Panel Discussion: State of VA NLP Research

Accomplishments, Opportunities, Challenges

September 10, 2015

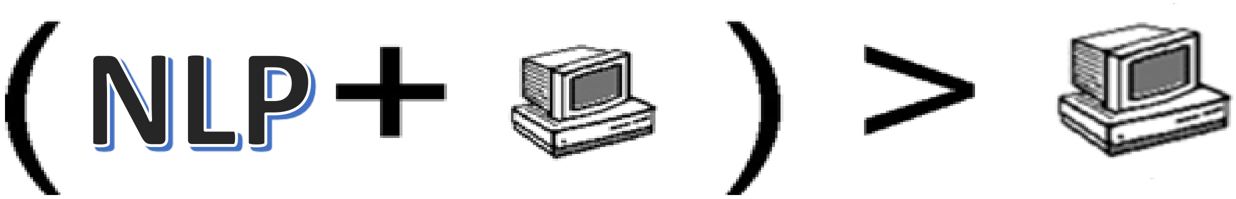
VA’s Unique Advantages I

- Size of intramural research community
  - Huge number of use cases
- Culture of collaboration across academic groups
  - Rewards through positive sum games
- Support for research-operational partnerships
  - Programs such as CREATE, QUERI, COIN

VA Unique Advantages II

- Parallels and synergies between NLP and qualitative research
- Every text project has national scope
- Tons of opportunities to drive transformation
  - Make manual chart review more productive

  - Make analysis of electronic data more powerful



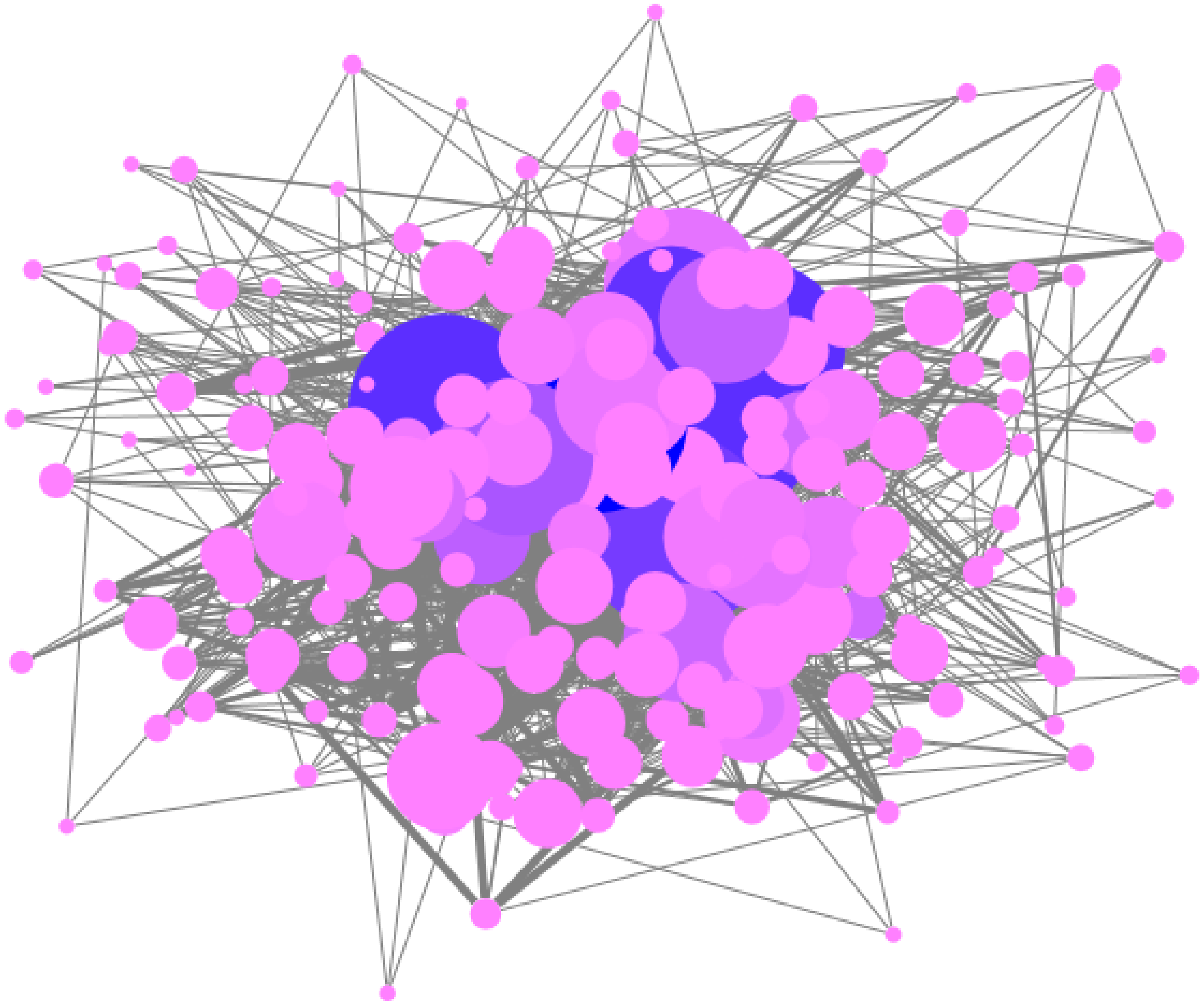
What has been accomplished

NLP Researchers New to VA Since 2009

An incomplete list

- Qing Zeng
- Hong Yu
- Wendy Chapman
- Dezon Finch
- Ruth Reeves
- Glenn Gobbel
- Scott Duvall
- Stephane Meystre
- Guy Divita
- Liz Workman
- Olga Patterson
- Lina Bouayad

Consortium for Healthcare Informatics Research (CHIR): Group Photo



CHIR social network graph based on co-authorship

Example Opportunity

- Sections and Templates

OBSecAnnot: An Automated Section Annotator for Semi-structured Clinical Documents

Le-Thuy T. Tran, Guy Divita, Andrew Redd, Marjorie Carter, Joshua Judd, Matthew H. Samore, Adi V. Gundlapalli

Novel Template Identification from VA Text Integration Utility Notes

Andrew M. Redd, Guy Divita, Adi V. Gundlapalli, Le-Thuy Tran, Mathew Samore

Example Challenge

- Create synthetic text documents to facilitate collaboration with partners outside of VA
- Generate documents that retain key distributional properties of real notes (e.g., concept co-occurrences)
- Leverage thousands of template types

Interactive & Active Assisted Annotation and Natural Language Processing of Medical Text

Michael E. Matheny, MD, MS, MPH

Associate Director, Advanced Fellowship in Medical Informatics
Tennessee Valley Healthcare System VA

Director, Center for Population Health Informatics
Vanderbilt University Departments of Biomedical Informatics, Medicine, and Biostatistics

CPHI
CENTER FOR POPULATION HEALTH INFORMATICS

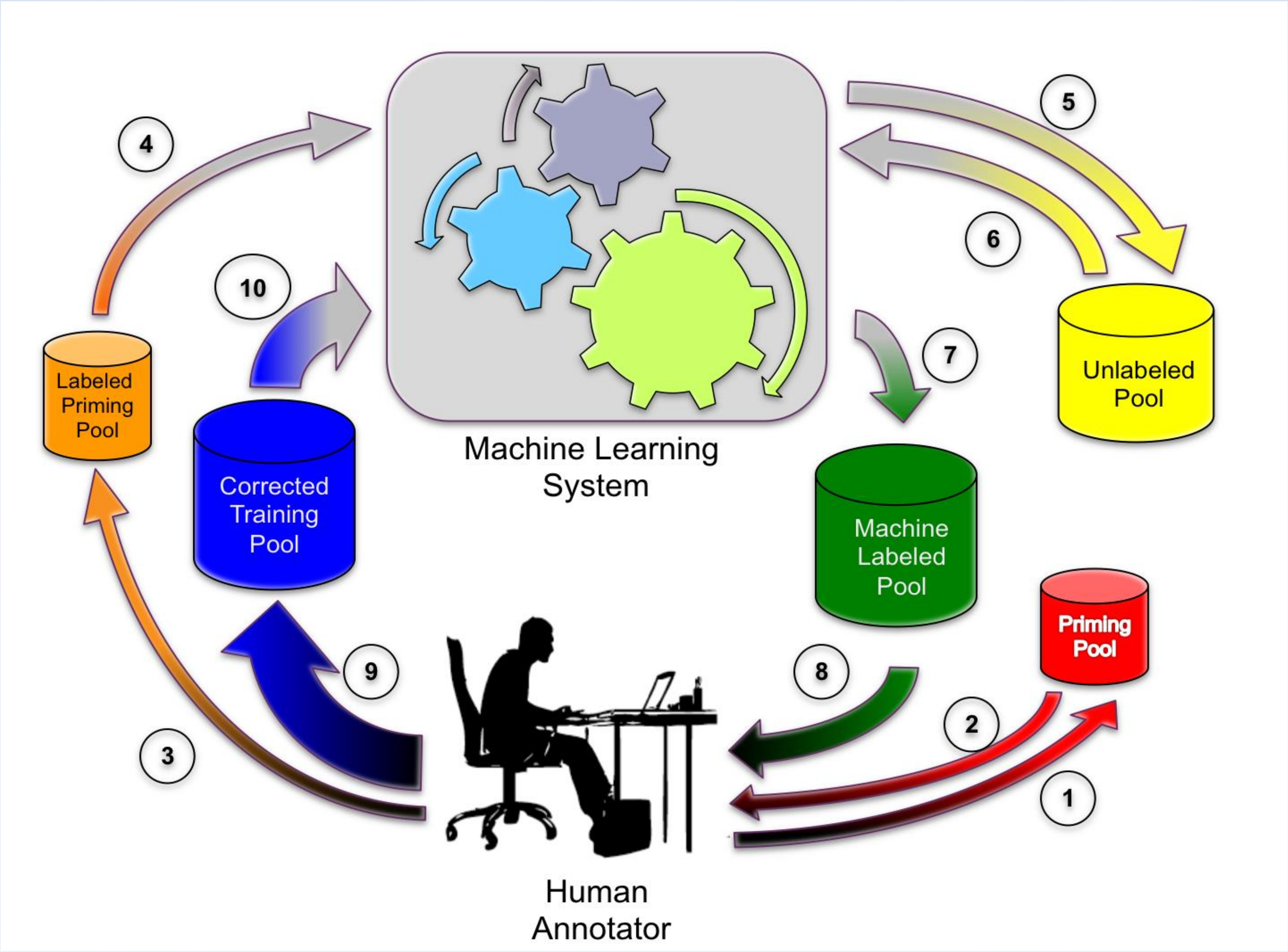
Objectives

To develop and validate natural language processing tools capable of near real-time processing employing active & interactive learning techniques

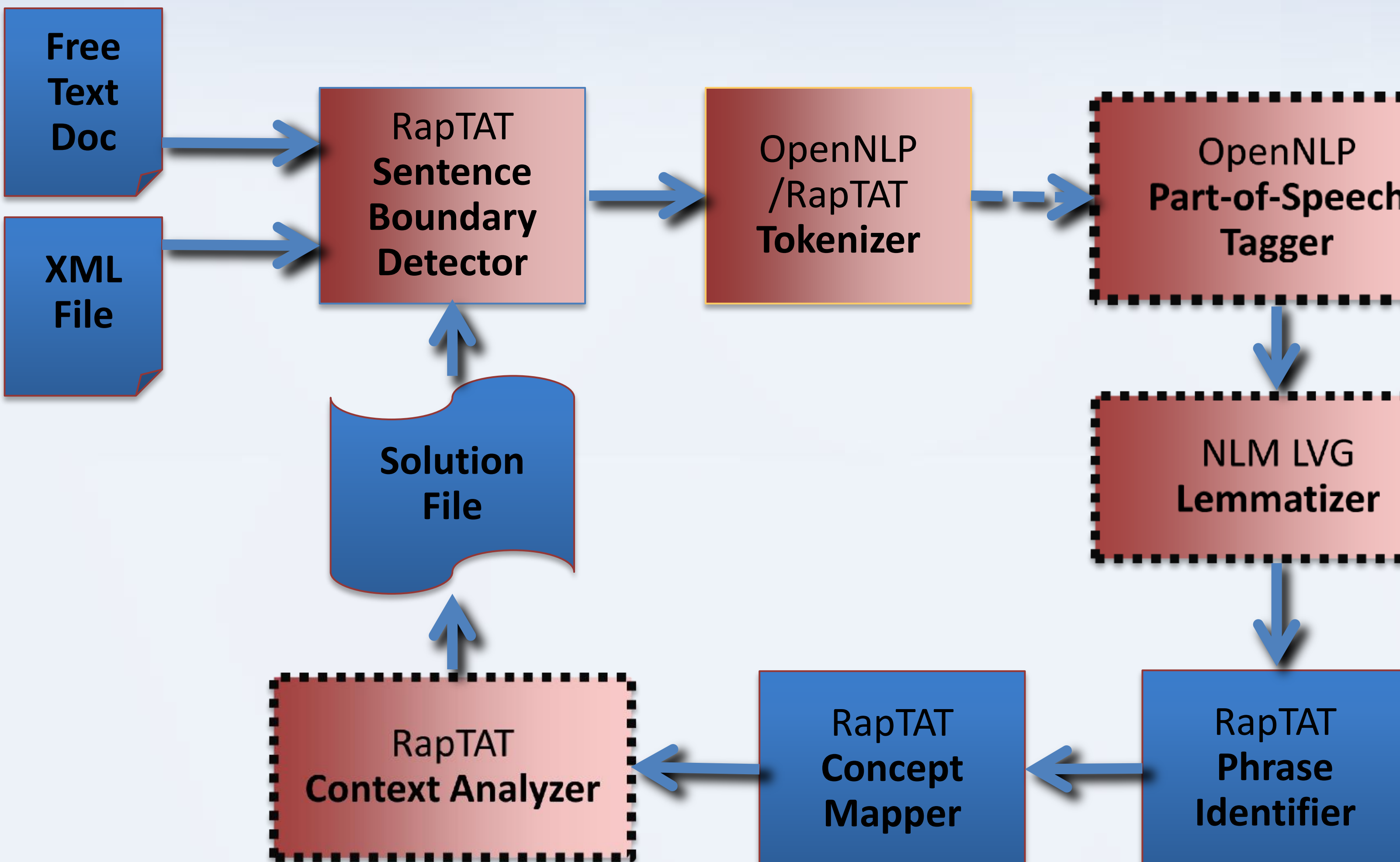
Use Cases:

- Heart Failure Quality Indicator Elements
- Acute Kidney Injury Risk Factors

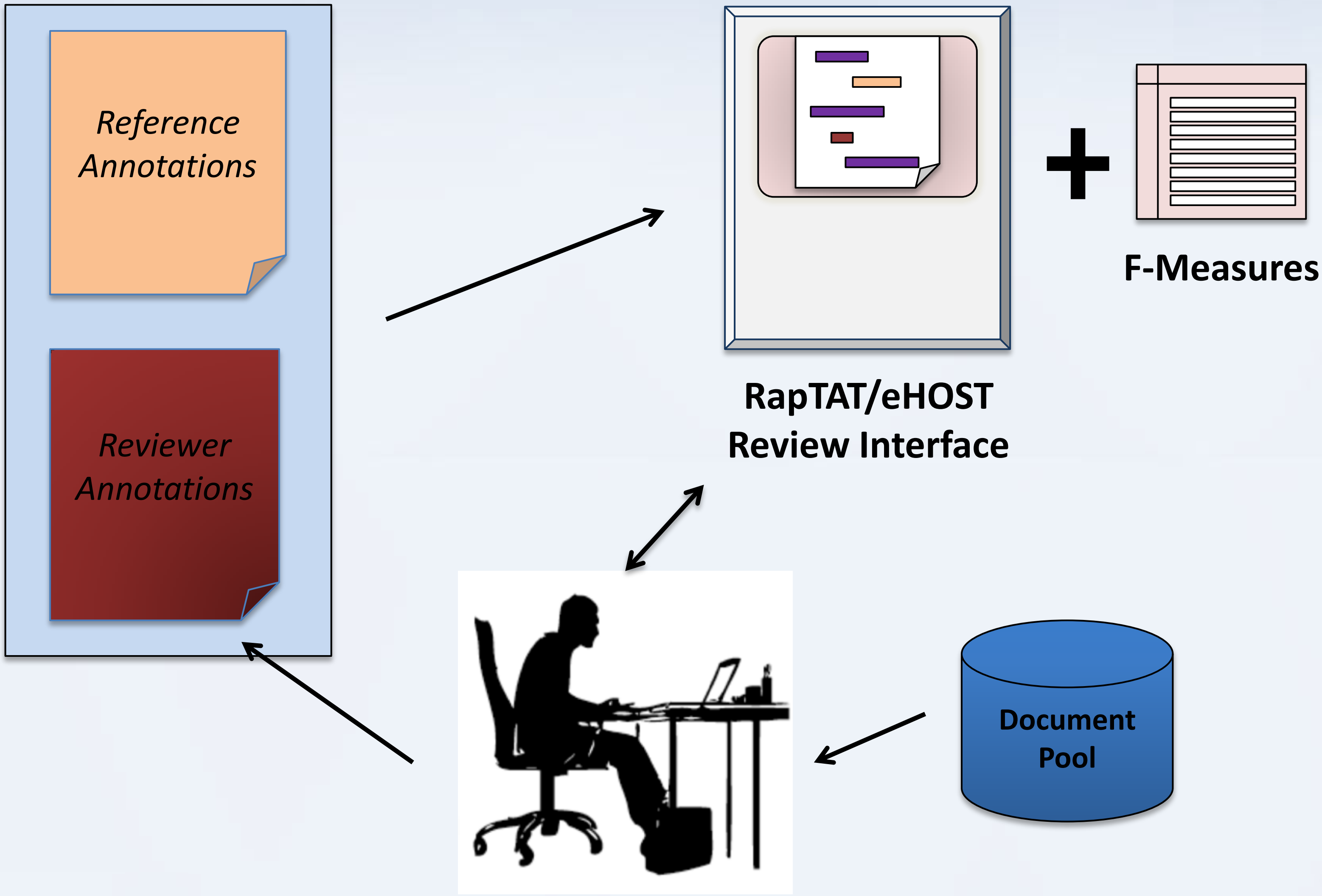
Online Assisted Annotation Workflow



RapTAT: Annotation/NLP Pipeline



RapTAT for Reviewer Training



Online Assisted Annotation

Heart Failure Use Case

Study Design: Docs & Schema

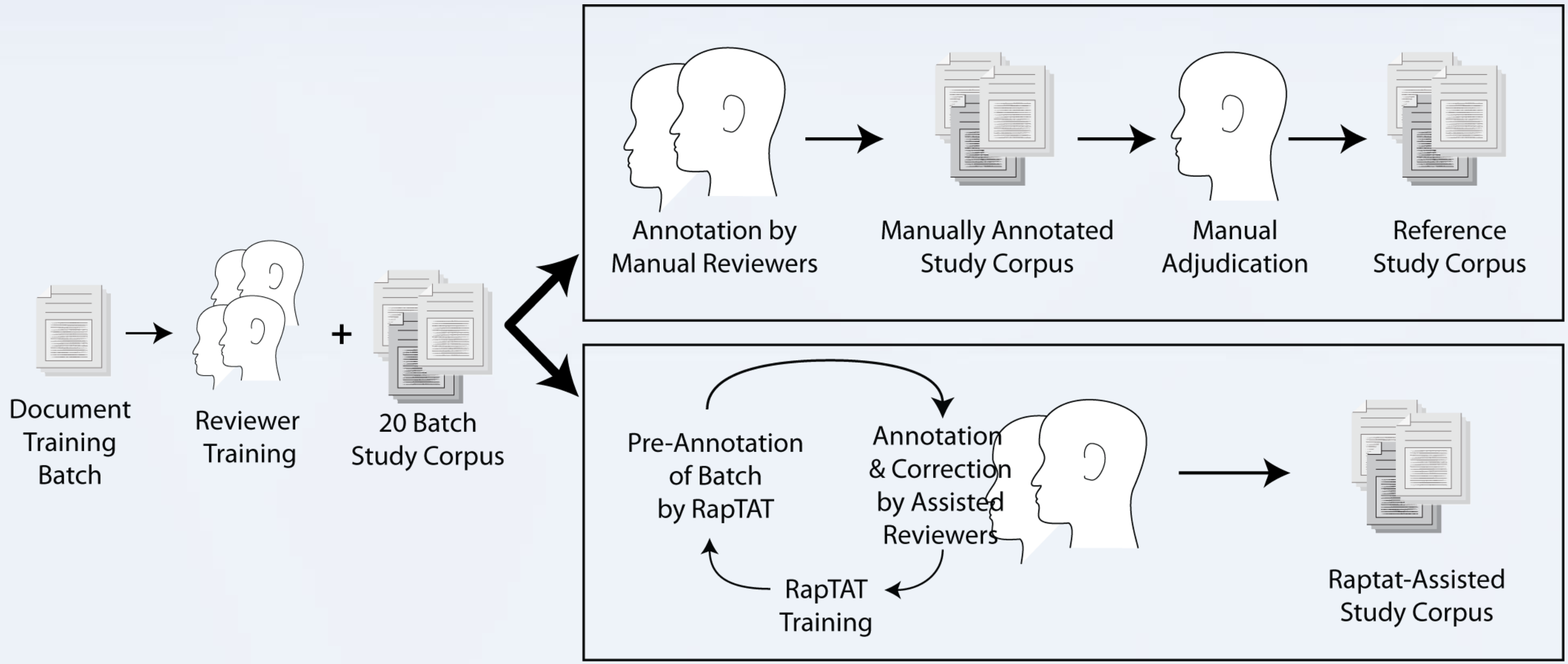
Documents:

Emergency Department and Inpatient provider and nursing notes as well as Primary care notes from 2007-2008 in 6 VA facilities (404 documents / 171 patients)

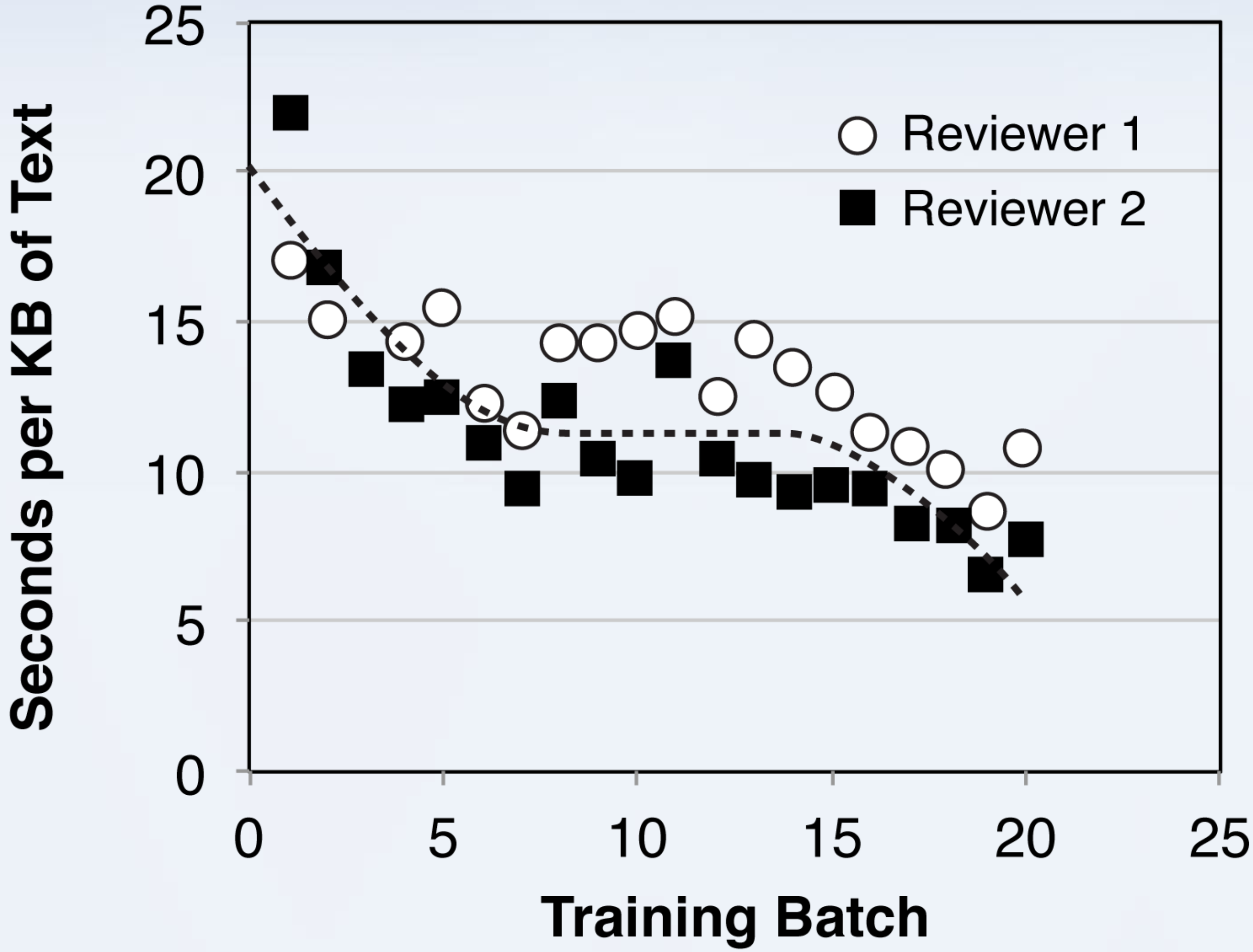
Annotation Schema:

Concept	# Documents containing concept	Example Phrases (Annotated Tokens in Bold)
Angiotensin Converting Enzyme Inhibitor	272	"**ACEI**," "**Altace**," "**Vaseretic**"
Angiotensin II Receptor Blocker	107	"**ARB**," "**Sartans**"
Ejection Fraction	201	"**LVEF**", "**Ejection fraction**"
Ejection Fraction Quantitation	197	"EF=**60-70%**," "EF is **about 30%**"
Left Ventricular Systolic Function/Dysfunction	79	"**Systolic dysfunction**," "**LV function**"
Left Ventricular Systolic Function Value	76	"Systolic function is **borderline normal**"
Documented Reason Not on ACE Inhibitor/ARB	40	"**Patient refuses** to take ACEI," "**Renal disease**"

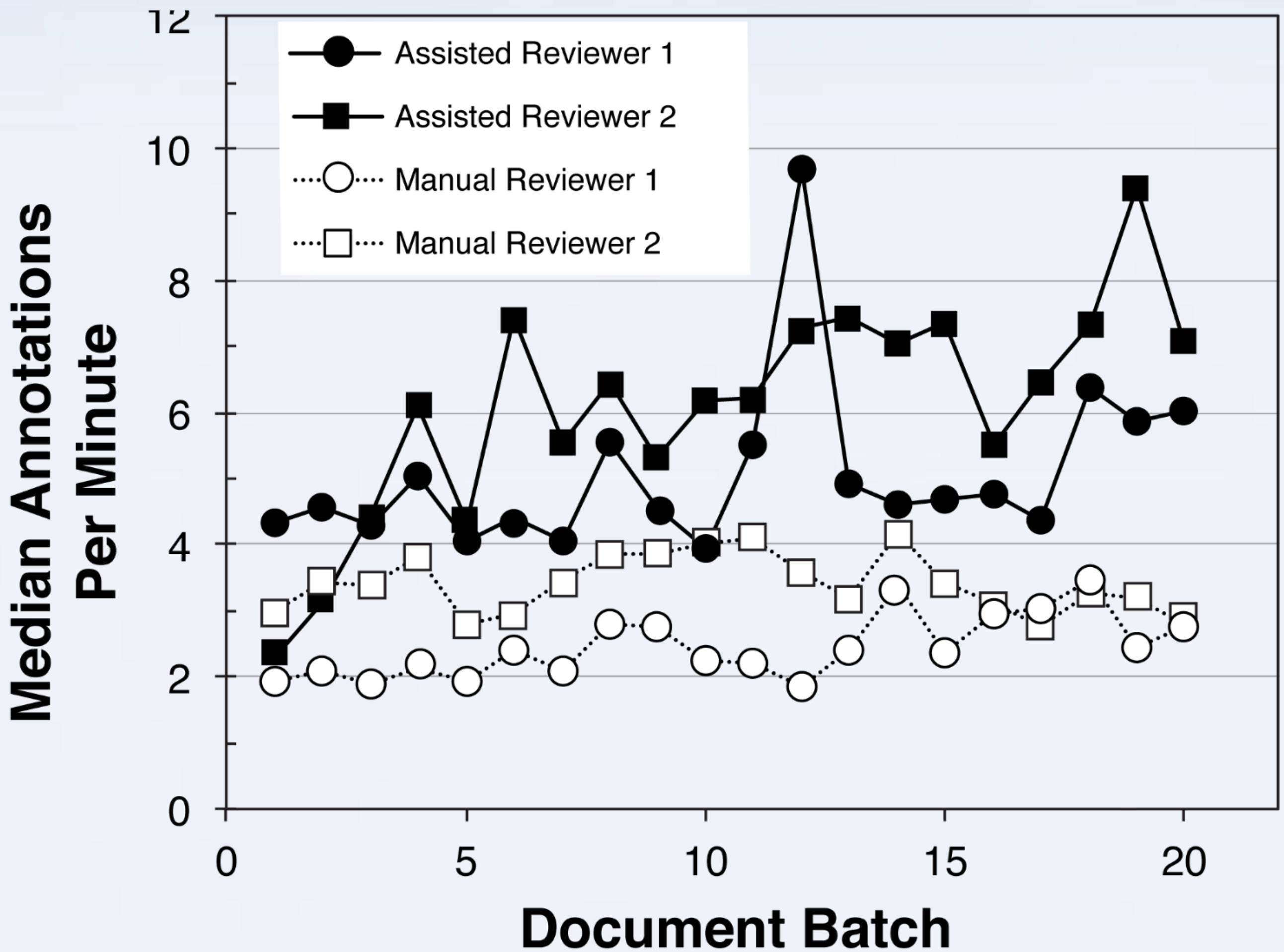
Study Design: Workflow



Assisted Annotation Time



Assisted Versus Manual Annotation Rate



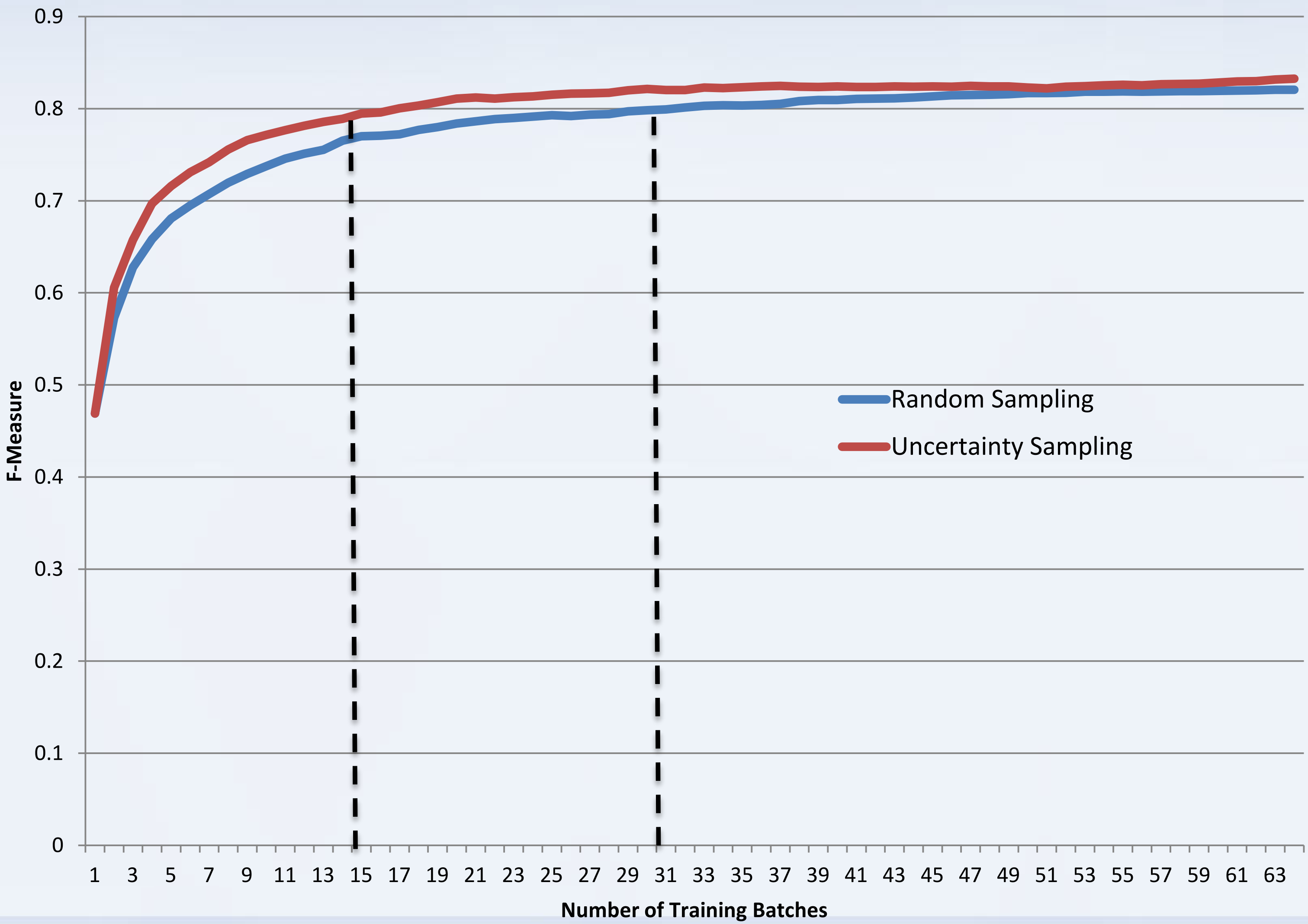
Inter-annotator agreement

Concept	Average IAA (95% Confidence Interval)	
	Manual	Assisted
Angiotensin Converting Enzyme Inhibitor	0.89 (0.86-0.93)	***0.93**** (0.91-0.96)
Angiotensin II Receptor Blocker	0.81 (0.72-0.89)	***0.97**** (0.95-1.00)
Ejection Fraction	0.86 (0.80-0.93)	***0.97**** (0.95-1.00)
Ejection Fraction Quantitation	0.90 (0.85-0.94)	0.88 (0.83-0.92)
Left Ventricular Systolic Function/Dysfunction	0.82 (0.73-0.91)	0.76 (0.62-0.89)
Left Ventricular Systolic Function Value	0.85 (0.78-0.93)	0.77 (0.64-0.90)
Reason Not on ACE Inhibitor/ARB	0.58 (0.46-0.70)	0.54 (0.45-0.64)
Total (Combined Over All Concepts)	0.85 (0.81-0.88)	***0.89**** (0.87-0.91)

RapTAT Performance

Concept	Performance Measure		
	Precision	*Recall*	*F*
Angiotensin Converting Enzyme Inhibitor	0.97	0.94	0.95
Angiotensin II Receptor Blocker	0.99	0.96	0.97
Ejection Fraction	0.96	0.95	0.96
Ejection Fraction Quantitation	0.77	0.82	0.80
Left Ventricular Systolic Function/Dysfunction	0.61	0.82	0.70
Left Ventricular Systolic Function Value	0.83	0.37	0.51
Reason Not on ACE Inhibitor/ARB	0.36	0.12	0.18
Total (Combined Over All Concepts)	0.87	0.82	0.85

Active Learning: Random vs Uncertainty Sampling



Near Real-Time NLP

Acute Kidney Injury Use Case

Study Design

- Documents: Stratified sample of emergency department, inpatient, and outpatient provider notes
- 14 Annotated Blocks (112 documents/block [1568 total]): 8 Training Blocks / 6 Testing Blocks
- Results for this presentation only on training data (final performance pending testing data evaluation)

Study Design: Annotation Schema

Clinical Variable	Attributes *	Examples
Renal Function Impairment*	Chronicity: Acute, Chronic, Acute-on-Chronic, Unstated	"Acute Kidney Injury", "Nephropathy", Renal Tubular Acidosis, "Nephrotic Range Proteinuria"
Anatomical Kidney Status	State: Solitary, Nonfunctioning, Atrophic, Surgically Removed, Other	"Hydronephrosis", "Nephrectomy", "Kidney Donor", "Renal Mass"
Renal Transplant Recipient		"Kidney Allograft", "Renal Rejection", "Renal Transplant Recipient"
Nephrology Care Delivery	Type: General, Transplant, Dialysis	"Hemodialysis", "Renal Consult", "Nephrology Clinic", "Renal Transplant Clinic"
NSAIDs		"Meloxicam", "Ketorolac","Celecoxib"
ACE Inhibitors		"Enalapril", "Mavik"
ARB		"Diovan", "Losartan"
Diuretic		"Furosemide", "Spirinolactone"
Diuresis		"diuresing", "forced diuresis"
Intake	Change: Increase, Decrease, Neutral Fluidity: Solid, Liquid, Both, Unstated Agency: Provider or Patient Initiated Delivery: IV, Oral, Unstated	"Fluid resuscitate", "no change in appetite", "fluid restriction", "intolerate of PO", "NPO"
Intravascular Volume	Status: Low, High, Normal	"Hypovolemia", "volume contraction", "isovolemic", "dry oral mucous membrane"
Weight Change	Status: Increase, Decrease, Neutral	"fluctuating weight", "cachectic", "weight loss", "weight gain"
Nausea/Vomiting/Diarrhea		"NVD", "N/V/D", "emesis"

* All variables include standard attributes: Assertion (Negation, Uncertainty), Time Frame (Past, Present, Future), Experiencer (Patient, Non-Patient)

NLP Performance Summary

Category	#	TP	FP	FN	Precision (PPV)	Recall (Sens.)	F-Measure
Drug Exposures							
• ACE Inhibitor	575	553	8	22	0.986	0.962	0.974
• ARB	149	137	0	12	1.000	0.919	0.958
• Diuretic	733	684	4	49	0.994	0.933	0.963
• NSAID	233	201	4	32	0.980	0.863	0.918
Fluid Status							
• Diuresis	118	83	6	35	0.933	0.703	0.802
• Intake	694	412	46	282	0.900	0.594	0.715
• Intravascular Volume Condition	527	432	12	95	0.973	0.820	0.890
• Nausea/Vomiting/Diarrhea	719	674	25	45	0.964	0.937	0.951
• Weight Change	221	130	14	91	0.903	0.588	0.712
Radiographic Media Exposure							
• Contrast	2095	1858	240	237	0.886	0.887	0.886
• Potential Contrast	439	255	65	184	0.797	0.581	0.672
Renal Status							
• Anatomical Kidney Status	57	9	4	48	0.692	0.158	0.257
• Nephrology Care Delivery	210	141	36	69	0.797	0.671	0.729
• Renal Function Impairment	449	368	44	81	0.893	0.820	0.855
Total Concept Performance	**7231**	**5661**	**341**	**1570**	**0.921**	**0.821**	**0.868**

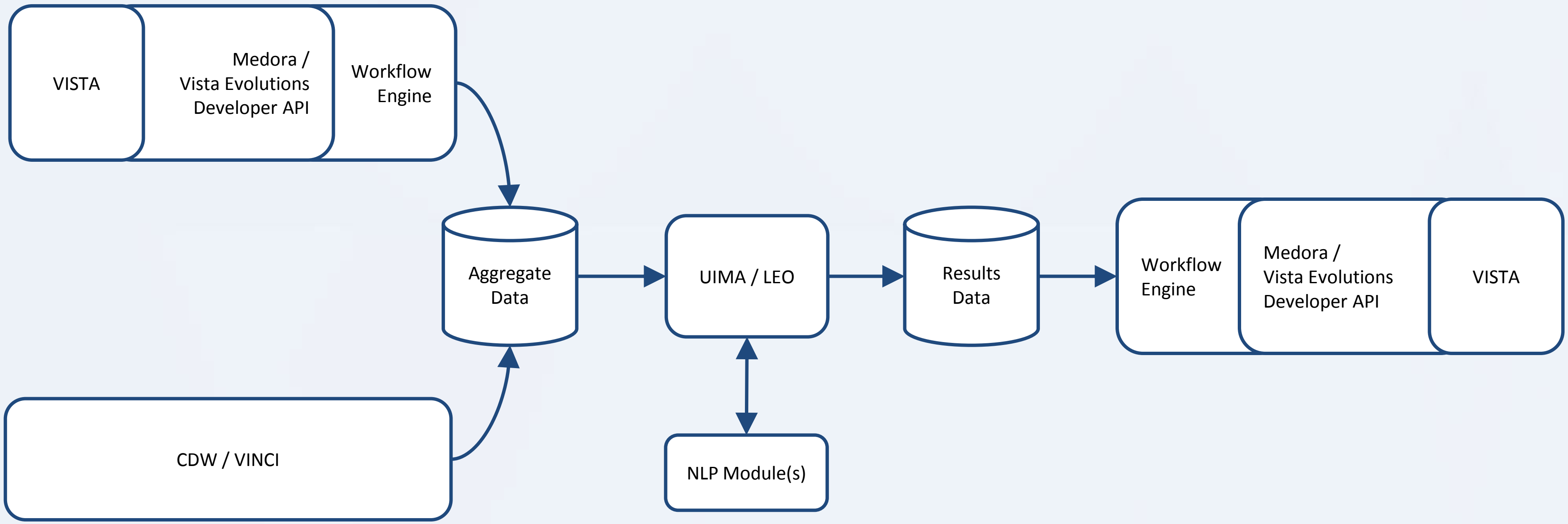
Concept Assertion Performance

		Algorithm			Recall (Sensitivity)
		Positive	Negative	Uncertain	92%
Reference Standard	Positive	1736	27	119	92%
	Negative	142	350	33	67%
	Uncertain	52	26	92	54%
Precision (PPV)		90%	87%	38%	85%

	TP	FP	TN	FN	Sensitivity	Specificity
Positive vs Negative	1849	167	358	167	92%	68%

Near Real-Time Deployment

- National Cardiac Catheterization Cohort:
  - 158,432 Patients (as of 09/2013)
  - 1,256,685 Documents (Filtered)



Processing Speed: ~1 sec/ document on an single machine installation = 86,400 / day

Conclusions

- Assisted annotation tools can reduce cost of NLP training samples – key requisite for all NLP tool performance assessment
- Focus on accuracy but awareness and optimization for performance in NLP tools necessary to allow near-real time implementation
- Focused NLP tasks, while less generalizable than generalized solutions, are likely to be the first implemented in point of care applications due to optimization and higher accuracy

Acknowledgements

Grant Funding

VA HSR&D IIR 11-292-(Matheny)

VA HSR&D IIR 13-052 (Matheny/Ho)

VA HSR&D CDA 08-020 (Matheny)

PCORI CDRN Phase 1 & 2 (pScanner Ohno-Machado)

VA HSR&D IIR 12-362 (Reeves)

RapTAT Lead Developer & Key Collaborator: Glenn Gobbel

- Nashville
  - Steven Brown
  - Fern FitzHenry
  - James Fly
  - T. Alp Ikizler
  - Shrimalini Jayaramaraja
  - Vincent Messina
  - Sharidan Parr
  - Ruth Reeves
  - Sanjib Saha
  - Edward Siew
  - Dax Westerman
  - Theodore Speroff
- Denver VA
  - Tom Maddox
  - Meg Plomondon
  - Tom Tsai
- Salt Lake VA
  - Wendy Chapman
  - Jennifer Garvin
  - Stephane Meystre
- Tampa VA
  - Steve Luther
  - Dezon Finch
  - James McCart

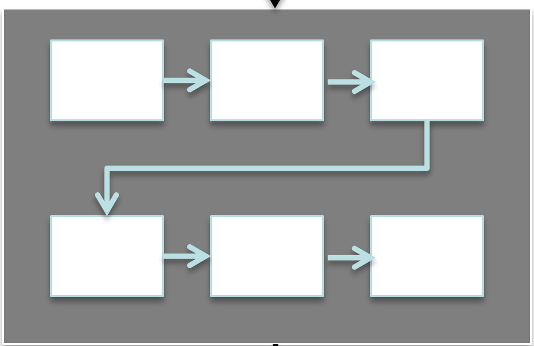
For more information contact:

michael.matheny@va.gov

michael.matheny@vanderbilt.edu

Applying NLP to Clinical Problems

"no family history of colon cancer"



(no (family history (of (colon cancer))))

C00289 negated

Gap

- Tell me how I can improve the quality of my colonoscopy exams
- Find patients with X
- Help me use the words in my report that justify fair billing codes
- Help me spend less time documenting
- Help my patient understand her report

Clinical NLP

Methodology Development

UNIVERSITY OF UTAH SCHOOL OF MEDICINE

Many interesting NLP research problems are far upstream from health care applications.

Document classification

Temporal reasoning

Part-of-speech tagging

Relationship identification

Named entity recognition

To have impact,
we must go beyond improving f-score
to create tools that can be applied to real-world problems

Clinical NLP

Methodology Development

Application of Methods to Domain Problem

Many new, interesting NLP research problems will arise when working on user-driven development of clinical applications.

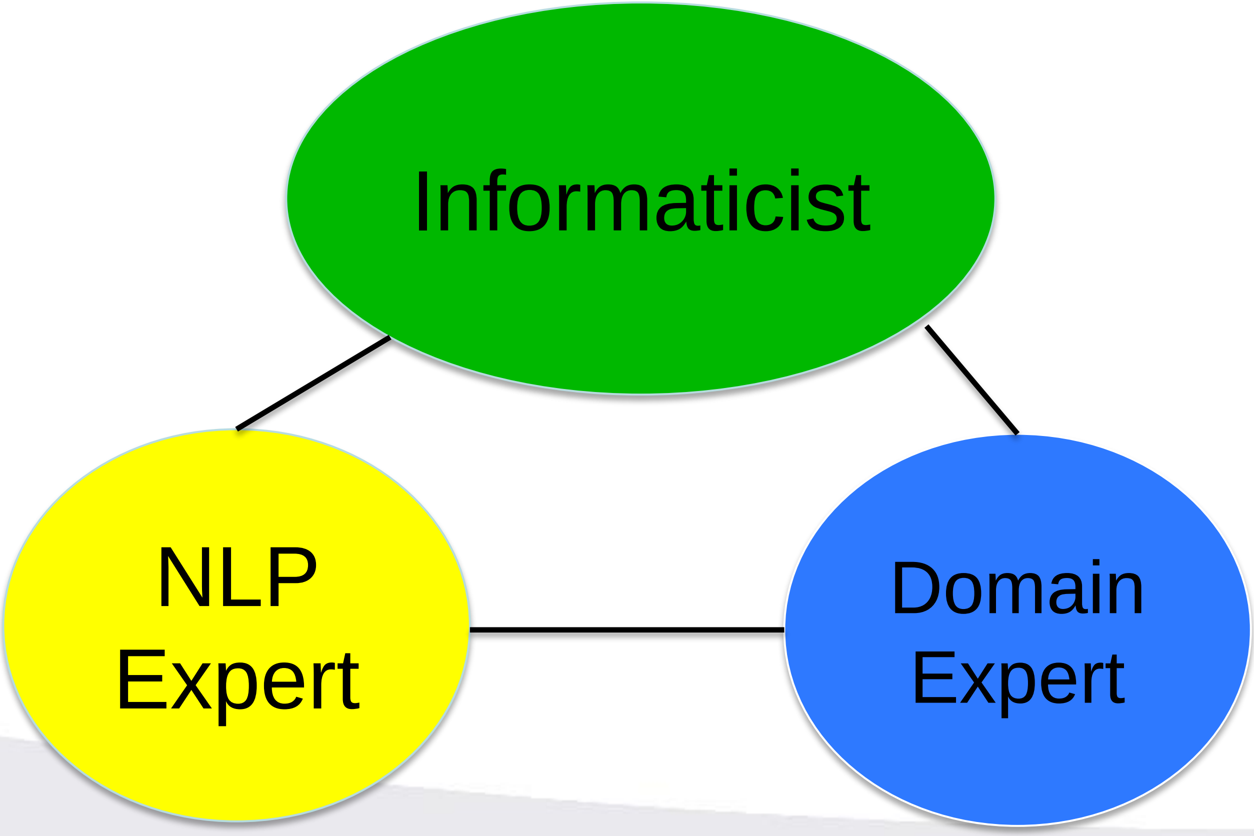
Clinical NLP

VA NLP

Methodology Development

Application of Methods to Domain Problem

Effective NLP in the VA = Partnership



IE-Viz

Information Extraction and Visualization

Knowledge Authoring

NLP Customizing

Classifier Developmen t

Visualization

UNIVERSITY OF UTAH
SCHOOL OF MEDICINE

Knowledge Authoring

Domain Schema Ontology

Linguistic representation of clinical elements

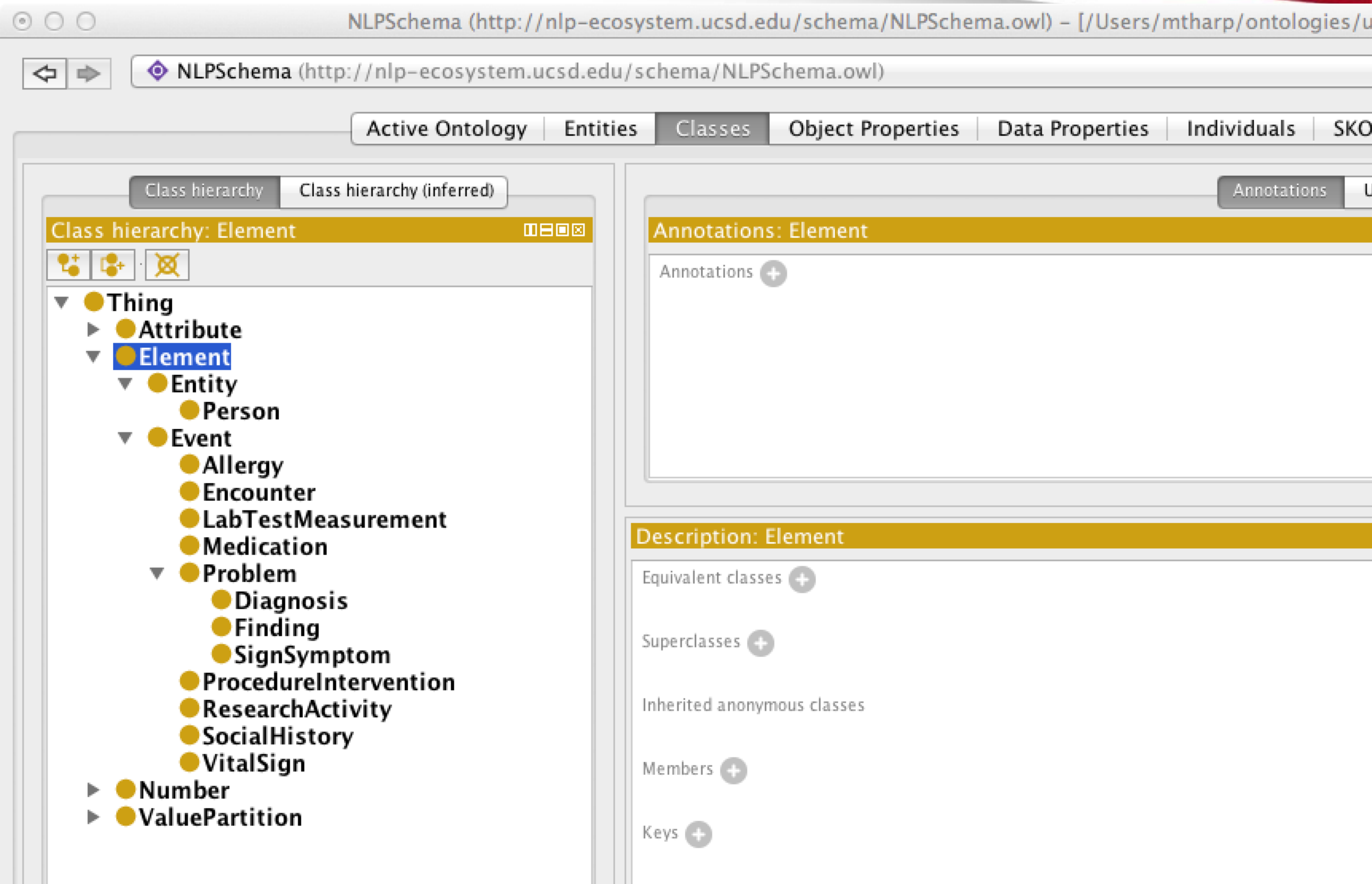
Modifier Ontology

Modifiers of clinical elements

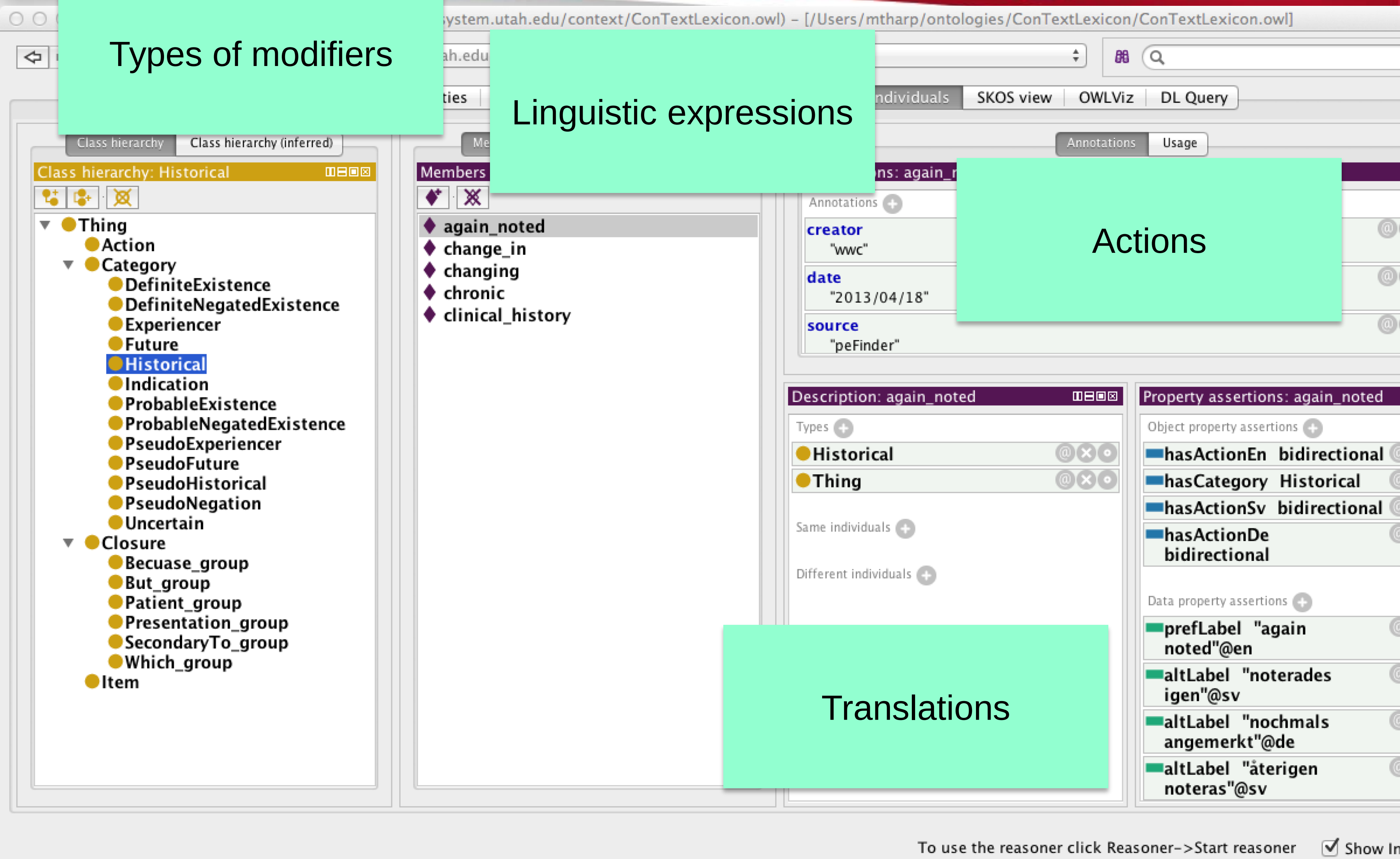
"Patient denies a family history of colon cancer" →

Disease: colon cancer
Experiencer: family
Negation: no
Historical: yes

Melissa Tharp, Bill Scuba, Danielle Mowery, Jan Wiebe, Harry Hochheiser


NLPSchema (http://nlp-ecosystem.ucsd.edu/schema/NLPSchema.owl) – [/Users/mtharp/ontologies/u
NLPSchema (http://nlp-ecosystem.ucsd.edu/schema/NLPSchema.owl)
Active Ontology
Entities
Classes
Object Properties
Data Properties
Individuals
SKO
Class hierarchy
Class hierarchy (inferred)
Class hierarchy: Element
Thing
Attribute
Element
Entity
Person
Event
Allergy
Encounter
LabTestMeasurement
Medication
Problem
Diagnosis
Finding
SignSymptom
ProcedureIntervention
ResearchActivity
SocialHistory
VitalSign
Number
ValuePartition
Annotations
Annotations: Element
Annotations
Description: Element
Equivalent classes
Superclasses
Inherited anonymous classes
Members
Keys

Modifier Ontology



Schema Ontology Imports Modifier Ontology

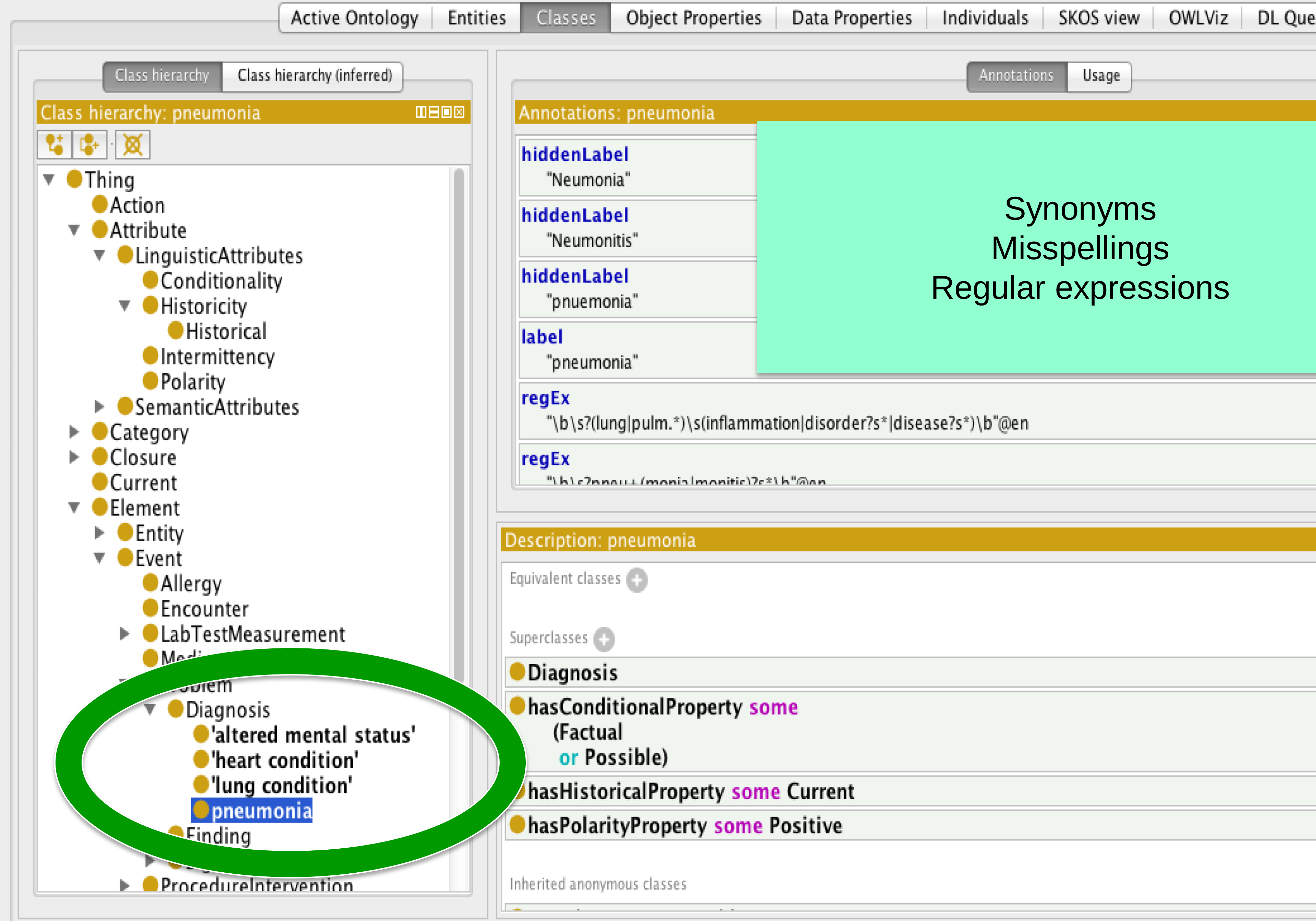
Medications
- Type
- Dose
- Frequency
- Route

Diagnosis
- Negation
- Uncertainty
- Severity
- History
- Experiencer

Consistent with other models:
Clinical element models, cTAKES type system, Common model, FHIR

Domain Ontology for NLP

- Instance of schema ontology
- Clinical elements from a particular domain



To use the reasoner click Reasoner->Sta

Knowledge Author

- Front end interface for users
- Back end
  - Schema ontology
  - Modifier ontology
- Output
  - Domain ontology
  - Schema for NLP system

African American Adult

Create Person Variable

Person Role

Patient

Family Member

Age From * 18

To * 120

Gender * Not Specified Male Female

Extra Fields

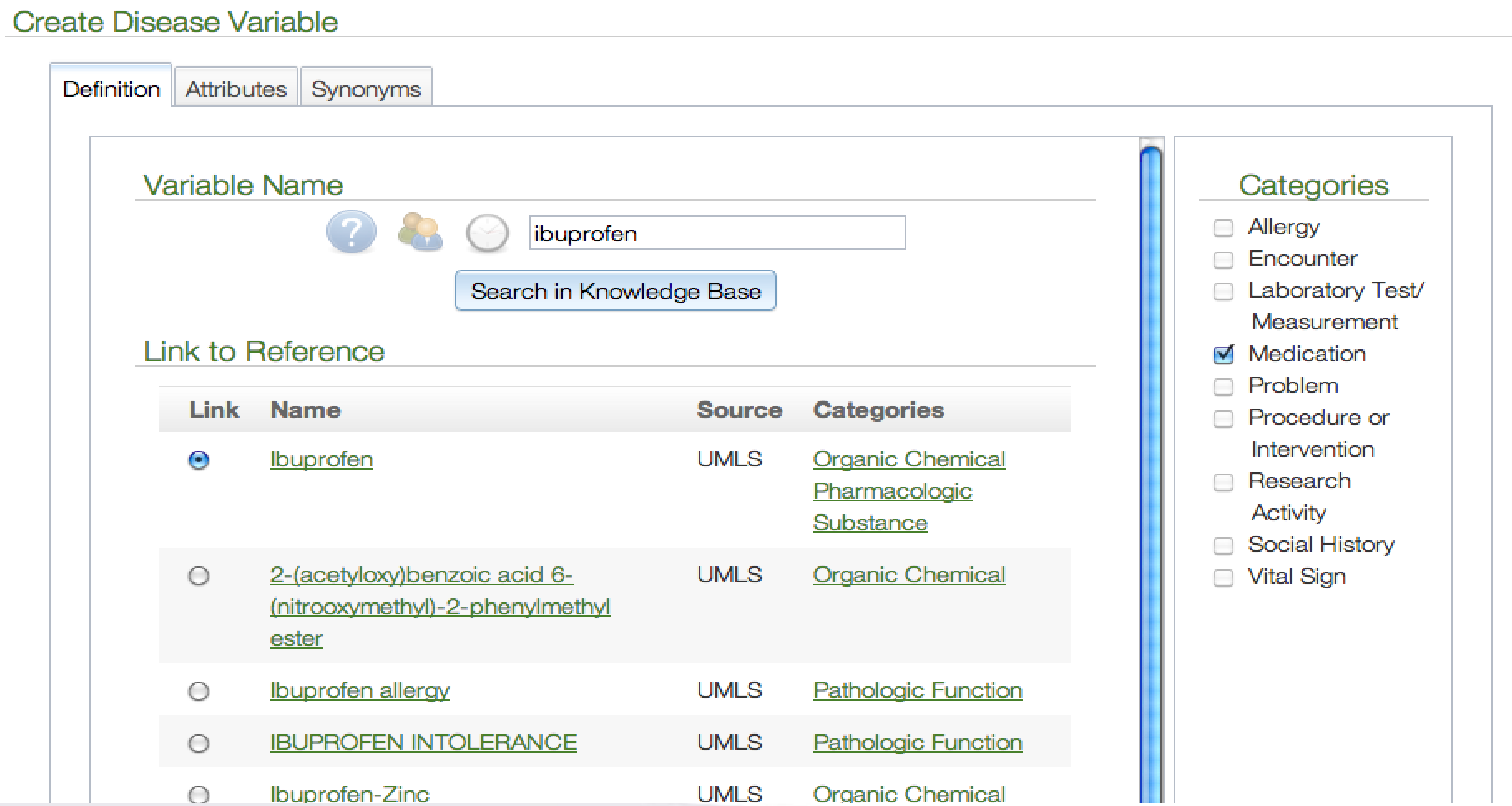
Race: African American

Death Date

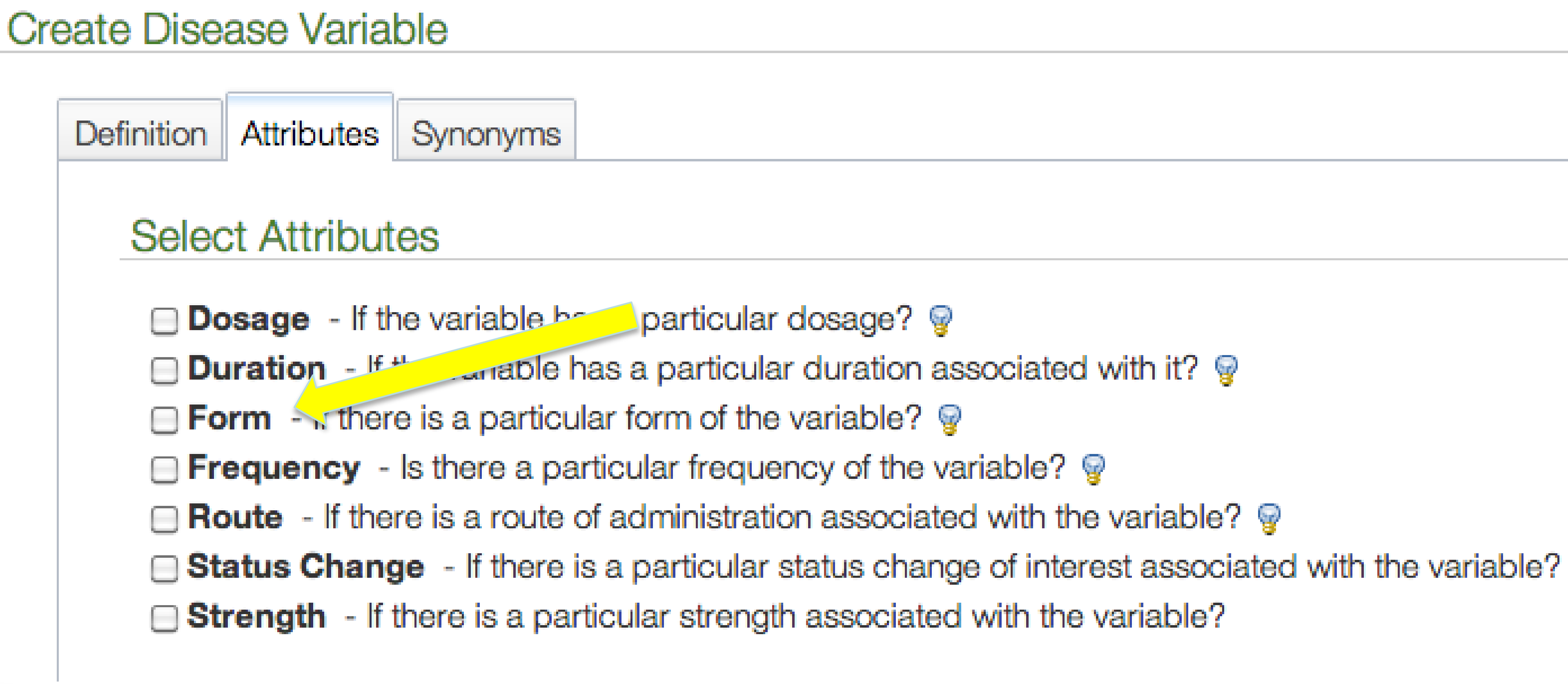
Birth Date

Ethnicity

Ibuprofen

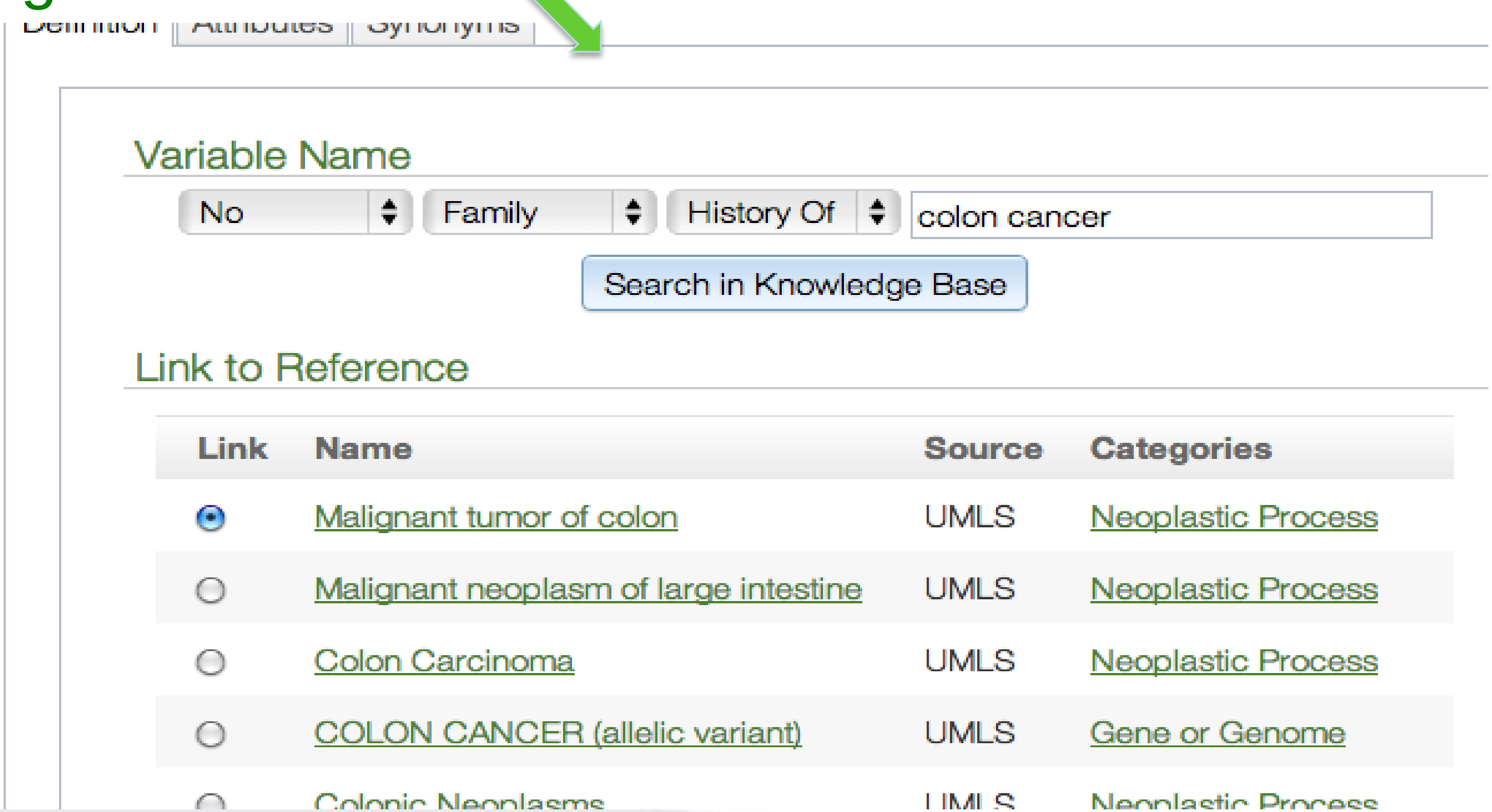

Create Disease Variable
Definition
Attributes
Synonyms
Variable Name
ibuprofen
Search in Knowledge Base
Link to Reference
Link
Name
Source
Categories
Ibuprofen
UMLS
Organic Chemical
Pharmacologic Substance
2-(acetyloxy)benzoic acid 6-(nitrooxymethyl)-2-phenylmethyl ester
UMLS
Organic Chemical
Ibuprofen allergy
UMLS
Pathologic Function
IBUPROFEN INTOLERANCE
UMLS
Pathologic Function
Ibuprofen-Zinc
UMLS
Organic Chemical
Categories
Allergy
Encounter
Laboratory Test/ Measurement
Medication
Problem
Procedure or Intervention
Research Activity
Social History
Vital Sign

Ibuprofen p.o. (per oral)



No family history of colon cancer

Linguistic modifiers



Definition | Attributes | Synonyms

Select Synonyms

- [x] colon cancer
- [x] cancer of colon
- [x] colonic cancer
- [x] malignant neoplasm of colon
- [x] malignant colon neoplasm
- [x] malignant colonic tumor
- [x] cancer of the colon
- [x] colon cancers
- [x] ca colon
- [x] malignant colonic neoplasm
- [x] cancer, colon
- [x] cancer colons
- [x] cancers, colon
- [x] colon ca
- [x] colon tumor, malignant
- [x] malignant colon tumor
- [x] cancer colonic
- [x] colon neoplasm, malignant
- [x] colonic cancers
- [x] cancers, colonic

Research Questions We Have Addressed

Knowledge Authoring

- Which modifiers are important for interpreting clinical text?
- Does the modifier ontology work with the ConText algorithm for other languages?
- How well do humans agree when annotating modifiers?
- Can we mine text to learn terms used as modifiers?
- Can we assist a non-NLP expert user in developing a domain knowledge base?
- Can we learn synonyms from text?

IE-Viz

Information Extraction and Visualization

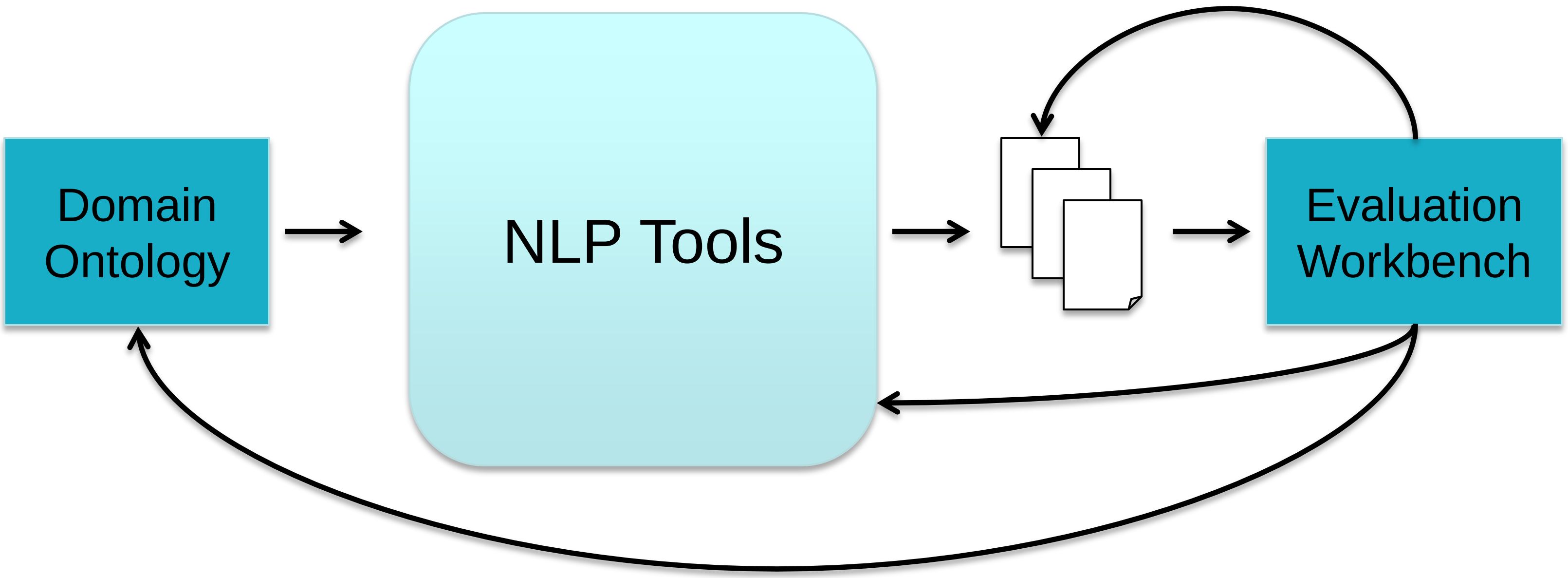
Knowledge Authoring

NLP Customizing

Classifier Developmen t

Visualization

NLP Customization

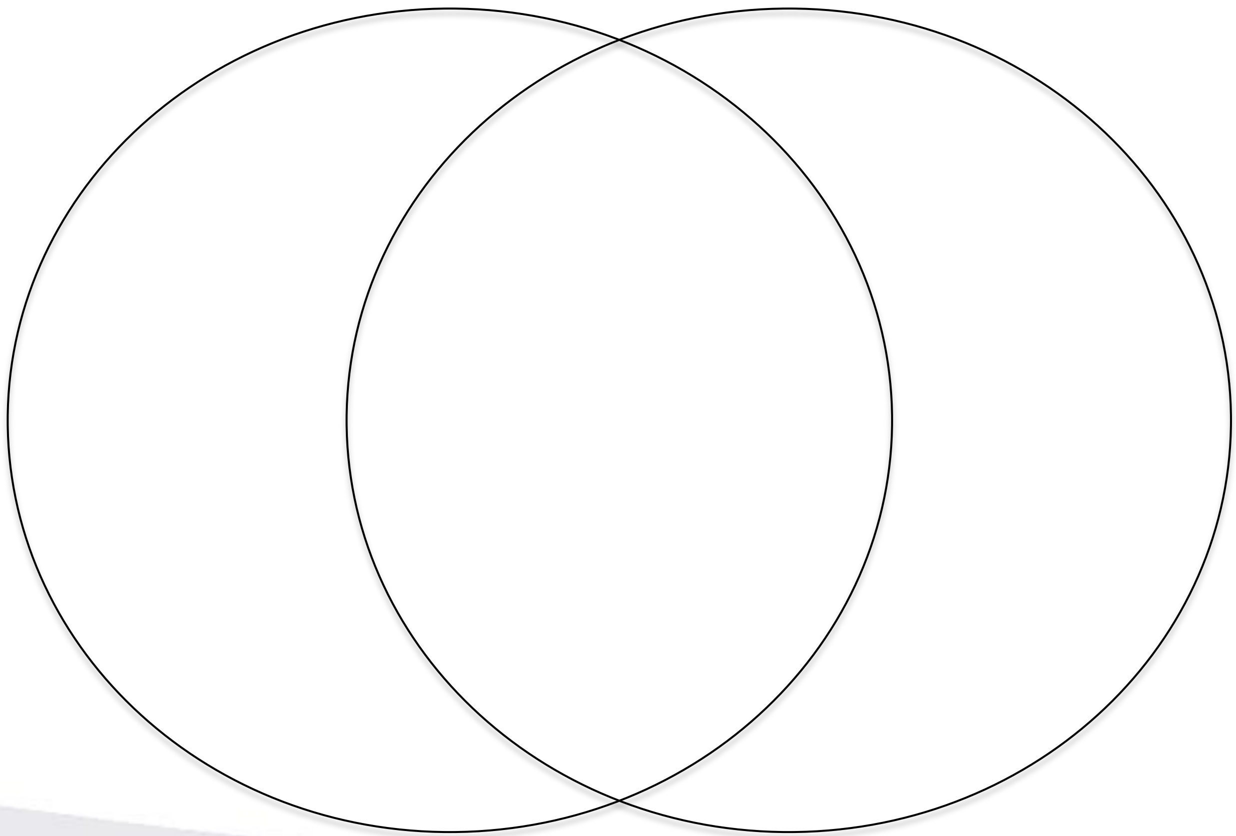


Use Case: Colonoscopy Quality Measures

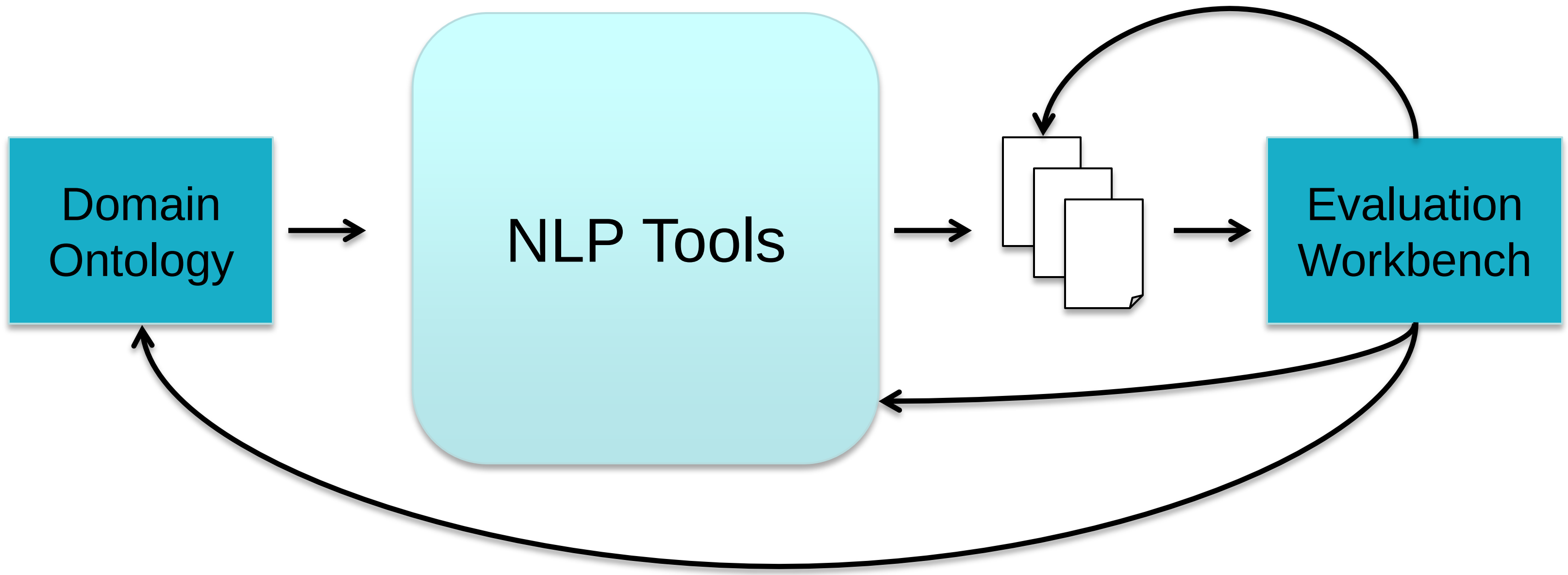
Samir Gupta
San Diego VA

Andrew Gawron
SLC VA

Overlapping Measures

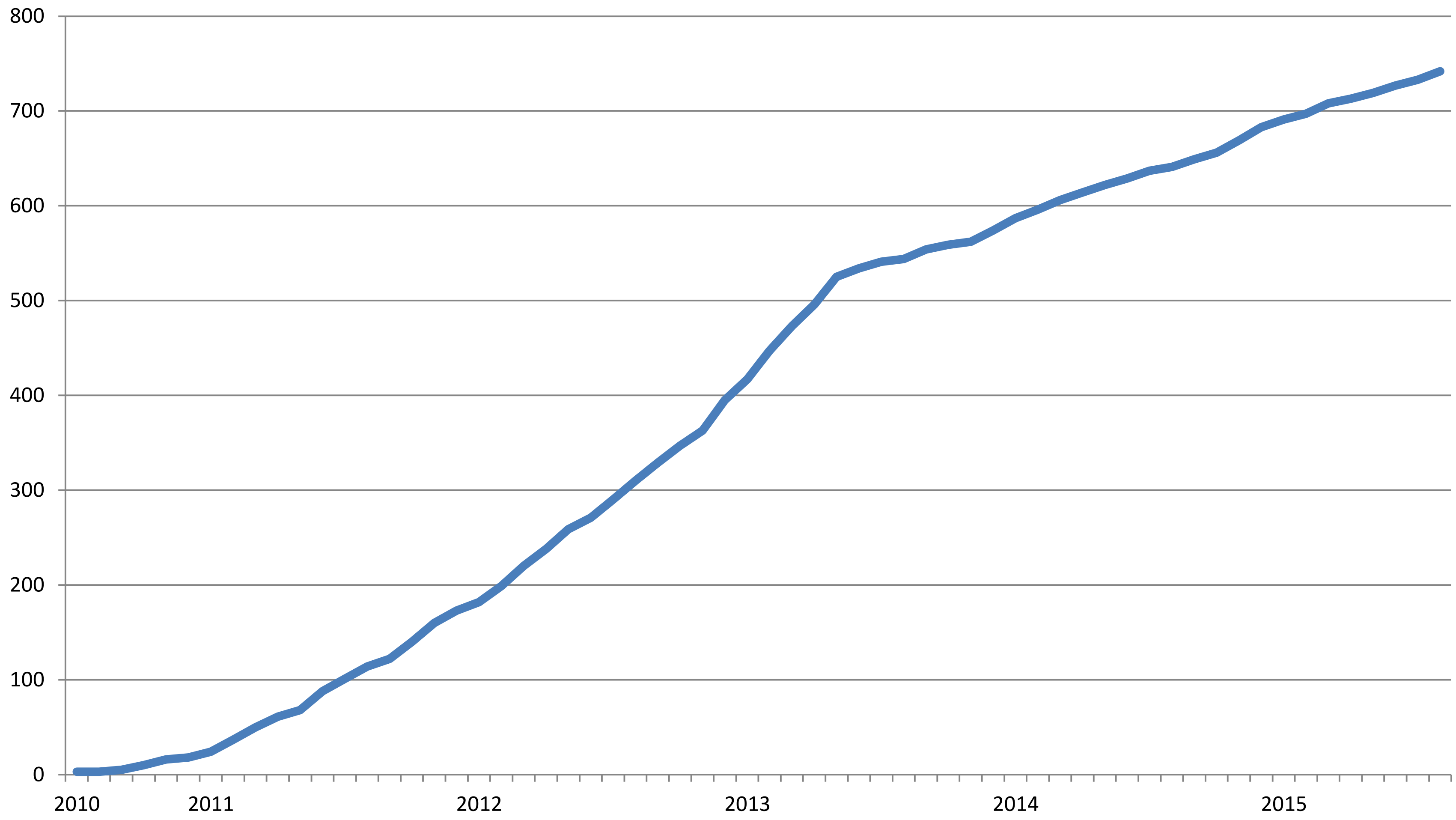


Collaboration



Shared knowledge representation enables collaboration

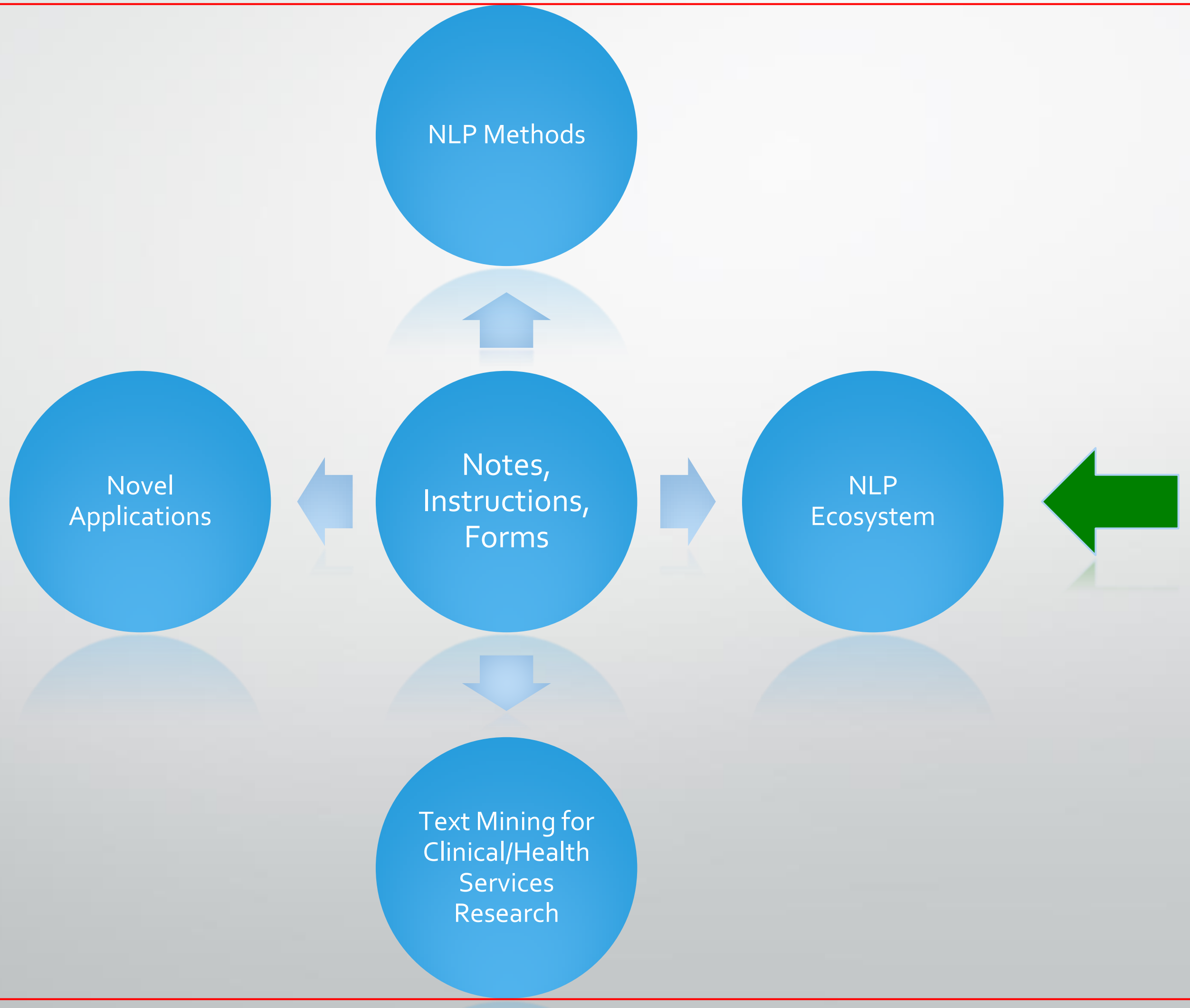
Cumulative Projects in VINCI using Text

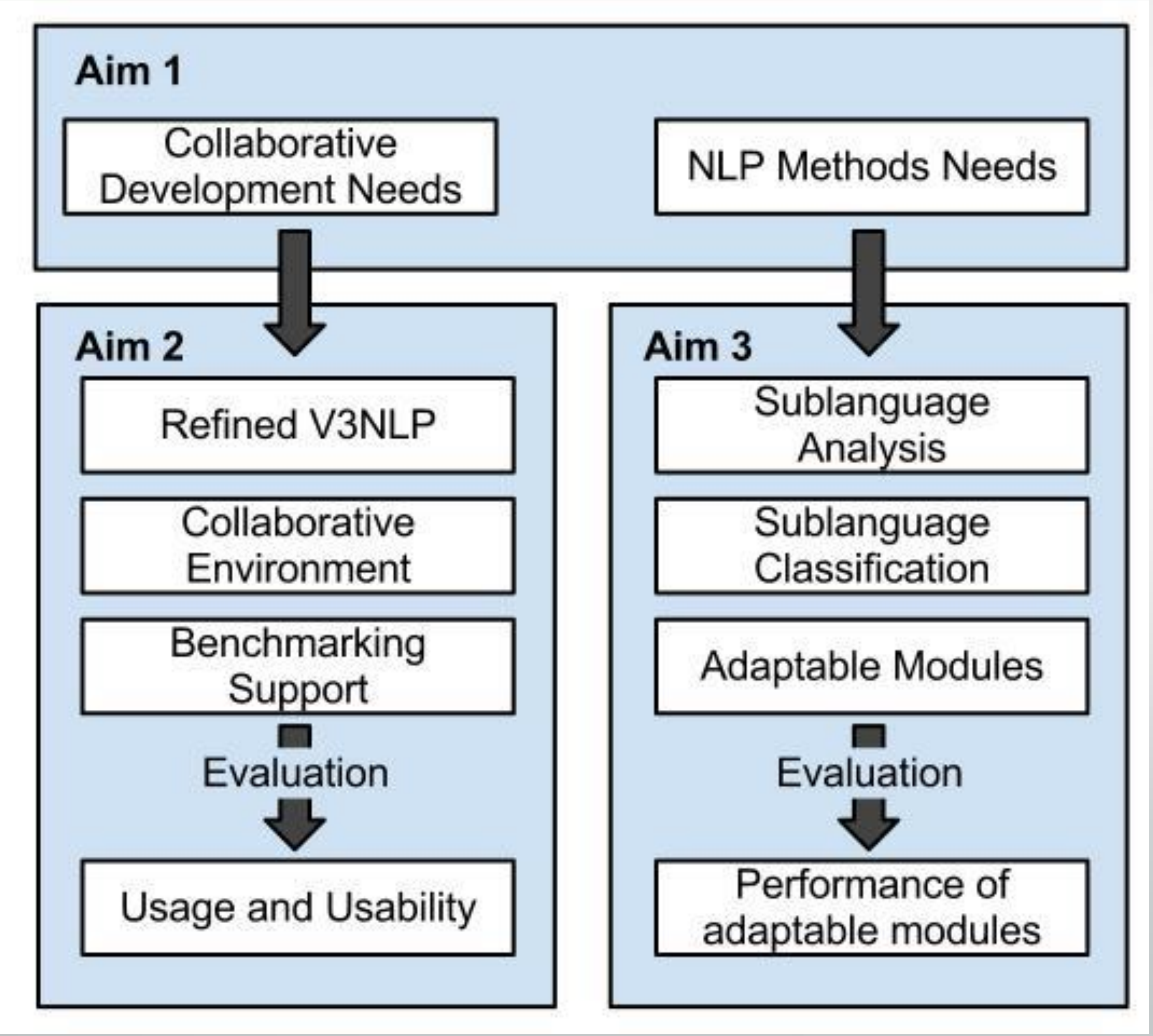


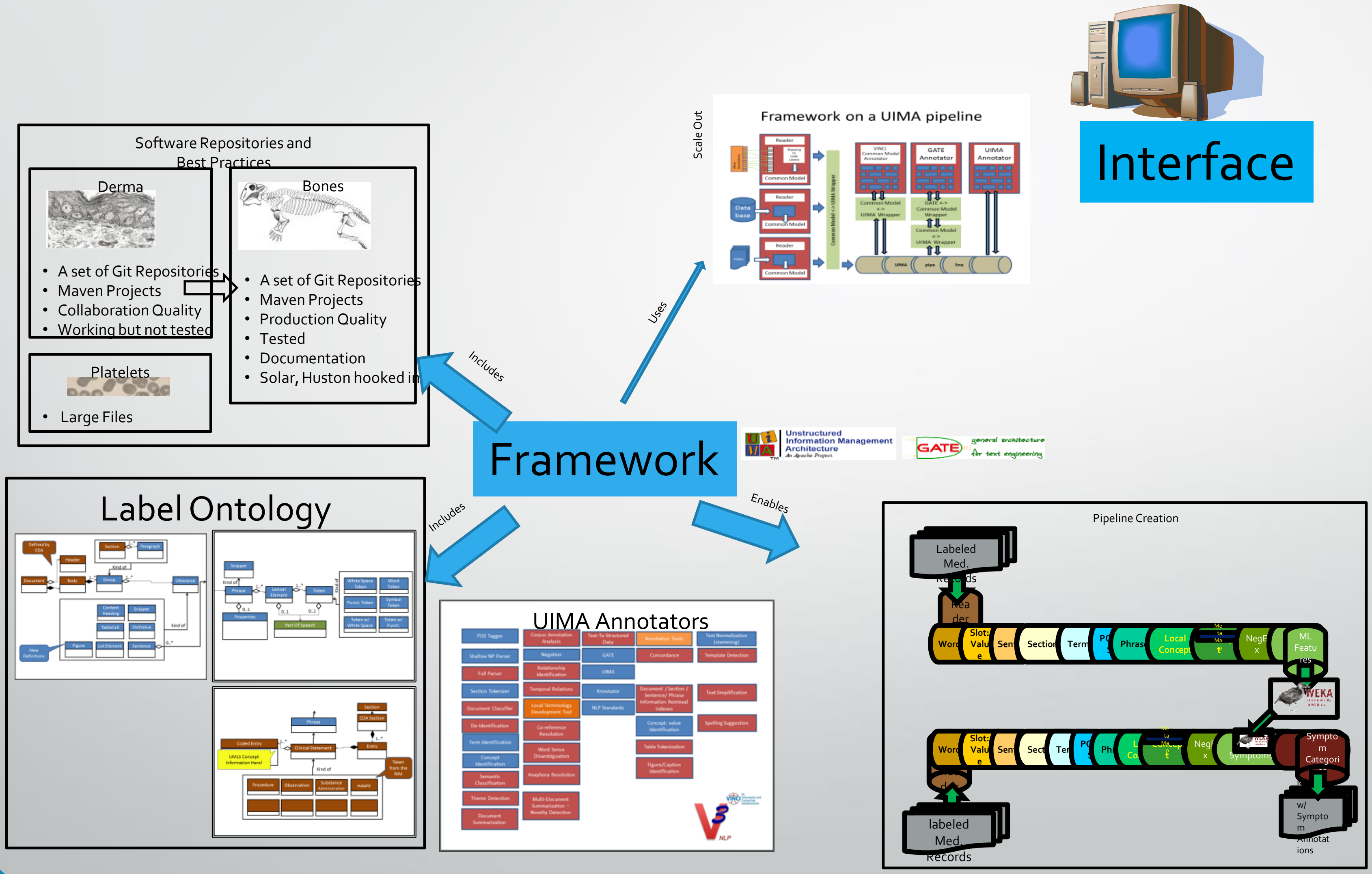
NLP Challenges

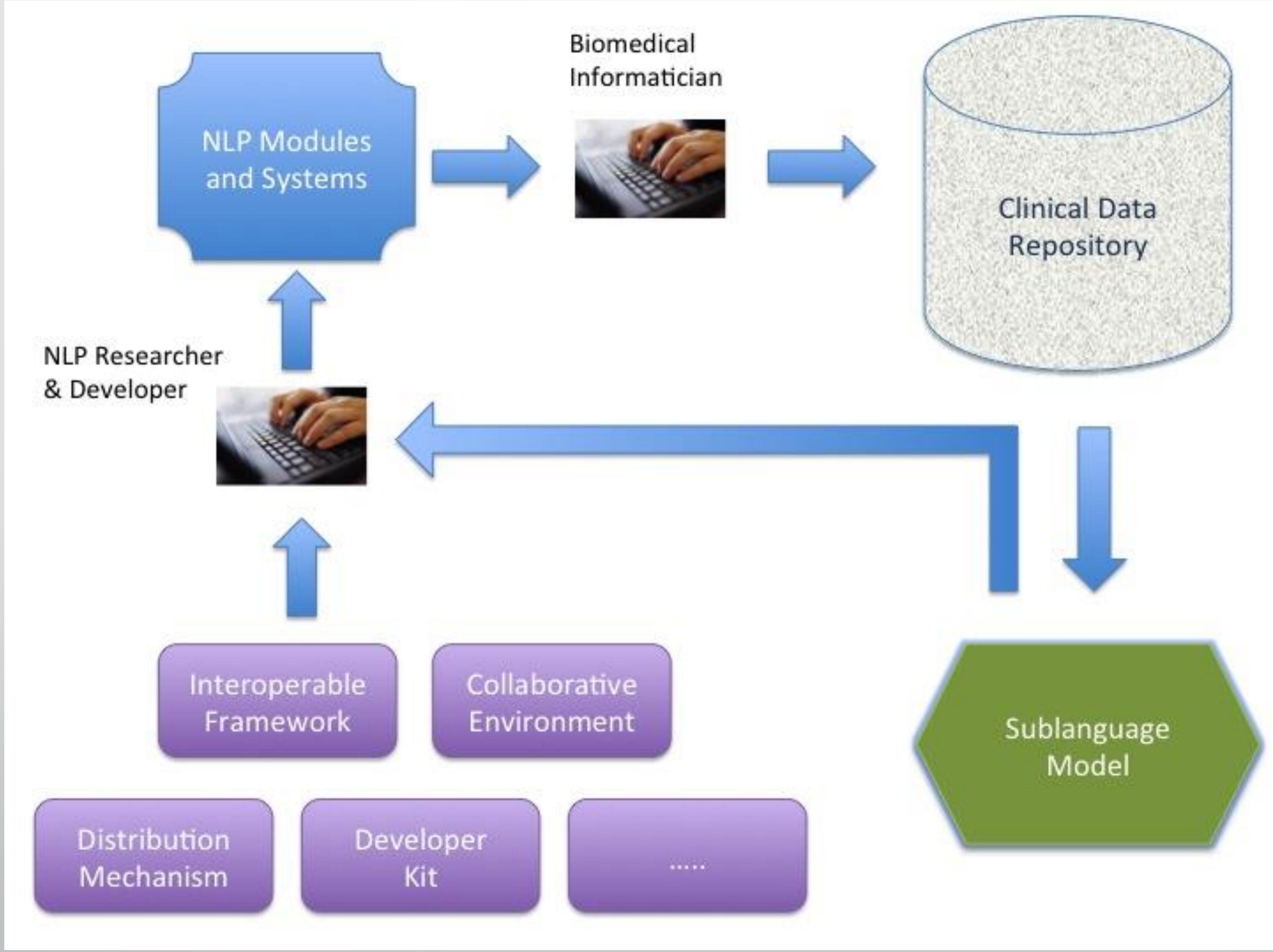
- Can't see the forest for the trees
  - Clinical NLP initially worked on local data in narrow domains but now needs to handle large and diverse data sets
- Every type of fruit requires a different ladder
  - Most use cases require tailoring and refinement of existing tools
  - Novel use cases emerge
- Fruits are not always low hanging
  - Some use cases are complex: stroke symptoms within 2 hours of ER admission or early signs dementia

My VA CLINICAL NLP Work




Aim 1
Collaborative Development Needs
NLP Methods Needs
Aim 2
Refined V3NLP
Collaborative Environment
Benchmarking Support
Evaluation
Usage and Usability
Aim 3
Sublanguage Analysis
Sublanguage Classification
Adaptable Modules
Evaluation
Performance of adaptable modules

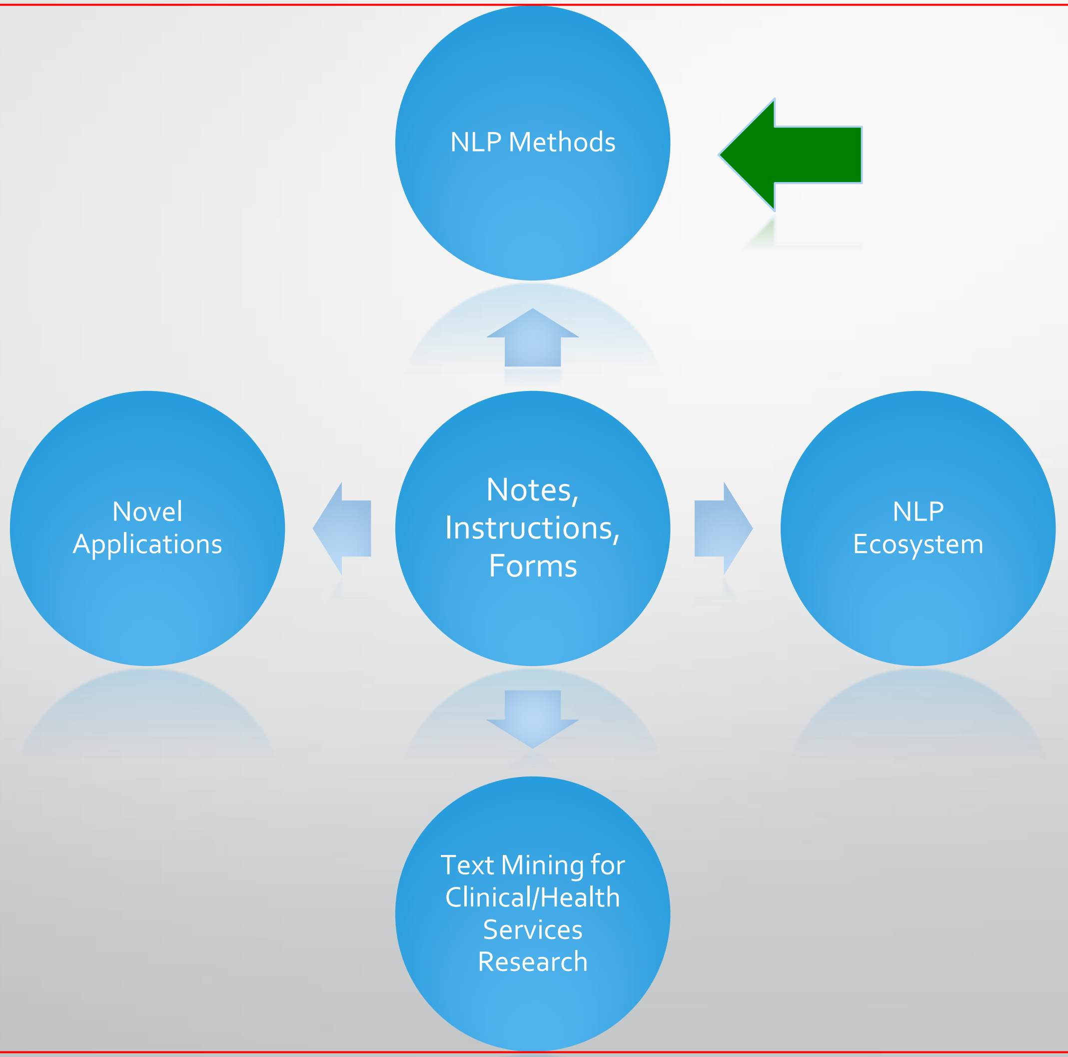

Interface
Framework on a UIMA pipeline
Scale Out
Uses
Framework
Unstructured Information Management Architecture
An Apache Project
GATE
general architecture for text engineering
Includes
Software Repositories and Best Practices
Derma
A set of Git Repositories
Maven Projects
Collaboration Quality
Working but not tested
Bones
A set of Git Repositories
Maven Projects
Production Quality
Tested
Documentation
Solar, Huston hooked in
Platelets
Large Files
Includes
Label Ontology
UIMA Annotators
Enables
Pipeline Creation
Labeled Med. Records
labeled Med. Records
Section
Local Concept
ML Features
WEKA
Symptom Categori
w/ Symptom Annotations


NLP Modules and Systems
Biomedical Informatician
Clinical Data Repository
NLP Researcher & Developer
Interoperable Framework
Collaborative Environment
Sublanguage Model
Distribution Mechanism
Developer Kit
.....

Progress So Far

- We have conducted interviews and workshops to assess strengths and weaknesses in current VA NLP development, and needs relating to ecosystem
- We created a prototype ecosystem cnlpecosystem.org
  - Sample data, Sample tools, VA NLP Bibilography
- Initial analysis of VA records indicate the existence of over a dozen sublanguage groups

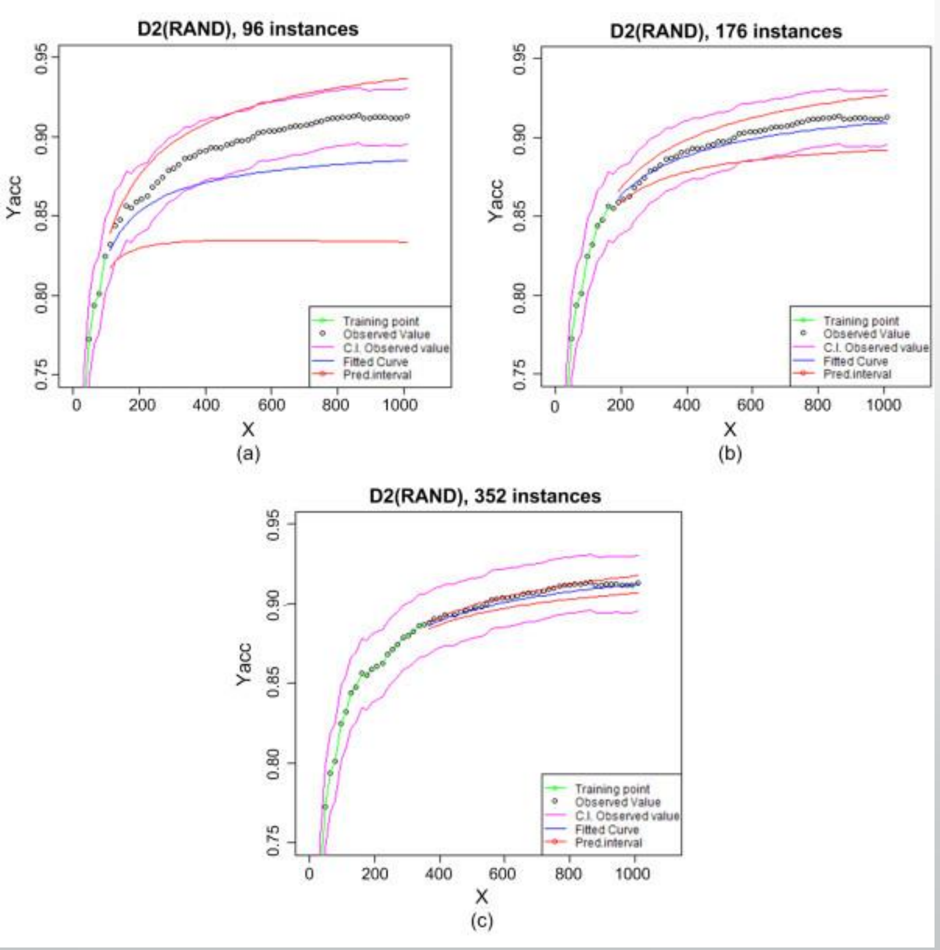
My VA CLINICAL NLP Work



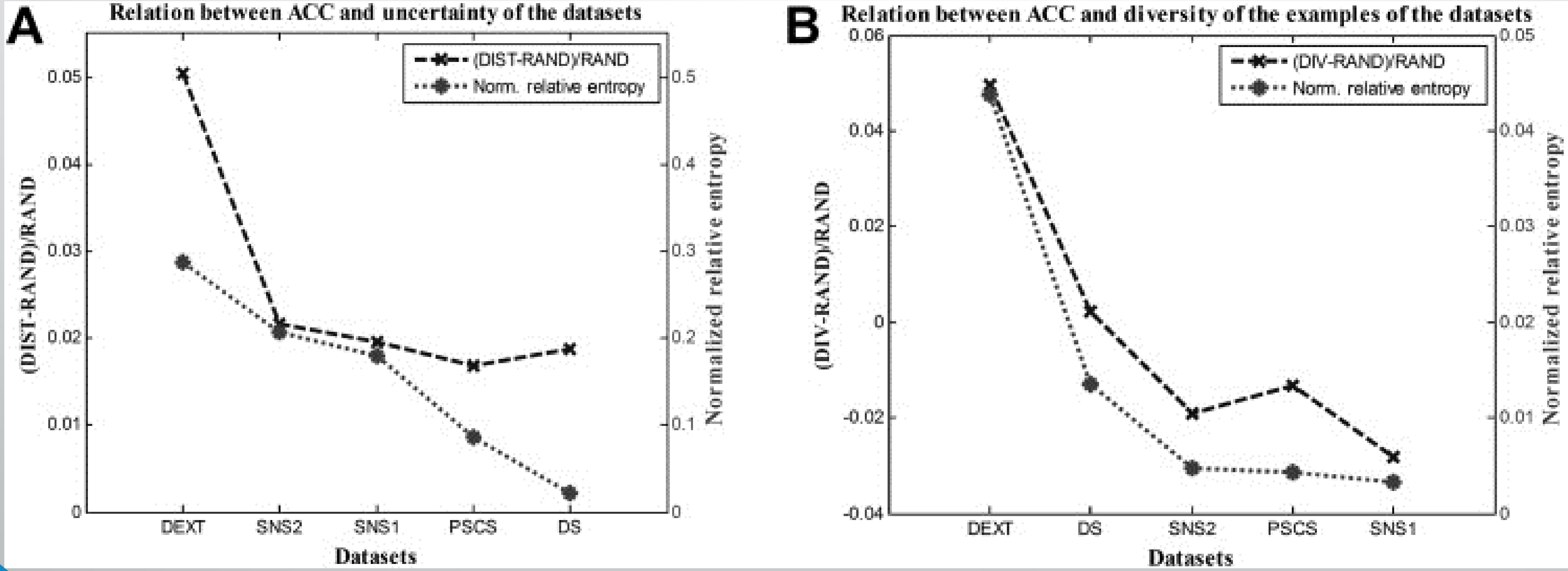
Novel Regular Expression Algorithm

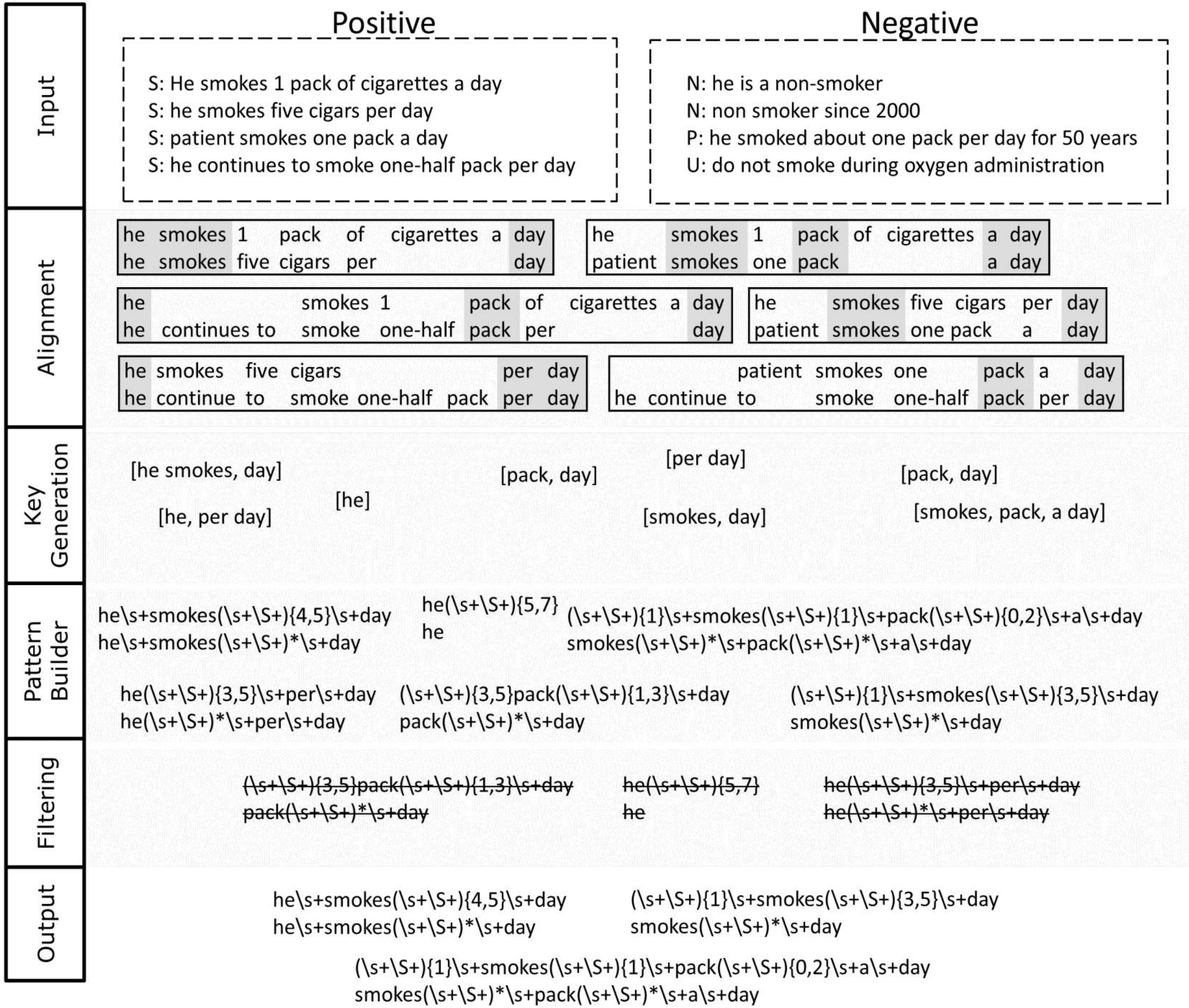
- Improve text classification and information extraction
  - REDCL for text classification
  - REDEX for information extraction

Sample Size Prediction

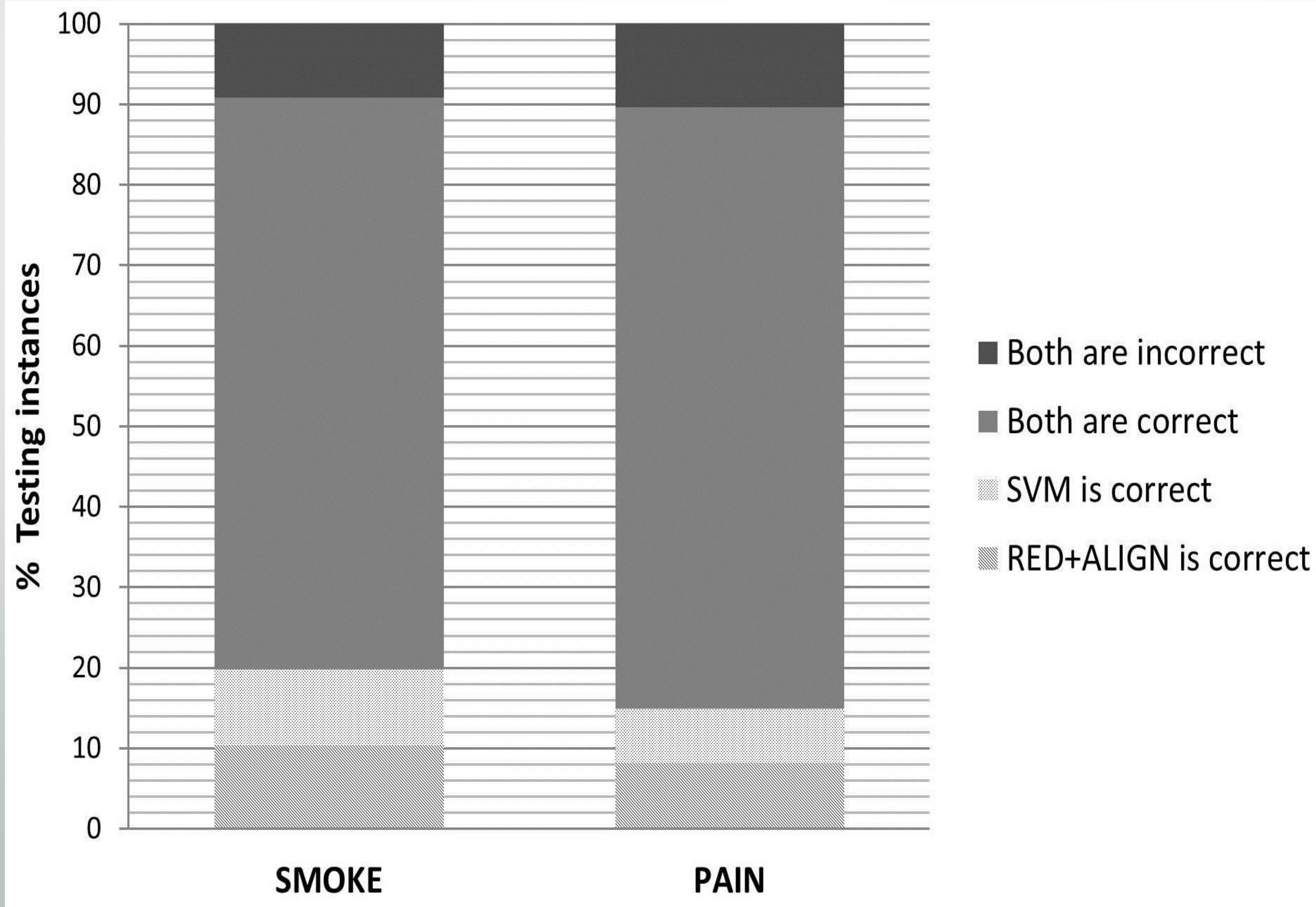


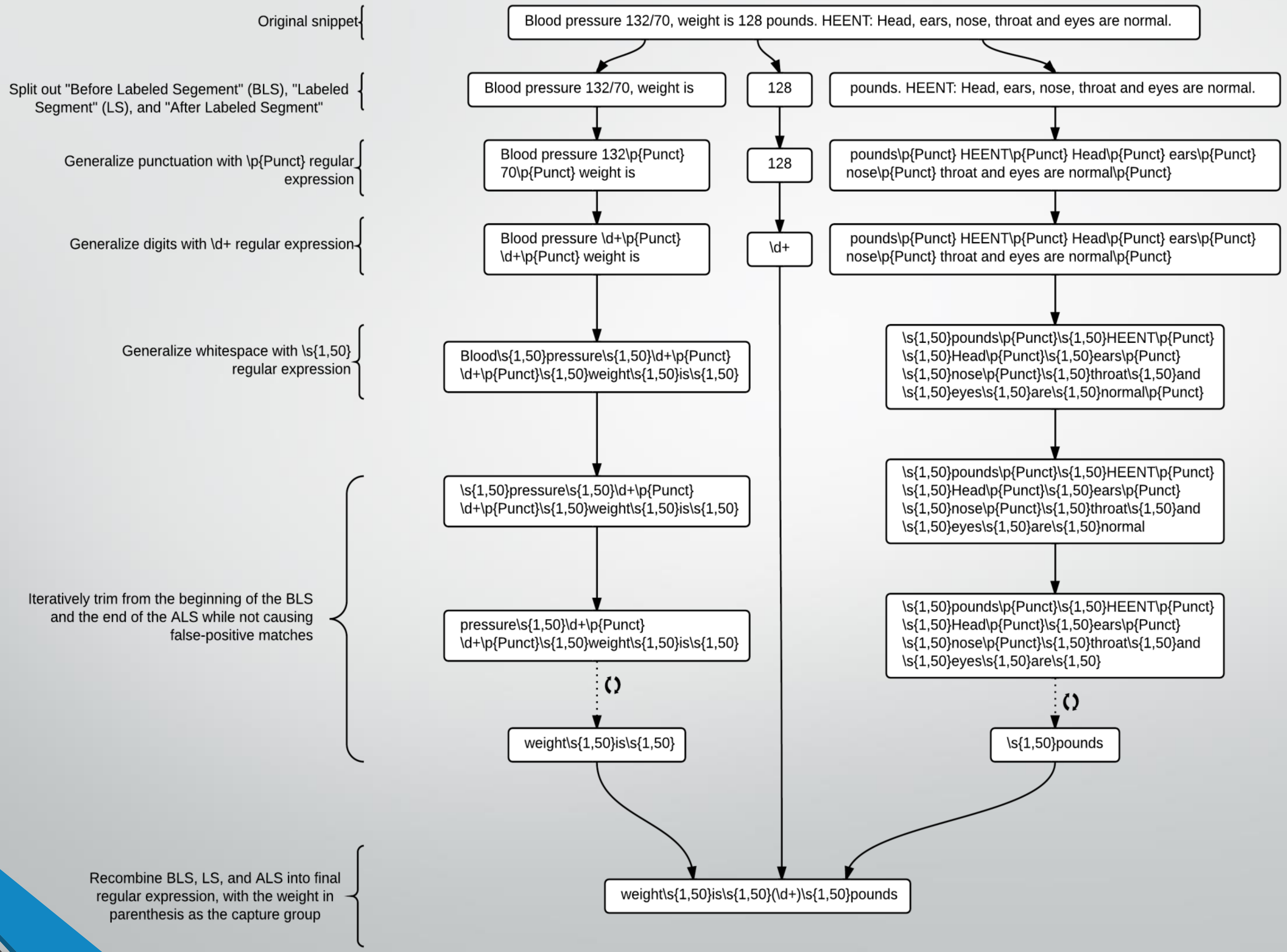
Active Learning




Input
Alignment
Key Generation
Pattern Builder
Filtering
Output
Positive
S: He smokes 1 pack of cigarettes a day
S: he smokes five cigars per day
S: patient smokes one pack a day
S: he continues to smoke one-half pack per day
Negative
N: he is a non-smoker
N: non smoker since 2000
P: he smoked about one pack per day for 50 years
U: do not smoke during oxygen administration
he smokes 1 pack of cigarettes a day
he smokes five cigars per day
he smokes 1 pack of cigarettes a day
patient smokes one pack a day
he smokes 1 pack of cigarettes a day
he continues to smoke one-half pack per day
he smokes five cigars per day
patient smokes one pack a day
he smokes five cigars per day
he continue to smoke one-half pack per day
patient smokes one pack a day
he continue to smoke one-half pack per day
[he smokes, day]
[he, per day]
[he]
[pack, day]
[per day]
[smokes, day]
[pack, day]
[smokes, pack, a day]
he\s+smokes(\s+\S+){4,5}\s+day
he\s+smokes(\s+\S+)*\s+day
he(\s+\S+){5,7}
he
(\s+\S+){1}\s+smokes(\s+\S+){1}\s+pack(\s+\S+){0,2}\s+a\s+day
smokes(\s+\S+)*\s+pack(\s+\S+)*\s+a\s+day
he(\s+\S+){3,5}\s+per\s+day
he(\s+\S+)*\s+per\s+day
(\s+\S+){3,5}pack(\s+\S+){1,3}\s+day
pack(\s+\S+)*\s+day
(\s+\S+){1}\s+smokes(\s+\S+){3,5}\s+day
smokes(\s+\S+)*\s+day
~~(\s+\S+){3,5}pack(\s+\S+){1,3}\s+day~~
~~pack(\s+\S+)*\s+day~~
~~he(\s+\S+){5,7}~~
~~he~~
~~he(\s+\S+){3,5}\s+per\s+day~~
~~he(\s+\S+)*\s+per\s+day~~
he\s+smokes(\s+\S+){4,5}\s+day
he\s+smokes(\s+\S+)*\s+day
(\s+\S+){1}\s+smokes(\s+\S+){3,5}\s+day
smokes(\s+\S+)*\s+day
(\s+\S+){1}\s+smokes(\s+\S+){1}\s+pack(\s+\S+){0,2}\s+a\s+day
smokes(\s+\S+)*\s+pack(\s+\S+)*\s+a\s+day

Comparison with SVM



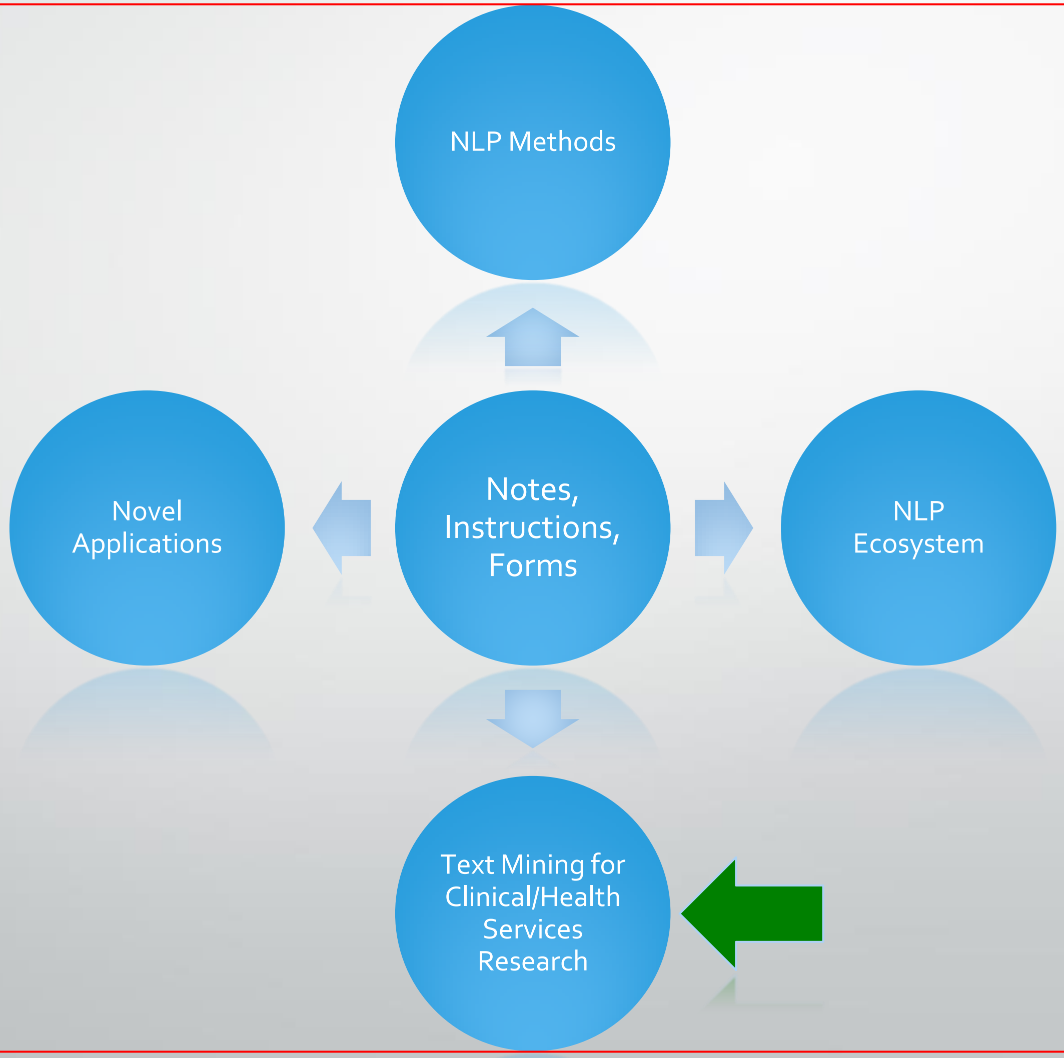


Original snippet
Blood pressure 132/70, weight is 128 pounds. HEENT: Head, ears, nose, throat and eyes are normal.
Split out "Before Labeled Segement" (BLS), "Labeled Segment" (LS), and "After Labeled Segment"
Blood pressure 132/70, weight is
128
pounds. HEENT: Head, ears, nose, throat and eyes are normal.
Generalize punctuation with \p{Punct} regular expression
Blood pressure 132\p{Punct} 70\p{Punct} weight is
128
pounds\p{Punct} HEENT\p{Punct} Head\p{Punct} ears\p{Punct} nose\p{Punct} throat and eyes are normal\p{Punct}
Generalize digits with \d+ regular expression
Blood pressure \d+\p{Punct} \d+\p{Punct} weight is
\d+
pounds\p{Punct} HEENT\p{Punct} Head\p{Punct} ears\p{Punct} nose\p{Punct} throat and eyes are normal\p{Punct}
Generalize whitespace with \s{1,50} regular expression
Blood\s{1,50}pressure\s{1,50}\d+\p{Punct} \d+\p{Punct}\s{1,50}weight\s{1,50}is\s{1,50}
\s{1,50}pounds\p{Punct}\s{1,50}HEENT\p{Punct} \s{1,50}Head\p{Punct}\s{1,50}ears\p{Punct} \s{1,50}nose\p{Punct}\s{1,50}throat\s{1,50}and \s{1,50}eyes\s{1,50}are\s{1,50}normal\p{Punct}
Iteratively trim from the beginning of the BLS and the end of the ALS while not causing false-positive matches
\s{1,50}pressure\s{1,50}\d+\p{Punct} \d+\p{Punct}\s{1,50}weight\s{1,50}is\s{1,50}
\s{1,50}pounds\p{Punct}\s{1,50}HEENT\p{Punct} \s{1,50}Head\p{Punct}\s{1,50}ears\p{Punct} \s{1,50}nose\p{Punct}\s{1,50}throat\s{1,50}and \s{1,50}eyes\s{1,50}are\s{1,50}normal
pressure\s{1,50}\d+\p{Punct} \d+\p{Punct}\s{1,50}weight\s{1,50}is\s{1,50}
\s{1,50}pounds\p{Punct}\s{1,50}HEENT\p{Punct} \s{1,50}Head\p{Punct}\s{1,50}ears\p{Punct} \s{1,50}nose\p{Punct}\s{1,50}throat\s{1,50}and \s{1,50}eyes\s{1,50}are\s{1,50}
weight\s{1,50}is\s{1,50}
\s{1,50}pounds
Recombine BLS, LS, and ALS into final regular expression, with the weight in parenthesis as the capture group
weight\s{1,50}is\s{1,50}(\d+)\s{1,50}pounds

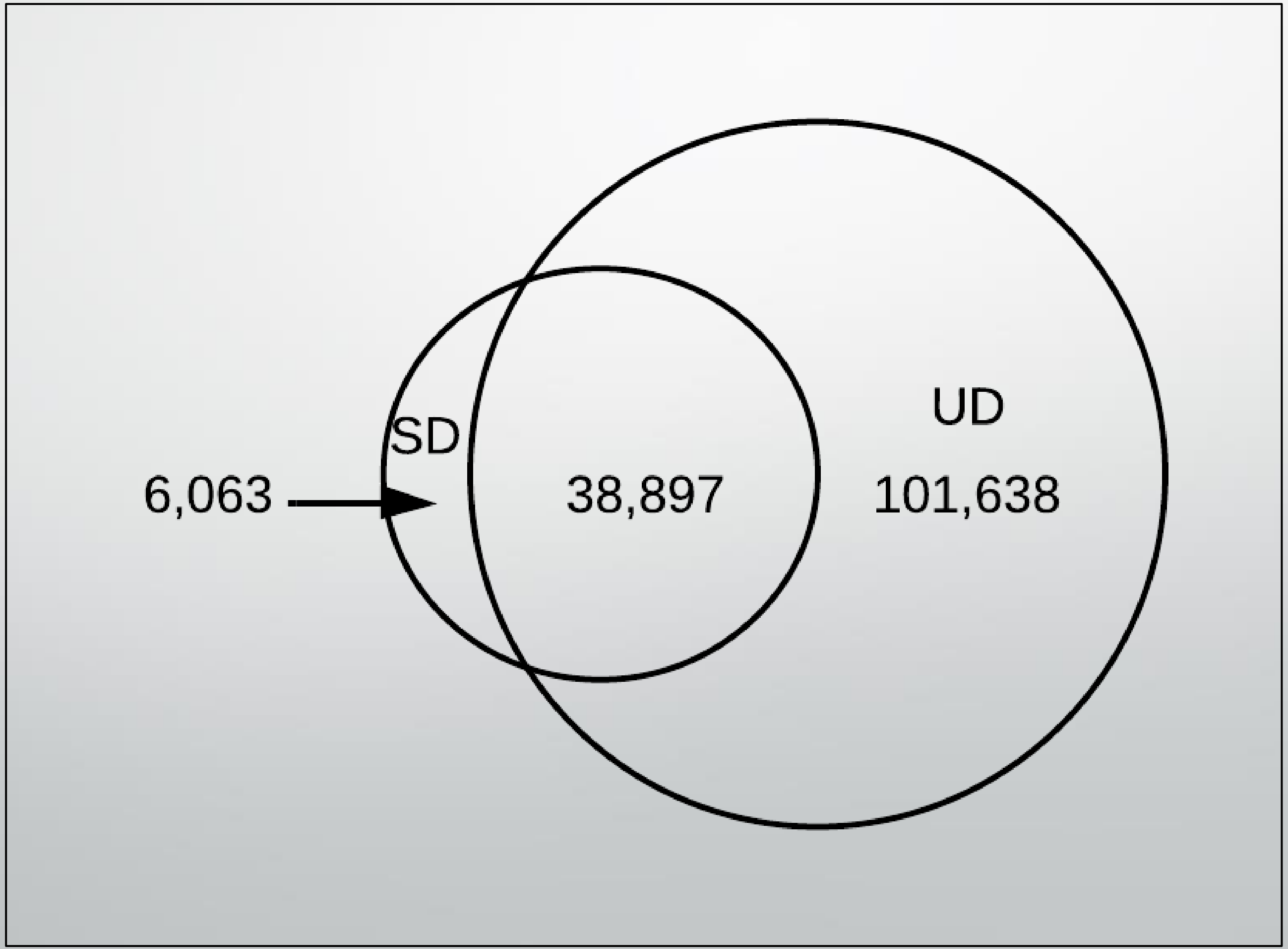
Sample Results

	Weight	METs	Temporal				KATZ	Pain
			Date	Time	Duration	Set		
Precision	0.99	0.89	0.97	0.98	0.93	1.00	0.99	0.88
Recall	0.98	0.92	0.97	0.97	0.96	0.83	1.00	0.97
F1-score	0.99	0.90	0.97	0.98	0.95	0.91	0.99	0.92
Accuracy	0.98	0.91	0.94	0.95	0.90	0.83	0.99	0.92
# Snippets	968	2701	493	289	169	18	1000	3862
# Reg Ex	98	812	128	54	41	7	35	572

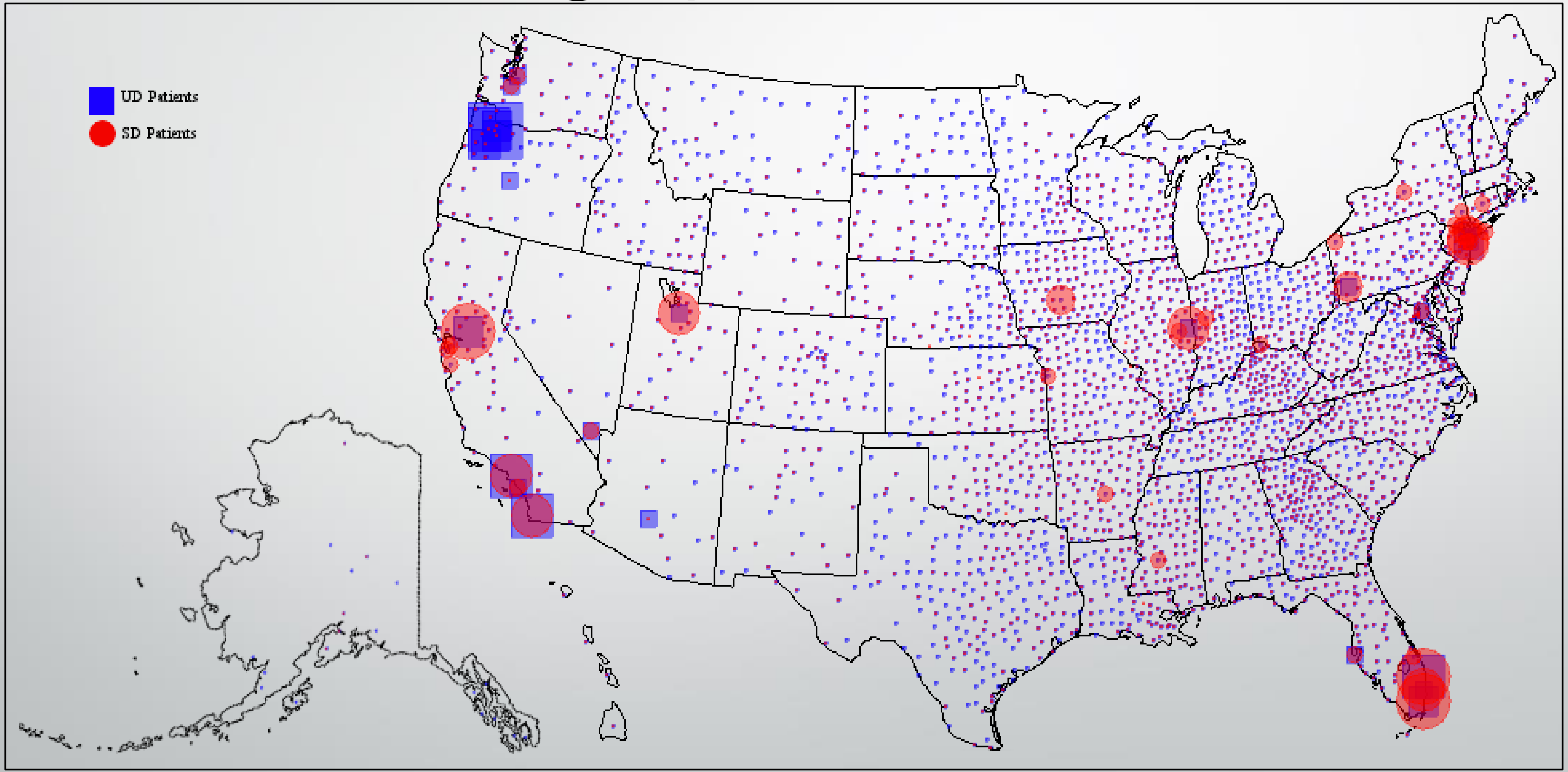
My WORK IN CLINICAL NLP



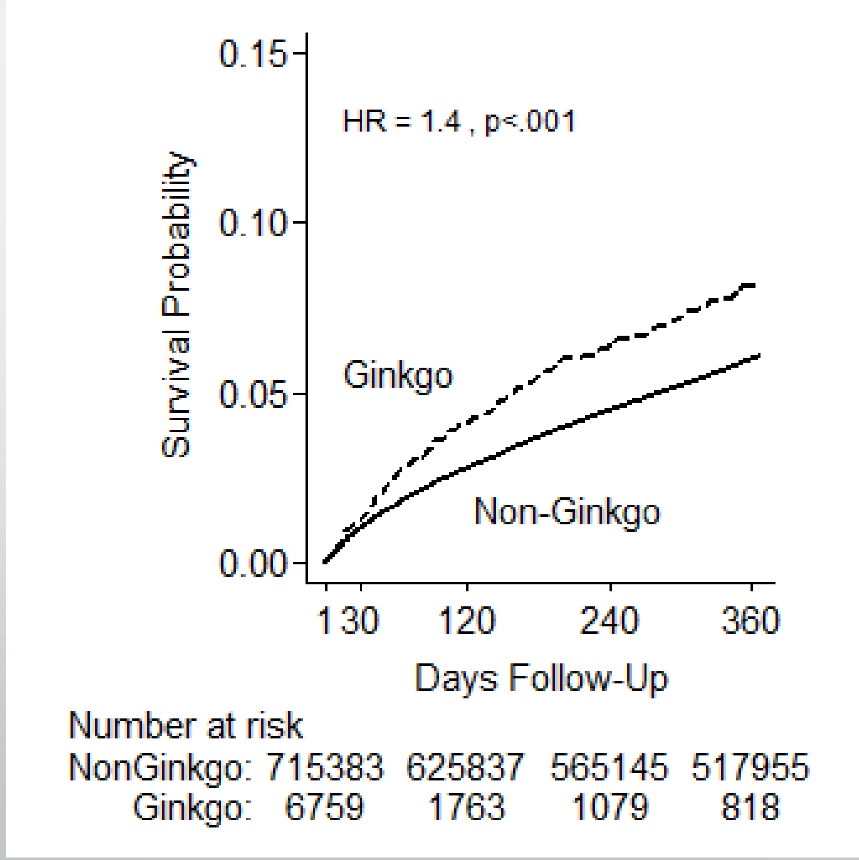
Cohort Sizes and Overlap



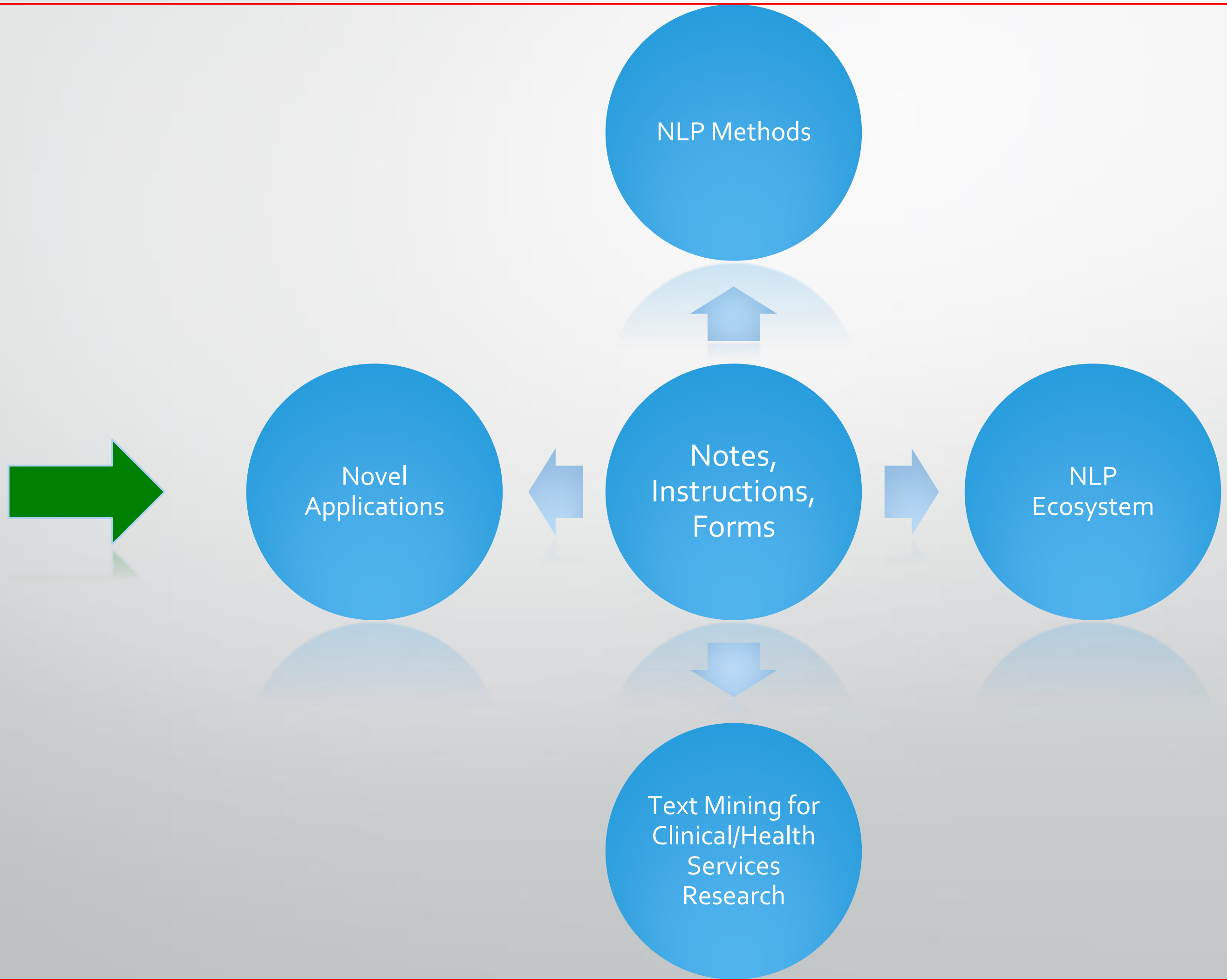
Geographic distribution



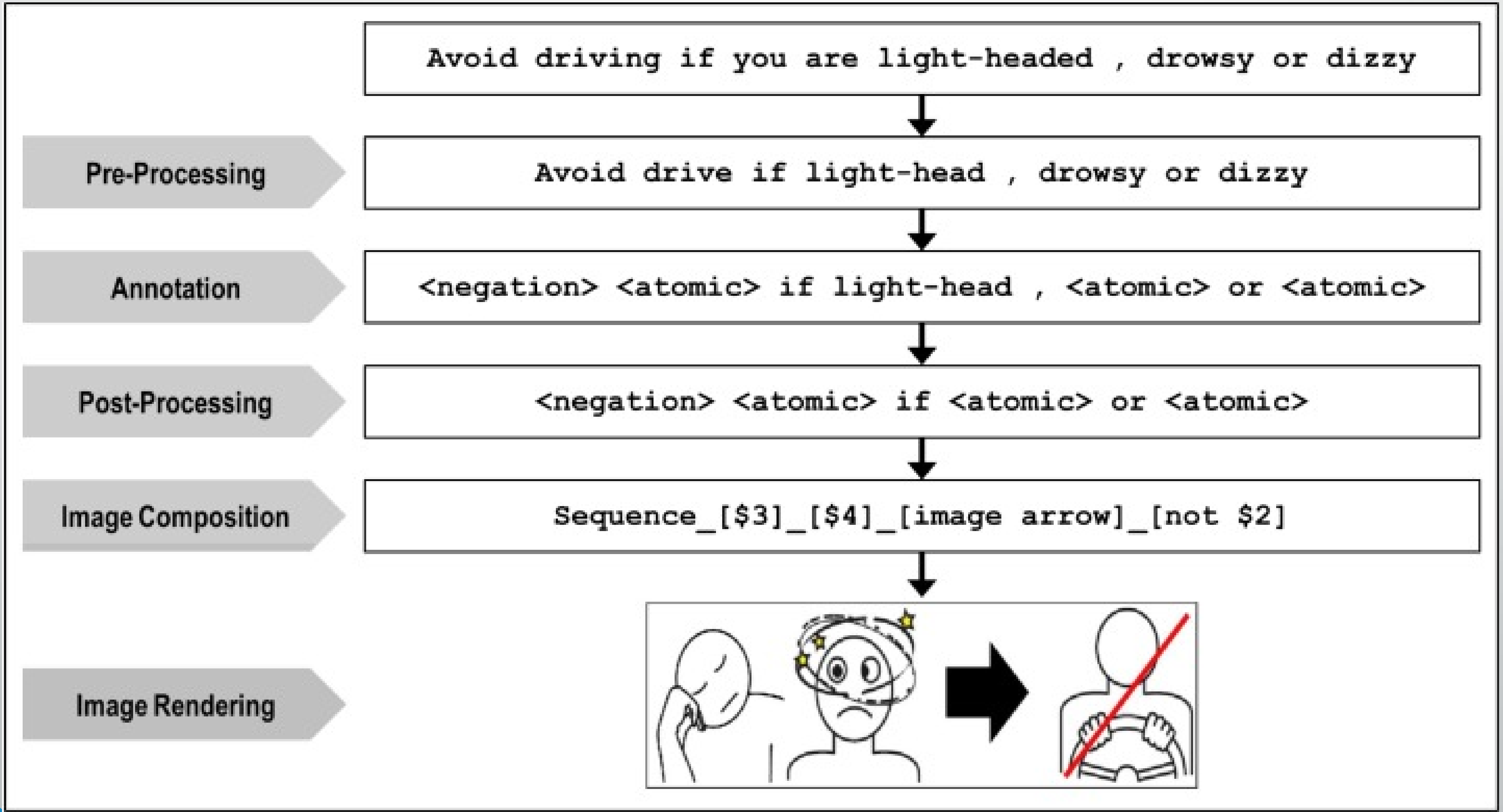
Gingko-Warfarin Bleeding Risk

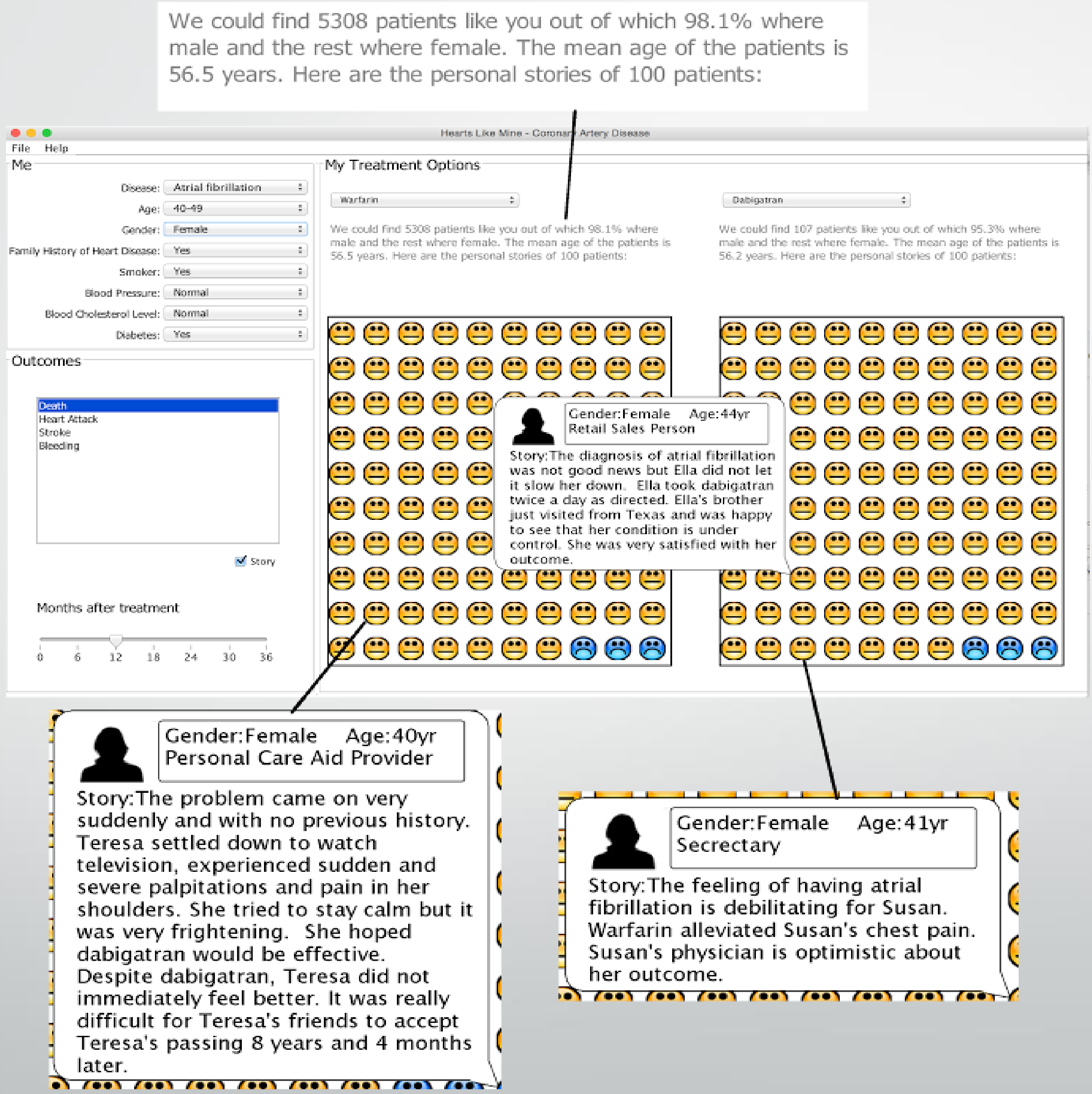


My WORK IN CLINICAL NLP

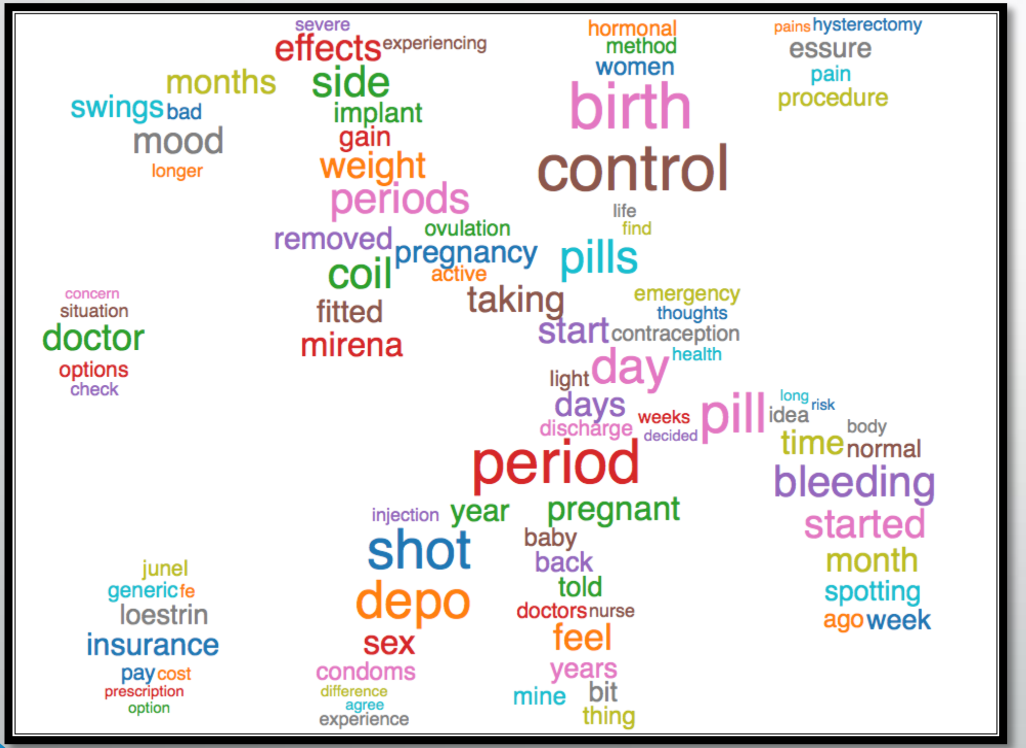


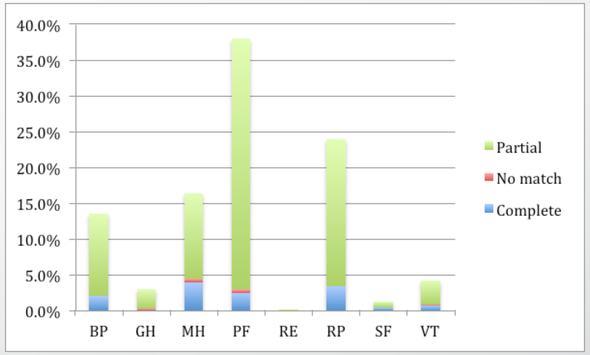
Example

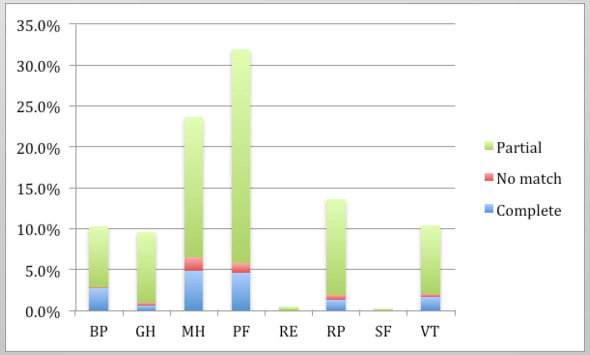

Avoid driving if you are light-headed , drowsy or dizzy
Pre-Processing
Avoid drive if light-head , drowsy or dizzy
Annotation
<negation> <atomic> if light-head , <atomic> or <atomic>
Post-Processing
<negation> <atomic> if <atomic> or <atomic>
Image Composition
Sequence_[$3]_[$4]_[image arrow]_[not $2]
Image Rendering


We could find 5308 patients like you out of which 98.1% where male and the rest where female. The mean age of the patients is 56.5 years. Here are the personal stories of 100 patients:
Hearts Like Mine - Coronary Artery Disease
File Help
Me
Disease: Atrial fibrillation
Age: 40-49
Gender: Female
Family History of Heart Disease: Yes
Smoker: Yes
Blood Pressure: Normal
Blood Cholesterol Level: Normal
Diabetes: Yes
Outcomes
Death
Heart Attack
Stroke
Bleeding
Story
Months after treatment
0 6 12 18 24 30 36
My Treatment Options
Warfarin
We could find 5308 patients like you out of which 98.1% where male and the rest where female. The mean age of the patients is 56.5 years. Here are the personal stories of 100 patients:
Dabigatran
We could find 107 patients like you out of which 95.3% where male and the rest where female. The mean age of the patients is 56.2 years. Here are the personal stories of 100 patients:
Gender:Female Age:44yr
Retail Sales Person
Story:The diagnosis of atrial fibrillation was not good news but Ella did not let it slow her down. Ella took dabigatran twice a day as directed. Ella's brother just visited from Texas and was happy to see that her condition is under control. She was very satisfied with her outcome.
Gender:Female Age:40yr
Personal Care Aid Provider
Story:The problem came on very suddenly and with no previous history. Teresa settled down to watch television, experienced sudden and severe palpitations and pain in her shoulders. She tried to stay calm but it was very frightening. She hoped dabigatran would be effective. Despite dabigatran, Teresa did not immediately feel better. It was really difficult for Teresa's friends to accept Teresa's passing 8 years and 4 months later.
Gender:Female Age:41yr
Secrectary
Story:The feeling of having atrial fibrillation is debilitating for Susan. Warfarin alleviated Susan's chest pain. Susan's physician is optimistic about her outcome.

Patient Information Needs & Patient Centered Outcome







Discussion

- We have developed new methods, tools, and applications.
- Our research has enabled clinical/health services research and led to knowledge discovery
- There is a lot of room for improvement
- We have not been incorporated into VA operation, but hope to…

CHOIR
Center for Healthcare Organization and Implementation Research

Patient-Centered Research

Systems for Helping Veterans Comprehend Electronic Health Record Notes

Hong Yu[1,2], Balaji Polepalli Ramesh[1,2], Jiaping Zheng[3], Jinying Chen[2], Louise Maranda[2], Cynthia Brandt[4,5], Kathleen Mazor[2], Donna Zulman[6,7], Thomas Houston[1,2]

[1]VA Bedford VAMC

[2]University of Massachusetts Medical School

[3]University of Massachusetts, Amherst

[3]VA Connecticut Health Care System

[4]Yale University

[5]VA Palo Alto Health Care System

[6]Stanford University

Background

- Patients reading their EHR notes has the potential to
  - ☐Enhance medical understanding
  - ☐Improve healthcare management and outcomes
- **Blue Button**: The Department of Veterans Affairs (VA) Blue Button enables Veterans to view, print, and download their EHRs, including clinical notes (e.g., progress notes).

The Challenge

- Physicians' notes are difficult to comprehend (Keselman et al)
- Many Veterans have limited health literacy (Schapira et al)

"The patient will be scheduled for a repeat **EGD** in one year for surveillance purposes of **Barrett's esophagus**. From a **GI** standpoint, we recommend to proceed with **bariatric surgery**. However, he will need to continue daily **PPI** administration to maximize acid reduction. Otherwise, there are no additional recommendations. The patient was treated with **myocardial infarction**."

The NoteAid System

- A system for helping patients comprehend electronic health record notes
- Automatically
  - Identifies clinically relevant concepts
  - Links concepts to their definitions and lay language
  - Links notes to other education material
- Funded by HSR&D (1I01HX001457) since May, 2015

Research

- Lay language resources and education materials
- NLP for translating EHR to lay language and linking EHR to education materials

Lay Language Resources

- Existing resources
  - Consumer Health Vocabulary (Zeng et al)
  - MedlinePlus, etc
  - However, we found 40%~60% EHR jargon do not appear in existing resources

Mining Lay Language from Wikipedia

Diabetes mellitus (**DM**), commonly referred to as **diabetes**, is a group of metabolic diseases in which there are high blood sugar levels over a prolonged period.[2] Symptoms of high blood sugar include frequent urination, $\boxed{\text{increased thirst}}$ and increased hunger. If left untreated, diabetes can cause many complications.[3] Acute complications include diabetic ketoacidosis and nonketotic hyperosmolar coma.[4] Serious long-term complications include cardiovascular disease, stroke, chronic kidney failure, foot ulcers, and damage to the eyes.[3]

Polydipsia

From Wikipedia, the free encyclopedia

For the term formerly used in reference to compulsive drinking of alcohol, see Dipsomania.

See also: Polydipsia in birds

Polydipsia is excessive thirst.[1] The word derives from the Greek *πολυδιψία*,[2] which is derived from *πολύς* (*polys*, "much, many") + *δίψα* (*dipsa*, "thirst"). An etymologically related term is dipsomaniac, meaning an alcoholic. Polydipsia is a nonspecific symptom in various medical disorders. It also occurs as an abnormal behaviour in animals.[3]

Methods, Evaluation, Results

- Entity Link Frequency (baseline)
- Similarity (word embedding)
- Wiki Pseudo Relevance Feedback
- Similarity+PRF

Methods	*MAP (Relaxed condition)*	*MAP (strict condition)*
ELF	0.6267	0.2401
ACS	0.6624	0.2383
PRF	0.6859	0.2519
AveR	0.6685	0.2433
RSC	**0.6900**	**0.2745**

Table 1: Mean Average Precision values for Relevance Feedback of 5

Prioritize Important EHR Content

EHR notes are frequently long and full of jargon, overwhelming patients.

The other possible etiologies for her **nephrotic syndrome** could be **FSGS**, **myeloma**, or an **interstitial nephritis**. I will send for urine eosinophil. I will also send for myeloma workup with a urine immunofixation electrophoresis.

Methods

Using CHV terms and CHV familiarity scores to train supervised machine learning models (e.g., SVM regression) to predict EHR term importance and term familiarity

	Ave. Precision (%)
TF*IDF	8.83
C-Value	10.23
Ranking by importance (UMLS)	19.42
Ranking by importance + unfamiliarity	23.31

Linking EHR to Education Materials

Patient remains in ICU with the following problems: Respiratory failure, hemodynamics, renal failure, status post liver transplant, atrial fib, infectious disease, nutrition.

Education Material:

Respiratory Failure
Deep Vein Thrombosis
Aspiration pneumonia
Pulmonary Hypertension
Kidney Failure
Atrial Fibrillation or Flutter
Liver Transplantation
Dialysis - Hemodialysis

System	P@10	MAP	Increase
Baseline	0	0.0091	-
CHV	5	0.0240	2.6
LDA	10	0.0489	5.4
Key (Wiki)	16	0.0851	9.4
Key (EHR)	16.5	0.0879	9.7
Key (Wiki+EHR)	18	0.1030	11.3
Instance Pruning	9.5	0.0316	3.5
Feature Augmentation	14	0.0684	7.5
Instance Weighting	21.5	0.1111	12.2

	Training Data					
	Wiki	EHR	Wiki+EHR	IW	IP	FA
Precision	16.27	35.92	33.79	47.59	40.00	46.60
Recall	26.88	34.09	33.18	34.41	6.02	28.86
F1	18.74	33.70	32.54	38.32	10.23	34.08

Evaluation 1: Self-Reported Comprehension

- 25 Lay people
- De-identified notes (not own notes)
- With or without NoteAid
- Self-reported comprehension
  - 1 to 5
  - 1: impossible to understand
  - 5: understand completely

Results

- Average self-reported comprehension scores

System	Note Alone	MedlinePlus	UMLS	Wiki	Hybrid
Score	2.95±0.67	4.12±0.33*	3.63±0.57*	3.85±0.47*	3.92±0.40*

*p<0.01, Non-parametric Mann-Whitney Wilcoxon signed-rank test

Evaluation 2: Paraphrasing

- **Data**: 5 de-identified progress notes
- **Subjects**: 40 subjects
- **Evaluation Process**:
  – Each subject was presented with 5 notes, one at a time, either with NoteAid (hybrid), or without.
  – Each subject paraphrased the main content of the note

Examples: Without NoteAid

- **Note** : The patient is doing okay today. He has some complaints of discomfort around his tracheostomy site but otherwise has no complaints of chest pain, shortness of breath, abdominal pain, nausea, or vomiting. OBJECTIVE : Vital Signs : Currently stable. Cardiac : Normal S1, S2. No murmurs, rubs, or gallops are present. Lungs : Clear to auscultation bilaterally. Abdomen : Soft, nontender, and nondistended. Normoactive bowel sounds are present. LABS : There are no new labs from today. ASSESSMENT AND PLAN : 1. Preoperative assessment from my prior note. 2. Hyponatremia likely to be SIADH related. Continue to monitor for now. 3. Leukocytosis. Likely to be reactive. No clear infectious source. 4. Thrombocytopenia. It is probably chronic in nature and appears to be stable. Although , we would continue to monitor this for now as well.
- **Paraphrase note 1**: "The patient seems to be doing alright despite some minor discomfort. His vital signs are stable and he hasn't had any labs today. The plan going forward is to keep monitoring the patient."
- **Paraphrase note 2**: "The patient appears to be doing well aside from minor discomfort around the site of where the tracheostomy was performed. Other than that, they do not appear to have any other discomfort. There are other apparent issues like **Hyponatremia** and **Leukocystosis** which appear to be a result of the procedure. The patient appears to be doing well, but should continue being monitored."

Examples: NoteAid

- **Note** : The patient is doing okay today. He has some complaints of discomfort around his tracheostomy site but otherwise has no complaints of chest pain, shortness of breath, abdominal pain, nausea, or vomiting. OBJECTIVE : Vital Signs : Currently stable. Cardiac : Normal S1, S2. No murmurs, rubs, or gallops are present. Lungs : Clear to auscultation bilaterally. Abdomen : Soft, nontender, and nondistended. Normoactive bowel sounds are present. LABS : There are no new labs from today. ASSESSMENT AND PLAN : 1. Preoperative assessment from my prior note. 2. Hyponatremia likely to be SIADH related. Continue to monitor for now. 3. Leukocytosis. Likely to be reactive. No clear infectious source. 4. Thrombocytopenia. It is probably chronic in nature and appears to be stable. Although , we would continue to monitor this for now as well

- **Paraphrase note**: "The patient is doing well although he has some discomfort around his tracheostomy area. The patient has stable vitals. Some conditions will continue to be monitored, including **low salt content** in the blood, **high white blood cell count**, and **low blood platelet count**."

Evaluation of Summary

Terms in 195 Summaries	With NoteAid	Without NoteAid
Number of medical jargon terms in lay language	**268**	68
Number of medical jargon terms copied and pasted	70	**169**

Patient Evaluation

- Recruit 3 patients, all college graduates
  - One Type I diabetes, two Type II diabetes
  - Being treated for 5~15 years
- Each patient presented with 1~2 own progress notes of his/her most recent clinical visit
  - Ask what they don't understand
  - Present Note+NoteAid
  - Think-aloud
  - A list of questions

Concepts and Abbreviations

- The 3 patients did not understand 5~10 concepts in his/her own progress note
  - E.g., FT4, auscultation, post-ablative hypothyroidism

Questions

- **Do you want to read your EHR notes?** Yes (3)
- **When do you want to read your notes?**
  - Before and after a clinical visit (3)
- **Is there a specific part of notes that you would like to read?**
  - All notes (2), current plan (1)
- **Would you be willing to share your notes with your friends or family members?**
  - No (2), spouse if asked (1)
- **Is NoteAid helpful?** Yes (3)
- **Do you want to use NoteAid?** Yes (3)

How to Improve NoteAid

- More concepts included by NoteAid (3)
- Easy-to-read definitions (3)
- More education material (3)

Accomplishments: Collaborations

- Collaborations with multiple VA and Universities
  - University of Massachusetts Medical School
  - University of Massachusetts, Amherst
  - VA Connecticut Health Care System
  - Yale University
  - VA Palo Alto Health Care System
  - Stanford University

Accomplishments: Publications

1. Polepalli Ramesh B, Houston T, Brandt C, Fang H, **Yu H**. 2013. Improving Patients' Electronic Health Record Comprehension with NoteAid, Studies in Health Technology and Informatics Vol.192: ***MEDINFO*** 2013, pp 714 - 718. DOI:10.3233/978-1-61499-289-9-714. **Best Student Paper Award**
2. Polepalli Ramesh B and **Yu H**. 2013. Systems for Improving Electronic Health Record Note Comprehension, In proceedings of Health Search and Discovery at ***ACM SIGIR*** 2013, pp 39-42.
3. Jagannatha, A., Chen, J. & Yu, H. 2015. Mining and Ranking Biomedical Synonyms from Wikipedia. ***EMNLP*** *Workshop on Health Text Mining and Information Analysis.*
4. Zheng, J. & Yu, H. 2015. Key Concept Identification for Medical Information Retrieval. *Empirical Methods for Natural Language Processing (**EMNLP**). Short paper*
5. Zheng, J. & Yu, H. 2015 Methods for linking EHR notes to education materials. ***Information Retrieval****. In Press.*
6. **Yu H,** Brandt C, Zulman D, and Houston T. 2015. Systems for helping Veterans Comprehend their own EHR notes. The 2015 ***HSR&D/QUERI*** National Conference.
7. Zheng, J. & Yu, H. 2015. Identifying Key Concepts from EHR Notes Using Domain Adaptation. ***EMNLP*** *Workshop on Health Text Mining and Information Analysis.*
8. Liu, W., Cai, S., Ramesh BP, Chiriboga G, Knight K, and **Yu H**. 2015. Translating Electronic Health Record Notes from English to Spanish: A Preliminary Study. ***ACL BioNLP*** (2015).
9. **Yu, H,** Makkapati, S**,** Maranda, L, Malkani, S. 2016. Mismatch between Patient Information-Seeking and Physician Expectation at a Diabetes Outpatient Clinic. ***Society of Behavioral Medicine****. Submitted.*
10. **Yu, H**. 2016. EHR Paraphrasing for NoteAid Evaluation. ***Society of Behavioral Medicine****. Submitted.*
11. Chen, J., Jagannatha, A. & **Yu H**. 2016. A Data Driven Approach for Mining and Prioritizing Medical Jargon from Electronic Health Record Notes. ***AAAI***. To be submitted

Accomplishments: Dissemination

- Best Student Paper Award
- Major Computer Science, Biomedical Informatics, and Behavior Science conferences
  - SIGIR, EMNLP, ACL, AAAI, MedInfo
- The 2015 *HSR&D/QUERI* National Conference
- HSR&D Cyberseminars
- Keynote at ICDM BioDM 2013
- Invited seminars at VA
  - Bedford VAMC, West Haven, CWM
  - eHMP
- Invited seminars at Universities
  - UMass-Boston, UMass-Lowell, UMass-Dartmouth, WPI
  - Peking University, University of Science and Technology Beijing

Opportunities

- VA is innovative
- VA is passionate
- VA is collaborative
- VA has excellent leadership
- VA is investing research and researchers
- VA is supportive
- VA is full of opportunities
  - My HealtheVet
    - Blue Button
  - eHMP

Challenges

- IRB application
- EHR data access
- Collaboration with operational partners
- Research fund spending
  - Hiring
  - Student support
  - Inflexibility
- A lot of paper work---fortunately VA is supportive

Thanks and Questions

- The NoteAid Prototype Systems:
  - http://www.clinicalnotesaid.org
- Contact
  - Hong.yu@va.gov or hong.yu@umassmed.edu

  - Hong Yu, MS, MA, and PhD
    Research Health Scientist,
    Bedford VAMC and

  Professor, Department of Quantitative Health Sciences
  Department of Computer Science (Adjunct)
  University of Massachusetts Medical School
  AS6-2071 ASC-QHS, 368 Plantation Street, Worcester, MA 01605

  508-856-3474/ Fax: 508-856-8993