# Natural Language Processing and Enhanced Clinical Decision Making Radiology and VINCI


Eliot Siegel, M.D.

Professor and Vice Chairman University of Maryland School of Medicine Department of Diagnostic Radiology and Nuclear Medicine

Chief Imaging VA Maryland Healthcare System


# Introduction

- Radiology is specialty with long history of research in NLP and Enhanced Decision Making but these may be critical in the continuing success of our specialty
- ACR Annual Meeting: "Quality is our Image"
  - But increased pressure to increase efficiency while not sacrificing quality
- SPIE 30th anniversary of PACS and research in computer aided detection and quantification of disease

# Content Based Image Retrieval

- Criteria include:
  - Grayscale and color
  - Texture
  - Point sets
  - Contours/curves
  - Surfaces
  - Regions and parts
    - Global shape, morphology, location and spatial

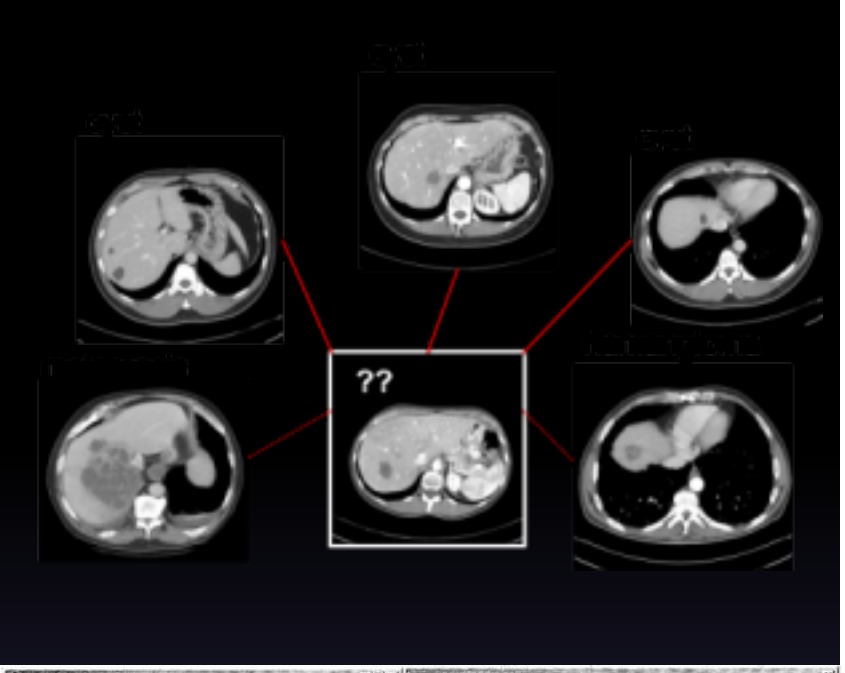






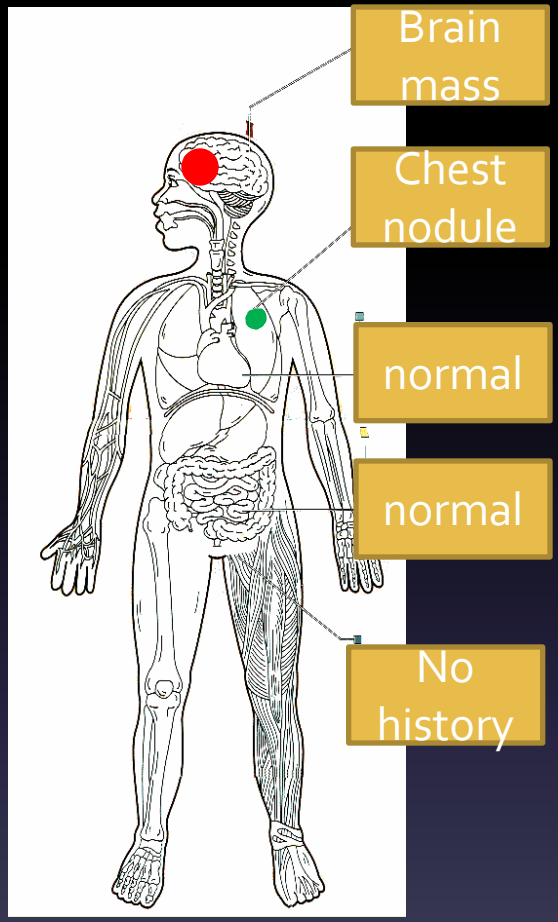
**Table 2. Main Characteristics of the Medical CBIR[a] Systems**

| Ref. | Descriptors | Similarity measures | Segmentation | Relevance feedback | Notes on modalities and datasets |
|---|---|---|---|---|---|
| 24 | General | Classifier-based | – | Yes | • 5,000 images from daily routine (a subset of imageCLEFmed)<br>• 20 categories (different organs, modalities, views) |
| 21a | General | Classifier-based | Manual | No | • 147 ROIs extracted from CT liver images<br>• 76 control vs. 71 pathological |
| 32 | General | Classifier-based | Manual | No | • 57 ROIs extracted from mammograms<br>• 37 benign vs. 20 malignant |
| 16 | Specialized | Classifier-based | Interactive | No | • 302 lung CT images<br>• 8 lung disease categories |
| 15 | Mixed | Classifier-based | – | No | • 8,725 images (CasImage[b] database)<br>• Medical retrieval task in imageCLEF 2004<br>• 26 query topics |
| 34 | Specialized | Procrustes | Manual | No | • NHANES II spine X-ray images<br>• 250 vertebra boundary profiles<br>• 10 categories of cervical and lumbar vertebra shapes |
| 17a | Mixed | Classifier-based | Manual | No | • 150 endoscopy images<br>• Several classes from endoscopic findings/diagnoses |
| 36a | Specialized | Classifier-based | – | No | • Brain MR images<br>• Hippocampus in schizophrenia (15 control vs. 15 patients)<br>• Corpus callosum in affective disorder (20 controls vs. 18 patients) |
| 33 | Specialized | Procrustes | Manual | No | • NHANES II spine X-ray images<br>• 2,812 vertebra boundary profiles<br>• No classification analysis |
| 58 | General | Classifier-based | Manual | Yes | • 76 Mammograms containing clustered microcalcifications<br>• Ground truth similarity obtained from human observer studies |
| 59a | General | Elastic deformation | – | No | • 90 cardiac ultrasound images<br>• View classification |
| 61 | Specialized | Elastic deformation | Interactive | No | • 100 intravascular ultrasound images containing calcium plaque structures<br>• Similarity-based retrieval used for improving registration |
| 14 | General | Vector distance | – | No | • Abdominal ultrasound images |
| 57 | General | Vector distance<br>Classifier-based | Interactive | No | • 70 brain MR images<br>• Hippocampus localization and identification<br>• 10 epileptic patients |
| 66 | General | Classifier | Automated | No | • Image categorization and retrieval on 1500 radiological images (IRMA project X-ray library).<br>• 17 radiological X-ray classes |
| 40a | General | Vector distance<br>Classifier-based | Manual | No | • fMRI activation contrast maps used as correlates of Alzheimer's disease<br>• 9 controls vs. 9 patients |
| 60 | Mixed | Vector distance<br>Textual | Automatic | No | • 300 VOIs extracted from 13 dynamic PET brain scans<br>• 2 tumor cases, 3 normal, 8 other neurological cases |
| 43 | Specialized | Graph matching | Manual | No | • 124 MR images<br>• No classification analysis<br>• Target application: indexing and fast search |
| 55 | General | Vector distance<br>Graph matching<br>Classifier-based | Automatic | Yes | • 1,617 radiographs from daily routine<br>• Classification based on image modality, body orientation, anatomic region, biological system |
| 56 | Mixed | Vector distance<br>Textual | – | No | • ImageCLEF 2005 medical retrieval tasks |
| 67 | General | Classifier | – | No | • Categorization: database of 12,000 radiographs (ImageClef 2007); 11,000 training, 1,000 testing; 116 different categories<br>• Retrieval: database of 66,000 images, 30 query topics (ImageClef 2008). Return ranked set of 1,000 images. |

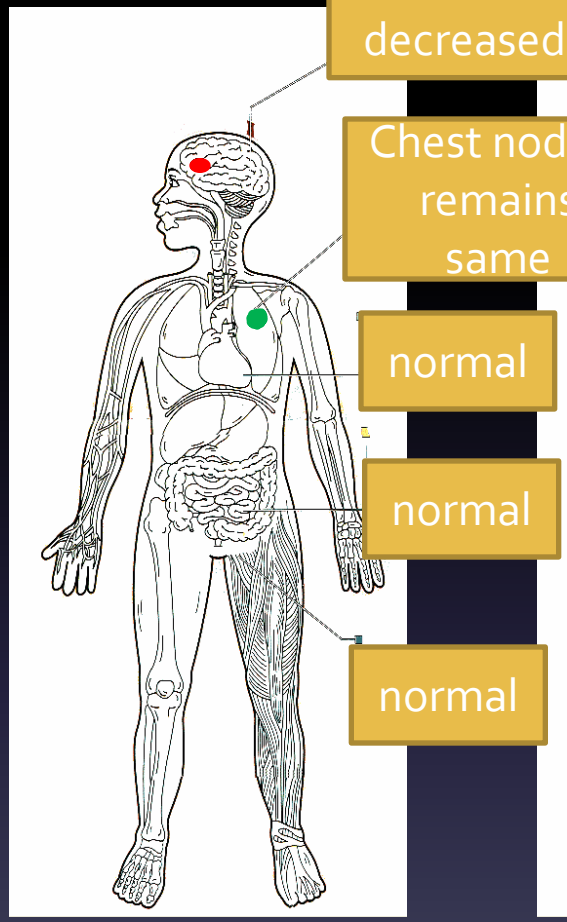
# Introduction

- SIIM meeting with Sam Dwyer presentation on next generation of PACS which will reply heavily on NLP and Enhanced Decision Support to a large extent from EMR and PHR
  - Relevant Clinical information which is rarely obtained in most places
  - Personalized medicine
  - Pathology correlation
  - Chart timelines
- How do we get there from here?

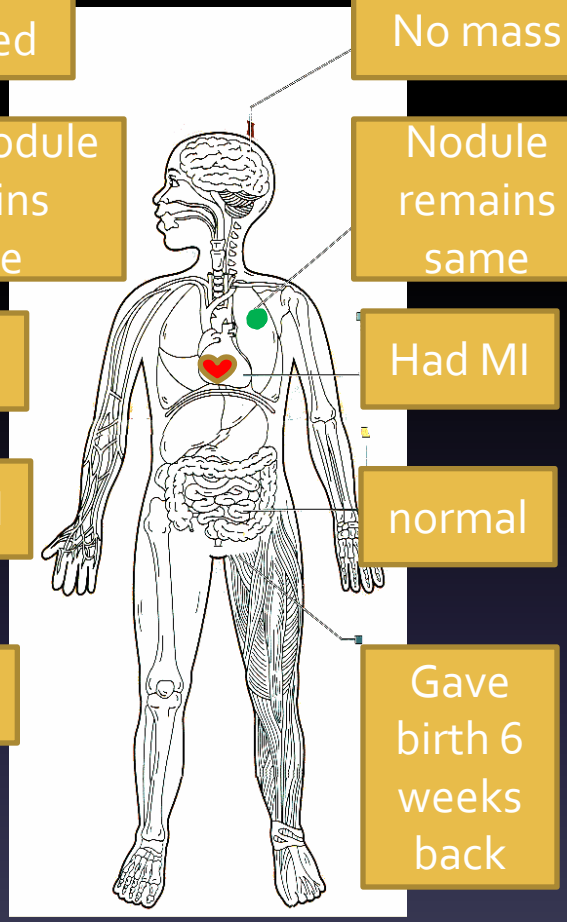
# Graphical sketch of a patient's radiological history




January 2000




May 2001




March 2002

# Outline

- Natural Language Processing and Enhanced Clinical Decision-Making in diagnostic imaging
- The Department of Veterans Affairs Vinci Project
  - An Overview and Opportunities
- IBM's Watson Deep Q/A software
  - Potential Medical Applications

# NLP in Radiology (Medical Imaging)

- Diagnostic Radiology reports, although text based, have a constrained vocabulary and limited number of concepts for each imaging modality
  - This combination has made radiology an ideal specialty to employ natural language processing and hundreds of articles have been written on the topic over the past more than 30 years

# Recent NLP in Radiology Articles

- Machine and radiology
- Automatic identification of critical follow-up recommendation sentences in radiology reports, Yetisgen-Yildiz, Gunn, Xia, Payne
- Natural language processing for lines and devices in portable chest x-rays
- Informatics in radiology: RADTF: a semantic search-enabled natural language process enabled radiology teaching file
- Informatics in radiology: Render; an online searchable radiology study repository
- Discerning tumor status from unstructured MRI reports—completeness of information in existing reports and utility of automated natural language processing

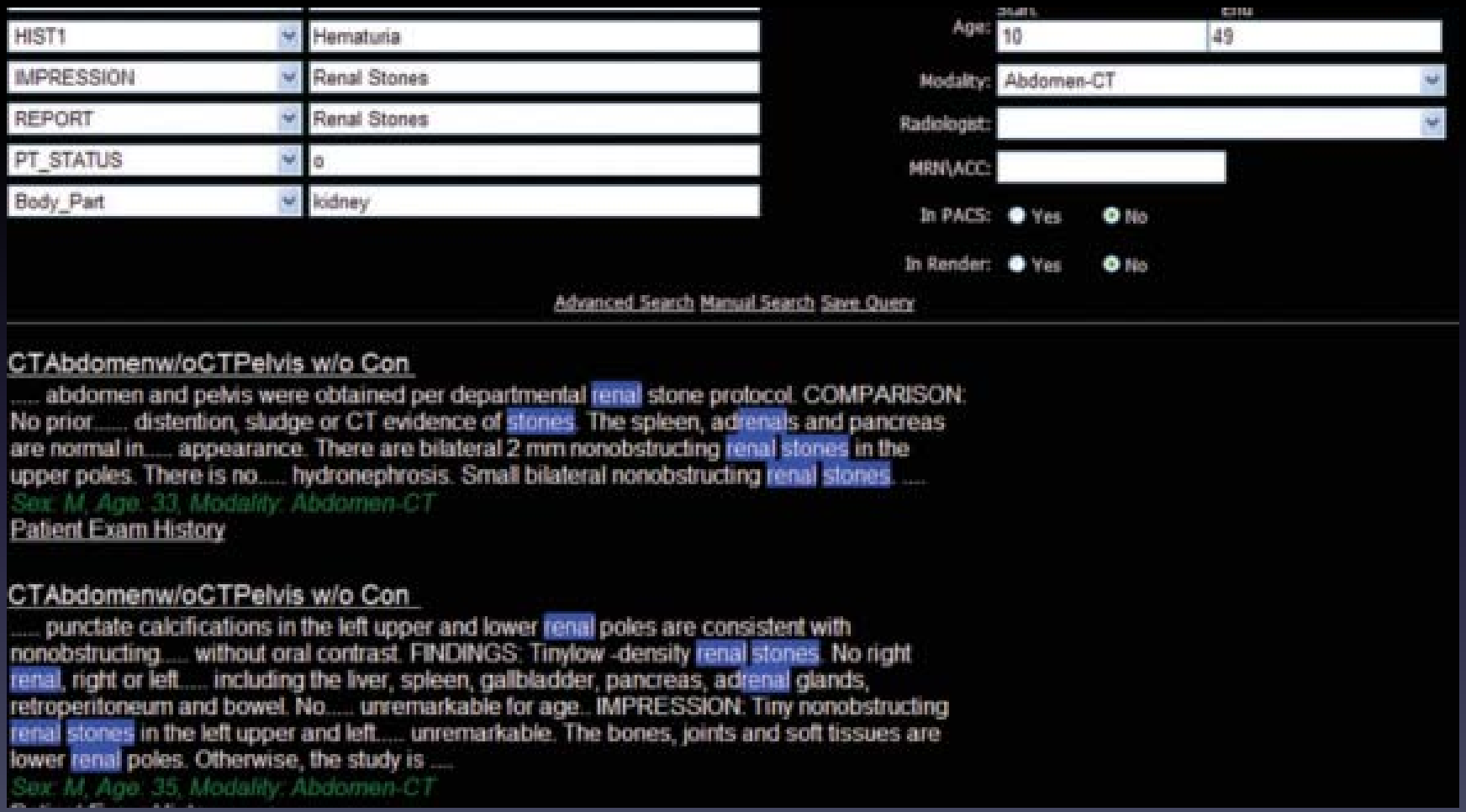
# Recent NLP in Radiology Articles

- Use of Radcube for extraction of finding trends in a large radiology practice
- Natural language processing using online analytic processing for assessing recommendations in radiology reports
- A study of lexical behavior of sentences in chest radiology reports
- Indexing anatomical phrases in neuro-radiology reports to the UMLS 2005AA
- Extracting information on pneumonia in infants using natural language processing of radiology reports
- Improved identification of noun phrases in clinical radiology reports using a high-performance natural language parser augmented with the UMLS specialist lexicon

# Four Fundamental Elements

- Identification of a study as positive or negative
- Recommendations
- Unexpected findings
- Indication and type of study for automatic protocol generation

# Mass General Dryer et al
# LEXIMER
# (Lexicon Mediated Entropy Reduction)



# Leximer Analysis of Millions of Reports

- **Overall Reports: 8070**

- Overall Positive Findings: 5819 (72.1%)
- Overall Negative Findings: 2251 (27.9%)
- Overall Positive Recommendations: 749 (9.3%)
- Overall Findings Positive with No Recommendations: 5151 (63.8%)
- Overall Findings Positive with Recommendations: 668 (8.3%)
- Overall Findings Negative with No Recommendations: 2167 (26.9%)
- Overall Findings Negative with Recommendations: 84 (1.0%)
- Time to Process: 3 seconds (Estimated Processing Rate: 10 million reports per hour)

**Patient:** IGNORE,TEST **MRN:** 0000006 **Ordering Physician:** Physician 1

Save/Complete Cancel

# HEAD CT

**Policy on CONTRAST: Scans are performed according to departmental protocols selected by the radiologist unless otherwise specified.**

**Special Considerations, Check If Appropriate**

**Contrast**

☐ Contrast MUST NOT BE USED because (Required):

☐ 3D
☐ Dissection Protocol (schedule chest, abdomen and pelvis in same time slot)
☐ Pregnant
☐ Reformats (sagital/coronal)
☐ Send additional report copies to:

☐ BUN/Creat (If Known)
☐ EVT Protocol (schedule abdomen and pelvis in same time slot)
☐ Radiation planning
☐ Stereotactic

**At least one box MUST be selected from either of the following groups**

**SIGNS / SYMPTOMS**

☐ Ataxia
☐ Convulsions
☐ Dementia
☐ Headache chronic with progressive worsening
☐ Headache sudden onset or Thunderclap
☐ Hyperprolactinemia
☐ Pain in face
☐ Swelling, mass or lump
☐ TIA with transient neurological disturbance
☐ Weakness-right/left/both

☐ Concussion mild or moderate acute, no neurological deficit
☐ Coordination changes, new or progressive
☑ Dizziness
☐ Headache migraine or chronic
☐ Hearing changes
☐ Mental Status change (after trauma)
☐ Speech changes
☐ Syncope/fainting
☐ Vision changes

**KNOWN DIAGNOSES (NOT Rule/out!)**

☐ Aneurysm
☐ Neoplasm CNS primary (specify)
☐ Neoplasm-Primary Unknown

☐ Intracranial hemorrhage
☐ Neoplasm non-CNS primary (specify)
☐ Sub-dural hemorrhage

Information for radiologist (Only 140 Characters Allowed):

Save/Complete Cancel

**Patient:** IGNORE,TEST **MRN:** 0000006 **Ordering Physician:** Physician 1

| Your Order: CT | | MR |
|---|---|---|
| 4 | | 8 |
| Proceed with Order | | Change Order |

Indicated 7-9 Marginal 4-6 Low utility 1-3

# HEAD CT

**Policy on CONTRAST: Scans are performed according to departmental protocols selected by the radiologist unless otherwise specified.**

**Special Considerations, Check If Appropriate**

**Contrast**

☐ Contrast MUST NOT BE USED because (Required):

☐ 3D
☐ Dissection Protocol (schedule chest, abdomen and pelvis in same time slot)
☐ Pregnant
☐ Reformats (sagital/coronal)
☐ Send additional report copies to:

☐ BUN/Creat (If Known)
☐ EVT Protocol (schedule abdomen and pelvis in same time slot)
☐ Radiation planning
☐ Stereotactic

**At least one box MUST be selected from either of the following groups**

**SIGNS / SYMPTOMS**

☐ Ataxia
☐ Convulsions
☐ Dementia
☐ Headache chronic with progressive worsening
☐ Headache sudden onset or Thunderclap
☐ Hyperprolactinemia
☐ Pain in face
☐ Swelling, mass or lump
☐ TIA with transient neurological disturbance
☐ Weakness-right/left/both

☐ Concussion mild or moderate acute, no neurological deficit
☐ Coordination changes, new or progressive
☑ Dizziness
☐ Headache migraine or chronic
☐ Hearing changes
☐ Mental Status change (after trauma)
☐ Speech changes
☐ Syncope/fainting
☐ Vision changes

**KNOWN DIAGNOSES (NOT Rule/out!)**

☐ Aneurysm
☐ Neoplasm CNS primary (specify)
☐ Neoplasm-Primary Unknown

☐ Intracranial hemorrhage
☐ Neoplasm non-CNS primary (specify)
☐ Sub-dural hemorrhage

Information for radiologist (Only 140 Characters Allowed):

# Any Hospital
## Department of Radiology

**Patient:** IGNORE,TEST **MRN:** 0000006 **Ordering Phys:** Physician 1

| Exam | Ordering Phys | Schedule Location/Time | Special Considerations | History |
|---|---|---|---|---|
| Cancel Exam HEAD (BRAIN) MRI | Physician 1 | Click Here to Schedule | No special considerations | -Dizziness 780.4 |

Start Again Print


**CT SCAN**
SELECT ONE


**MRI**
SELECT ONE


**Ultrasound**
SELECT ONE


**Nuclear Medicine**
SELECT ONE

**Bone Densitometry**
SELECT ONE

**Mammography**
SELECT ONE


**Plain Film**
SELECT ONE

**Spine,Bone & Procedures**
SELECT ONE

**Nuclear Cardiology Exams**
SELECT ONE

**FOR CONSULTATION PLEASE CALL:**

| | |
|---|---|
| MUSCULOSKELETAL | 726-7717 |
| CHEST | 724-4254 |
| ABDOMINAL | 726-8396 |
| NEURORADIOLOGY | 726-8320 |
| NUCLEAR | 726-8350 |
| PEDIATRICS | 724-4207 |

# Any Hospital
## Department of Radiology

**Patient:** IGNORE,TEST  **MRN:** 0000006  **Ordering Phys:** Physician 1

| Exam | Ordering Phys | Schedule Location/Time | Special Considerations | History |
|---|---|---|---|---|
| Cancel Exam HEAD (BRAIN) MRI | Physician 1 | Click Here to Schedule | No special considerations | -Dizziness 780.4 |

Start Again | Print

**CT SCAN**
SELECT ONE

**Nuclear Medicine**
SELECT ONE

**Plain Film**
SELECT ONE

**MRI**
SELECT ONE
- SELECT ONE
- ABD/PELVIS MRI
- CHEST MRI
- EXTREMITY MRI
- HEAD (BRAIN) MRI
- SPINE MRI
- FACE OR SINUS MRI
- PELVIS SOFT TISSUE MRI
- PELVIS (BONE) MRI
- KIDNEY/ADRENAL (GU) MRI
- LIVER/PANCREAS/SPLEEN MRI
- NECK MRI
- BREAST MRI
- MR ARTHROGRAM

**Ultrasound**
SELECT ONE

**Mammography**
SELECT ONE

**Nuclear Cardiology Exams**
SELECT ONE

**FOR CONSULT**

| | |
|---|---|
| MUSCULOSKELETA | 726-7717 |
| CHEST | 724-4254 |
| ABDOMINAL | 726-8396 |
| NEURORADIOLOGY | 726-8320 |
| NUCLEAR | 726-8350 |
| PEDIATRICS | 724-4207 |

**Patient:** IGNORE,TEST **MRN:** 0000006 **Ordering Physician:** Physician 1

Save/Complete Cancel

# EXTREMITY MRI

**Patients must be cooperative and able to hold still for 1 hour to have an MRI Scan**
**If sedation is required call 4-XRAY**

**Special Considerations, Check If Appropriate**

- ☐ With Contrast
- ☐ Head/Neck MRA
- ☐ Send additional report copies to:
- ☐ Intercranial aneurysm clip
- ☐ Ear implant or prosthesis
- ☐ Metallic foreign body
- ☐ Pregnant
- ☐ MRA
- ☐ With & w/o Contrast
- ☐ Pacing device
- ☐ Artificial heart valve
- ☐ Employment as metal worker
- ☐ Claustrophobia
- ☐ MRCP

**EXAM REQUESTED** **Pick only ONE of the Following**

| | |
|---|---|
| ○ **Left** | ◉ **Right** |
| ○ Shoulder | ○ Arm |
| ○ Elbow | ○ Forearm |
| ○ Wrist | ○ Hand |
| ○ Hip | ○ Thigh |
| ◉ Knee | ○ Leg |
| ○ Ankle | ○ Foot |

**At least one box MUST be selected from either of the following groups**

**SIGNS / SYMPTOMS**

- ☐ Bone Pain
- ☐ Deformity
- ☐ Instability
- ☐ Joint Pain
- ☐ Limited movement
- ☐ Swelling or mass

**KNOWN DIAGNOSES (NOT Rule/out!)**

- ☑ Arthritis
- ☐ Aseptic necrosis
- ☐ Dislocation
- ☐ Fracture specify location
- ☐ Neoplasm - Musculoskeletal Primary (specify)
- ☐ Neoplasm - Non-musculoskeletal Primary (specify)
- ☐ Neoplasm - Primary Unknown
- ☐ Nonunion
- ☐ Osteomyelitis

**ABNORMAL PREVIOUS EXAMINATIONS**

- ☐ Abnormal x-ray

Information for radiologist (Only 140 Characters Allowed):

**Patient**: IGNORE,TEST **MRN**: 0000006 **Ordering Physician**: Physician 1

| Your Order: MR | | X_Ray | CTA | MRA | CT |
|---|---|---|---|---|---|
| 3 | | 9 | 4 | 4 | 3 |
| Proceed with Order | | Change Order | | | |

Indicated 7-9 Marginal 4-6 Low utility 1-3

# EXTREMITY MRI

**Patients must be cooperative and able to hold still for 1 hour to have an MRI Scan**
**If sedation is required call 4-XRAY**

**Special Considerations, Check If Appropriate**

- ☐ With Contrast
- ☐ Head/Neck MRA
- ☐ Send additional report copies to:
- ☐ Intercranial aneurysm clip
- ☐ Ear implant or prosthesis
- ☐ Metallic foreign body
- ☐ Pregnant
- ☐ MRA
- ☐ With & w/o Contrast
- ☐ Pacing device
- ☐ Artificial heart valve
- ☐ Employment as metal worker
- ☐ Claustrophobia
- ☐ MRCP

**EXAM REQUESTED** **Pick only ONE of the Following**

| | |
|---|---|
| ○ **Left** | ◉ **Right** |
| ○ Shoulder | ○ Arm |
| ○ Elbow | ○ Forearm |
| ○ Wrist | ○ Hand |
| ○ Hip | ○ Thigh |
| ◉ Knee | ○ Leg |
| ○ Ankle | ○ Foot |

**At least one box MUST be selected from either of the following groups**

**SIGNS / SYMPTOMS**

- ☐ Bone Pain
- ☐ Deformity
- ☐ Instability
- ☐ Joint Pain
- ☐ Limited movement
- ☐ Swelling or mass

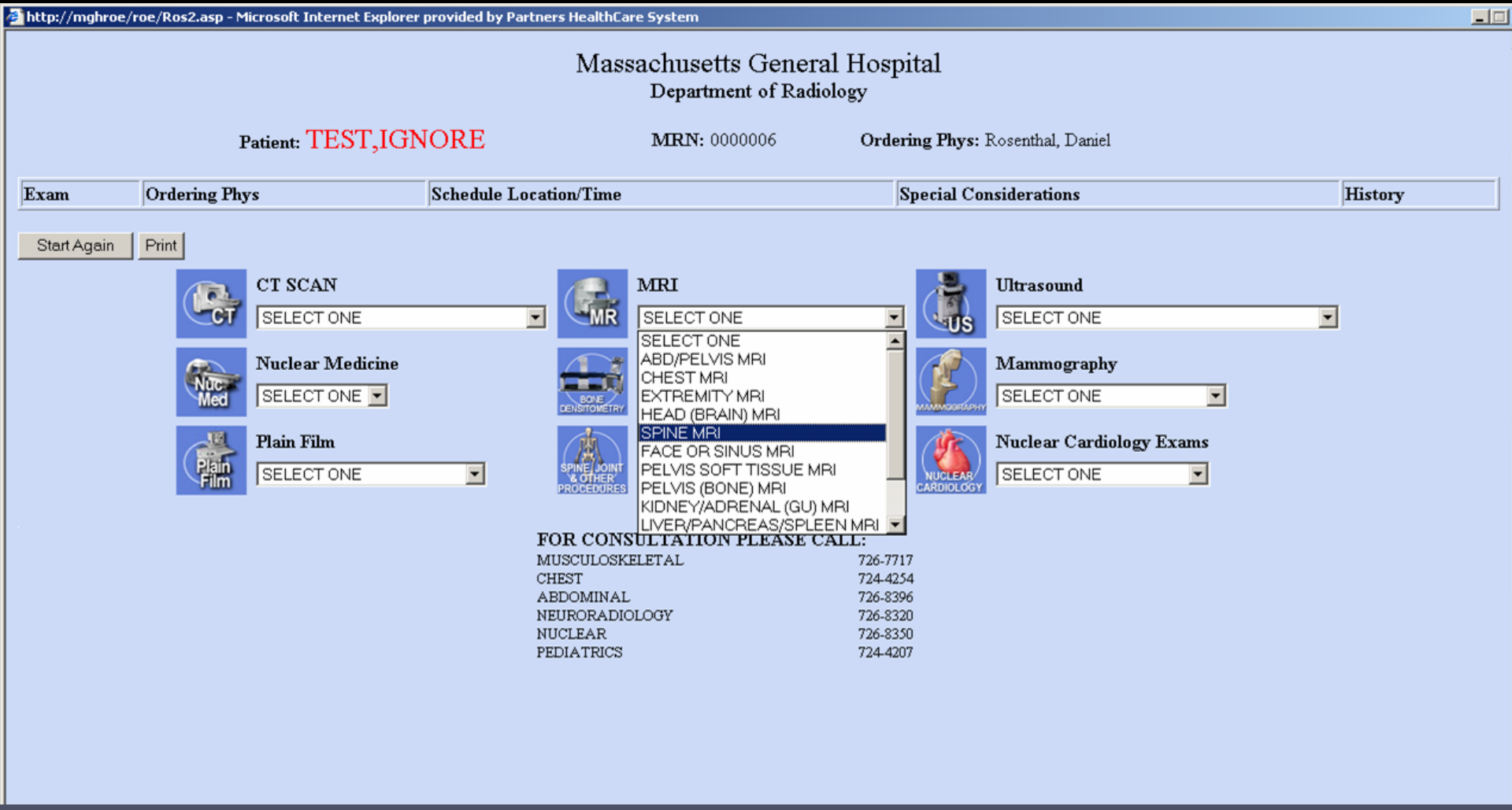
**KNOWN DIAGNOSES (NOT Rule/out!)**

- ☑ Arthritis
- ☐ Dislocation
- ☐ Neoplasm - Musculoskeletal Primary (specify)
- ☐ Neoplasm - Primary Unknown
- ☐ Osteomyelitis
- ☐ Aseptic necrosis
- ☐ Fracture specify location
- ☐ Neoplasm - Non-musculoskeletal Primary (specify)
- ☐ Nonunion

**ABNORMAL PREVIOUS EXAMINATIONS**

- ☐ Abnormal x-ray

# A “Low utility” exam



Requisition Form - Microsoft Internet Explorer provided by Partners Health

**Patient:** TEST,IGNORE **MRN:**

| Your Order: MR | | X_Ray | CT |
|---|---|---|---|
| 3 | | 3 | 2 |
| Proceed with Order | | Change Order | |

Indicated 7-9 Marginal 4-6 Low utility 1-3

# SPINE MRI

- ○ Cervical ○ Thoracic
- ◉ L-S ○ Sacrum
- ○ SI Joints ○ Limited complete (for mets)

**At least one box MUST be selected from either of the following groups**

**SIGNS / SYMPTOMS**

- ☐ Abnormal extremity reflexes ☐ Abnormal extremity sensation
- ☑ Back pain ☐ Back pain following trauma
- ☐ Back pain, prior surgery ☐ Extremity weakness (paraplegia)
- ☐ Neck pain ☐ Neck pain following trauma
- ☐ Radiculopathy ☐ Sciatic leg pain
- ☐ Swelling, mass or lump

**KNOWN DIAGNOSES (NOT Rule/out!)**

- ☐ Cauda Equina syndrome
- ☐ Congenital spine malformation (specify)
- ☐ Demyelinating disease with spinal cord syx (type)
- ☐ Demylinating disease without spinal cord syx
- ☐ Disk disease
- ☐ Neoplasm - Primary Unknown

| Your Order: MR | | CT | X_Ray |
|---|---|---|---|
| 9 | | 2 | 2 |
| Proceed with Order | | Change Order | |

Indicated 7-9 | Marginal 4-6 | Low utility 1-3

**Patients must be cooperative and able to hold still for 1 hour to have an MRI Scan**
**If sedation is required call 4-XRAY**

**Special Considerations, Check If Appropriate**

- ☐ With Contrast
- ☐ Head/Neck MRA
- ☐ Send additional report copies to:
- ☐ Intercranial aneurysm clip
- ☐ Ear implant or prosthesis
- ☐ Metallic foreign body
- ☐ Pregnant
- ☐ MRA
- ☐ With & w/o Contrast
- ☐ Pacing device
- ☐ Artificial heart valve
- ☐ Employment as metal worker
- ☐ Claustrophobia
- ☐ MRCP

**EXAM REQUESTED** **Pick only ONE of the Following**

- ○ Cervical ○ Thoracic
- ◉ L-S ○ Sacrum
- ○ SI Joints ○ Limited complete (for mets)

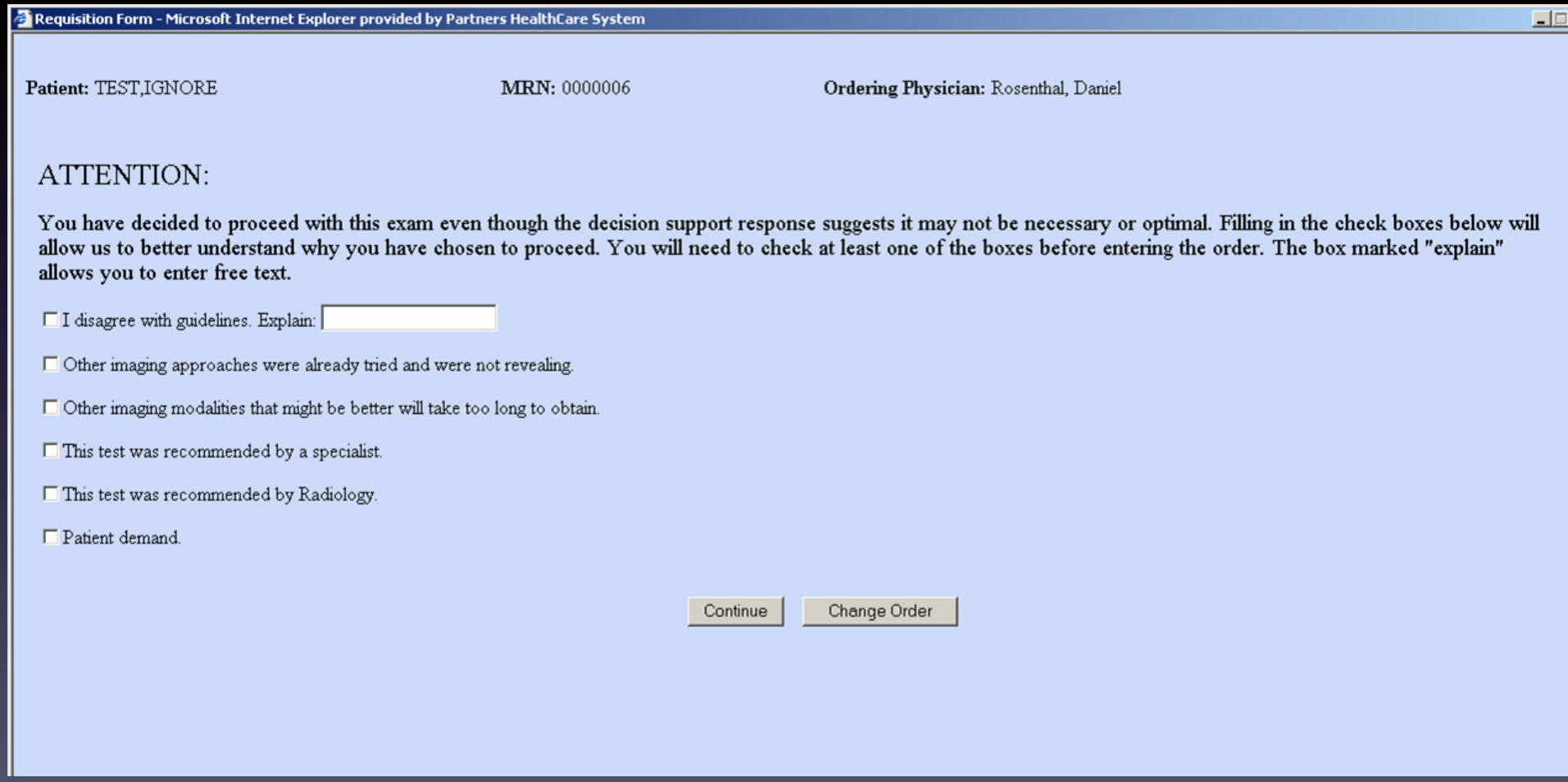
**At least one box MUST be selected from either of the following groups**

**SIGNS / SYMPTOMS**

- ☑ Abnormal extremity reflexes
- ☑ Back pain
- ☐ Back pain, prior surgery
- ☐ Neck pain
- ☐ Radiculopathy
- ☐ Swelling, mass or lump
- ☐ Abnormal extremity sensation
- ☐ Back pain following trauma
- ☐ Extremity weakness (paraplegia)
- ☐ Neck pain following trauma
- ☐ Sciatic leg pain

Fix the problem:
Changed order
with additional history

# Proceed on Red: Reasons



This study is reviewed with Dr Smith. Standard protocol was used to obtain an MRI of the brain with MRA of the circle of Willis and DWI imaging.

Dizziness and recurrent syncope. Please evaluate the posterior circulation. Comparison is to a CT of the head performed 3 September 99. Comparison is also to a CT performed the day after the MRI on 5 September 1909. Bilateral subdural hemorrhages are present. The right sided subdural hemorrhage appears improved when compared to the prior CT. It has a component extending further posteriorly than appreciated on the CT, appearing to involve the occipital lobe on the right side. The left subdural hemorrhage is worse than it appeared on the initial CT. There is extensive subarachnoid hemorrhage better appreciated on MRI than on CT.

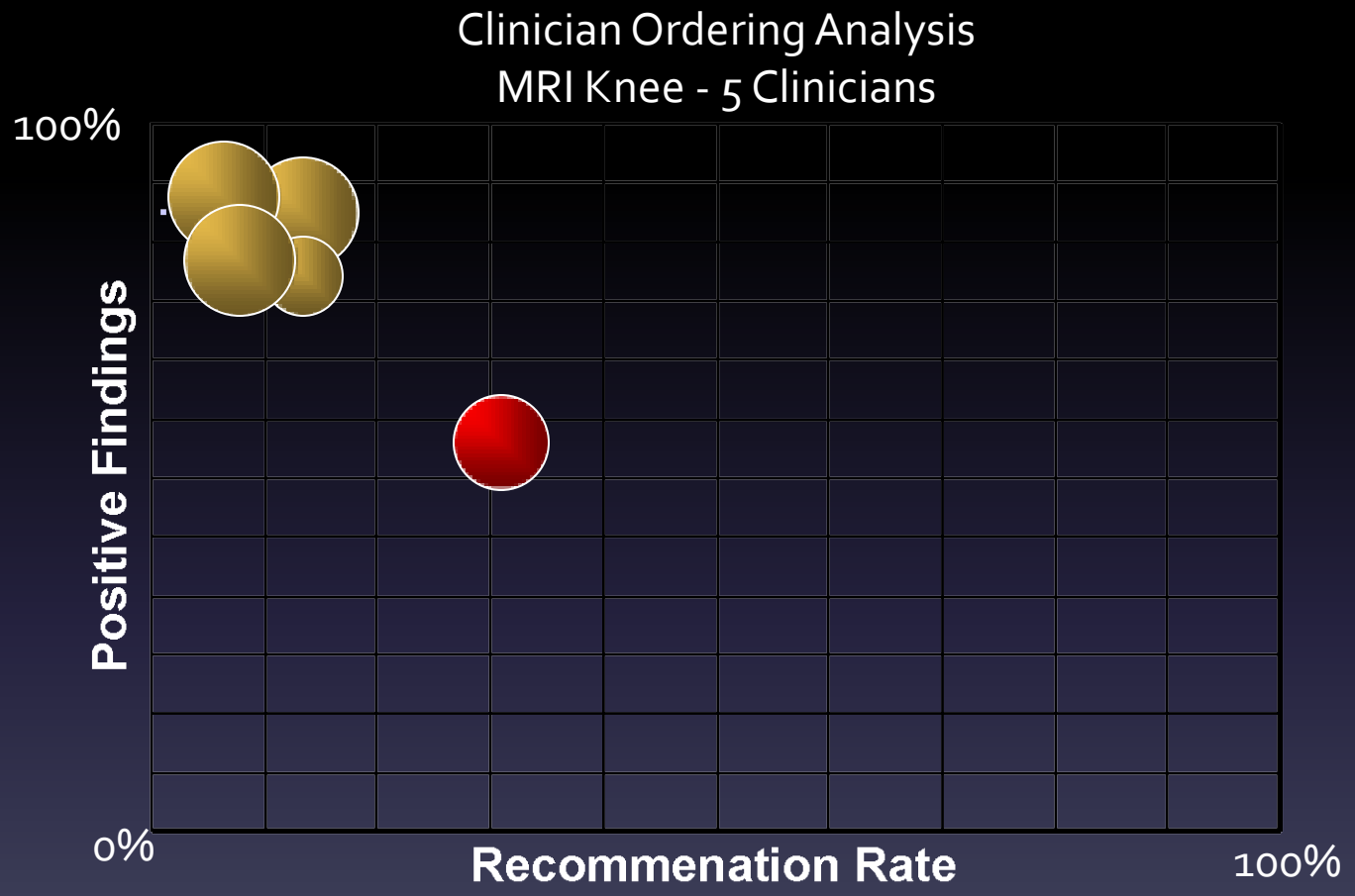
There is no evidence of tentorial subdural hematoma. The subsequent CT did show such a bleed, this must have occurred in the interval between studies. DWI imaging of the brain parenchyma is normal in appearance. There is no evidence of acute infarction. The circle of Willis was imaged with particular attention to the posterior circulation. The right vertebral artery appears prominent. The procedure circulation appears entirely normal. Because imaging was centered on the procedure circulation, the MCA's are not completely evaluated. The ventricular system and CSF spaces do not show evidence of abnormal dilation. The visualized extracranial structures are normal in appearance.

Impression. No evidence of acute infarction on diffusion weighted imaging. Bilateral subdural hemorrhages with subarachnoid hemorrhage. The posterior circulation appears entirely normal. A follow up MRI of the brain is recommended within 7 days to assess progression of hemorrhage.

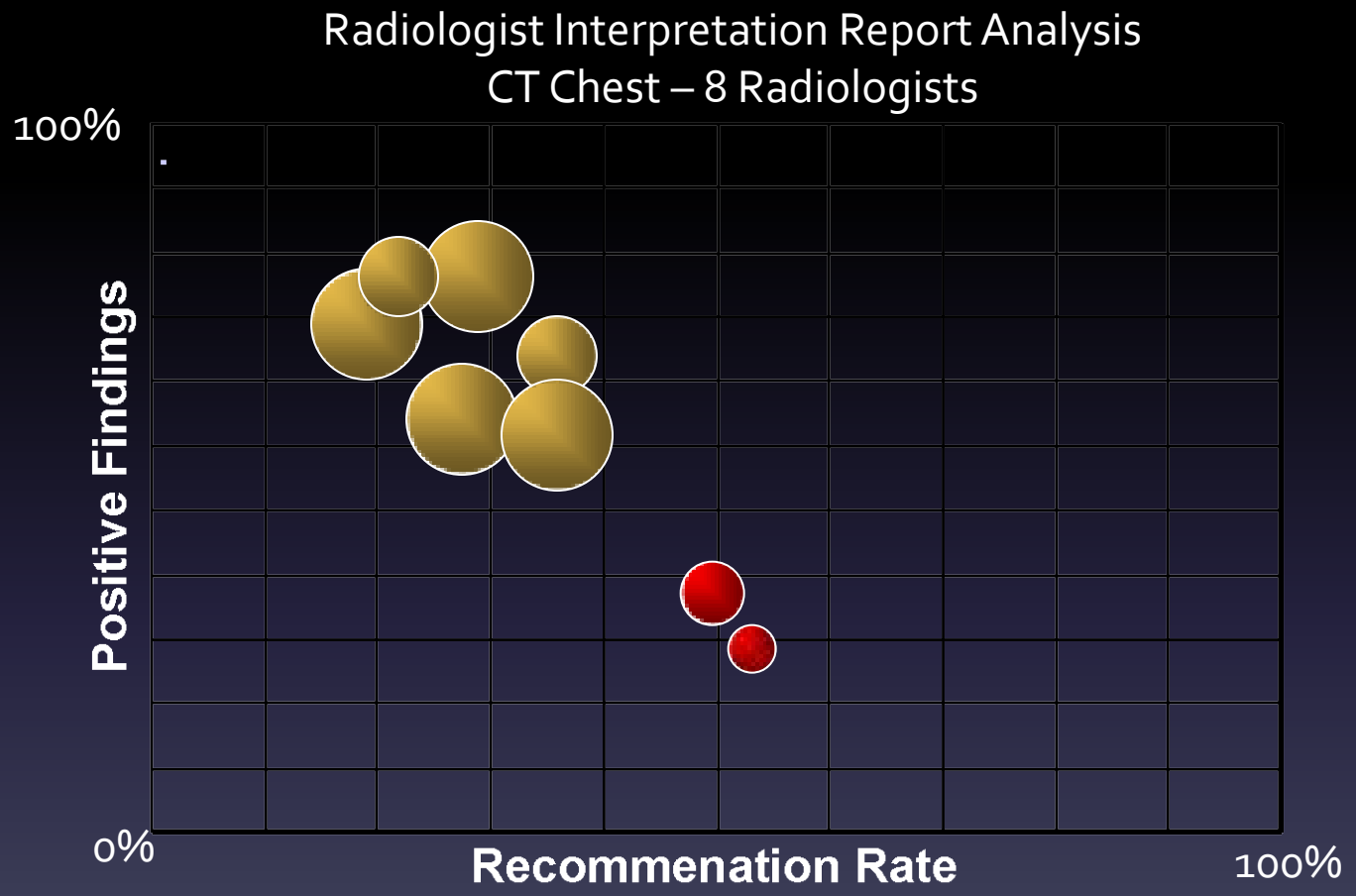
Bilateral subdural hemorrhages with subarachnoid hemorrhage.

A follow up MRI of the brain is recommended within 7 days to assess progression of hemorrhage.

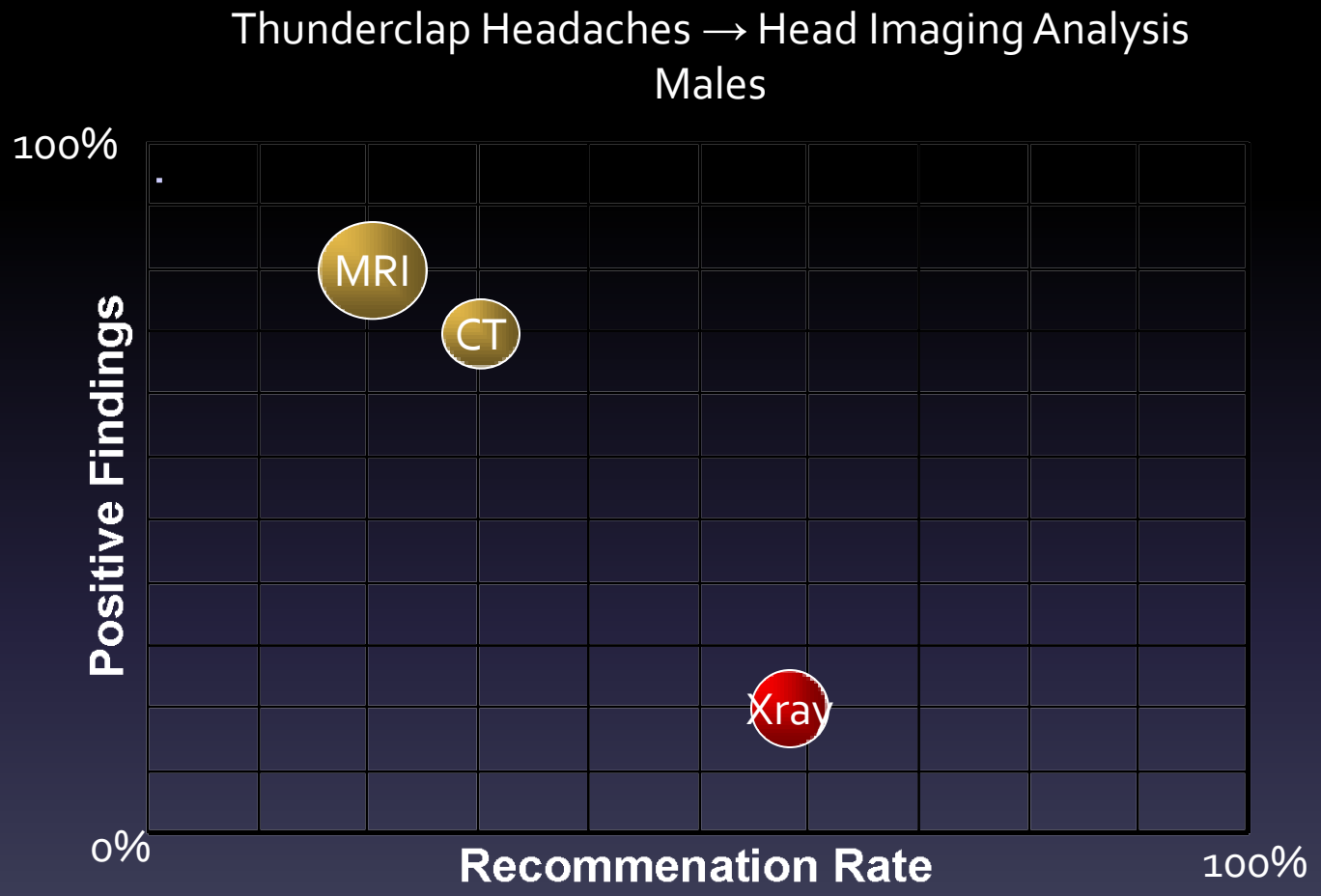
# Diagnostic Decisions Evaluation Comparison of 5 Ordering Physicians



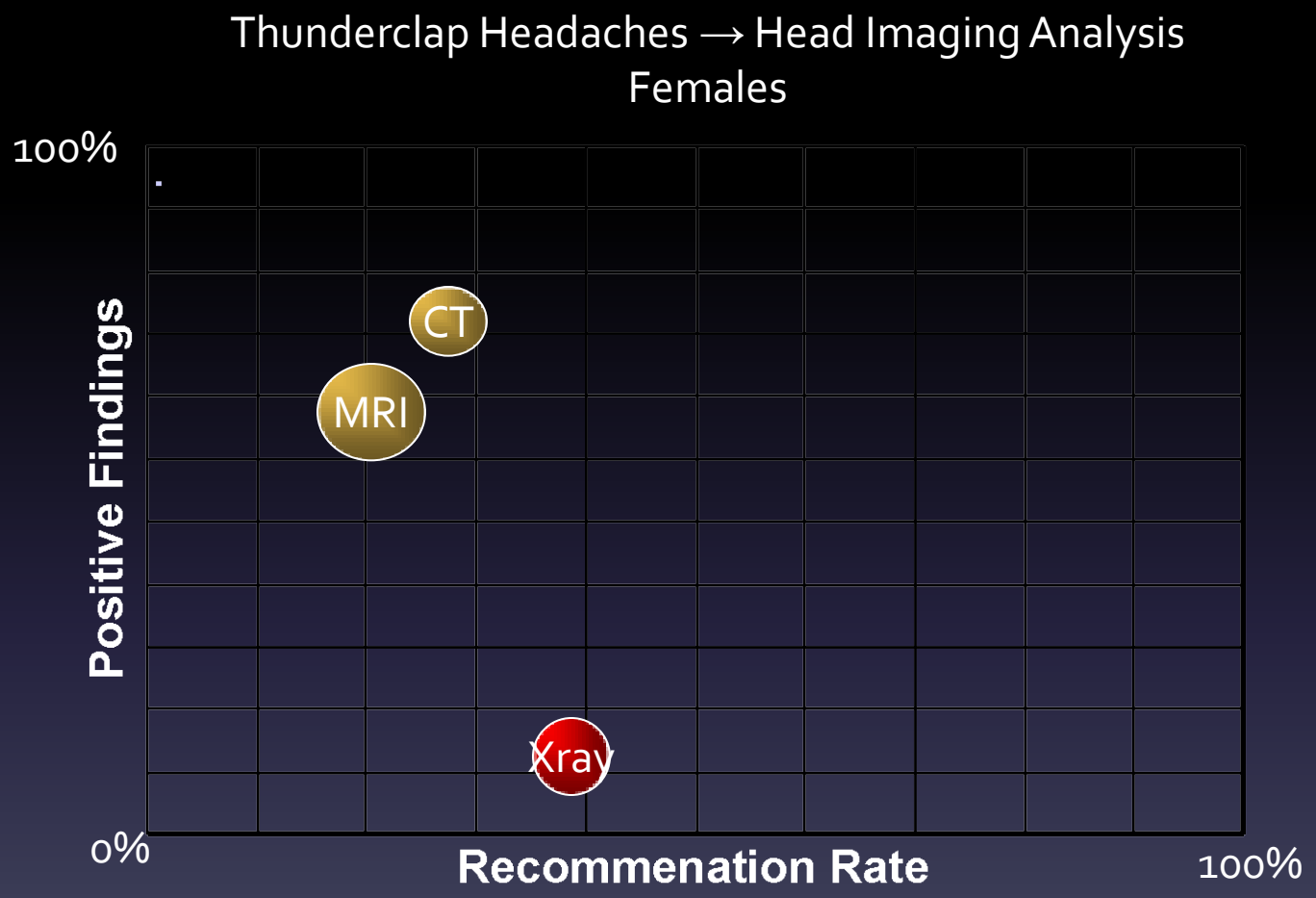
# Diagnostic Decisions Evaluation Comparison of Radiologists



# Indication → Examination Evaluation Comparison of Exams Males



# Indication → Examination Evaluation

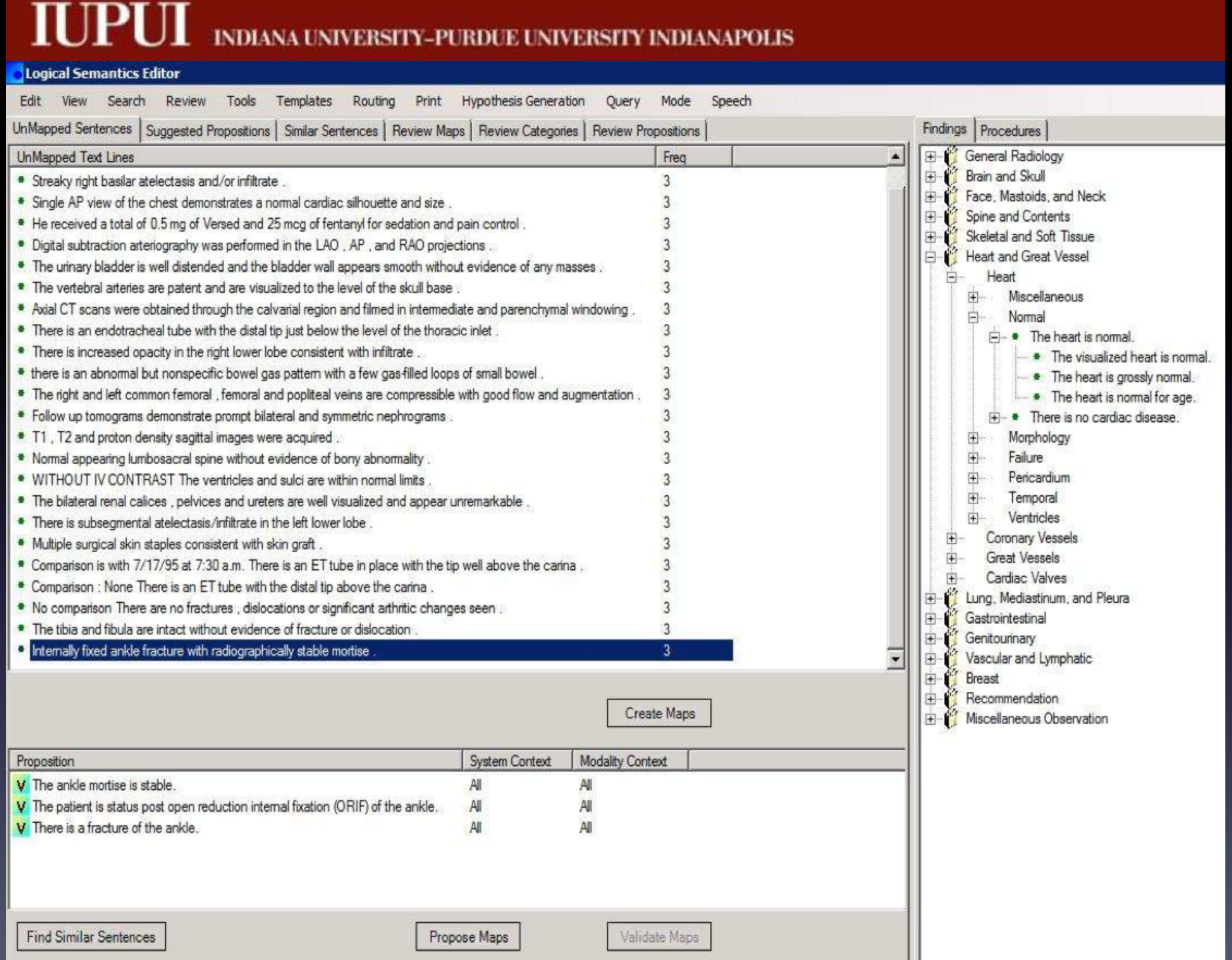


# EDM for Evaluation of Indication for Automated Protocoling

- Once examination has been selected choosing the best technique, e.g. with or without contrast, which MRI sequences to use given the specific indication

# EDM for Unexpected Findings

- Findings that fall through the "cracks" represent a major challenge in radiology
- Closing the communication loop is currently done manually at our facility
  - Most do not close this loop

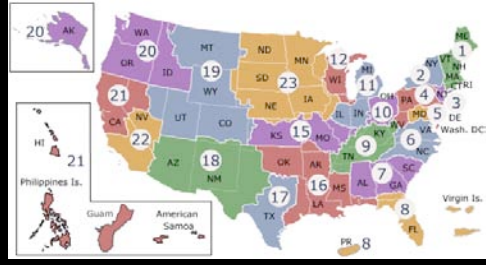

IUPUI
INDIANA UNIVERSITY–PURDUE UNIVERSITY INDIANAPOLIS
Logical Semantics Editor
Edit View Search Review Tools Templates Routing Print Hypothesis Generation Query Mode Speech
UnMapped Sentences | Suggested Propositions | Similar Sentences | Review Maps | Review Categories | Review Propositions
UnMapped Text Lines | Freq
Streaky right basilar atelectasis and/or infiltrate . 3
Single AP view of the chest demonstrates a normal cardiac silhouette and size . 3
He received a total of 0.5 mg of Versed and 25 mcg of fentanyl for sedation and pain control . 3
Digital subtraction arteriography was performed in the LAO , AP , and RAO projections . 3
The urinary bladder is well distended and the bladder wall appears smooth without evidence of any masses . 3
The vertebral arteries are patent and are visualized to the level of the skull base . 3
Axial CT scans were obtained through the calvarial region and filmed in intermediate and parenchymal windowing . 3
There is an endotracheal tube with the distal tip just below the level of the thoracic inlet . 3
There is increased opacity in the right lower lobe consistent with infiltrate . 3
there is an abnormal but nonspecific bowel gas pattern with a few gas-filled loops of small bowel . 3
The right and left common femoral , femoral and popliteal veins are compressible with good flow and augmentation . 3
Follow up tomograms demonstrate prompt bilateral and symmetric nephrograms . 3
T1 , T2 and proton density sagittal images were acquired . 3
Normal appearing lumbosacral spine without evidence of bony abnormality . 3
WITHOUT IV CONTRAST The ventricles and sulci are within normal limits . 3
The bilateral renal calices , pelvices and ureters are well visualized and appear unremarkable . 3
There is subsegmental atelectasis/infiltrate in the left lower lobe . 3
Multiple surgical skin staples consistent with skin graft . 3
Comparison is with 7/17/95 at 7:30 a.m. There is an ET tube in place with the tip well above the carina . 3
Comparison : None There is an ET tube with the distal tip above the carina . 3
No comparison There are no fractures , dislocations or significant arthritic changes seen . 3
The tibia and fibula are intact without evidence of fracture or dislocation . 3
Internally fixed ankle fracture with radiographically stable mortise . 3
Create Maps
Proposition | System Context | Modality Context
The ankle mortise is stable. All All
The patient is status post open reduction internal fixation (ORIF) of the ankle. All All
There is a fracture of the ankle. All All
Find Similar Sentences
Propose Maps
Validate Maps
Findings | Procedures
General Radiology
Brain and Skull
Face, Mastoids, and Neck
Spine and Contents
Skeletal and Soft Tissue
Heart and Great Vessel
Heart
Miscellaneous
Normal
The heart is normal.
The visualized heart is normal.
The heart is grossly normal.
The heart is normal for age.
There is no cardiac disease.
Morphology
Failure
Pericardium
Temporal
Ventricles
Coronary Vessels
Great Vessels
Cardiac Valves
Lung, Mediastinum, and Pleura
Gastrointestinal
Genitourinary
Vascular and Lymphatic
Breast
Recommendation
Miscellaneous Observation

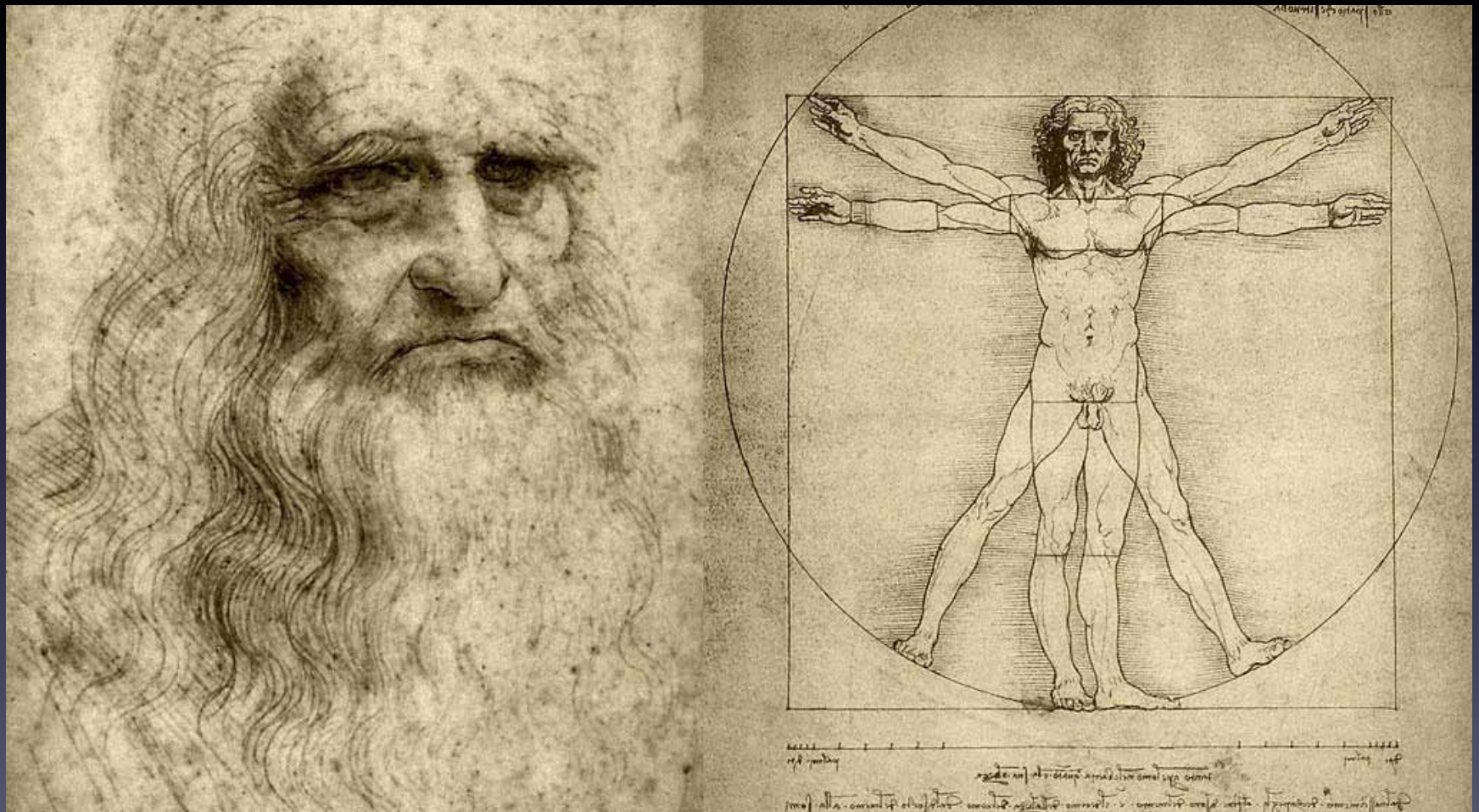
# Brief History of Radiology Decision Support Tools Chuck Kahn

- ICON
  - Rule based expert system for differential diagnosis of lung disease
    - E.g. If a patient with Hodgkin's disease has a pleural effusion and no lymphadenopathy, there is a moderate probability that the effusion is caused by an infectious process
    - Uses critiquing approach responding with evidence supporting or refuting tentative physician diagnosis
  - PHOENIX
    - Contains 54 common clinical problems such as head trauma or pulmonary embolism and generates recommended work-up using flowcharts
  - ROENTGEN
    - Uses case based reasoning to help plan radiation therapy protocols
  - ISIS
    - Uses case based decision support to help physicians select diagnostic imaging procedures based on actual cases abstracted from patient records and from published texts

# Mining the EMR for Data in Support of Diagnostic Imaging VA's VINCI Project

# VA Informatics and Computing Infrastructure (VINCI)



# VHA Facilities

- 163 Hospitals
- 800 Clinics
- 135 Nursing Homes
- 43 Domiciliaries
- 180,000 Healthcare Professionals

- Serving 6 Million Veterans per Year

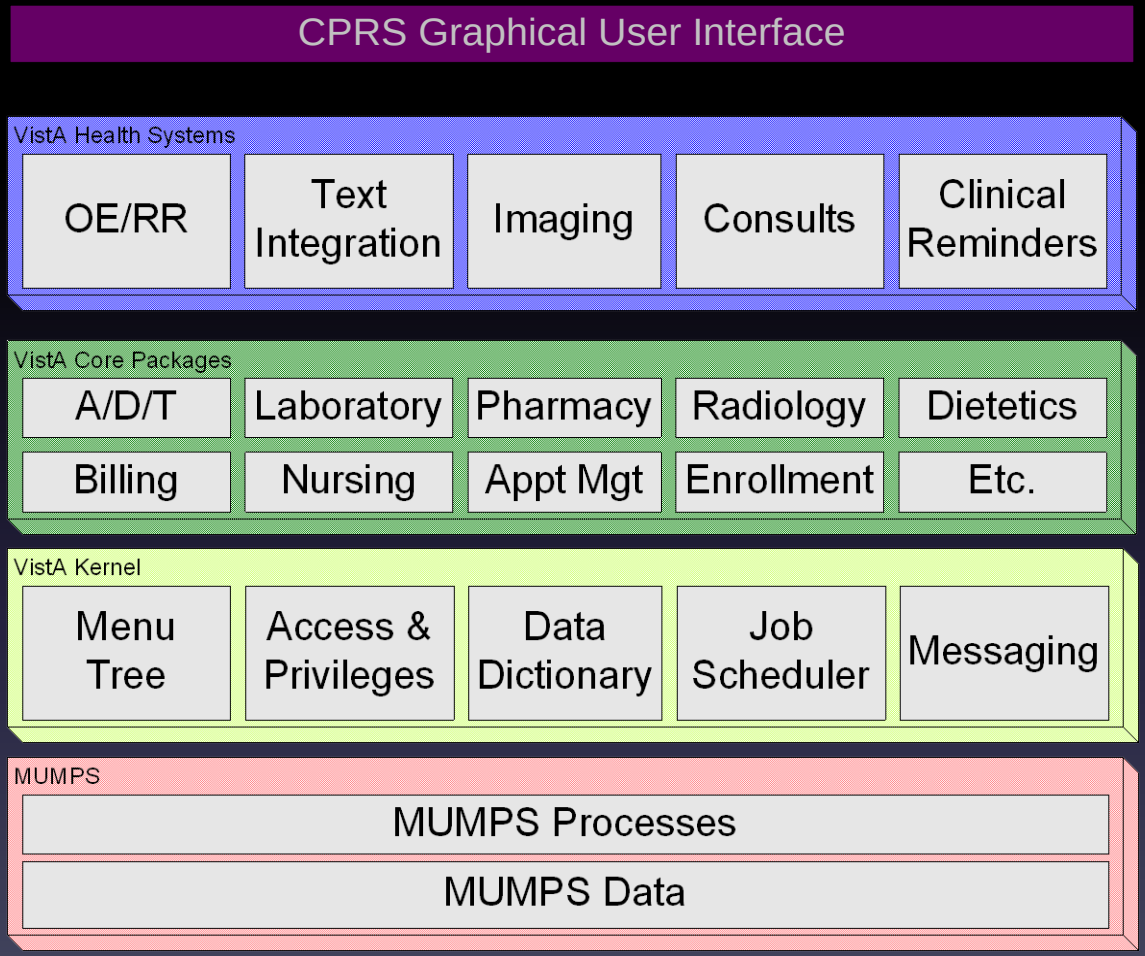
Courtesy of Kevin Meldrum

# VHA Volume

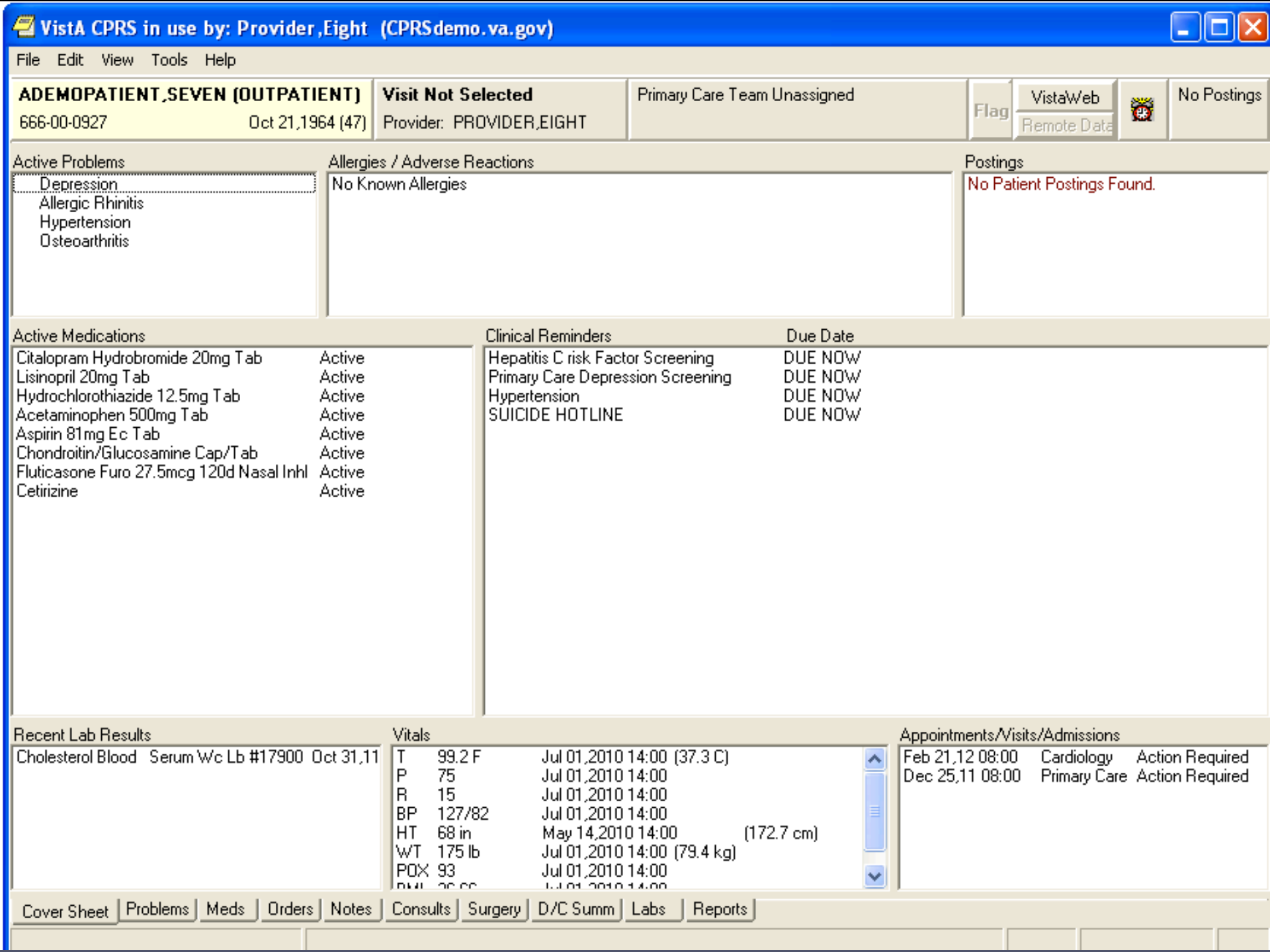
- Hospital sizes vary from 100 to 1000 beds.
- Outpatient visits vary at the facility from 30,000 to 450,000 visits per year
- 46.5 million outpatient visits per year
- 564,000 inpatient admissions per year
- 167 million prescription-months filled

Courtesy of Kevin Meldrum

# Layered Approach




Courtesy of Kevin Meldrum


VistA CPRS in use by: Provider,Eight (CPRSdemo.va.gov)
File Edit View Tools Help
ADEMOPATIENT,SEVEN (OUTPATIENT)
666-00-0927 Oct 21,1964 (47)
Visit Not Selected
Provider: PROVIDER,EIGHT
Primary Care Team Unassigned
Flag
VistaWeb
Remote Data
No Postings
Active Problems
Depression
Allergic Rhinitis
Hypertension
Osteoarthritis
Allergies / Adverse Reactions
No Known Allergies
Postings
No Patient Postings Found.
Active Medications
Citalopram Hydrobromide 20mg Tab Active
Lisinopril 20mg Tab Active
Hydrochlorothiazide 12.5mg Tab Active
Acetaminophen 500mg Tab Active
Aspirin 81mg Ec Tab Active
Chondroitin/Glucosamine Cap/Tab Active
Fluticasone Furo 27.5mcg 120d Nasal Inhl Active
Cetirizine Active
Clinical Reminders Due Date
Hepatitis C risk Factor Screening DUE NOW
Primary Care Depression Screening DUE NOW
Hypertension DUE NOW
SUICIDE HOTLINE DUE NOW
Recent Lab Results
Cholesterol Blood Serum Wc Lb #17900 Oct 31,11
Vitals
T 99.2 F Jul 01,2010 14:00 (37.3 C)
P 75 Jul 01,2010 14:00
R 15 Jul 01,2010 14:00
BP 127/82 Jul 01,2010 14:00
HT 68 in May 14,2010 14:00 (172.7 cm)
WT 175 lb Jul 01,2010 14:00 (79.4 kg)
POX 93 Jul 01,2010 14:00
Appointments/Visits/Admissions
Feb 21,12 08:00 Cardiology Action Required
Dec 25,11 08:00 Primary Care Action Required
Cover Sheet
Problems
Meds
Orders
Notes
Consults
Surgery
D/C Summ
Labs
Reports

VistA CPRS in use by: Provider,Eight (CPRSdemo.va.gov)

File Edit View Action Options Tools Help

**ADEMOPATIENT,SEVEN (OUTPATIENT)** | **Visit Not Selected** | Primary Care Team Unassigned | Flag | VistaWeb | Remote Data | No Postings

666-00-0927 Oct 21,1964 (47) | Provider: PROVIDER,EIGHT

Last 100 Signed Notes (Total: 8)

- All signed notes
  - Jul 01,10 PRIMARY
  - May 21,10 PRIMAR
  - May 14,10 PRIMAR
  - May 11,09 PRIMAR
  - May 04,09 PRIMAR
  - Jun 25,08 PRIMAR
  - Apr 22,08 PRIMAR
  - Apr 15,08 PRIMAR

Templates

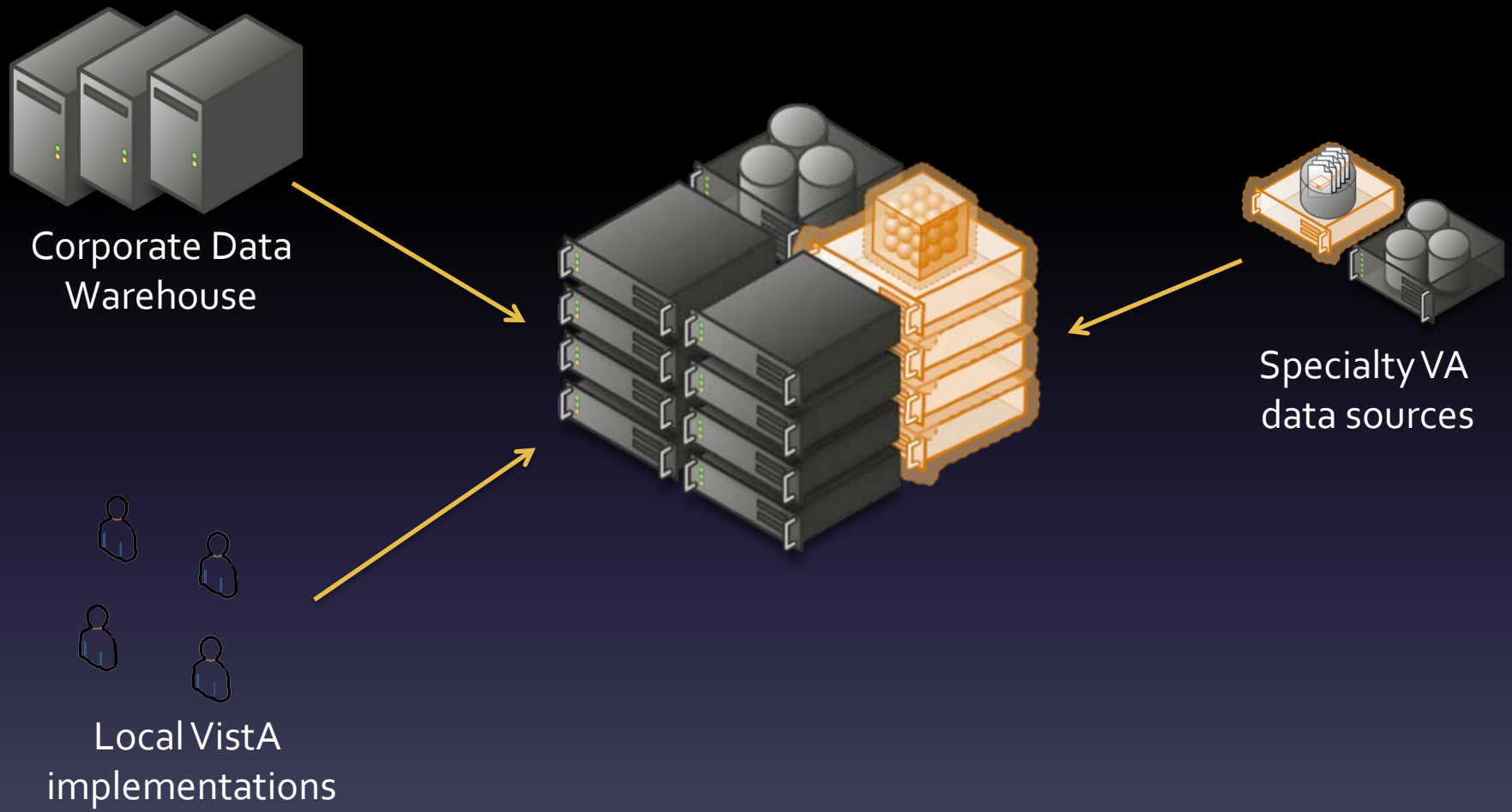
New Note

Visit: 07/01/10 PRIMARY CARE GENERAL NOTE, GENERAL MEDICINE, ONE PROVIDER (Jul 01,10@13:00)

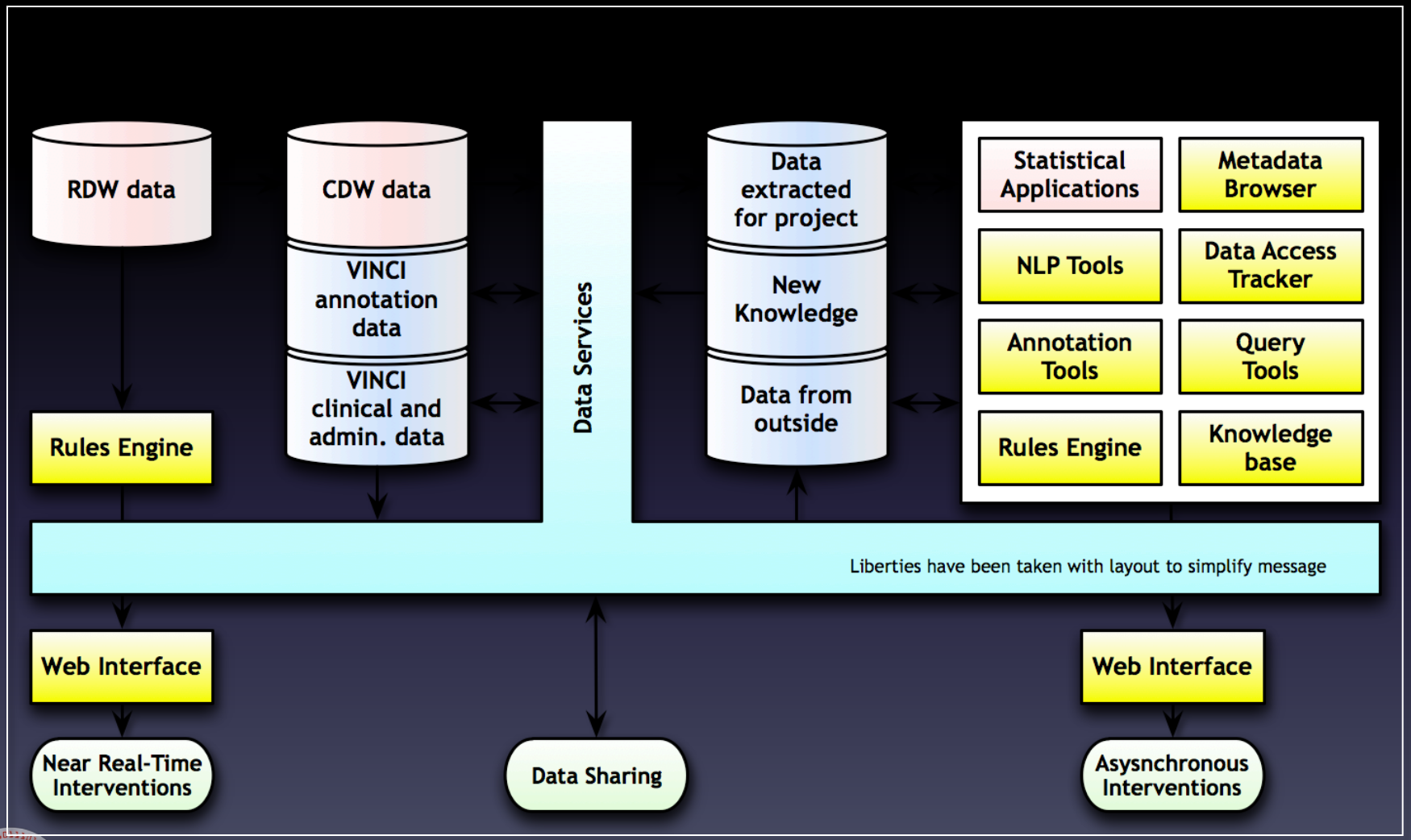
```
 LOCAL TITLE: PRIMARY CARE GENERAL NOTE
STANDARD TITLE: PRIMARY CARE NOTE
DATE OF NOTE: JUL 01, 2010@13:00     ENTRY DATE: MAR 01, 2011@09:22:19
      AUTHOR: PROVIDER,ONE          EXP COSIGNER:
     URGENCY:                            STATUS: COMPLETED

CC
F/U Depression and lab work.

Subjective
The patient went by the lab this morning and got the labs ordered for today.

Depression
The patient indicates that his symptoms have improved significantly, but not as
much as he expected.  He is still sleeping a lot (about 12 hours per day) and
finds it hard to concentrate on looking for work.  He is eating better and has
gained a few pounds.  He denies suicidal ideation.  His PHQ-9 score is 16 today.

Allergies
NKDA

PMH
Depression
Hypertension
Allergic rhinitis
Acute sinusitis
Osteoarthritis

PSurgHx
None

FamHx
Mother died of MI at 64 years of age.

SocHx
```

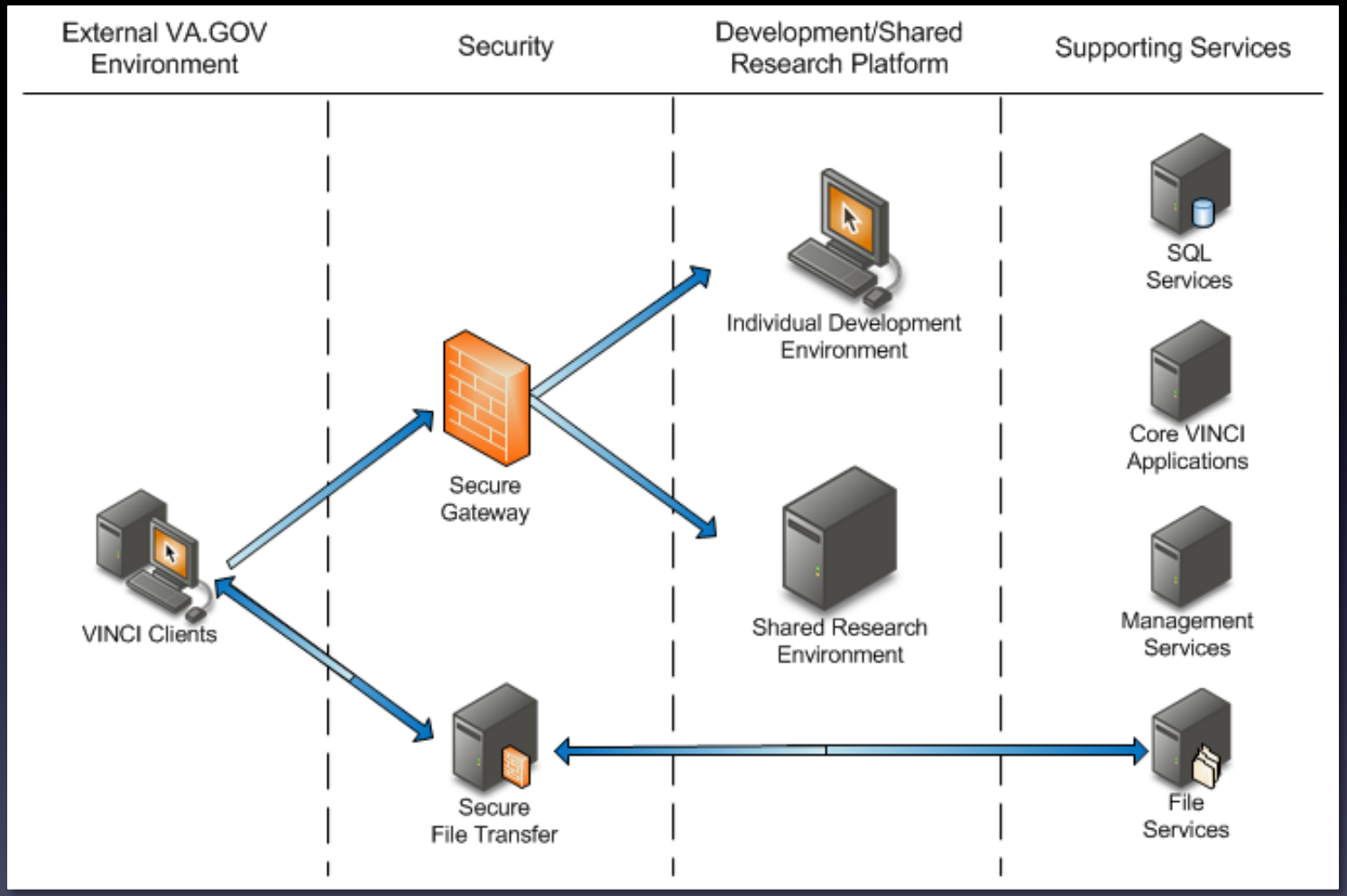
Cover Sheet | Problems | Meds | Orders | Notes | Consults | Surgery | D/C Summ | Labs | Reports

# VINCI


Corporate Data Warehouse
Specialty VA data sources
Local VistA implementations

# A Research / IT Partnership




External VA.GOV Environment
Security
Development/Shared Research Platform
Supporting Services
SQL Services
Individual Development Environment
Core VINCI Applications
Secure Gateway
Shared Research Environment
Management Services
VINCI Clients
Secure File Transfer
File Services

# Data Extraction Research Model

- Research Study Team uploads cohort to a secure site or VINCI creates cohort per requirements
- Research Study Team completes the Data Selection Forms on correspondence site
- NDS approved data domains are extracted and provided
- Analysis performed by research project team staff

# External Data Can Be Uploaded

- Research Study Team may upload other data for analysis into project database or workspace
- Secure data upload process
- Optional direct database upload
- VINCI data managers work on behalf of research team to upload data from other data providers

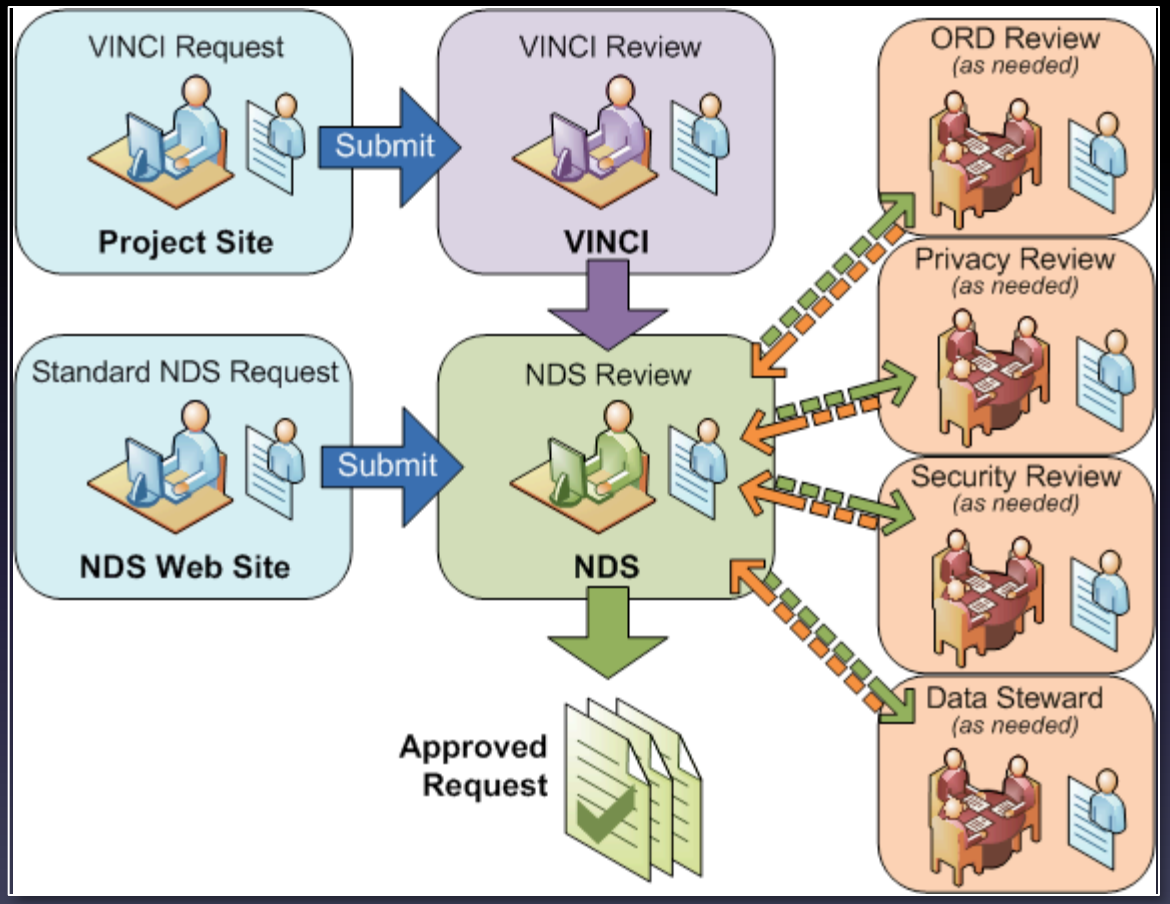
# Data Processing/Analysis tools

- SQL Server as primary data store
- Multiple high performance servers
- Most data queries performed in SQL Server
- Accessible by all analysis software
- SSIS, SSRS, SSAS
- High speed intra-server network will allow distributed queries

# SAS

- High performance SAS server
  - 2 TB of RAM, 64 cores, 2 TB SSD
  - Launch grid jobs
- SAS grid – very large data analysis work
  - 10 high performance servers
  - Most advanced SAS implementation in VA
  - Dedicated SAN
  - Additional SAS modules
  - SAS knowledge base SharePoint site
- Dedicated SAS administrator

# Data Approvals



# Data Access & Security

- Access groups created based on IRB and NDS approved research team
- Only research team members have access to the data
- Data stored on secure VINCI servers
- Regular data backup and archiving
- Workspace vs. collaboration site
- Project work can be performed in VINCI
- Export final result & publication

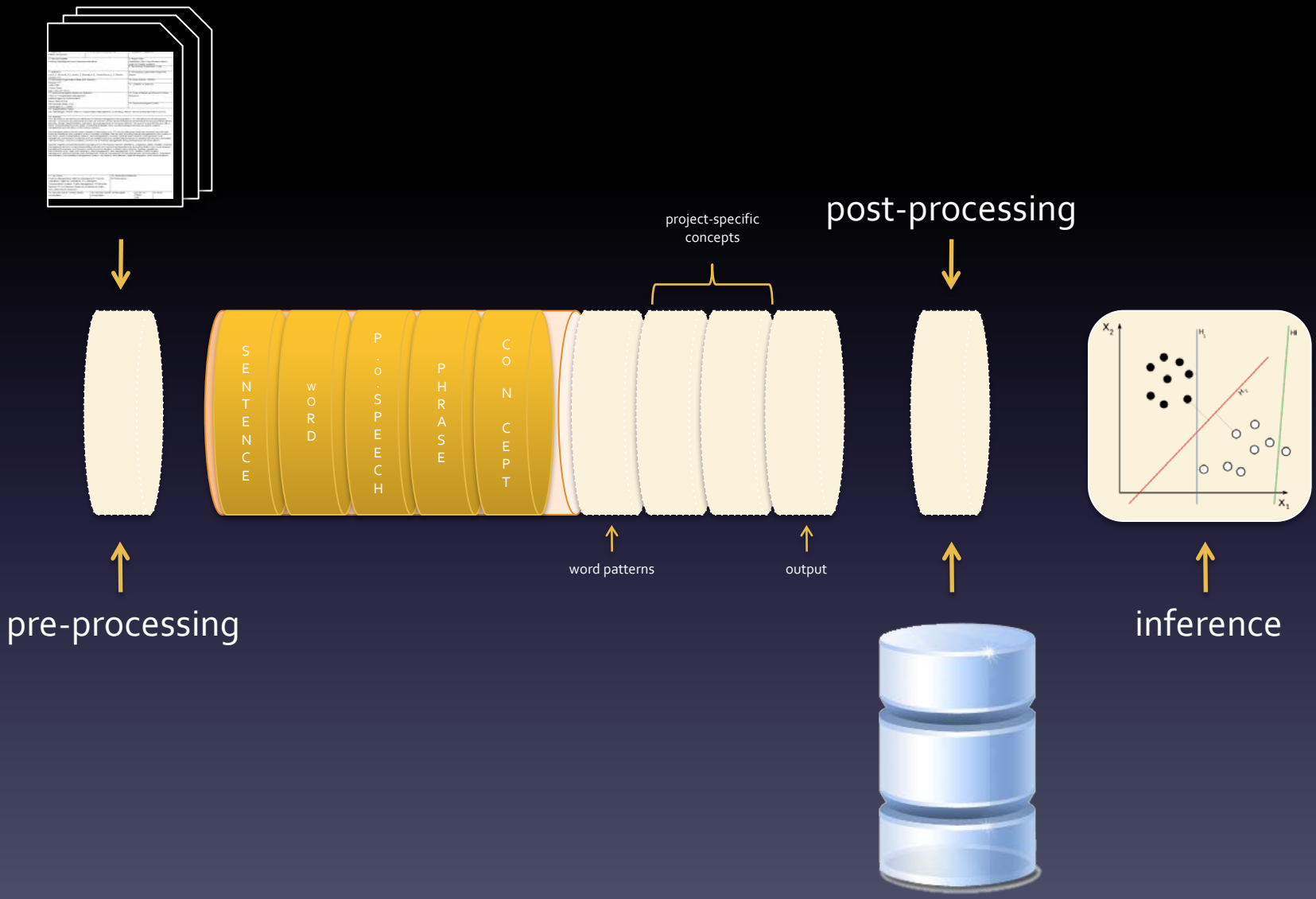
# Data Type Examples

- Patient demographic information
- Vital signs
- Diagnoses and procedures from patient encounters
- Outpatient pharmacy data
- Laboratory values
- Immunizations
- Microbiology reports
- Text notes (including progress notes, discharge summaries, radiology reports)
- VA Decision Support System (DSS) in database tables linked with other VINCI data or as the original SAS files
- VHA Medical SAS datasets

# Applications of VINCI Data in VA

- Increasing demands for more detailed clinical data
  - Quality Measures
  - Evidence-based medicine
  - Phenotyping for genomic-related analysis
  - Biosurveillance
- The majority of EMR data is free text

# An NLP Pipeline



# Pipeline Issues

Base modules have trouble with:

- Templates
- Incomplete Sentences
- Jargon

# Next Generation Data Mining and VINCI


# IBM Jeopardy Software

- Deep Q/A is unique and exciting because it represents a fundamentally new approach that creates tools to rapidly mine a dynamic and non-predefined database
- Represents a potential fundamental change in opportunities for Artificial Intelligence applications in medicine
- But in some ways Watson is a “special needs” student
- How does one train a system that is so remarkable at Jeopardy! questions and apply to medicine?

- Watson can process 500 gigabytes, the equivalent of a million books, per second
- Hardware cost has been estimated at about $3 million
- 80 TeraFLOPs , 49th in the Top 50 Supercomputers list
- Content was stored in Watson's RAM for the game because data stored on hard drives too slow to process

# Deep Q/A

- Massively parallel, component based pipeline architecture
- Uses extensible set of structured and unstructured content sources
- Uses broad range of pluggable search and scoring components

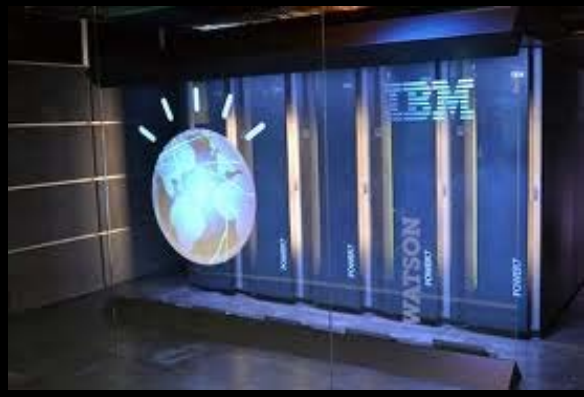
# Deep Q/A

- These allow integration of many different analytic techniques
- Input from scorers is weighed and combined using machine learning to generate a set of ranked candidate answers and associated confidence values
- Each answer is linked to its supporting evidence

# Deep Q/A

- Does not map question to database of answers
- Represents software architecture for analyzing natural language content in both questions and knowledge sources
- Discovers and evaluates potential answers and gathers and scores evidence for those answers using unstructured sources such as natural language documents and structured sources such as relational and knowledge databases

# Hardware

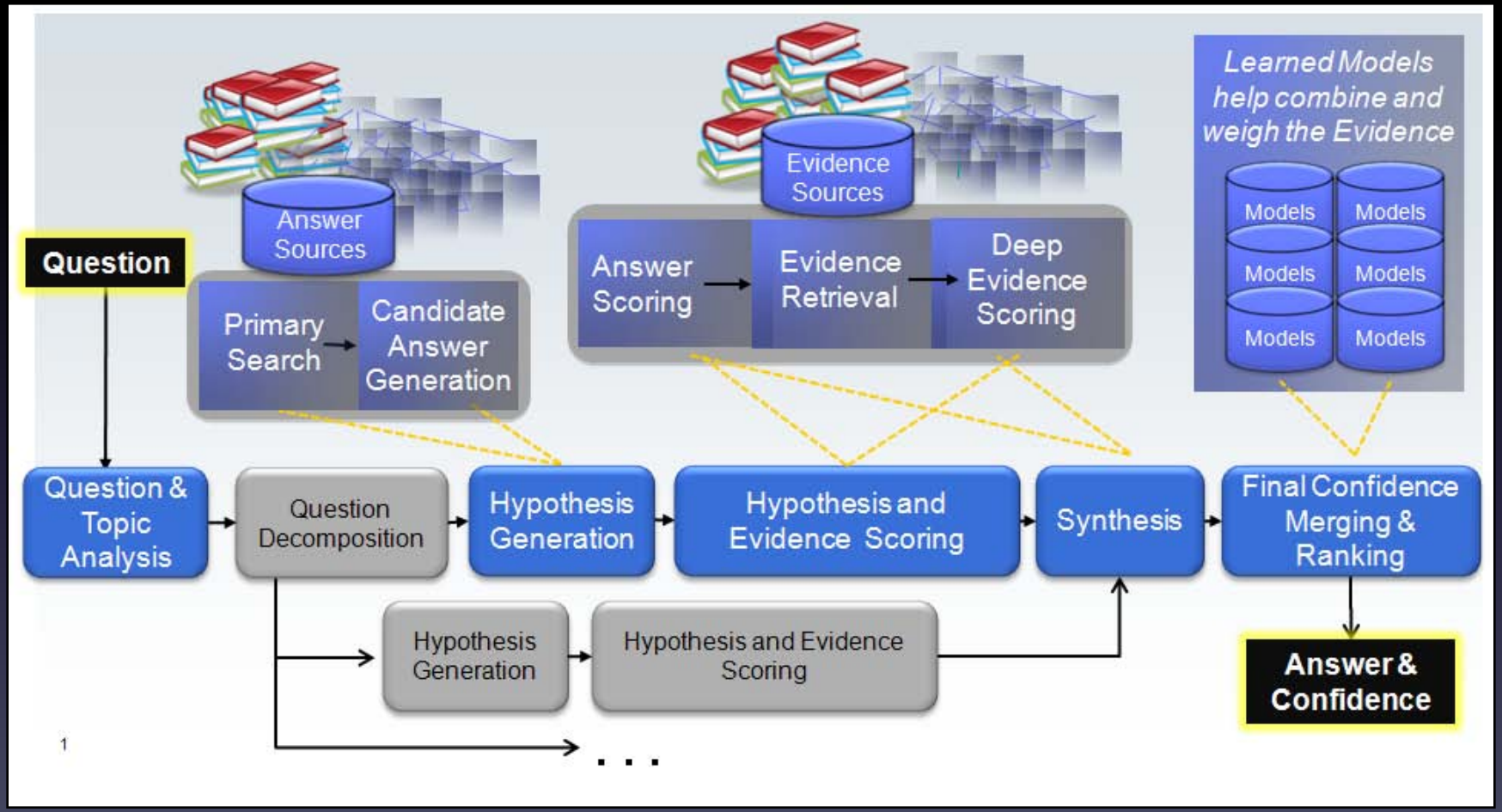



- Cluster of ninety IBM Power 750 servers (plus additional I/O, network and cluster controller nodes in 10 racks) with a total of 2880 POWER7 processor cores and 16 Terabytes of RAM
- Each Power 750 server uses a 3.5 GHz POWER7 eight core processor, with four threads per core
- The POWER7 processor's massively parallel processing capability is an ideal match for Watson's IBM DeepQA software which is embarrassingly parallel (that is a workload that is easily split up into multiple parallel tasks)

# Software

- Watson's software was written in both Java and C++ and uses Apache Hadoop framework for distributed computing
- Apache UIMA (Unstructured Information Management Architecture) framework
- IBM's DeepQA software and SUSE Linux Enterprise Server 11 operating system
- "More than 100 different techniques are used to analyze natural language, identify sources, find and generate hypotheses, find and score evidence, and merge and rank hypotheses."

# High Level View of DeepQA Architecture



# The Science Behind an Answer

# Deep QA Process

- Analyzes input question and generates many possible candidate answers through broad search of volumes of content
- Hypothesis is formed based on considerate of each candidate answer in context of original question and topic
  - For each of these, DeepQA spawns independent thread attempting to prove it
  - Searches content sources for evidence supporting or refuting each hypothesis
  - Applies hundreds of algorithms for each evidence hypothesis pair that dissects and analyzes along different dimensions of evidence

# Can We Use Deep Q/A to Mine VINCI Data?

- IRB approval to mine >28 million patients over 12 years
- Cannot export data due to difficulties de-identifying patient PHI
- In what format to present data to Deep Q/A software?
  - Currently using screen scraping text information from various aspects of patient record
  - Want to export using HL7 CDA or other standard, reproducible method
- VA has plans to collect DNA samples from 1,000,000 of its patients and can cross correlate this with clinical and lab and other data over period of time
- Initial project proposed has been evaluation of factors that predispose to metabolic syndrome/diabetes type II
- Also planning to use these data for diagnostic imaging specific queries including mining for recommendations, positive studies, etc. but also questions such as impact of imaging studies that involve radiation on cancer rates

# Conclusion

- Radiology/diagnostic imaging has historically and will continue to be a rich subspecialty not only for image processing and computer aided diagnosis but also for Natural Language Processing and Enhanced Clinical Decision Making from the EMR
- These two techniques can be utilized on very large databases such as VINCI for research and clinical support purposes and have the potential to have a major impact on research as well as our day to day decision making in medicine