



# Natural Language Processing – The Silver Bullet?

Lynn Kosegi - Director of Health Information Services, M\*Modal

Jürgen Fritsch - Chief Scientist, M\*Modal

# Agenda

- Promise of electronic records
- Documentation challenges
- Narrative documentation and discrete data
- NLP – the silver bullet?
- From narrative to data to actionable information

# Part I: Trends & Challenges in Healthcare Documentation

# The EHR and Meaningful Use

The promise...

- Achieving Meaningful Use of EHR will –
  - Improve patient care
  - Reduce costs
  - Improve population health

How?

From HIM Professionals...

**“MTs are editing in Epic”**

**“Residents are texting”**

**“DOCTORS WON'T USE IT”**

# The Doctor's Reality

**“I can't see the patient story”**

**“I CAN'T TELL WHAT'S NEW”**

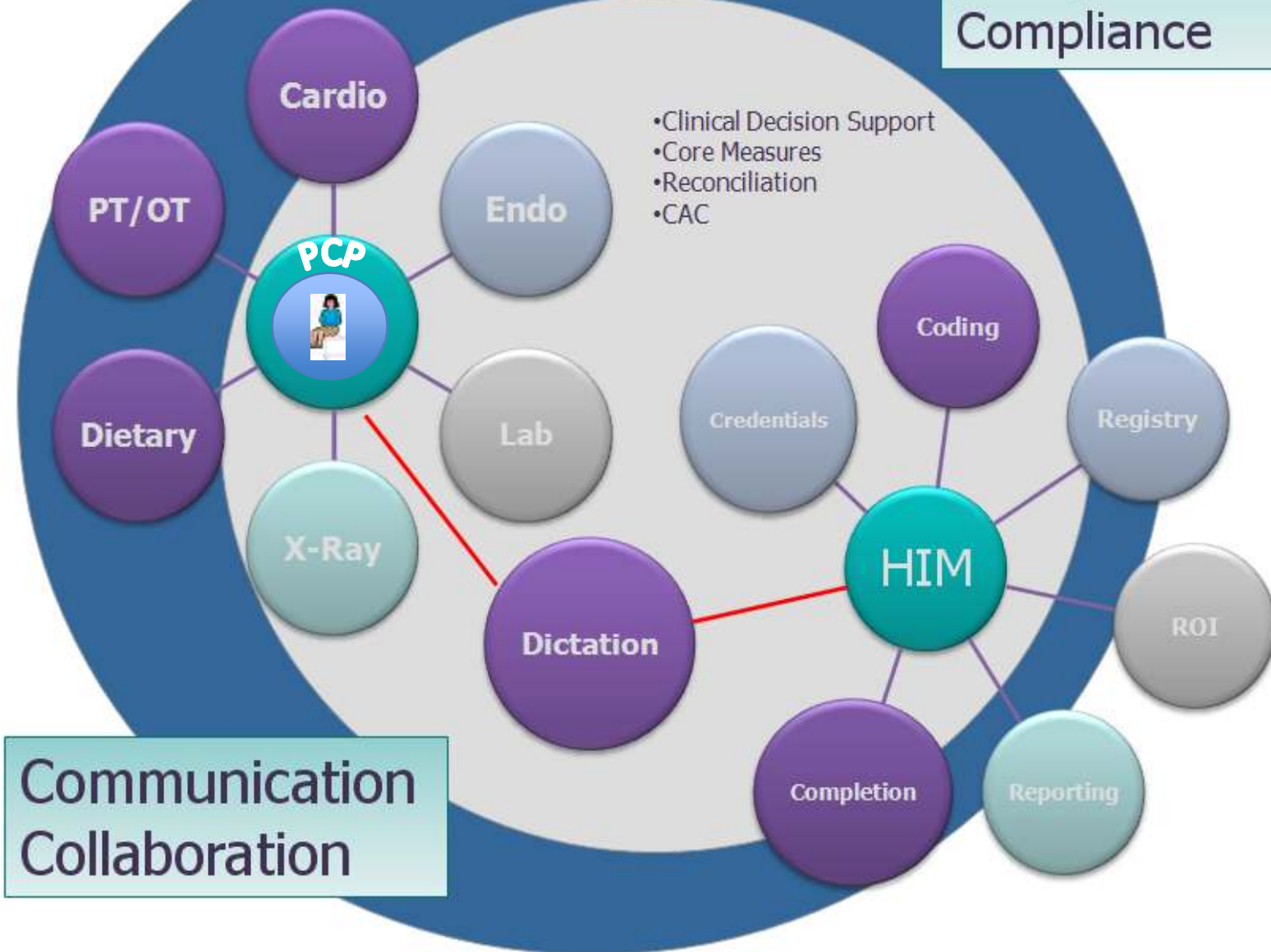
“I can't tell how the patient feels”

“I see 2 fewer patients a day”

**“I'm under-coding and  
under-billing”**

Health Information  
Ecosystem

Quality  
Safety and Security  
Compliance



Communication  
Collaboration

# Industry Stressors

- Increased demand for information
  - ICD-10
  - Meaningful Use
  - Accountable care
- More difficult methods of information capture
  - Direct data entry
  - Removing the ability to capture thoughts and intuition

# HIM – Need for Data

ICD – 10 -

- More specificity
- Extensive training
- Decreased productivity – 50%
  - Learning curve
  - Hybrid phase
- Loss of “super bills”
- Loss of staff – attrition

# ICD-10 Challenges

- Deadline is **2013**
- Requires granularity and specificity
- Improve capture of quality and outcomes
- Better capture for population health statistics and case mix index
- More appropriate reimbursement

# ICD-10 Challenges

**Problem:** Unspecified documentation leads to unspecified codes

- Lost revenue
- Inaccurate capture of severity - case mix
- Increased audit risk

**Solution:** Improved documentation

- Accurate
- Comprehensive
- Complete
- Compliant

# GEMs

## General Equivalence Mappings

- No direct mappings
  - ~ 17,000 ICD-9 codes to >140,000 ICD-10
  - one-to-many relationships
- The way we code is different

# ICD-9 to ICD-10



“Diabetes type 2”  
“Uncontrolled blood sugars”  
“Spilling ketones”  
“Insulin drip”

## DISCHARGE DIAGNOSIS:

Bronchioloalveolar carcinoma  
Diabetes type 2

## HOSPITAL COURSE:

Patient underwent lobectomy for removal of non-small cell lung carcinoma of the upper left lobe. Patient's blood sugars were uncontrolled. He was spilling ketones in his urine and required an insulin drip for 24 hours post op without loss of consciousness. Blood sugars in control at time of discharge.

## ICD-9:

- 250 – Diabetes mellitus
- 250.00 – Diabetes mellitus without mention of complication, unspecified type uncontrolled
- 250.12 - Diabetes with ketoacidosis, unspecified type, uncontrolled (without mention of coma,

## ICD-10:

- Based on manifestation of diabetes
- No more “unspecified”
- No more “uncontrolled”
- No more ambiguity
  - “without mention of coma”
  - “type 2 or unspecified”

# The Problem with Prok...

## Major Diagnoses

1. *High-risk AML*

...status post RIC and double umbilical cord stem cell transplant on day +13.

## Surgical/procedure History

2. *Graft-versus-host disease prophylaxis*

...dose was decreased to 200 mg b.i.d. **He did not hold his dose this morning**, so we will recheck tomorrow morning.

## Compliance

3. *History of pulmonary and hepatic lesions.*

...appeared while he was neutropenic post induction chemotherapy.

## Relevant history

4. *Nicotine dependence.*

...continues to smoke cigarettes. He was seen by nicotine

...but did not pursue any interventions. Given a

...prescription for nicotine gum **which he filled but has yet to try**. He

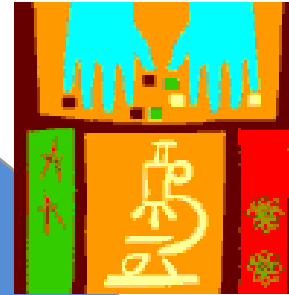
...is encouraged to decrease, if not quit, his smoking at this time,

...but he politely declined.

## Social history

# Quest for Semantic Interoperability





# Narrative Documentation

- *Human* understanding
- Ease of use for physicians
- Quality and compliance
- Usefulness
- Enterprise-wide efficiency
- *Options, flexibility, and scale*

.....across the continuum of care

# The Challenges with Narrative

- Accessibility
  - How to make accessible the data buried in narrative documentation?
  - E.g. need to identify & maintain problem lists
- Actionability
  - How to derive actions from narrative documentation?
    - Coding for billing
    - Automation of CDS and standardized care plans

*Is NLP the silver bullet?*

# Part II: The Role of NLP

# Definition

- Natural Language processing (NLP) is a field of computer science and linguistics concerned with the interactions between computers and human (natural) languages.
- A general term used to describe the use of computers to process information expressed in natural (i.e. human) languages.
- The use of linguistic analysis to infer meaning from narrative text

# Definition

- Natural Language processing (NLP) is a field of computer science and linguistics concerned with the **interactions** between computers and human (natural) languages.
- A general term used to describe the use of computers to **process information** expressed in natural (i.e. human) languages.
- The use of linguistic analysis to **infer meaning** from narrative text

# In other words...

- It's a tough problem
- Far from being solved

## Bearman Cartoons



© 6/24/09 beartoons.com

## Artificial Intolerance



bearmancartoons@yahoo.com

# But...

- NLP has evolved substantially in recent years
- Useful applications in limited domains
- If carefully deployed, maintained and improved in the context of other necessary technologies

# Flavors of NLP

- Rule-Based
  - Traditional approach based on handwritten rules, grammars and parsers
  - Pros: Explicit representation of expert knowledge, results are interpretable
  - Cons: Difficult to scale, no automatic learning
- Statistical
  - More recent approach using machine learning algorithms
  - Pros: Enables learning from data & feedback
  - Cons: Harder to adjust to changing requirements

# NLP in Healthcare – The Hype

- Understands like a “clinician”?
- Extracts information?
- Ensures Meaningful Use compliance?
  - Quality and outcomes reporting
    - Core measures, PQRS
  - Patient summary
  - Reconciliation of problems and medications
- Drives Computer Assisted Coding (CAC)?
- Improves documentation quality & efficiency?

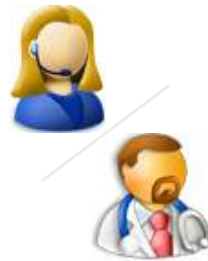
# NLP – The Silver Bullet?

- NO!
- NLP is a useful technology tool
- NLP makes mistakes
- There is no single, general purpose NLP – always specific to a particular task
- NLP is needed but not sufficient
- Needs a lot of supporting infrastructure

# NLP Platform Infrastructure

- Connectivity to relevant data sources
- Management of various versions and configurations of NLP engines
- Management of various versions of clinical vocabularies and ontologies
- Feedback loops to enable learning
- Workflow

# Traditional SR + NLP Approach



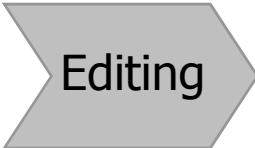
MRN: 000000  
 DOS: 11/11/2001  
 CHIEF COMPLAINT:  
 Fatigue  
 SUBJECTIVE:  
 Patient is a 25 year old woman complaining of feeling frequently fatigued. She reported also occasional dizziness. Sleeping difficulties and morning headaches.  
 OBJECTIVE:  
 Recent bout with the flu.  
 PHYSICAL EXAMINATION:  
 Vital signs are normal with a blood pressure of 120/80, pulse 62, temperature 98.6 degrees, weight 108 pounds.  
 ASSESSMENT:  
 Although flu symptoms were in remission, patient has not fully recovered yet.  
 PLAN:  
 Place patient on Biaxin for the next two weeks. The patient will call us if there is no improvement, any worsened or new symptoms.

Extracted  
 Clinical  
 Facts

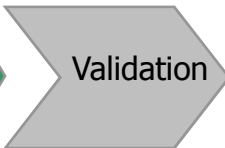
Dictation  
 Recording



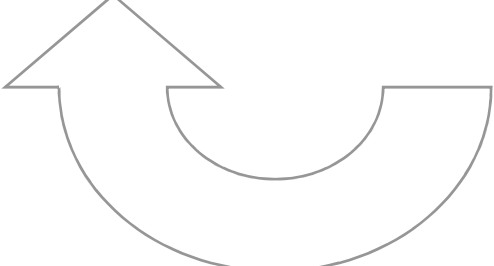
Text



Text



Extracted  
 Clinical  
 Facts



Feedback:  
 corrected textual  
 draft documents



# Problems with this Approach

- Multi-step review/validation workflow
  - First reviewing the SR draft for textual correctness
  - Then later reviewing the NLP derived clinical facts
- Missing automated NLP feedback loop
- ‘Extraction’ of clinical facts
  - Losing information & interpretability
  - Narrative context needs to be preserved
  - Natural language is more expressive than any clinical data model

# The Natural Language Gap

- Clinical narrative is much more expressive than any discrete, coded representation

...likes to run a couple of miles a week his past medical is significant for insulin-dependent diabetes mellitus that he has had since 1997 and a history of migraines and some asthma surgical history is significant for wisdom tooth extraction medications include lantus humalog and topamax allergies are no known drug allergies...

...likes to run a couple of miles a week

`code=120.2` **Past Medical History**

Significant for insulin-dependent diabetes mellitus that he has had since 1997 and a history of migraines and some asthma.

`code=324` **Past Surgical History**

Significant for wisdom tooth extraction.  
`code=346.9` `code=493.9` `modifier="SOME"`  
`code=62972009`

`code=680.0` **Current Medications**

lantus, humalog, and topamax.  
`code=RX10202204`

`code=450.4` **Allergies**

No known drug allergies...  
`code=416098002` `modifier="ABSENCE-OF"`

# Speech Understanding Approach



Hunt. He had a fall in June 2005, when he fell down some stairs landing on his right elbow since that time and swelling in his right elbow. He also has been complaining of shoulder for the last three months. He has sharp severe pain at night. Has difficulty sleep overhead activity.

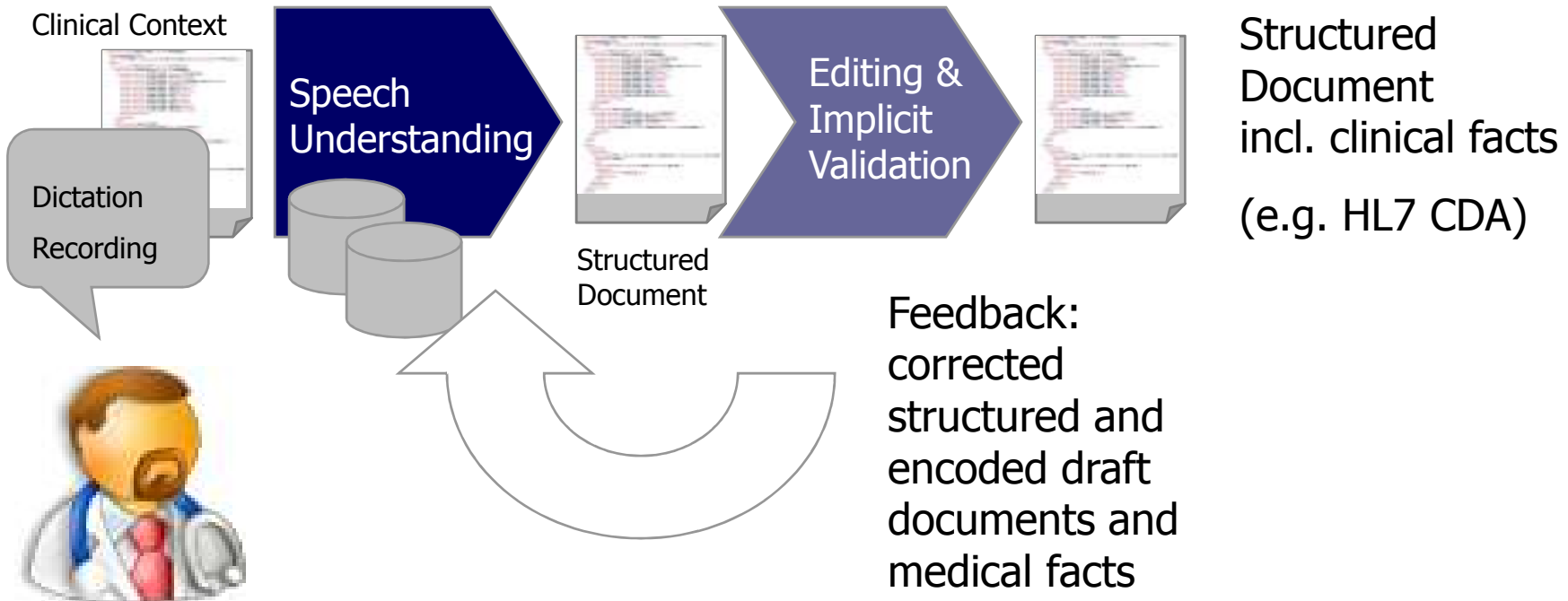
## Past Medical History

### PRIOR PROBLEMS:

Description	Modifier	Original Dictation
hypertension		Significant for hypertension,
phrase	Hypertension	gastroesophageal reflux disease,
concept	Hypertension	sleep apnea,
type	implicit	arthritis in both his knees.
code	997.91	
codeSystem	ICD9-CM	

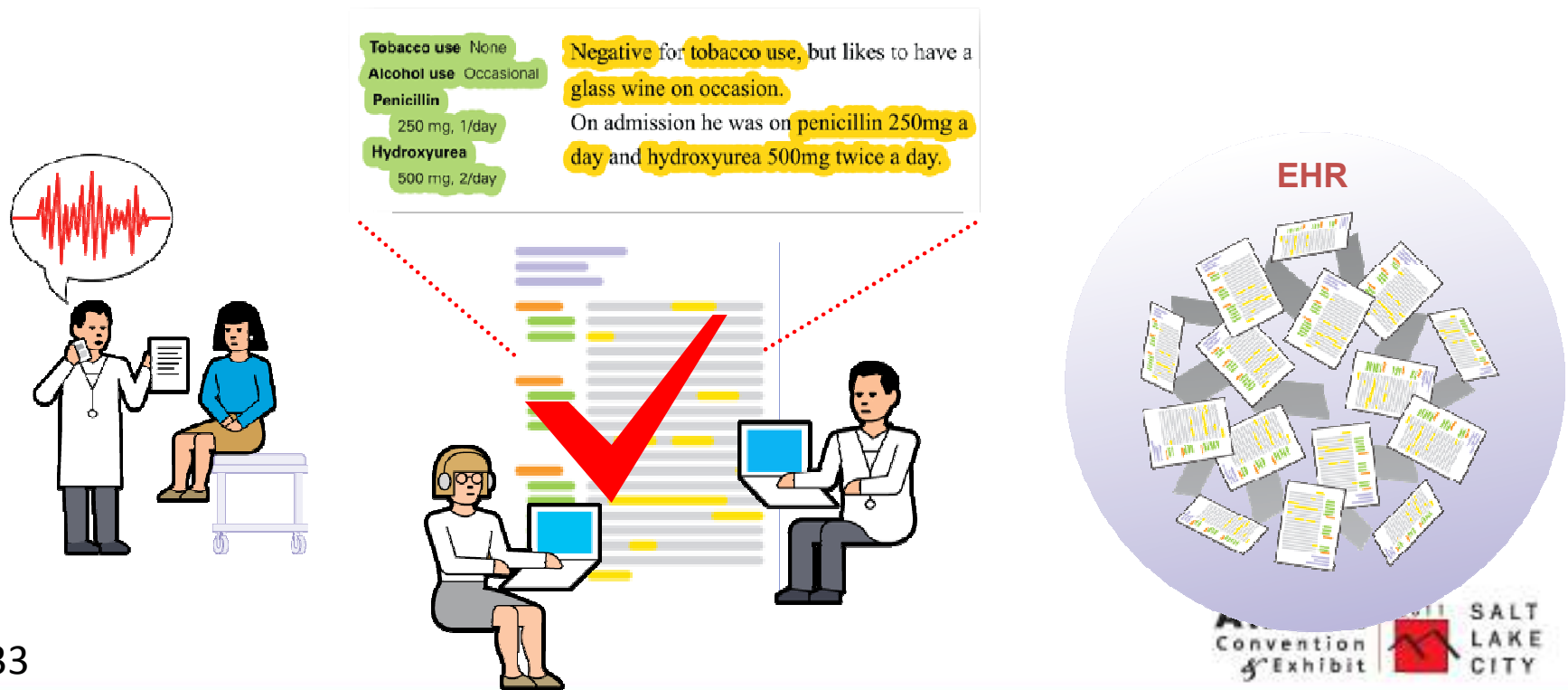
## Past Surgical History

Significant for right knee arthroscopy x2 and a left knee arthroscopy x1.



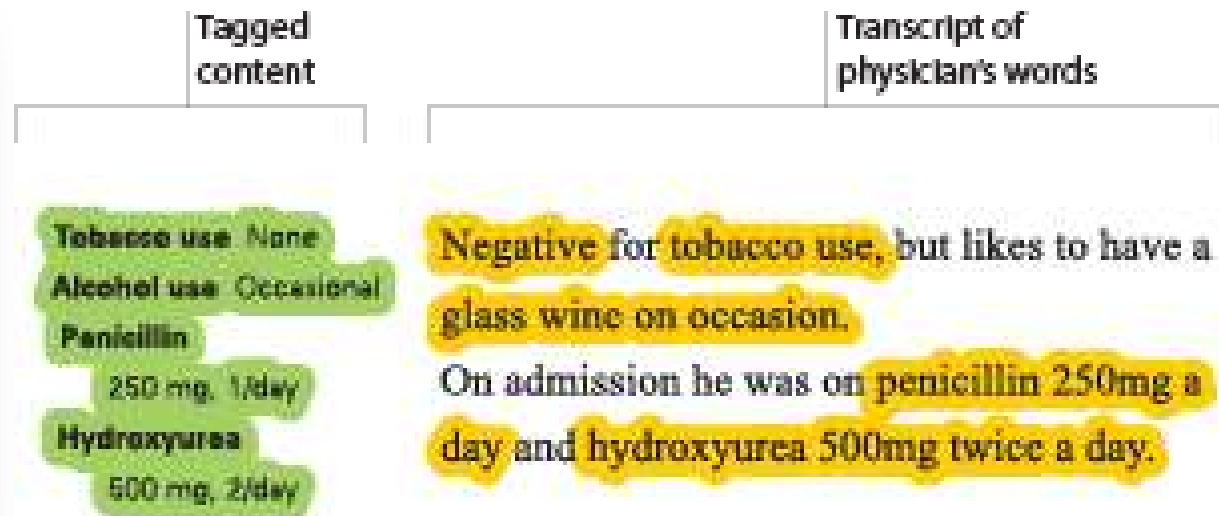
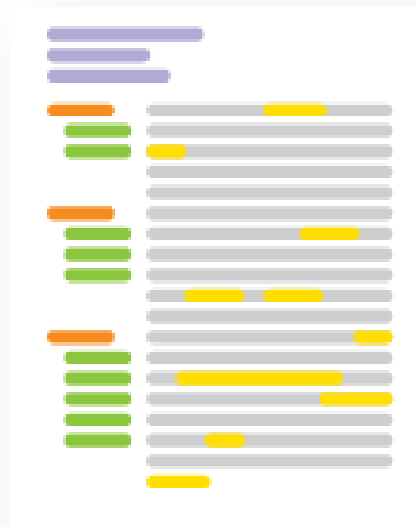
# Speech Understanding Approach

Transform dictation **directly** into structured clinical documents while annotating & encoding data depending on the care givers and organizations needs



# Meaningful Clinical Documents

- Meaningful Clinical Documents are those that blend free form text and fully structured documentation



# Technology-Driven Workflow

- Validation of NLP-annotated data in a single step
- Reconciliation between data captured in structured form and NLP-derived data
- Derivation of actionable information in real-time and appropriate presentation to guide workflow & decisions

# Bigger Picture

- NLP needs to be embedded in scalable platform that manages & deals with
  - Evolving and task-dependent NLP technologies
  - Evolving/changing ontologies and clinical vocabularies
  - Configuration needs to suit a particular task
  - Feedback loop to enable continuous, automatic learning/improvements
  - Technology-driven workflow

# Meaningful Use – Problem List Mgmt

The screenshot displays the M\*Modal EMR interface for a patient named Bob Smith. The top navigation bar includes 'Home', 'Search', 'Patients', 'Reports', and 'Outcomes'. The patient's name and gender (Male) are visible. Below the navigation, there are tabs for 'Documents (10)' and 'Problem List (18)'. The 'Problem List' tab is active, showing a table of medical conditions with columns for ICD-9-CM code, description, date, status, verified, and resolved. The row for 'Depressive disorder' (code 311) is highlighted, and a context menu is open over it with the option 'Discharge Summary' selected. To the right, a 'Discharge Summary' window is open, showing patient information (Bob Smith, MUM, 10000, Sex: Male, Created On: December 19, 2010) and a table of discharge diagnoses. The table lists problems with their review dates, status, codes, descriptions, and comments. Below the table, there are sections for 'Procedures' (listing thoracentesis and duplex ultrasound) and a 'Summary' section with a paragraph of text.

ICD-9-CM	DESCRIPTION	DATE	STATUS	VERIFIED	RESOLVED
599.00	Urinary incontinence	11/20/10 (5 months ago)	Current Past	<input checked="" type="checkbox"/>	<input type="checkbox"/>
481.21	Obstructive chronic bronchitis with exacerbation	11/20/10 (5 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
272.4	Other and unspecified hyperlipidemia	12/15/10 (5 months ago)	Current Past	<input checked="" type="checkbox"/>	<input type="checkbox"/>
311	Depressive disorder	12/19/10 (4 months ago)	Current Past	<input checked="" type="checkbox"/>	<input type="checkbox"/>
511.9	Unspecified pleural effusion	12/19/10 (4 months ago)	Current Past	<input checked="" type="checkbox"/>	<input type="checkbox"/>
518.81	Acute respiratory failure	12/19/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
241.0	Nontoxic unimodular goiter	12/23/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
401.9	Unspecified essential hypertension	12/26/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
255.9	Unspecified disorder of adrenal glands	12/26/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
512.1	Iatrogenic pneumothorax	12/26/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
294.1	Dementia in conditions classified elsewhere	12/26/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
788.3	Urinary incontinence	12/26/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
285.1	Acute posthemorrhagic anemia	12/30/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
281.9	Unspecified deficiency anemia	12/30/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
783.41	Failure to thrive	12/30/10 (4 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
648.24	Anemia	01/03/11 (3 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
496	Chronic airway obstruction	01/03/11 (3 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
799.3	Unspecified debility	01/03/11 (3 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
782.3	Edema	01/03/11 (3 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
162.5	Malignant neoplasm of lower lobe	01/03/11 (3 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>
203.9	Unspecified protein-calorie malnutrition	01/03/11 (3 months ago)	Current Past	<input type="checkbox"/>	<input type="checkbox"/>

### Discharge Summary

Patient: Bob Smith  
 Birthdate: [REDACTED]  
 Author: [REDACTED]  
 MUM: 10000  
 Sex: Male  
 Created On: December 19, 2010

#### Discharge Diagnosis(es)

Review Date	Status	Code	Description	Comment
2010/12/19	Active	511.9	Unspecified pleural effusion	Secondary to recent radiofrequency ablation of lung mass.
2010/12/19	Active	783.41	Failure to thrive	
2010/12/19	Active	311	Depressive disorder	
2010/12/19	Chronic	281.9	Unspecified deficiency anemia	
2010/12/19	Resolved	518.81	Acute respiratory failure	
2010/12/19	Active	496	Chronic airway obstruction	
2010/12/19	Active	203.9	Unspecified protein-calorie malnutrition	

#### Procedures

Include

- Thoracentesis by Interventional Radiology
- Duplex ultrasound of bilateral lower extremities

#### Summary

HOSPITAL COURSE: Briefly, Dr. Smith is a very pleasant, [REDACTED]-year-old, white male who was evaluated in the pulmonary clinic on day of admission for possible right lower lobe infiltrate versus effusion. Approximately 1 month ago, patient underwent a radiofrequency ablation for an isolated right-sided bronchogenic carcinoma, which resulted in a complicated hospital course because of an iatrogenic pneumothorax and tube thoracostomy. He, since discharge from that hospitalization, has gradually become more somnolent, withdrawn. He is less interested in activities that typically he enjoys. He is not eating and drinking as well as he had been. He also has had significant urinary retention after having an indwelling Foley during that hospitalization. His family has been concerned about his withdrawal, and so after seeing

# Quality Measures Reporting

M\*Modal CDA Content Server - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://datamining.mmodal.com/CsGuiHosted/index.html

Most Visited Getting Started Latest Headlines The New York Times ... JIRA Twiki Paychex

M\*Modal CDA Content Server

**M\*Modal** Home Search Patient Reports Advanced

Preferences juergen@rsna2 Logout

Home

- Critical Findings
  - Critical Findings, NC
  - Pneumothorax**
  - Pneumothorax, NC
  - Leaking Aortic Aneurysm
  - Leaking Aortic Aneurysm, NC
  - Ectopic Pregnancy
  - Ectopic Pregnancy, NC
  - Pulmonary Embolism
  - Pulmonary Embolism, NC
  - Deep Vein Thrombosis
  - Deep Vein Thrombosis, NC
- PQRI Measures
  - Number 10 Stroke CT
  - Follow-ups
  - Mammography

**Document List**

60 total results. Displaying page 1 of 6

**CT SOFT TISSUE NECK W & W/O**  
Date: Feb 28, 2006 Physician: CHRISTIANA MEIER  
SEASON JESKA

**CT CHEST W/O**  
Date: Jul 31, 2008 Physician: TOYA WITASZEK  
IN AMADOR

**CT ANGIOGRAM CHEST**  
Date: Oct 03, 2007 Physician: JARROD SHERWOOD  
DEB PERILLI

**CT CHEST W**  
Date: Apr 23, 2007 Physician: LEO HALLAWAY  
ROGER HANZLIK

**CT CHEST W/O**  
Date: Sep 30, 2005 Physician: CHRISTIANA MEIER  
TERESE SKLAR

**XR RIBS BIL**  
Date: Mar 10, 2005 Physician: CLAUD WITTENBURG  
PHUNG BARTLOME

**XR CHEST 2 VIEWS**  
Date: Apr 28, 2008 Physician: RICKY RAMBIN  
EMELINE BRANYON

**XR CHEST 2 VIEW**  
Date: Aug 17, 2005 Physician: CHRISTIANA MEIER  
EZEKIEL SAVITCH

**XR CHEST 2 VIEWS**  
Date: Mar 21, 2008 Physician: JARROD SHERWOOD

**CT SOFT TISSUE NECK W & W/O**

Patient: SEASON JESKA MRN:133614  
Birthdate: September 15, 1967 Sex: Female  
Author: CHRISTIANA MEIER Created On: February 28, 2006

**CT SCAN OF THE NECK:**

**INDICATION:**  
Neck, jaw and face pain. Right throat tightening.

**TECHNIQUE:**  
5 mm axial sections were performed through the neck with IV contrast. The patient received 100 cc of Isovue-300 nonionic contrast for the increased safety factor.

**FINDINGS:**  
The pneumothorax and oropharynx is normal. There is no pharyngeal mucosal space abnormality. The parapharyngeal space is normal. The parotid and submandibular glands are normal. The intrinsic tongue musculature is normal. The glottis and subglottic region is normal. There is no abnormality in the thyroid. The supraclavicular and axillary regions are normal. Study of the maxilla and mandible show normal bone mineralization. The condylar heads are smooth. There is no obvious arthritic change. There are no abnormalities noted in the skull base. The occipital condylar

**Breakdown of all cases by physician**

Physician	Cases
216946011138833211448010	1
216946011138833211448011	2
216946011138833211448012	3
216946011138833211448013	4
216946011138833211448014	5
216946011138833211448015	6
216946011138833211448016	7
216946011138833211448017	8
216946011138833211448018	9
216946011138833211448019	10
216946011138833211448020	11
216946011138833211448021	12
216946011138833211448022	13
216946011138833211448023	14
216946011138833211448024	15
216946011138833211448025	16
216946011138833211448026	17
216946011138833211448027	18
216946011138833211448028	19
216946011138833211448029	20
216946011138833211448030	21
216946011138833211448031	22
216946011138833211448032	23
216946011138833211448033	24
216946011138833211448034	25
216946011138833211448035	26
216946011138833211448036	27
216946011138833211448037	28
216946011138833211448038	29
216946011138833211448039	30
216946011138833211448040	31
216946011138833211448041	32
216946011138833211448042	33
216946011138833211448043	34
216946011138833211448044	35
216946011138833211448045	36
216946011138833211448046	37
216946011138833211448047	38
216946011138833211448048	39
216946011138833211448049	40
216946011138833211448050	41
216946011138833211448051	42
216946011138833211448052	43
216946011138833211448053	44
216946011138833211448054	45
216946011138833211448055	46
216946011138833211448056	47
216946011138833211448057	48
216946011138833211448058	49
216946011138833211448059	50

**Breakdown of my cases by age**

Age Group	Cases
60-69	7
70-79	5

Done

1:53 PM

# Search & Discovery

The screenshot displays the M\*Modal web application interface. At the top, there is a search bar containing the text "<neoplasm>". Below the search bar, the interface is divided into several sections:

- Left Panel:** Contains a pie chart and a list of filters including "Age Group", "Responsible Provider Type", "Sex", "Work Type", and "Practice Priority". Below the chart are dropdown menus for "Date Range" (set to "All") and "Provider" (set to "All").
- Search Results Panel:** Displays a list of search results. The first result is highlighted in grey and is titled "US TRANSVAGINAL". Other results include "CT SINUS AXIAL & CORONAL W/O", "US THYROID", "US BREAST LT", "MRI LUMBAR SPINE W/O", "MRI FOOT RT W/O", "US PELVIC AND TRANSVAG PROBE", and "US PELVIC AND TRANSVAG". Each result includes the date, physician name, and patient name.
- Right Panel:** Displays the details for the selected "US TRANSVAGINAL" report. It includes patient information (SOPHIA CRAPSE, DOB: February 8, 1998, Author: JARROD SHERWOOD), MRN (27341), and creation date (March 28, 2007). The report title is "TRANSABDOMINAL AND TRANSVAGINAL PELVIC ULTRASOUND". The "HISTORY" section notes "Heavy periods with clots." The "FINDINGS" section describes an enlarged uterus (12.9 x 9.1 x 10.5 cm) with multiple focal hypoechoic uterine masses compatible with numerous uterine leiomyomata. The "IMPRESSION" section concludes with "Enlarged uterus with multiple uterine leiomyomata." and "The endometrial stripe is distorted by multiple uterine leiomyomata with some thickening, possible hyperplasia. Other differential would include endometrial leiomyoma, polyp, hyperplasia, or even carcinoma. Recommend appropriate GYN followup." A note at the bottom states "Please check glucose level as part of GYN followup to rule out diabetic correlations" with a "Submit" button.

# Conclusion

- NLP by itself is not a silver bullet to make narrative data accessible and actionable
- NLP is a useful technology tool, but by itself doesn't scale
- Natural language is super-expressive, NLP is not perfect -> validation and feedback loops are important
- Keep an eye on new developments
- NLP needs to be embedded in an open, scalable application platform

# Thank you!

Questions?