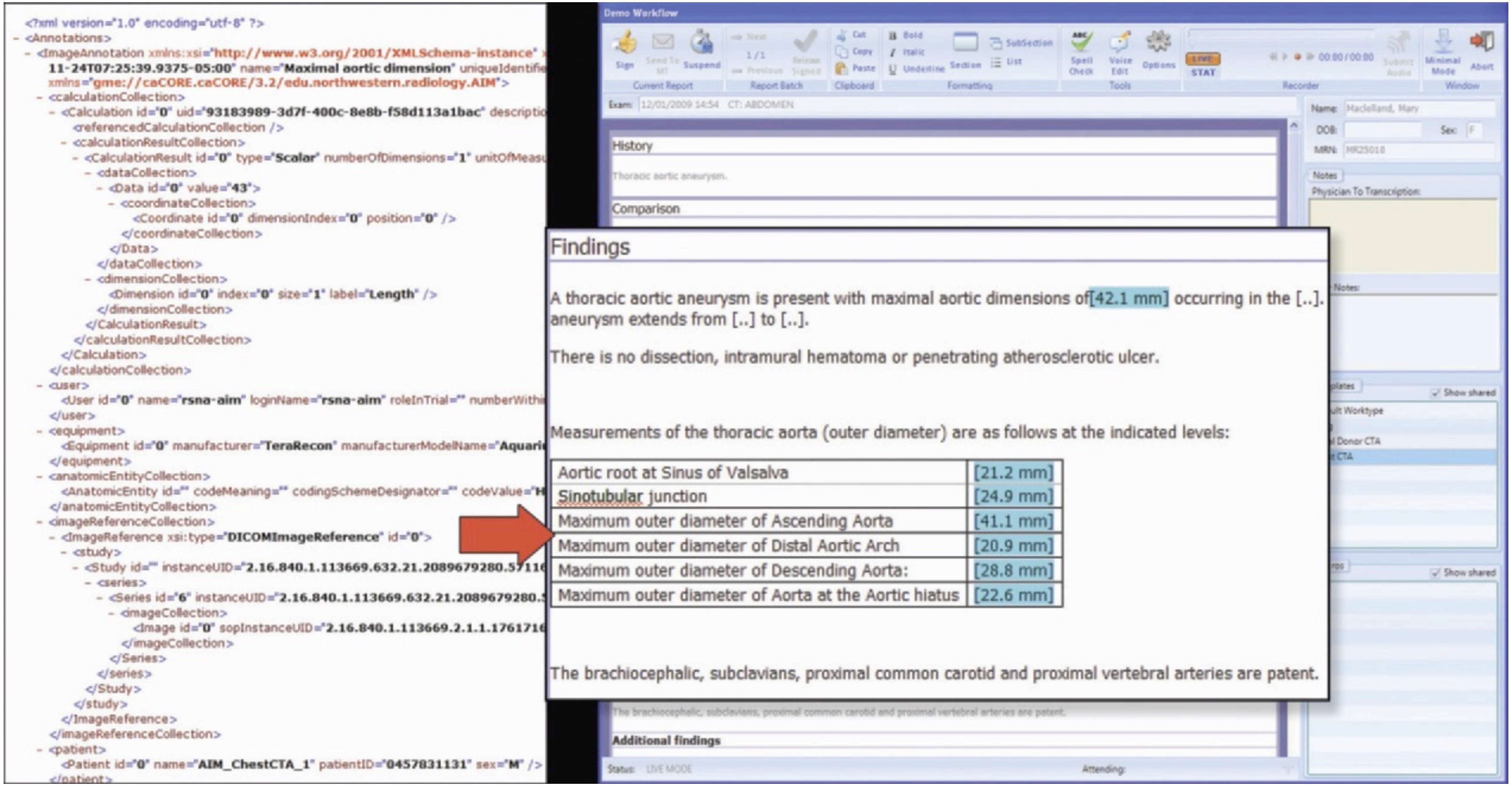
Annotation Image Markup (AIM) is a structured text-based language that captures annotations and markups in a standardized format using **eXtensible Markup Language (XML).**

https://wiki.nci.nih.gov/display/AIM/Annotation+and+Image+Markup+-+AIM

Zimmerman SL. Automated structured reporting of imaging findings using the AIM standard and XML. RadioGraphics, 2011.



Zimmerman SL. Automated structured reporting of imaging findings using the AIM standard and XML. RadioGraphics, 2011.



David Geffen School of Medicine

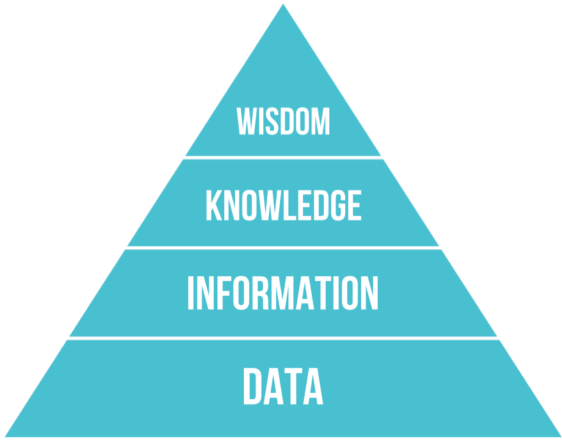
Zimmerman SL. Automated structured reporting of imaging findings using the AIM standard and XML. RadioGraphics, 2011.

UCLA Health Radiology

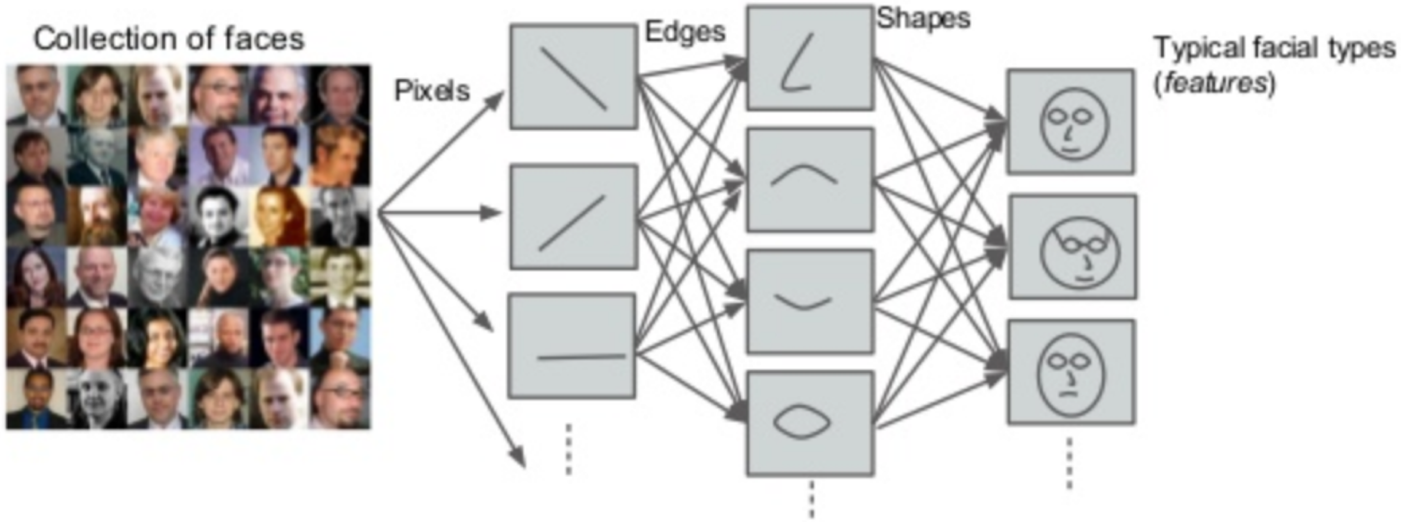
Key Takeaways

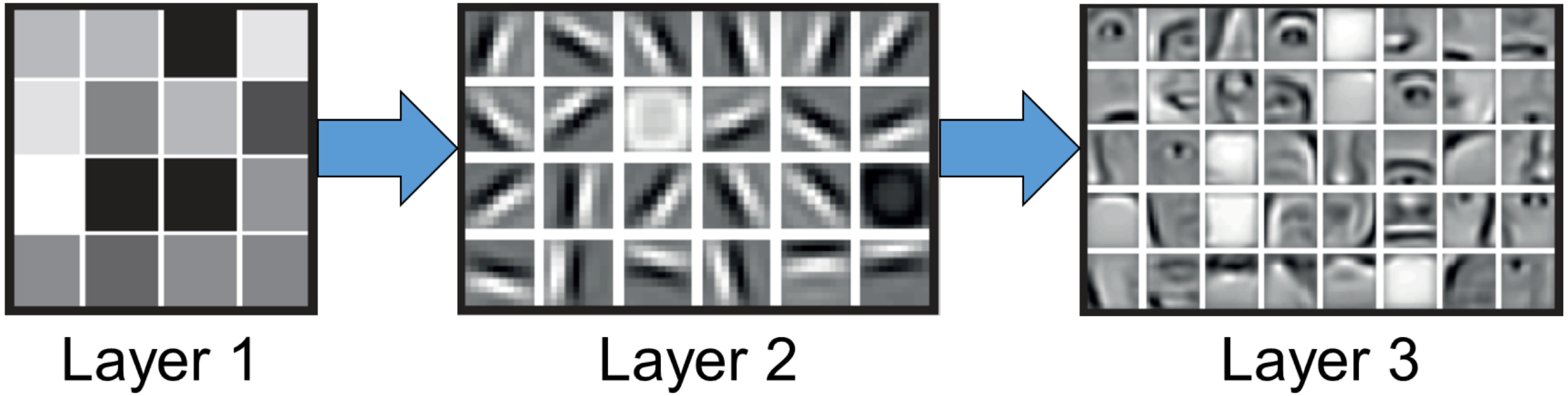
- Images and radiology reports are unstructured data
    - Images are comprised of pixel data that need to be normalized, registered, and segmented to extract relevant features
    - Using natural language processing and/or structured reporting, information from radiology reports can be represented in a form that is amenable to mining
- Structured information can be combined using standardized markup languages such as Annotation Image Markup (AIM) for storing or sharing

Knowledge



Representation is Key to Knowledge

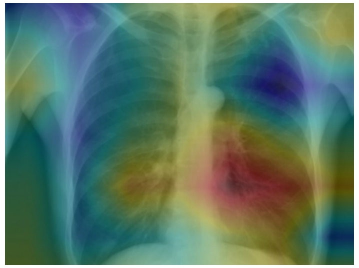




IEEE SPECTRUM

Stanford Algorithm Can Diagnose Pneumonia Better Than Radiologists

Source: https://spectrum.ieee.org/the-human-os/biomedical/diagnostics/stanford-algorithm-can-diagnose-pneumonia-better-than-radiologists



engadget

For a dollar, an AI will examine your medical scan

Zebra-Med's tech helps radiologists find heart, liver, bone and other diseases.

Source: https://www.engadget.com/2017/10/27/for-a-dollar-an-ai-will-examine-your-medical-scan/

The NEW ENGLAND JOURNAL of MEDICINE

Predicting the Future — Big Data, Machine Learning, and Clinical Medicine

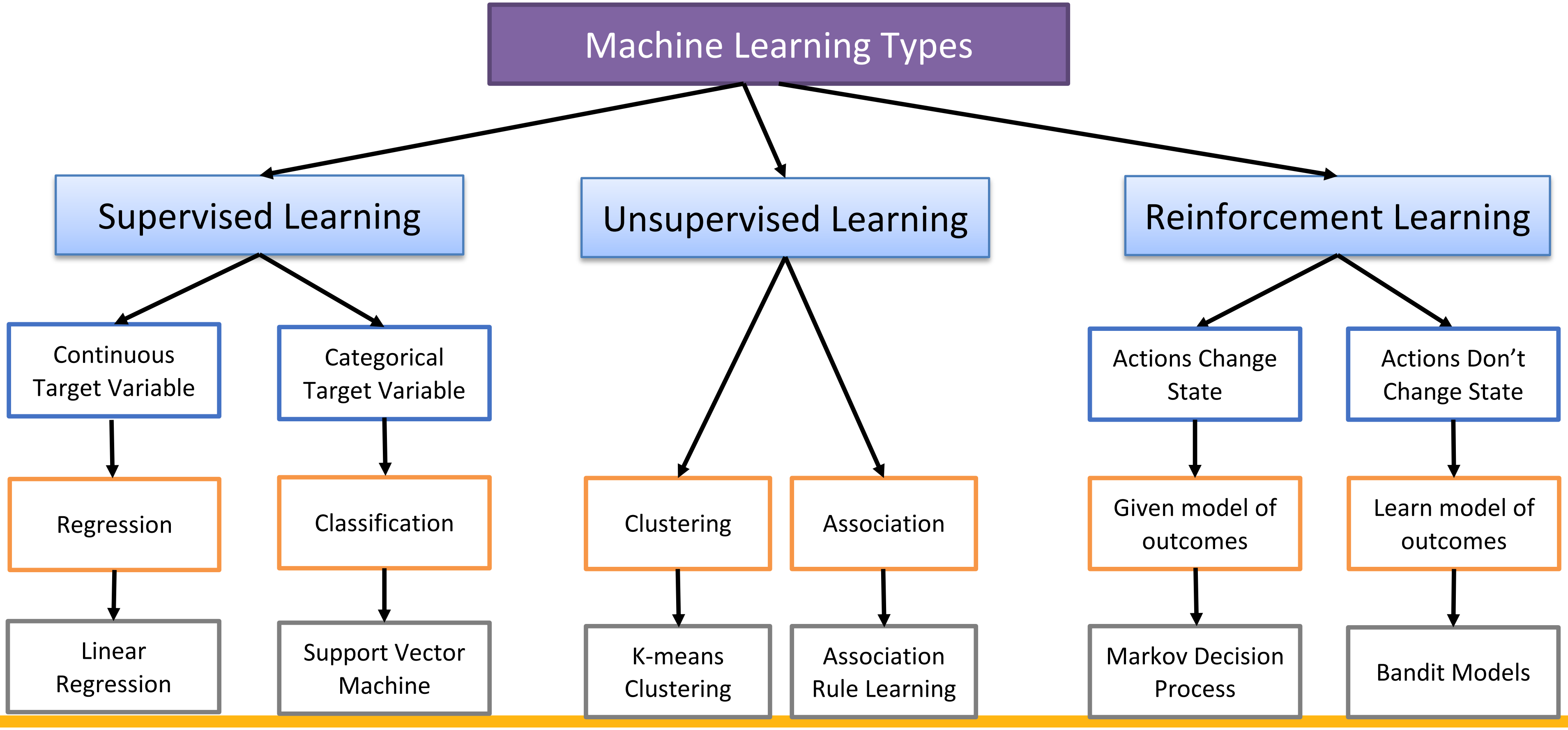
Ziad Obermeyer, M.D., and Ezekiel J. Emanuel, M.D., Ph.D.

Source: http://www.nejm.org/doi/full/10.1056/NEJMp1606181

Improving the Use of Imaging

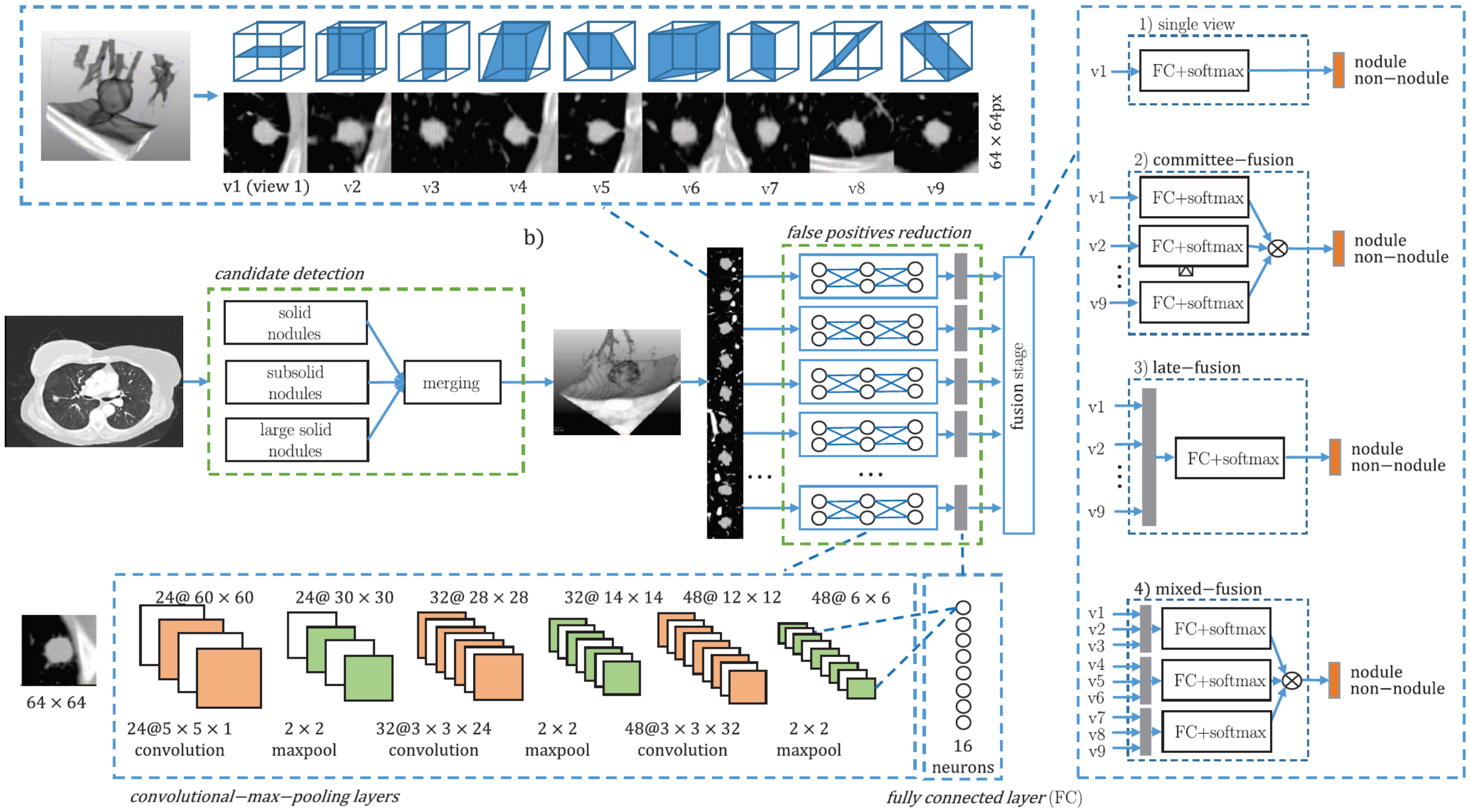
Wrong study requested → Poor study acquisition → Poor study interpretation → Poor study documentation

A Variety of Approaches Exist

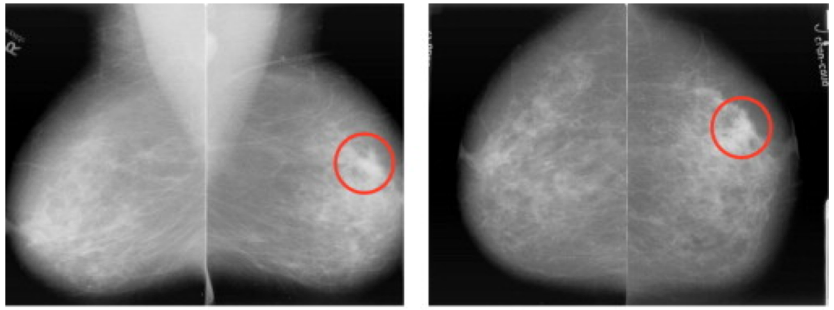


Inspired from: http://en.proft.me/2015/12/24/types-machine-learning-algorithms/

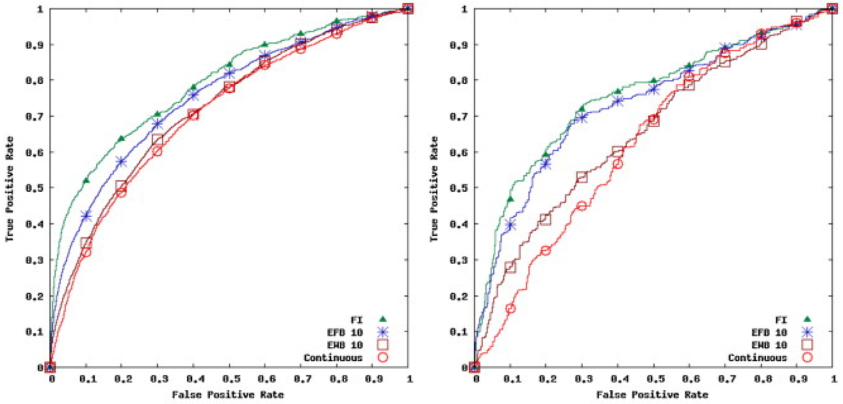
CADe: Computer Aided Detection



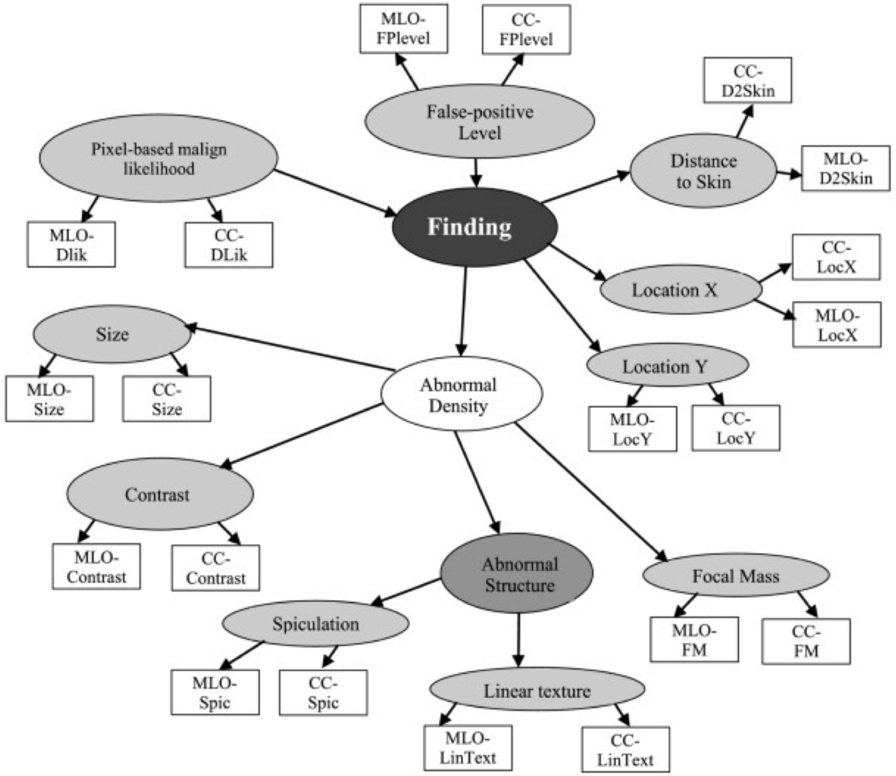
CADx: Computer Aided Diagnosis



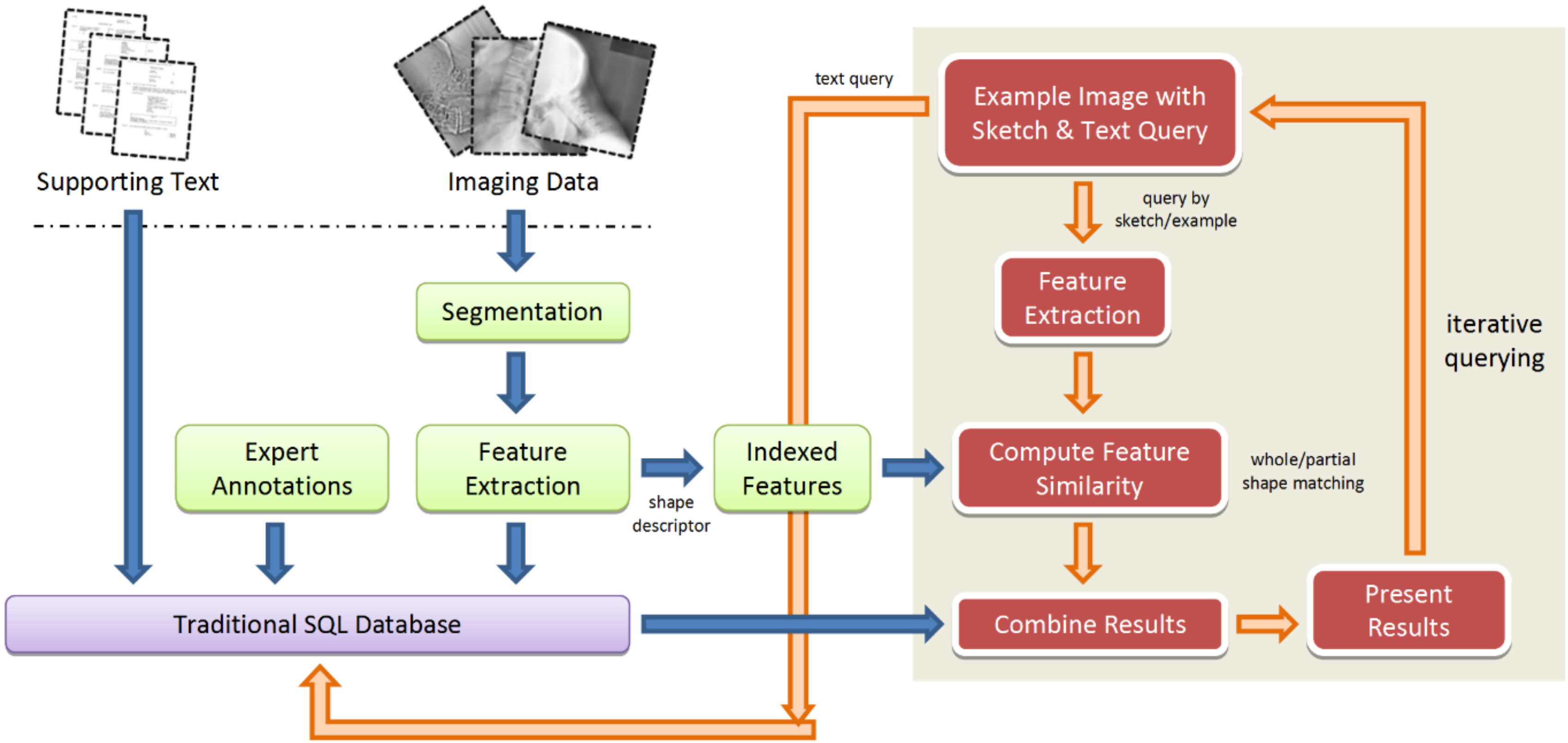
(a) MLO (b) CC



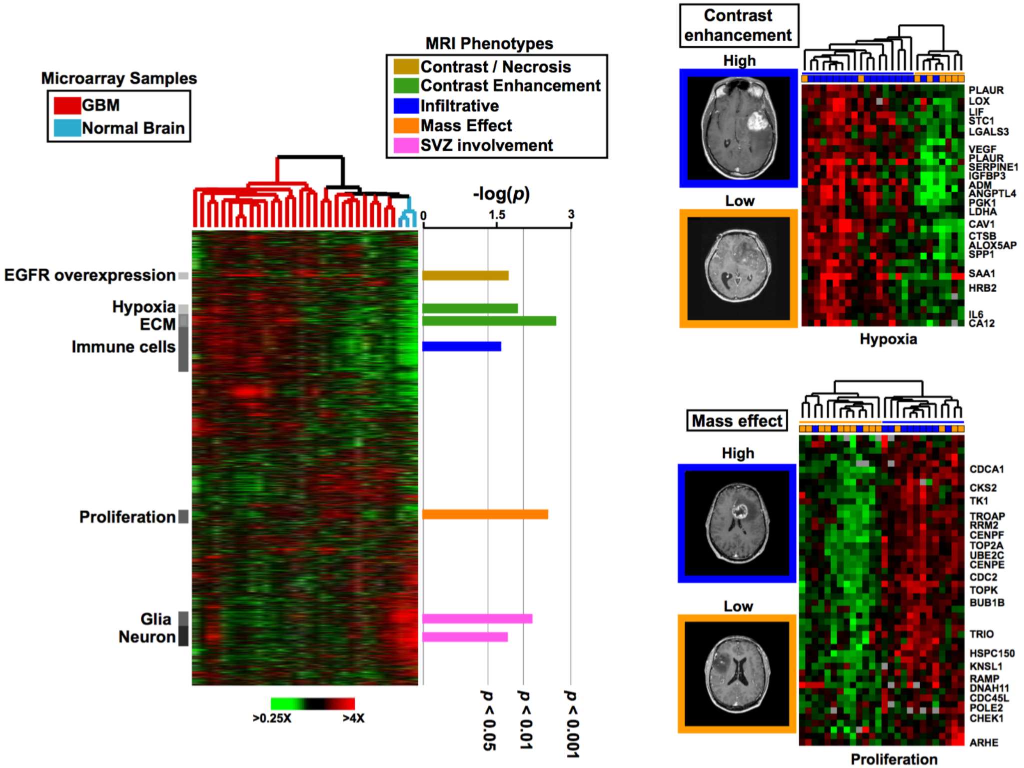
(a) Link level (b) Patient level



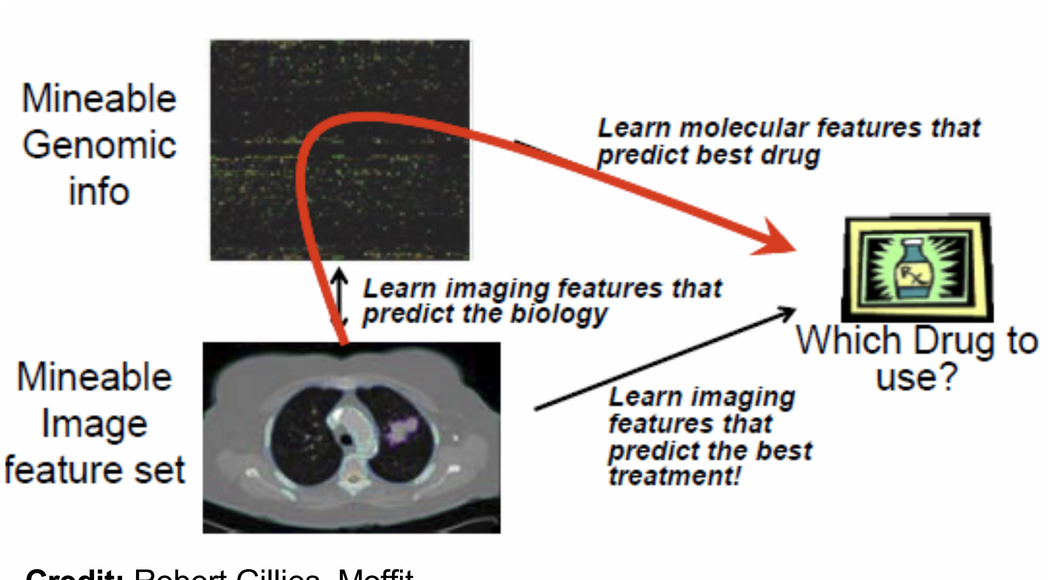
Content-based Image Retrieval



Radiogenomics



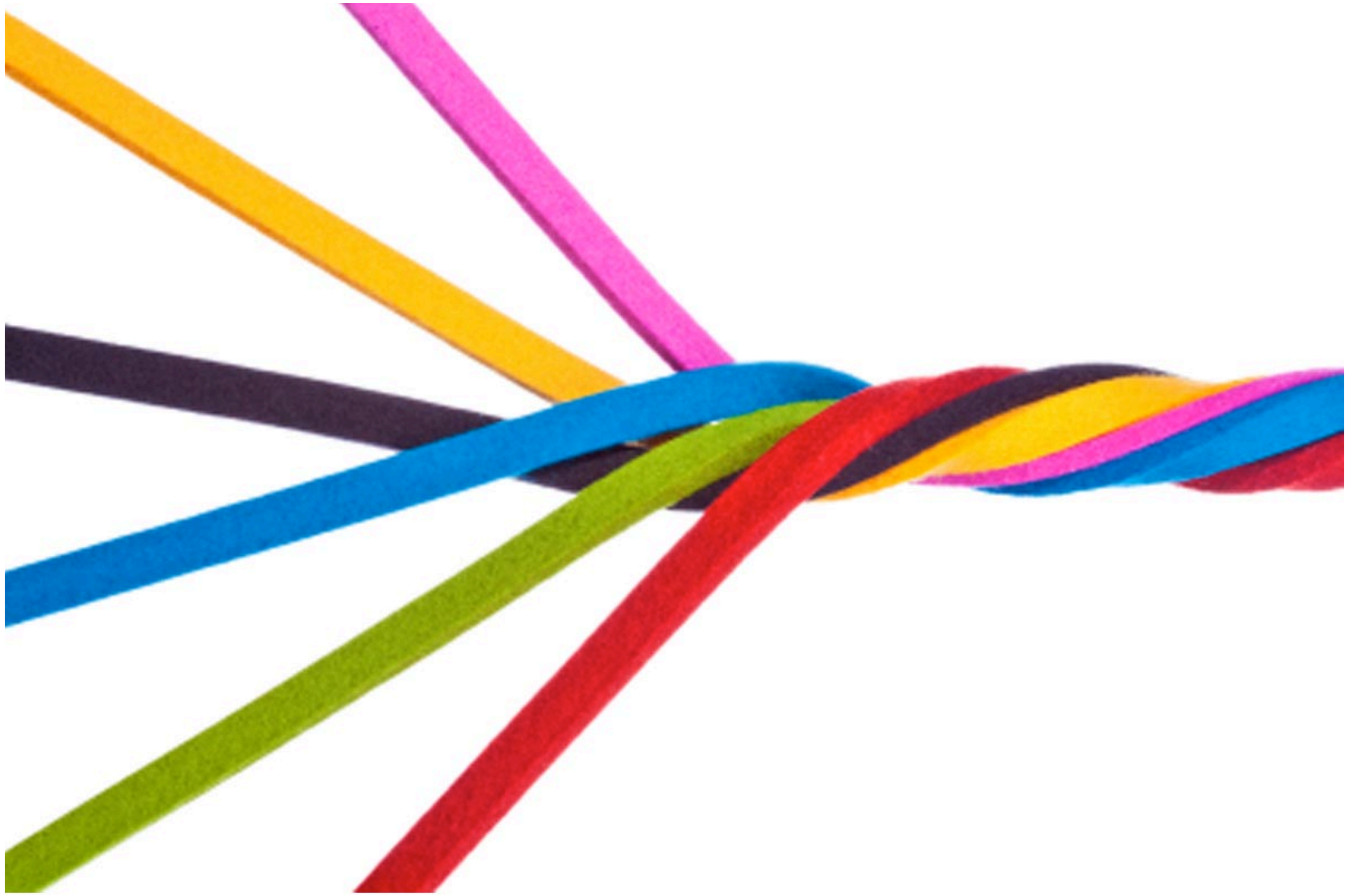
Source: Diehn M et al, PNAS, 2008



Credit: Robert Gillies, Moffit

Takeaways

- Knowledge is attained by finding the appropriate representation for a given task
  - Machine and deep learning are different forms of representations for integrating biomedical information for clinical decision support
  - No one representation is the optimal solution for all tasks
- A variety of tasks exist in imaging
  - Computer-aided detection/diagnosis (CADe/CADx), content-based image retrieval (CBIR), radiomics/radiogenomics
- More on this later… Fabien Scalzo's lecture (9/17)



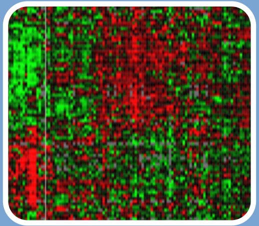
Integrated Diagnostics (IDx)

- Radiological Sciences | Pathology & Lab Medicine
  - http://idx.mednet.ucla.edu
- **Mission:** Provide an infrastructure for capturing, curating and retrieving validated patient datasets across clinical, imaging, pathology and molecular data sources for the purposes of improving early detection, diagnosis and treatment of cancer
  - Clinical & outcomes data
  - Medical imaging
  - Histologic features
  - Molecular features, depending upon research questions
  - Collected longitudinally

Objectives

Enable hypothesis generation

- Capture highly curated data and annotations
- Support open ended queries and real-time analytics

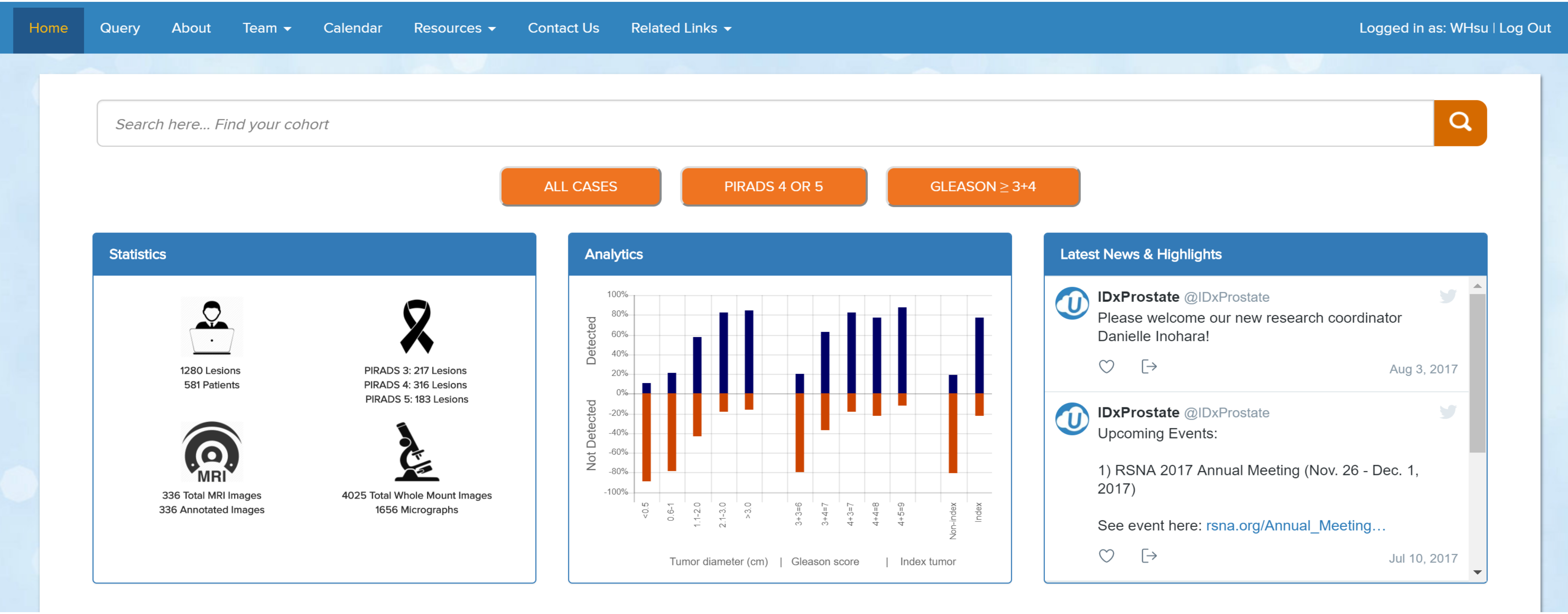


Integrate with molecular diagnostics

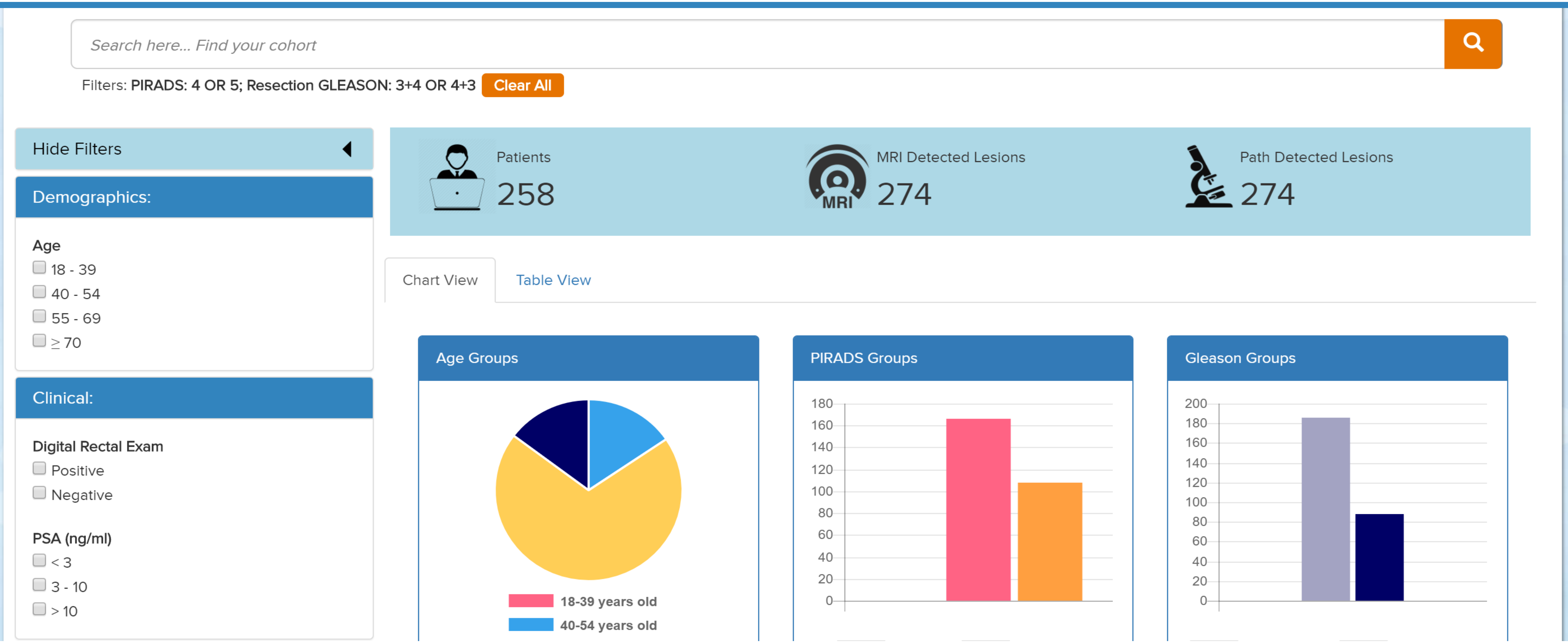
- Link findings across clinical, imaging, pathologic, and molecular data
- Characterize disease heterogeneity and evolution

Provide tailored clinical decision support

- Generate new evidence for appropriate use of imaging
- Improve early detection, diagnosis, and treatment



IDx Prostate Data Portal: Front Page



Charts View

David Geffen School of Medicine

UCLA Health Radiology

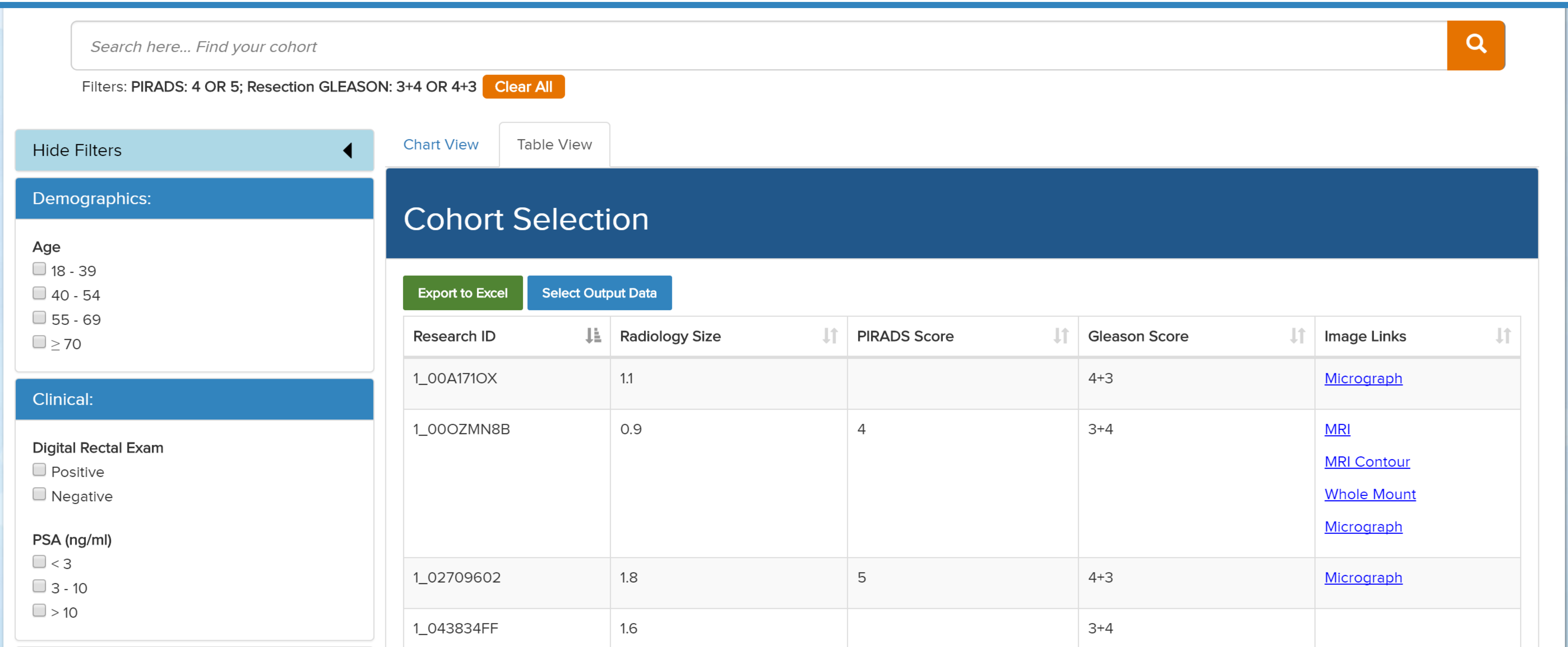
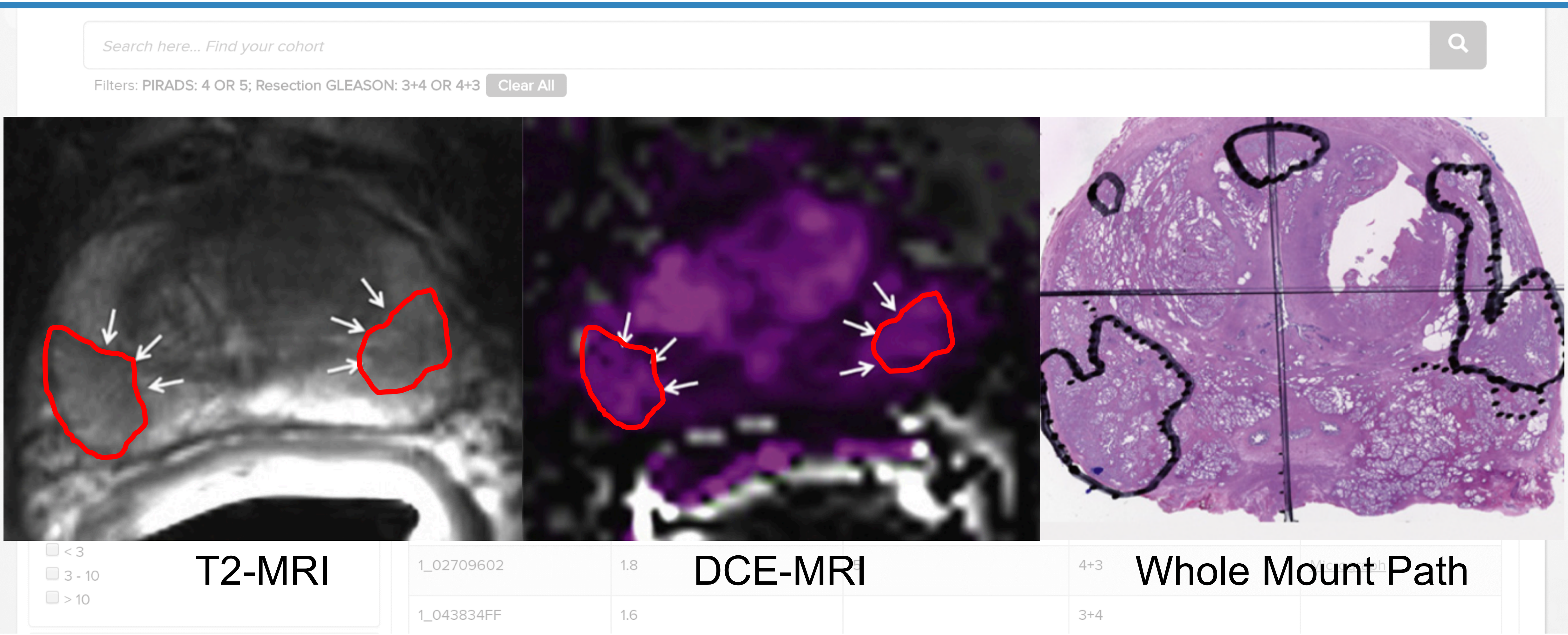


Table View

David Geffen School of Medicine

UCLA Health Radiology



Annotated Datasets

David Geffen School of Medicine

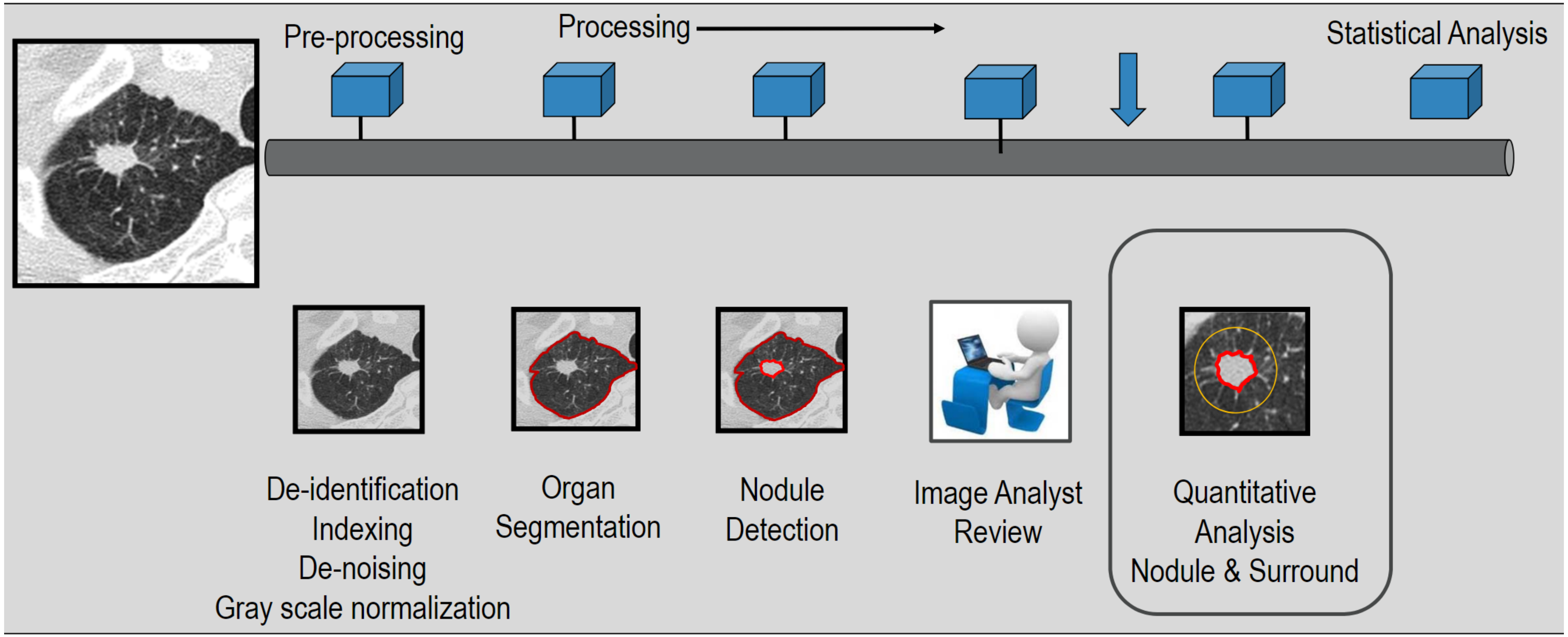
UCLA Health Radiology

Integrated Diagnostics

- Project launched in March 2015

	Cases	Imaging	Specimens	Annotations	Clinical
Prostate	665 pts	MR	Banked: 164 pts	PI-RADS, prostate contour, MRI lesion, whole mount lesion	PSA, digital rectal exam, diagnosis, recurrence
Lung	1,982	CT	Banked: 528	LungRADS, nodule, emphysema, interstitial lung disease, coronary artery calcium	Risk factors, comorbidities, pulmonary function test, diagnosis, treatments
Breast (Athena)	49,037	MR, MG (FFDM, DBT)	SNP: 364 WES: 13	BI-RADS, breast density	Risk factors, comorbidities, diagnosis, treatments
Liver	449	CT/US biopsy (121)	Banked: 71	--	Labs (AFP, bilirubin, Creatinine, etc), diagnosis, comorbidities, treatments
Kidney	804	CT	Banked: 35	Lesion attenuation, other organs/structures attenuation (reference)	Diagnosis, treatments, history

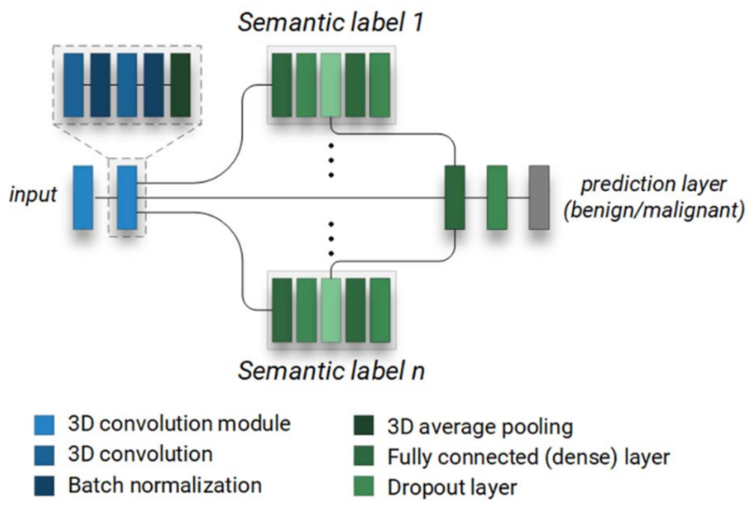
IDx Lung: Nodule Characterization



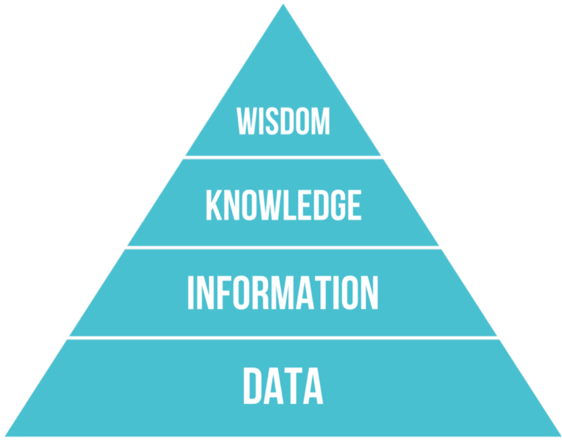
- Pipeline accepts only cases of < 2 mm slice thickness; the remainder must be manually segmented.

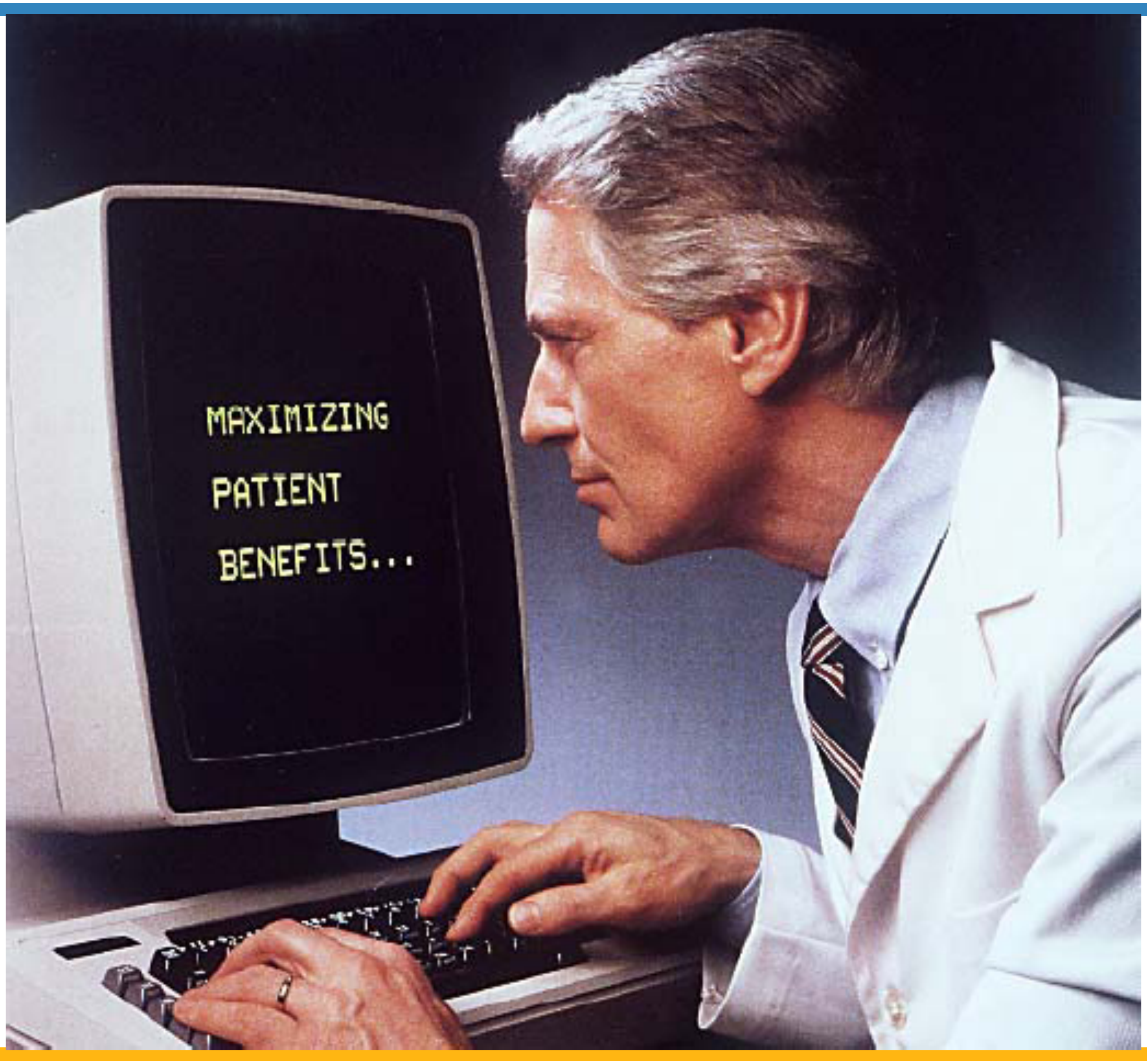
IDx Lung: Towards Improved Risk Predictions

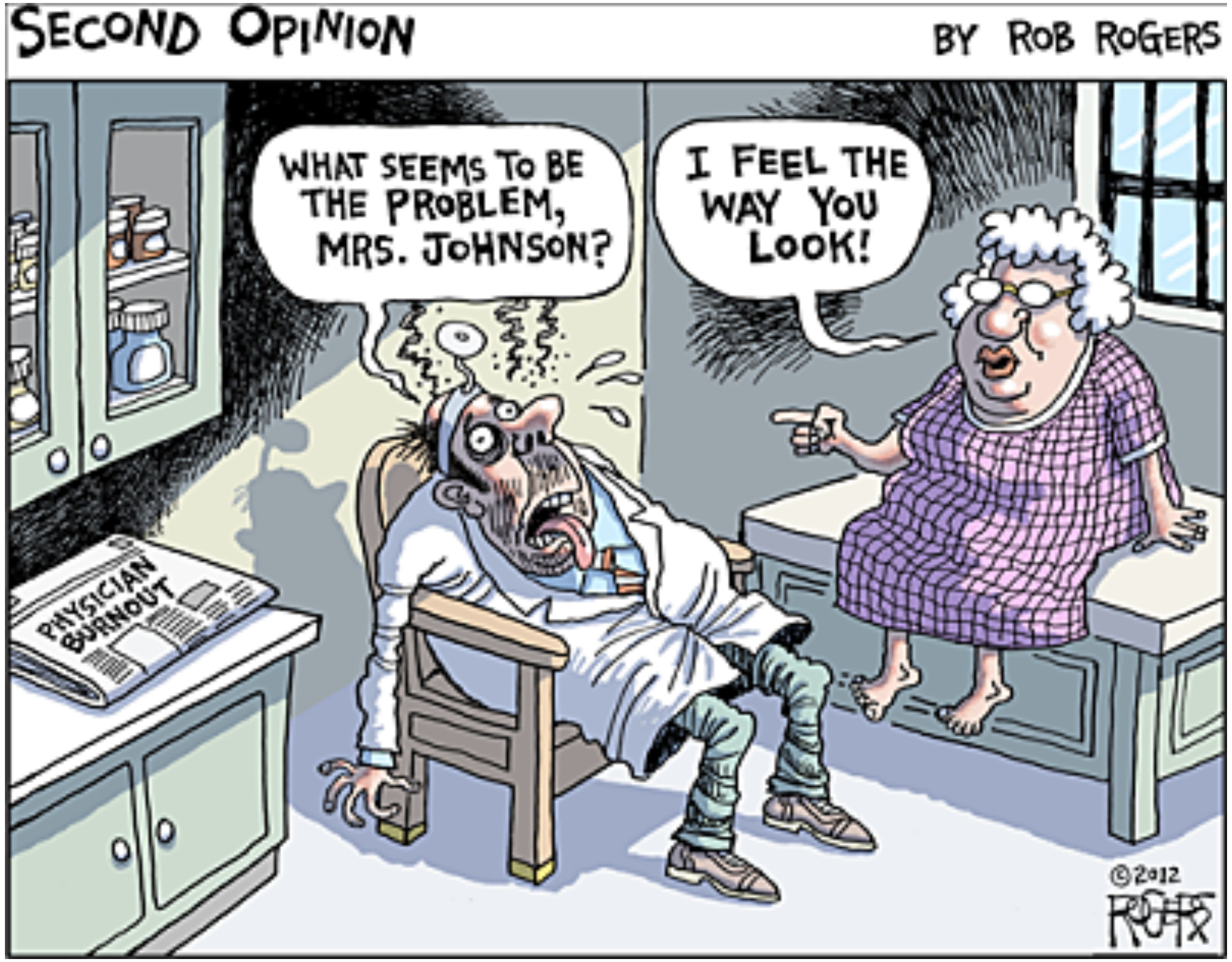
- Classification task: Cancer vs. No cancer
- Train CNN with semantic labels to make output more interpretable to humans
- Trained on LIDC dataset
- Assess classification performance for semantic labels & diagnostic prediction
- Comparing CNN to hierarchical semantic network:
  - 3D single CNN AUC = 0.847 (± 0.024)
  - 3D hierarchical semantic network AUC = 0.856 (± 0.026)
  - Mean difference = 0.005 (95% CI: 0.0051-0.0129); p = 0.009



Concluding Thoughts: Towards Wisdom?




MAXIMIZING
PATIENT
BENEFITS...


SECOND OPINION
BY ROB ROGERS
WHAT SEEMS TO BE THE PROBLEM, MRS. JOHNSON?
I FEEL THE WAY YOU LOOK!
PHYSICIAN BURNOUT
©2012
ROGERS

A STAT INVESTIGATION

IBM pitched its Watson supercomputer as a revolution in cancer care. It's nowhere close

By CASEY ROSS @caseymross and IKE SWETLITZ @ikeswetlitz / SEPTEMBER 5, 2017

Machine Learning and Prediction in Medicine — Beyond the Peak of Inflated Expectations

Jonathan H. Chen, M.D., Ph.D., and Steven M. Asch, M.D., M.P.H.

VIEWPOINT

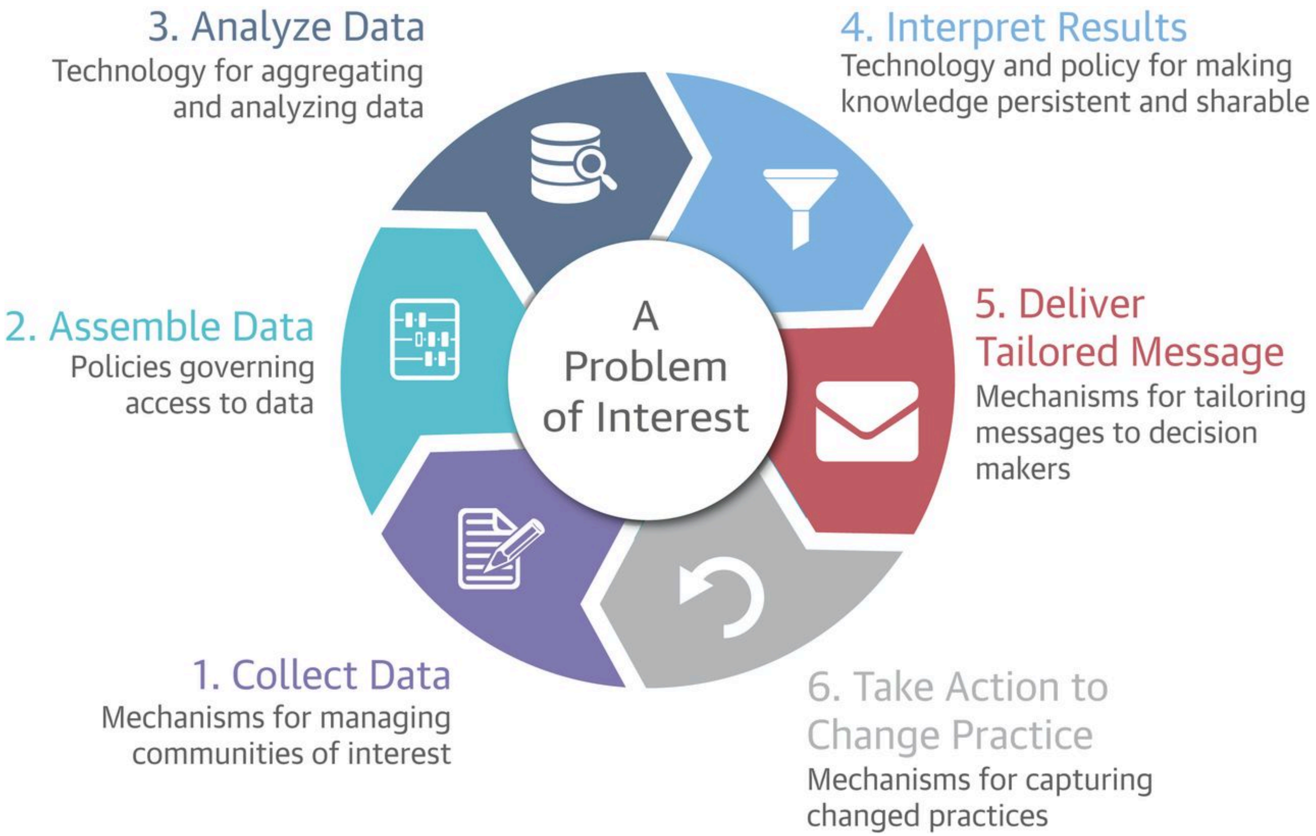
Unintended Consequences of Machine Learning in Medicine

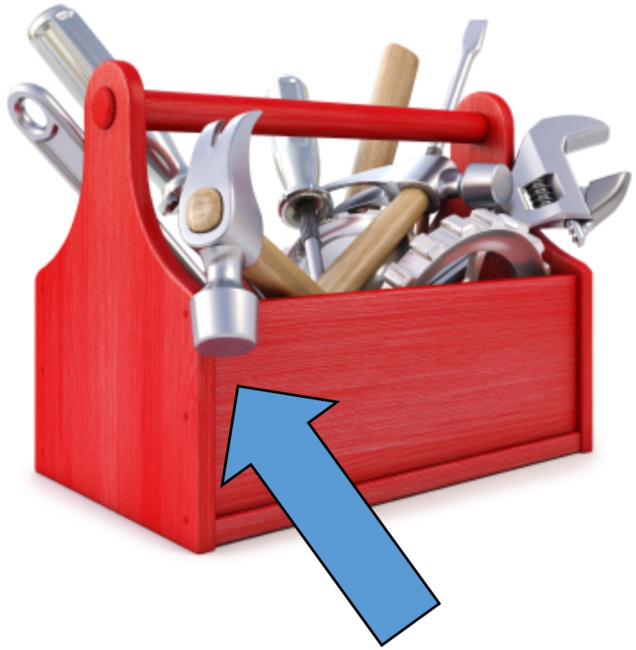
Federico Cabitza, PhD
Department of Informatics, University of Milano-Bicocca, Milan, Italy; and IRCCS Istituto Ortopedico Galeazzi, Milan, Italy.

Over the past decade, machine learning techniques have made substantial advances in many domains. In health care, global interest in the potential of machine learning has increased; for example, a deep learning algorithm has shown high accuracy in detecting diabetic retinopathy.[1] There have been suggestions that machine learning will drive changes in health care within a few years, specifi-

the expense of other elements that are more difficult or impossible to easily describe. Relying on ML-DSS requires considering digital data as reliable and complete representations of the phenomena that these data are supposed to render in a discrete and trustworthy form. This may be a problem when the clinical context is not represented, particularly if physicians lose awareness of

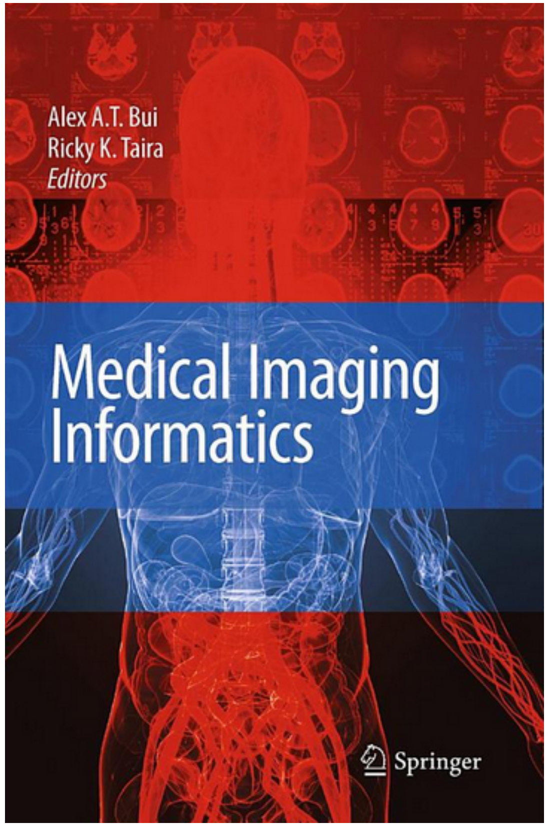
A Learning Health System

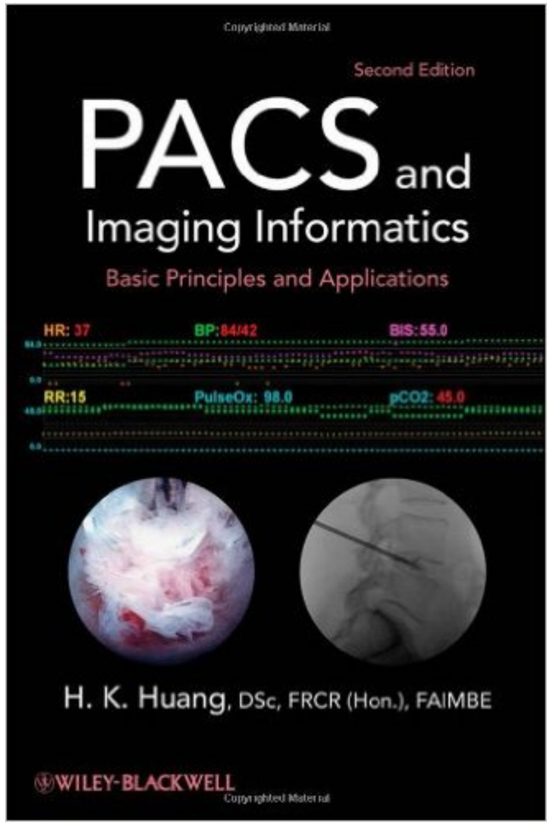


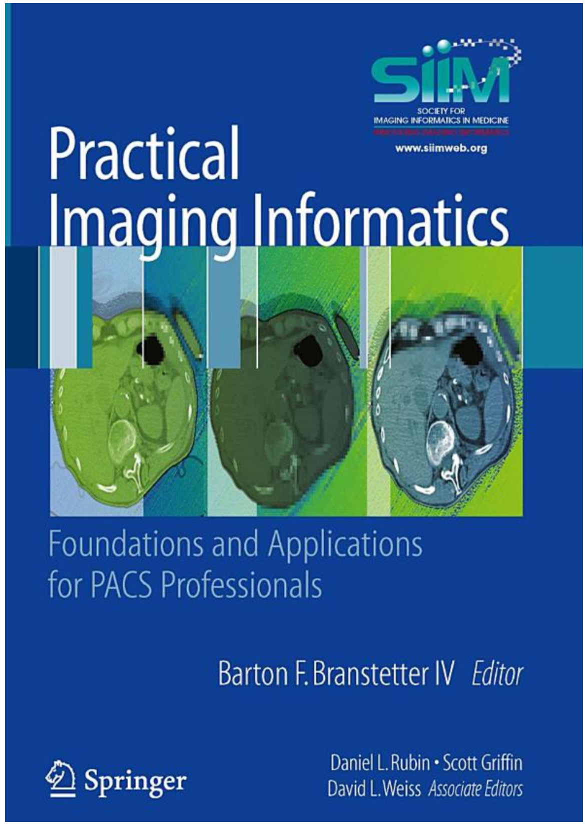


Data integration, machine learning

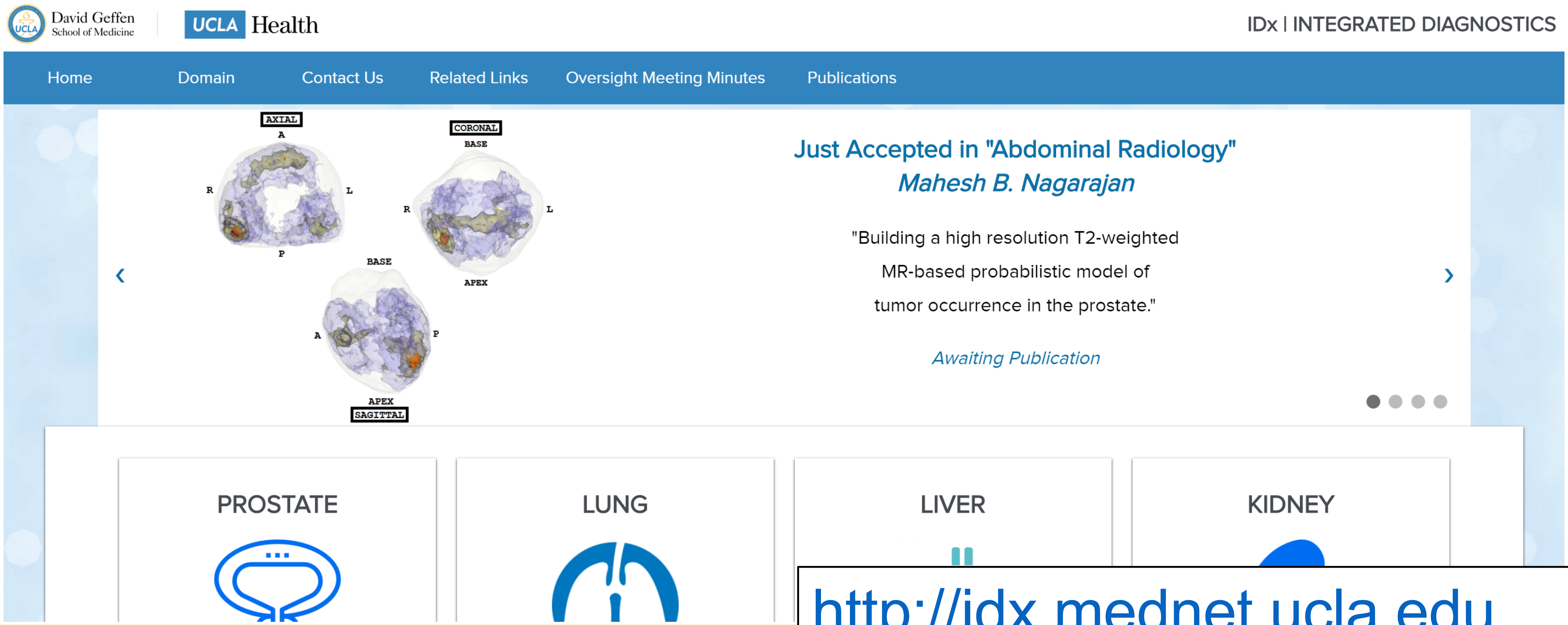
Resources


Alex A.T. Bui
Ricky K. Taira
Editors
Medical Imaging Informatics
Springer


Second Edition
PACS and Imaging Informatics
Basic Principles and Applications
H. K. Huang, DSc, FRCR (Hon.), FAIMBE
WILEY-BLACKWELL

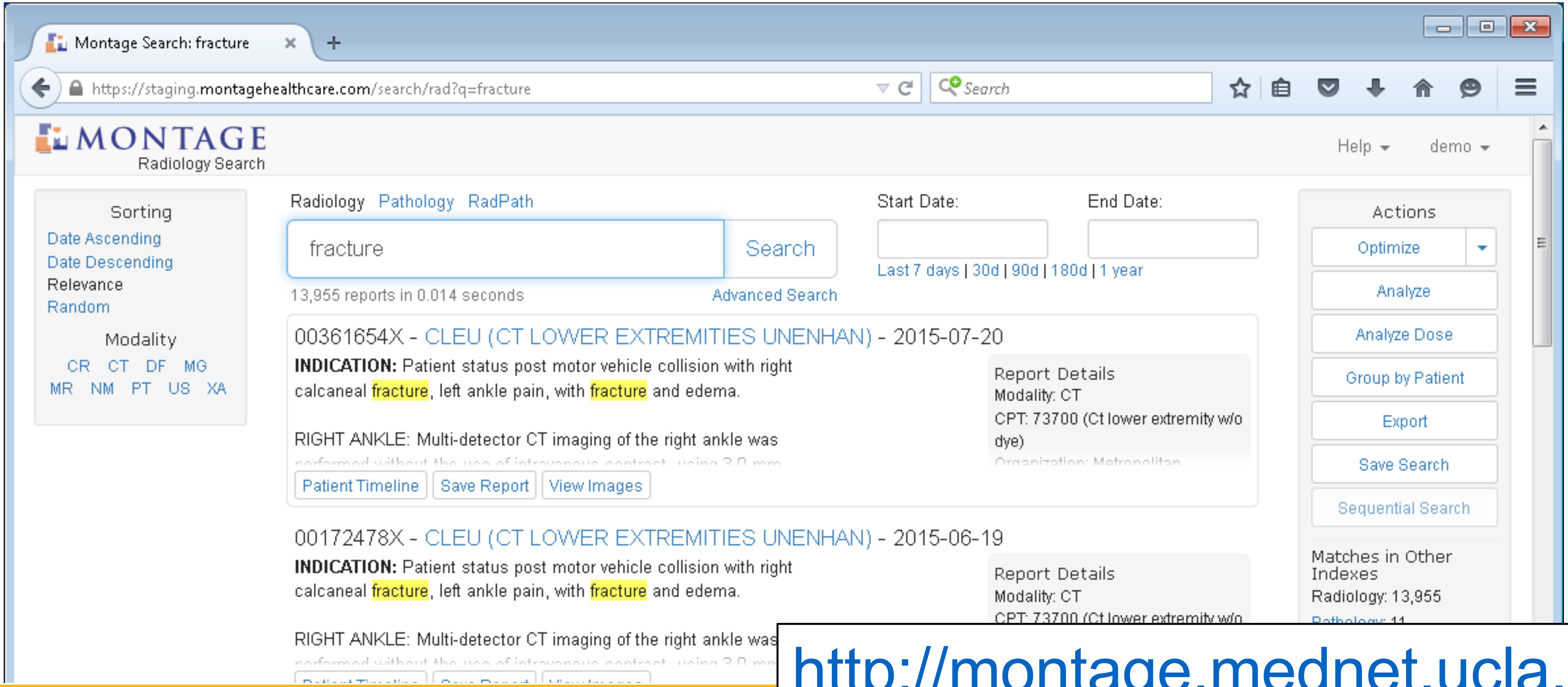

SIIM
www.siimweb.org
Practical Imaging Informatics
Foundations and Applications for PACS Professionals
Barton F. Branstetter IV Editor
Springer
Daniel L. Rubin • Scott Griffin
David L. Weiss Associate Editors

Resources

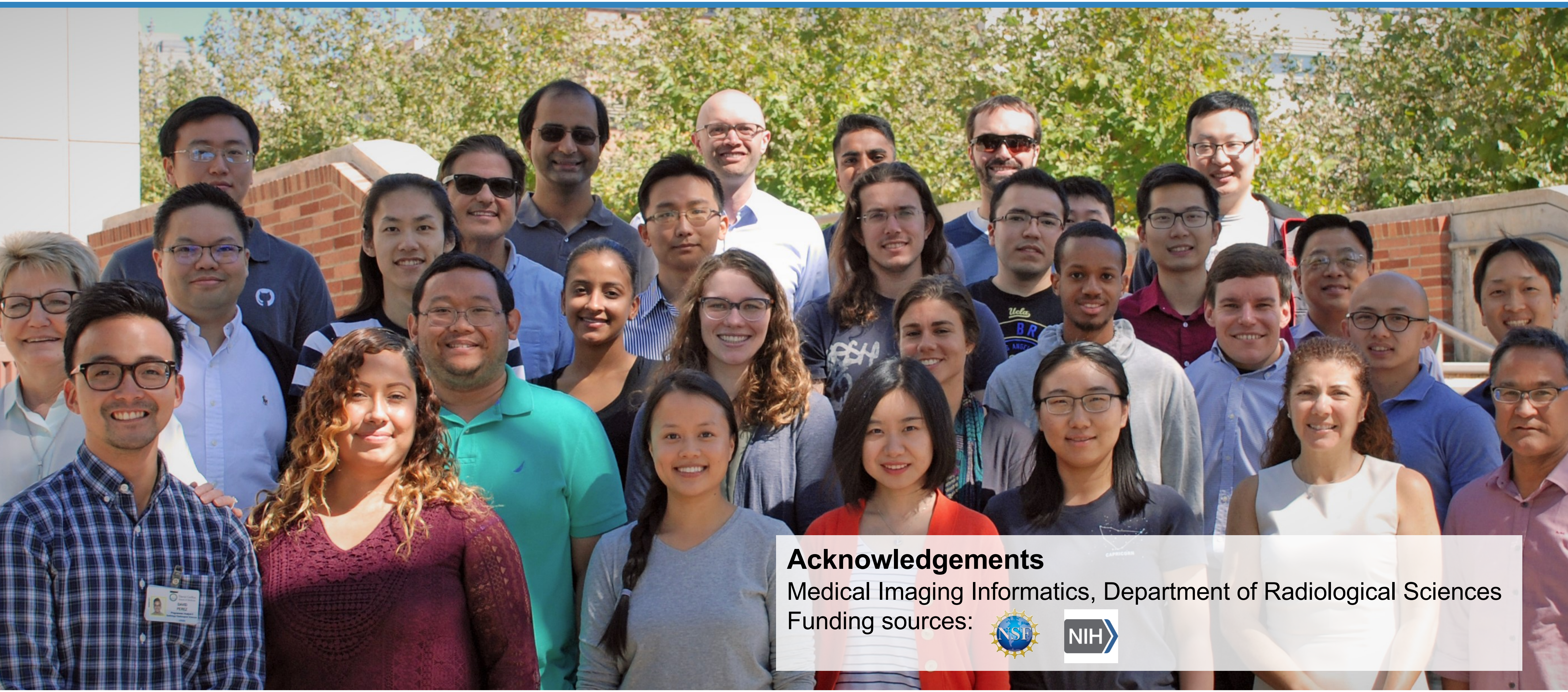


http://idx.mednet.ucla.edu

Resources



http://montage.mednet.ucla.edu



Acknowledgements

Medical Imaging Informatics, Department of Radiological Sciences

Funding sources: NSF NIH