



Chart Abstraction on VINCI: tools and services

Olga Patterson

Cyberseminar 12/13/2018

Disclaimer

The contents of this presentation do not represent the views of the Department of Veterans Affairs or the United States Government.

Acknowledgement

Work described in this presentation was supported using resources and facilities at the VA Salt Lake City Health Care System with funding from VA Informatics and Computing Infrastructure (VINCI), VA HSR HIR 08-204.

1. Overview of chart abstraction
2. Performing chart review using ChartReview
3. VINCI Annotation Services
4. Questions and discussion

Outline

Jeff Scehnet

Kevin Malohi

VINCIservices@va.gov

- Concierge
- Data Provisioning
- Compliance
- Feasibility
- Recruitment
- **Annotation / Chart Review**
- Natural Language Processing
- Analytics and Data Services
- Application Development

VINCI Services Team

What is your primary role in VA?

- A. Researcher/Investigator
- B. Research staff (coordinator, statistician, analyst...)
- C. Administrator, manager, or policy-maker
- D. Other
- E. Not affiliated with VA

Poll Question #1

Which best describes your chart review experience?

- A. I have not participated in any chart review activities
- B. I have collaborated on a project that used chart review
- C. I have performed chart review as an abstractor/annotator
- D. I have managed at least one chart review project
- E. I have led (planned, designed, conducted) chart review projects

Poll Question #2

Which best describes your familiarity with VINCI and VA data?

- A. I have never used VINCI or VA data for research
- B. I have used VA data for research but not in VINCI
- C. I have used VINCI and VA data
- D. I am an experienced user of VINCI and VA data

Poll Question #3

I am assuming that you are:

- VA-affiliated researcher or research staff
- Familiar with chart abstraction methodology
- Familiar with VA CDW
- Familiar with VINCI
- Planning to use VINCI ChartReview for chart abstraction

Target Audience

- Introductory HSRD cyberseminar was presented on 6/4/2018

https://www.hsrp.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=2484

Chart Abstraction (Chart review, Medical record review, Chart annotation)

- A research methodology of data collection for retrospective investigation

Overview

• Annotation

- **Class** - assigned meaning to data

Label = Concept = annotation class = annotation type \approx semantic type

- **Span** - a pointer to start and stop points in a text
- **Features** - attributes of the Class and their **values**

- Generated by human, machine, or human+machine.

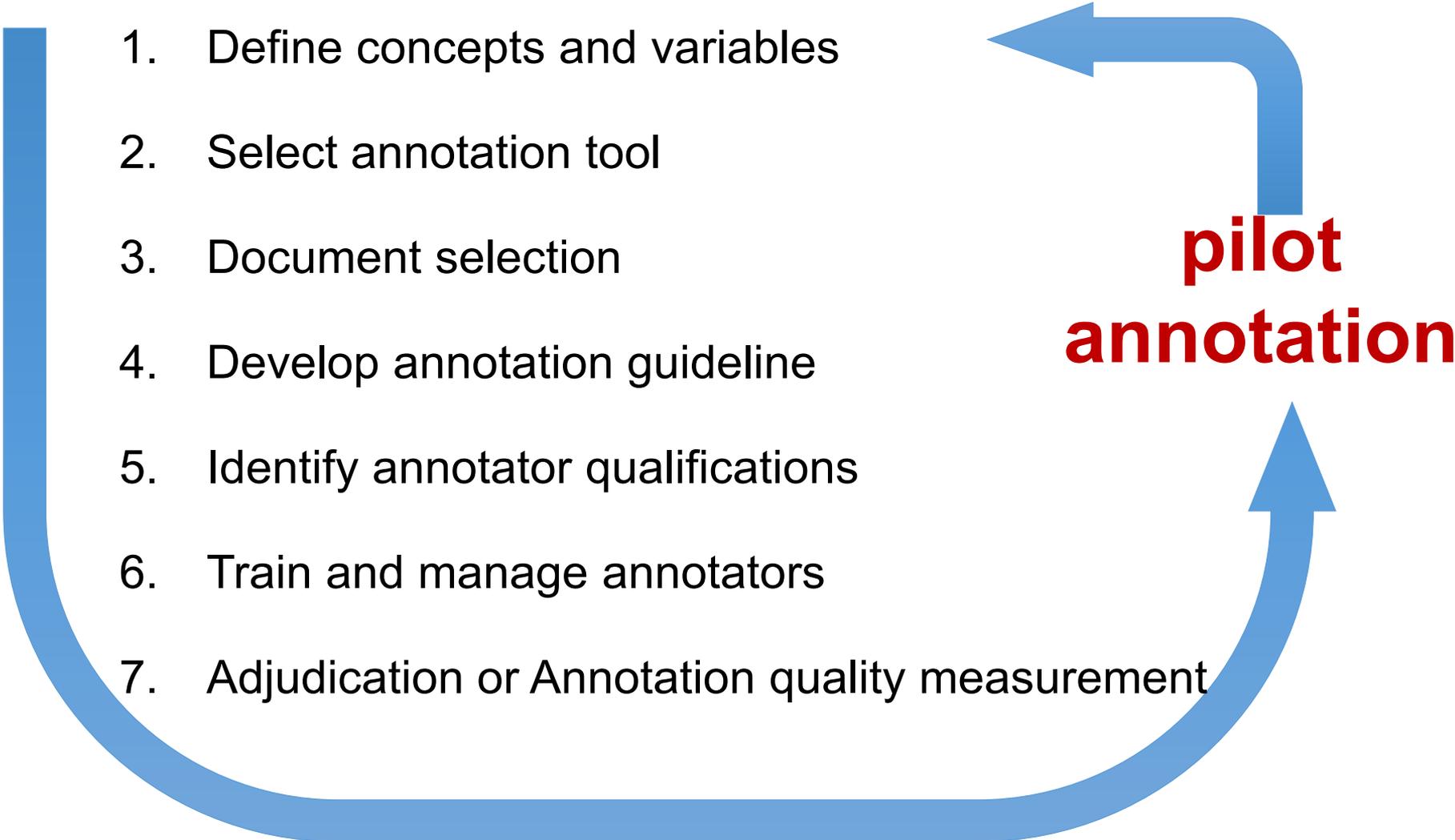
0123456789012345678901234567890123

The CXR shows LLL consolidation.

Annotation

Class: **Finding**
Span: **14:30**
Feature: **Assertion = Present**

Chart Abstraction through Annotations

1. Define concepts and variables
 2. Select annotation tool
 3. Document selection
 4. Develop annotation guideline
 5. Identify annotator qualifications
 6. Train and manage annotators
 7. Adjudication or Annotation quality measurement
- 
- pilot
annotation**

Annotation Project Workflow

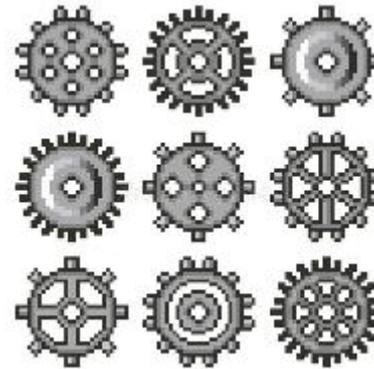
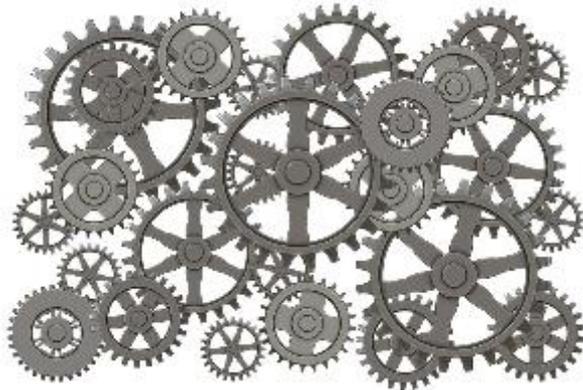
- Markables – elements to be annotated
 - Annotation type = label to be assigned to a segment of text
 - Relationship = link between instances of annotations
 - Features = Attributes of annotation types and relationships
- Annotation Schema
 - Complete list of all markables for the project

What to annotate

Annotation **complexity** inversely proportional to annotation **consistency**

- Working memory limit is 4* to 7**
- Consistency of annotation is the key!

Project size vs number of projects



* Cowan N. The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Curr Dir Psychol Sci.* 2010;19(1):51-57.

** Miller GA. The Magical Number Seven, Plus or Minus Two Some Limits on Our Capacity for Processing Information. *Psych Review.* 1956.

Scope of the annotation project

1. Concept definition
 - Diagnosis, lab test, action, event...
2. Variable definition – values that the concepts can have
 - Diagnosis: explicitly mentioned or inferred
 - Lab test: exact numeric value or range or direction
 - Action: planned or occurred
 - Event: explicitly mentioned or inferred
3. Level of the annotation
 - patient, event, document, or instance
4. Annotation boundaries for instance annotations

Annotation type definition

Concept sheet - formalized concept/variable definition

Concept	Variable definition	Source	Range of values
Procedure indication	Explicitly stated indication for procedure.	Colonoscopy report	Screening High Risk/diagnostic Treatment

Example of a concept sheet for a study on quality of colonoscopy procedures

Operationalization !!!

No vague definitions!

Words to avoid:

“Any evidence of ...”

“... including but not limited to ...”

! Tip !

Annotation type definition

Defining annotation level

Instance Level

- A specific phrase in a document is being identified
- A document may have 0 to many instances (mentions, phrases)
- Instances in a single document may not always agree

Document Level

- The contents of the whole document are being labeled

Event Level

- Information in several documents is used to label patient on a specific date

Patient Level

- The total of all instances and/or documents are combined to reach a final patient classification
- A patient may have 0 to many documents, events, and instances

Annotation type definition

Instance boundary delimitation

Example: Annotation type = Polyp removal method
What is the phrase to annotate?

Polyp was removed with cold snare

Polyp was removed with cold snare

Polyp was removed with cold snare

Annotation type definition

- Annotation schema
 - Annotation types = classes, labels
 - Attributes = features, modifiers
 - Relationships
- Formal step-by-step description of the annotation process
 - Must have:
 - Background information
 - Annotation instructions with examples specific to the annotation tool and to the documents in the corpus
 - Examples
 - what to annotate
 - what not to annotate

Annotation guideline

- eHost - <https://code.google.com/p/ehost/>
- ChartReview - <http://department-of-veterans-affairs.github.io/ChartReview/>
- BRAT - <http://brat.nlplab.org/>
- RapTAT – <https://code.google.com/p/raptat/>
- Knowtator – <http://knowtator.sourceforge.net/>
- GATE Teamware - <https://gate.ac.uk/sale/tao/splitch25.html>
- WebAnno - <https://webanno.github.io/webanno/>
- Prodigy - <https://prodi.gy>
- XTrans - <https://www ldc.upenn.edu/language-resources/tools/xtrans>
- WordFreak - <http://wordfreak.sourceforge.net>
- Djangology - <https://sourceforge.net/projects/djangology/>
-

Annotation tools

VINCI ChartReview



Annotation

Task

Done Save Hold/Next Submit/Next Task Info

Task Name: CR_TEST_Synth_Pres (197617)
Principal: CR_SYNTH_Smoking_Patients_20170717
Description: no detailed description

Items

- 1.1. Smoked Current_Smoker
1.2. gender
1.3. age
2. Smoking intensity
2.1. frequency
3. Cigar Use
3.1. Does the patient have a history of cigar Use?

Previous Skip Next/Complete

1.1. Has the patient ever smoked:

- Yes
No
(Clear Answer)

Annotations

All Annotations Wrap

Table with columns: Span, ClinicalEle..., Class, Type, Creation D..., ClinicalEle...
Row: TOBACCO USE DISORDER, CR_SYNTH_Sr 800000085674, Current_Smoke, New, 2017-08-09T16:42:12Z, CR_SYI...

Certainty: High

Tobacco product type:

Cigarettes

CR_SYNTH_Smoking_Patients_20170717

Summary

Done
Classify:
ScrSSN: 111000111
Race: White
Gender: M
DOB: 1950-01-01

CR_SYNTH_Smoking_Documents_20170717

List (5 of 5)

Detail: Right Description Wrap Auto-select Search Detail View

Table with columns: ReferenceDateTime, TIUDocumentSID, TIUDocumentDefinition
Rows: Nov 1, 2012 12:00:00 AM 800000000099 ADMIN_NOTE
Oct 11, 2012 12:00:00 AM 800000032581 Clinic Visit Note
Nov 29, 2012 12:00:00 AM 800000045678 Nursing Shift Note Assessment
Nov 21, 2012 12:00:00 AM 800000078432 Followup Summary
Nov 20, 2012 12:00:00 AM 800000085674 Followup Summary

View in new tab Done Popup Annotation Text

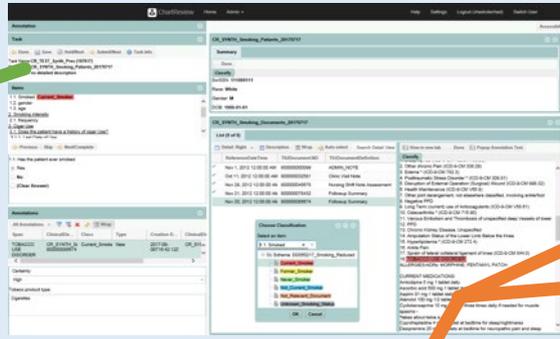
Classify:
2. Other chronic Pain (ICD-9-CM 338.29)
3. Edema * (ICD-9-CM 782.3)
4. Posttraumatic Stress Disorder * (ICD-9-CM 309.81)
5. Disruption of External Operation (Surgical) Wound (ICD-9-CM 998.32)
6. Health Maintenance (ICD-9-CM V85.9)
7. Other joint derangement, not elsewhere classified, involving ankle/foot
8. Negative PPD
9. Long Term (current) use of Anticoagulants (ICD-9-CM V58.61)
10. Osteoarthritis * (ICD-9-CM 715.90)
11. Venous Embolism and Thrombosis of unspecified deep Vessels of lower
12. PPD
13. Chronic Kidney Disease, Unspecified
14. Amputation Status of the Lower Limb Below the Knee
15. Hyperlipidemia * (ICD-9-CM 272.4)
16. Ankle Pain
17. Sprain of lateral collateral ligament of knee (ICD-9-CM 844.0)
18. TOBACCO USE DISORDER
ALLERGIES/ADRS: MORPHINE, FENTANYL PATCH
CURRENT MEDICATIONS:
Amlodipine 5 mg 1 tablet daily
Ascorbic acid 500 mg 1 tablet daily with iron - uses twice daily
Aspirin 81 mg 1 tablet daily
Atenolol 100 mg 1/2 tablet daily
Cyclobenzaprine 10 mg 1 tablet three times daily if needed for muscle spasms -
*takes about twice a day
Cyproheptadine 4 mg 1 tablet at bedtime for sleep/nightmares
Desipramine 25 mg 5 tablets at bedtime for neuropathic pain and sleep

Choose Classification dialog box showing a tree view with 'Current_Smoker' selected.

ChartReview

VINCI

<http://vhacdwdwhweb04:8080/chart-review>



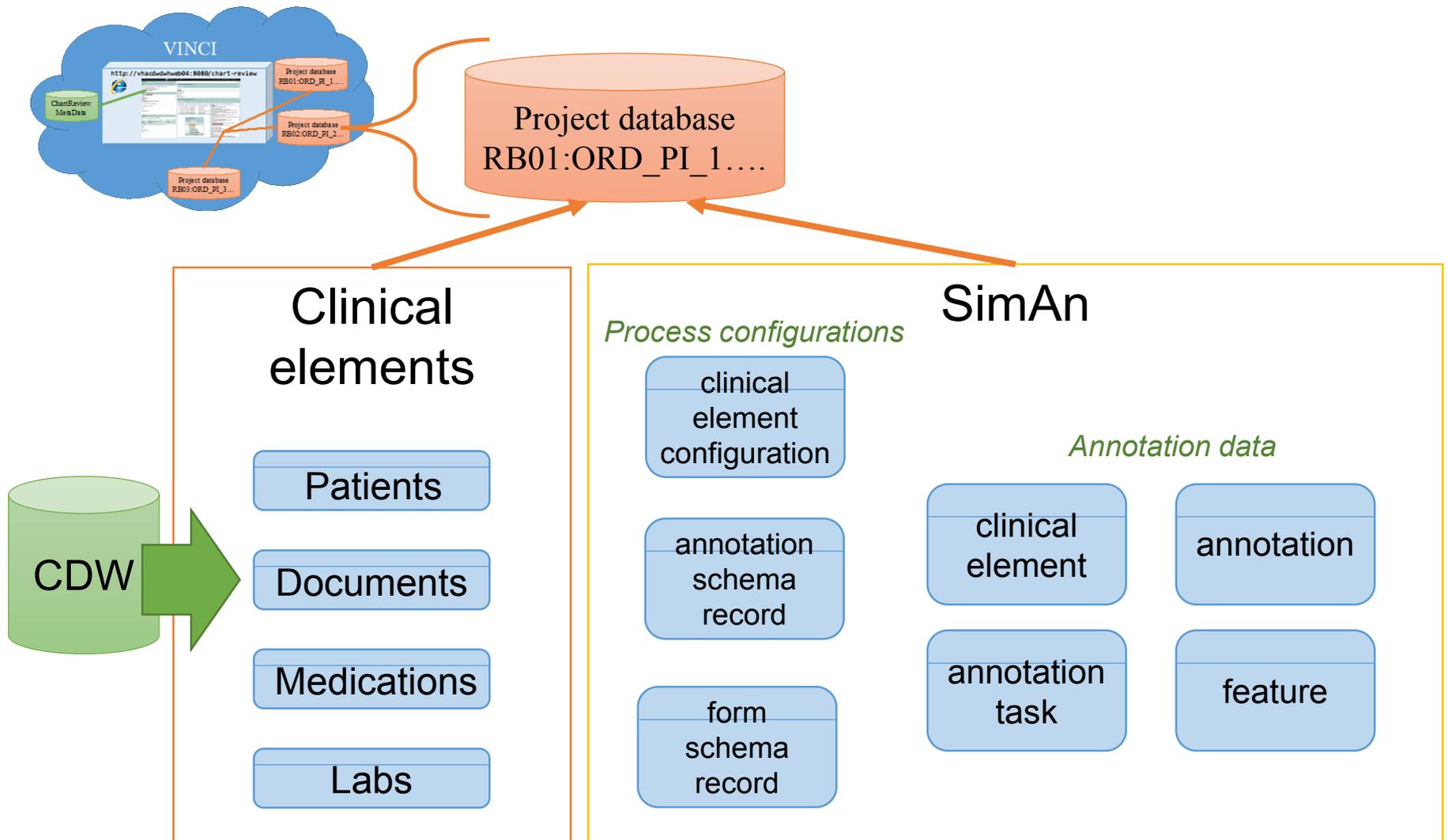
ChartReview
MetaData

Project database
RB01:ORD_PI_1....

Project database
RB02:ORD_PI_2...

Project database
RB03:ORD_PI_3...

ChartReview architecture

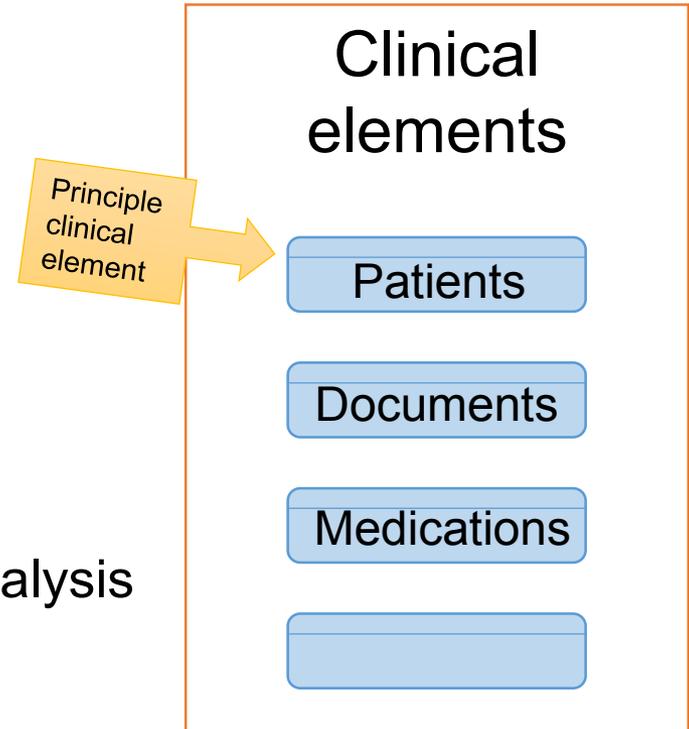


* SimAn - Simple Annotation Schema

ChartReview architecture

Clinical element

- Data element to be viewed during chart abstraction and/or annotation.
- Examples:
 - Lab values
 - Notes
 - Radiology reports
 - ICD-9/10 Codes
 - Medications
 - Vital status
 - Etc.



Principle clinical element

- Clinical element that serves as a unit of analysis
- Most frequently -- Patient
- Links all other clinical elements

ChartReview terminology

Project

- Assigned workspace
- Associated with VINCI research database
 - i.e. (ORD_PI_201801020D)

Process

- Individual abstraction/annotation undertaking for a specific research project
- Defined by: schema, clinical elements

Task

- The individual unique item to be reviewed within the process.
- Corresponds to [principle clinical element](#)

ChartReview Terminology

Annotation Schema

- Defines markables
- Example:

Annotation type= Smoking

- Attribute = smoking status
- Value data type = Option (current, past, never)

Annotation type = Patient Date of Birth

- Attribute = Date
- Value data type = date in range (01/01/1910 - 01/01/2001)

Form Schema

- Abstraction instrument
- Example:

1. Does the patient have associated co-morbidities: YES/NO

If Yes, then select from the following:

- Chronic Heart Failure
- Chronic Kidney Disease
- Diabetes
- Etc.

ChartReview Terminology

- IRB Approval
- DART Approval
 - Request Real SSN
 - Request CDW production domains
 - Request TIU Text Notes
 - Ensure HIPAA Waiver
- Contact VinciServices@va.gov

Does your study require a HIPAA Waiver?

Yes

No

IDENTIFIERS [\(More about identifiers\)](#)

Real SSN

Scrambled SSN

Identifiable data but no real or scrambled SSNs

Regulatory Requirements

Jeff Scehnet

Kevin Malohi

VINCIservices@va.gov

- Concierge
- Data Provisioning
- Compliance
- Feasibility
- Recruitment
- **Annotation / Chart Review**
- Natural Language Processing
- Analytics and Data Services
- Application Development

VINCI Services Team

VINCI Services Annotation team

Lead: Olga Patterson, PhD

Annotation manager: Lacey Lewis, MPH

- Range of services

- Education and training
- Process definition and guideline development
- Contracting annotators
- Full chart abstraction process

VINCI
sponsored

Collaborative
project

VINCI Annotation Services

What topic would you like to see in the next cyberseminar related to annotation and chart abstraction in VINCI:

(select all that apply)

- A. Detailed training on overall annotation methodology
- B. Measuring annotations quality and output interpretation
- C. Practical SQL for data analysis from SimAn tables
- D. Other topic (if selected, please email your suggestion)
- E. Nothing, I already know everything that I need to know

Poll Question #4

The seminar will end promptly at 1pm MT (3pm ET)

If you have a question that was not answered, or suggestion, please email it to:

VinciServices@va.gov

Subject: ChartReview cyberseminar

Discussion

< [Clinical Element Configuration List](#)

Show Clinical Element Configuration

General Configuration Columns Content Template

Name	Lab Element
Description	Simple lab element based on example database.
Active	true
Title Field	LAB_NAME
Description Field	DESCRIPTION
Type	LIST
All Elements By Patient Id Query	select lab.lab_date, lab_test_lookup.lab_name, lab.result, concat(lab_test_lookup.lab_name, ' ', cast(lab.result as char)) as description, lab.id, lab.lab_performed_by from lab, lab_test_lookup, patient where lab.lab_performed_id = lab_test_lookup.id and lab.patient_id = patient.id and patient.id = ?
Single Element Query	select lab.lab_date, lab_test_lookup.lab_name, lab.result, concat(lab_test_lookup.lab_name, ' ', cast(lab.result as char)) as description, lab.id, lab.lab_performed_by from lab, lab_test_lookup, patient where lab.lab_performed_id = lab_test_lookup.id and lab.patient_id = patient.id and lab.id = ?
Created By	admin
Created Date	2014-03-05 16:28:16.0

Edit

Chart-Review: Configuration

Step 2 - Process Parameters

Process Name Clinical Element Group

Process Steps

usertask1

Single Step ChartReview

Schema

AQUAL

Clinical Elements

 Include/exclude all clinical elements

Include	Hidden	Clinical Element	Position
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Example Lab Element	1 <input type="text"/>
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Lab Element	2 <input type="text"/>
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Lab Simple	3 <input type="text"/>
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Lab Simpler	4 <input type="text"/>
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Patient Element	5 <input type="text"/>
<input checked="" type="checkbox"/>	<input type="checkbox"/>	TIU Document	6 <input type="text"/>
<input checked="" type="checkbox"/>	<input type="checkbox"/>	TIU Document Plain	7 <input type="text"/>
<input checked="" type="checkbox"/>	<input type="checkbox"/>	TIU Document simple	8 <input type="text"/>

Task Name Annotation Group Task Pre-Annotation Group

Task Principal Clinical Element

Example Lab Element

Chart-Review: Process Creation

Step 3 - Users and Clinical Elements In This Process

Task Creation Query

Enter the query to create process tasks. ⓘ

```
select id from patient;
```

Users

- admin
- project1
- project2

Other

Assignment style ⓘ

By Process

Previous

Create Tasks and Finish

Chart-Review: Task Creation Query



Annotation

Task

Done Save Hold/Next Submit/Next Task Info

Task Name: CR_TEST_Synth_Pres (197617)
Principal: CR_SYNTH_Smoking_Patients_20170717
Description: no detailed description

Items

- 1.1. Smoked Current_Smoker
1.2. gender
1.3. age
2. Smoking intensity
2.1. frequency
3. Cigar Use
3.1. Does the patient have a history of cigar Use?

Previous Skip Next/Complete

1.1. Has the patient ever smoked:

- Yes
No
(Clear Answer)

Annotations

All Annotations Wrap

Table with columns: Span, ClinicalEle..., Class, Type, Creation D..., ClinicalEle...
Row 1: TOBACCO USE DISORDER, CR_SYNTH_Sr 800000085674, Current_Smoke, New, 2017-08-09T16:42:12Z, CR_SYI...

Certainty: High

Tobacco product type:

Cigarettes

CR_SYNTH_Smoking_Patients_20170717

Summary

Done

Classify

ScrSSN: 111000111
Race: White
Gender: M
DOB: 1950-01-01

CR_SYNTH_Smoking_Documents_20170717

List (5 of 5)

Detail: Right Description Wrap Auto-select Search Detail View

Table with columns: ReferenceDateTime, TIUDocumentSID, TIUDocumentDefinition
Rows: Nov 1, 2012 12:00:00 AM ADMIN_NOTE, Oct 11, 2012 12:00:00 AM Clinic Visit Note, Nov 29, 2012 12:00:00 AM Nursing Shift Note Assessment, Nov 21, 2012 12:00:00 AM Followup Summary, Nov 20, 2012 12:00:00 AM Followup Summary

View in new tab Done Popup Annotation Text

Classify

- 2. Other chronic Pain (ICD-9-CM 338.29)
3. Edema * (ICD-9-CM 782.3)
4. Posttraumatic Stress Disorder * (ICD-9-CM 309.81)
5. Disruption of External Operation (Surgical) Wound (ICD-9-CM 998.32)
6. Health Maintenance (ICD-9-CM V85.9)
7. Other joint derangement, not elsewhere classified, involving ankle/foot
8. Negative PPD
9. Long Term (current) use of Anticoagulants (ICD-9-CM V58.61)
10. Osteoarthritis * (ICD-9-CM 715.90)
11. Venous Embolism and Thrombosis of unspecified deep Vessels of lower
12. PPD
13. Chronic Kidney Disease, Unspecified
14. Amputation Status of the Lower Limb Below the Knee
15. Hyperlipidemia * (ICD-9-CM 272.4)
16. Ankle Pain
17. Sprain of lateral collateral ligament of knee (ICD-9-CM 844.0)
18. TOBACCO USE DISORDER
ALLERGIES/ADRs: MORPHINE, FENTANYL PATCH
CURRENT MEDICATIONS:
Amlodipine 5 mg 1 tablet daily
Ascorbic acid 500 mg 1 tablet daily with iron - uses twice daily
Aspirin 81 mg 1 tablet daily
Atenolol 100 mg 1/2 tablet daily
Cyclobenzaprine 10 mg 1 tablet three times daily if needed for muscle spasms -
*takes about twice a day
Cyproheptadine 4 mg 1 tablet at bedtime for sleep/nightmares
Desipramine 25 mg 5 tablets at bedtime for neuropathic pain and sleep

Choose Classification dialog box showing a tree view with 'Current_Smoker' selected under 'Schema: DD050217_Smoking_Reduced'.

Chart-Review: User Interface

1. South BR, Shen S, Barrus R, DuVall SL, Uzuner O, Weir C. Qualitative analysis of workflow modifications used to generate the reference standard for the 2010 i2b2/VA challenge. *AMIA Annu Symp Proc.* 2011;2011:1243-1251.
2. Forbush TB, Shen S, South BR, Duvall SL. What a catch! traits that define good annotators. *Stud Health Technol Inform.* 2013;192:1213.
3. Warner JL, Anick P, Drews RE. Physician inter-annotator agreement in the Quality Oncology Practice Initiative manual abstraction task. *J Oncol Pract.* 2013;9(3):e96-e102.
4. Matt V, Matthew H. The retrospective chart review: important methodological considerations. *J Educ Eval Health Prof.* 2013;10:12.
5. Liu K, Chapman W, Hwa R, Crowley RS. Heuristic Sample Selection to Minimize Reference Standard Training Set for a Part-Of-Speech Tagger. *J Am Med Informatics Assoc.* 2007;14(5):641-650.
6. Ogren P V, Savova G, Chute CG. Constructing Evaluation Corpora for Automated Clinical Named Entity Recognition. In: *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08).* Marrakech, Morocco; 2008.
7. Chapman WW, Dowling JN, Hripcsak G. Evaluation of training with an annotation schema for manual annotation of clinical conditions from emergency department reports. *Int J Med Inf.* 2008;77(2):107-113.

Suggested reading