



SOCIETY OF
ACTUARIES®

2019 **ANNUAL
MEETING**
& EXHIBIT

October 27-30
Toronto, Canada

Session 176: The Future of Medical Records in Underwriting

[SOA Antitrust Compliance Guidelines](#)

[SOA Presentation Disclaimer](#)

The Future of Medical Records in Underwriting

Dr. Timothy Meagher
David Moore, FSA, MAAA
Dae Won Kim
Moderator: June Quah, FSA, FCIA

Wednesday, October 30th, 2019



SOCIETY OF ACTUARIES

Antitrust Compliance Guidelines

Active participation in the Society of Actuaries is an important aspect of membership. While the positive contributions of professional societies and associations are well-recognized and encouraged, association activities are vulnerable to close antitrust scrutiny. By their very nature, associations bring together industry competitors and other market participants.

The United States antitrust laws aim to protect consumers by preserving the free economy and prohibiting anti-competitive business practices; they promote competition. There are both state and federal antitrust laws, although state antitrust laws closely follow federal law. The Sherman Act, is the primary U.S. antitrust law pertaining to association activities. The Sherman Act prohibits every contract, combination or conspiracy that places an unreasonable restraint on trade. There are, however, some activities that are illegal under all circumstances, such as price fixing, market allocation and collusive bidding.

There is no safe harbor under the antitrust law for professional association activities. Therefore, association meeting participants should refrain from discussing any activity that could potentially be construed as having an anti-competitive effect. Discussions relating to product or service pricing, market allocations, membership restrictions, product standardization or other conditions on trade could arguably be perceived as a restraint on trade and may expose the SOA and its members to antitrust enforcement procedures.

While participating in all SOA in person meetings, webinars, teleconferences or side discussions, you should avoid discussing competitively sensitive information with competitors and follow these guidelines:

- **Do not** discuss prices for services or products or anything else that might affect prices
- **Do not** discuss what you or other entities plan to do in a particular geographic or product markets or with particular customers.
- **Do not** speak on behalf of the SOA or any of its committees unless specifically authorized to do so.
- **Do** leave a meeting where any anticompetitive pricing or market allocation discussion occurs.
- **Do** alert SOA staff and/or legal counsel to any concerning discussions
- **Do** consult with legal counsel before raising any matter or making a statement that may involve competitively sensitive information.

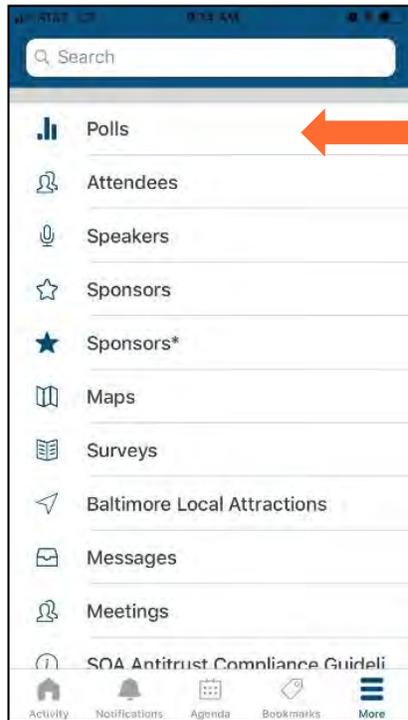
Adherence to these guidelines involves not only avoidance of antitrust violations, but avoidance of behavior which might be so construed. These guidelines only provide an overview of prohibited activities. SOA legal counsel reviews meeting agenda and materials as deemed appropriate and any discussion that departs from the formal agenda should be scrutinized carefully. Antitrust compliance is everyone's responsibility; however, please seek legal counsel if you have any questions or concerns.

Presentation Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the participants individually and, unless expressly stated to the contrary, are not the opinion or position of the Society of Actuaries, its cosponsors or its committees. The Society of Actuaries does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented. Attendees should note that the sessions are audio-recorded and may be published in various media, including print, audio and video formats without further notice.

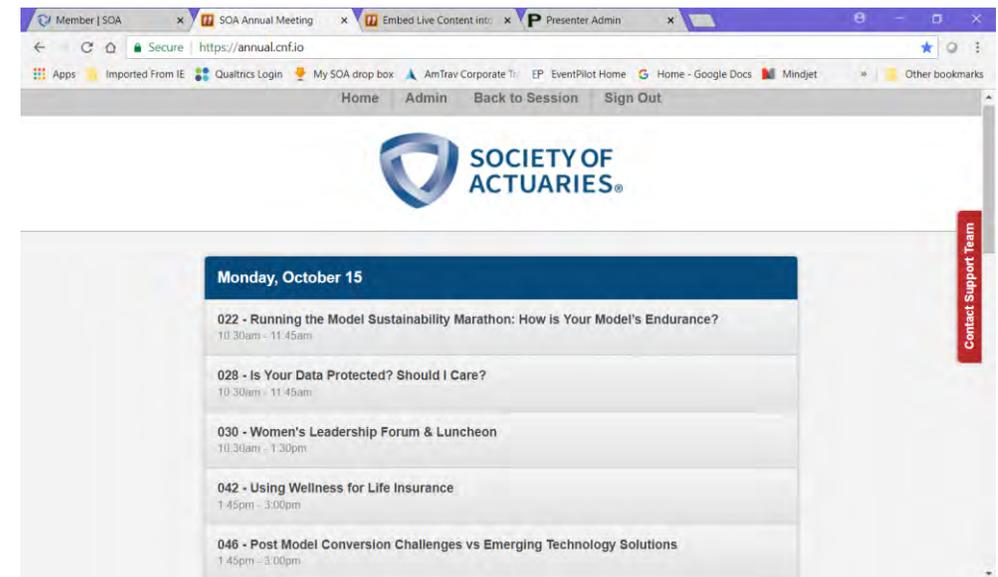
To Participate, look for Polls in the SOA Event App or visit annual.cnf.io in your browser

Find The Polls Feature Under **More**
In The Event App or Under This
Session in the Agenda



or

Type annual.cnf.io In Your Browser



NOT IF, BUT HOW

The Future of Medical Records in Underwriting

Dr. Tim Meagher, VP & Chief Medical Officer
October 30, 2019

What is a Medical Record?

- Repository of personal medical information
- Contains diagnostic and prognostic (predictive) information



John Smith

55 years old, dentist

1. Presenting complaint:

Chest pain for 2 weeks

2. Past Medical History:

High blood pressure for 10 years ; Diabetes for 5 years

3. History of Presenting Illness:

Pain provoked by exercise; squeezing quality, radiating to neck, perspiration +++

4. Social History

- Dentist, works 50 hrs/week
- Smoke 10 cigs/ day x 20 years; 1-2 glasses wine most days
- Aerobic exercise 30-45 mins. 1x/week
- No illicit drug use
- Rx: Vasotec 5mgs/day; allergic to penicillin

5. Family History

- Father died 'stroke', age 72
- Mother, 80; breast cancer age 50
- 4 sibs; B 50, high blood pressure, B 48 A/W, S 46 A/W, S 44 Breast cancer age 42
- PGF died age 50; PGM died old age.
- MGF died age 75; MGM died age 55 ? Cancer
- 3 children, M 22, F 21, F 19- all A/W
- Paternal uncle 'heart attack' age 60; maternal aunt breast cancer, died age 58

6. Functional Enquiry (Review of Systems)

- CVS, RS, GI, GU, MSK, Endo, Neuro, Psych, Heme,

7. Physical Examination

- BP 148/100, HR 78/min, RR 15, Weight 218 lbs.
- General: A and O x3; slightly anxious
- HHN: normal
- CVS: JVP normal . S1, S2 N, no S3, S4 or murmurs. Peripheral pulses N. No edema
- RS: lungs clear
- GI: abdomen N
- Remainder exam N

8. Clinical Impression (diagnosis)

- Angina Pectoris

9. Differential Diagnosis

- Musculoskeletal pain
- Anxiety

10. Investigations

- Bloods, ECG, CXR, exercise ECG

11. Treatment (Rx)

- Metoprolol 50 mgs bid; ASA 300 mgs daily

12. Test Results

- ECG abnormal; CXR N; glucose elevated; Exercise ECG positive

13. Progress Notes

- **Clinic Visit 1**
 - Discuss abnormal test results; propose coronary angiogram
- **Clinic Visit 2**
 - Coronary angiogram: LAD 90% obstruction; drug-eluting stent placed
- **Clinic Visit 3**
 - Pain free, exercising

Value of various components

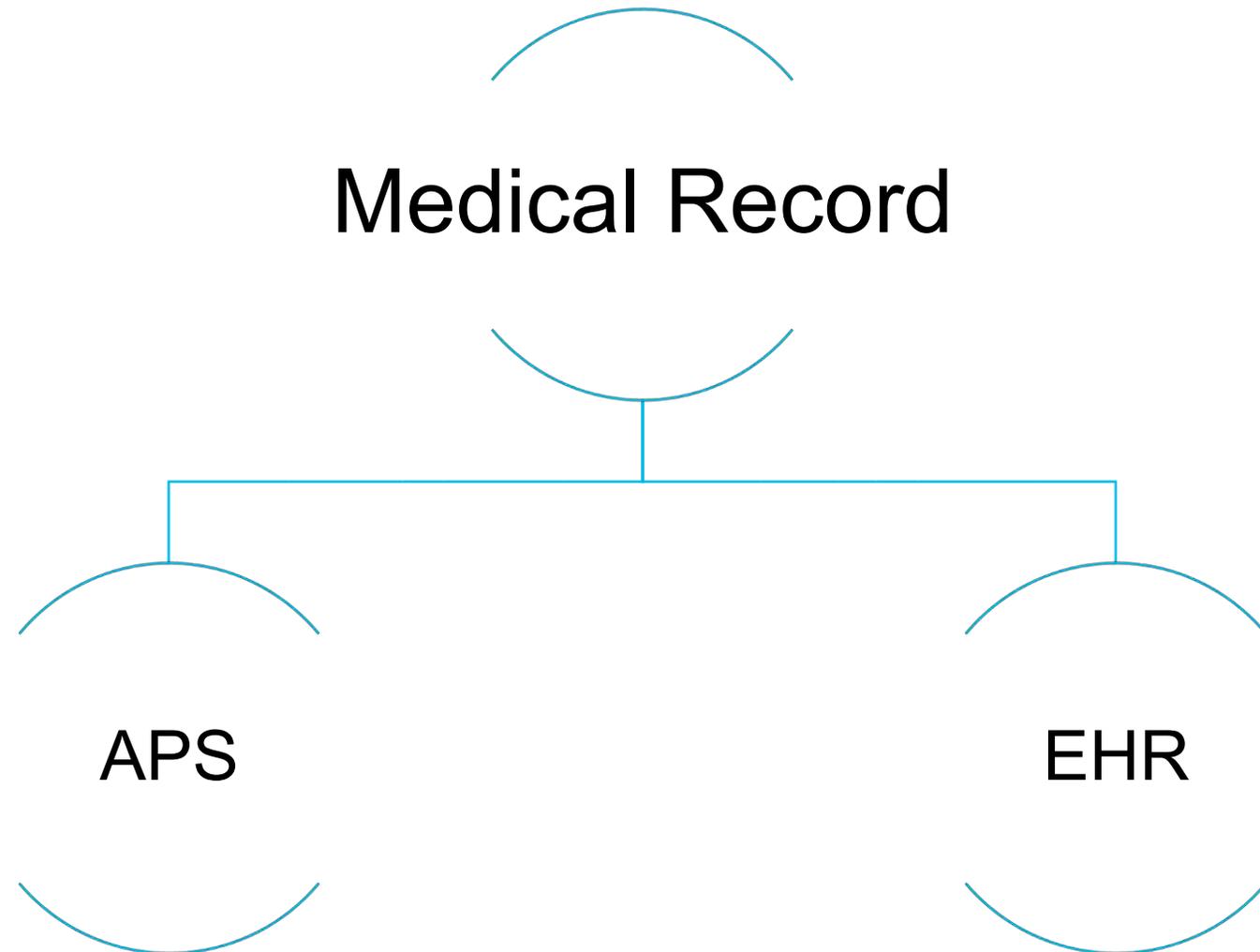
- History provides **70%** of diagnoses; physical examination **30%**
- Physical examination is on the wane
- Investigations taking larger role (imaging, genetics)
 - Encroaching on value of history/physical

The World Before EHR's

- Patient 'chart'
- Messy, incomplete, immobile and illegible (often!)



'Medical Records' come in different shapes



When you order an APS, you may get:

1. A single **handwritten** page that summarizes the medical chart, (or portions thereof), or
2. A **photocopy/pdf** of the medical chart, (or portions thereof), or
3. A **photocopy/pdf** of an EHR

1078
wants to do BLE
the fully lined & leg part 1078 m
nos - wmm; 9 med; mmm

1079 = 120770 1078 = 68 med
wmm - C
wmm - wmm

1080 - wmm; by mmm 9; the 1078
lulu

SEP 23 2006

1081 dry cry 1 H. h
nos - wmm; 9 med; mmm
lulu - 9 med = 813; 1078 = 37

1082 wmm stable; wmm - C

EHR versus Traditional Patient Chart

	Traditional Patient Chart	EHR
History	+	+
Physical Examination	+	+
Diagnoses	+	+
Laboratory Results	+	+++
Completeness	+	++++
Legibility	+	++++
Prescription record	+	++++
Inter-professional communication	+	++++
Safety*	+	+++
Research	+	++++

*fewer medication errors, CPG adherence, etc.

NOT IF, BUT HOW

Thank you

Case Study: EHR Pilot Projects

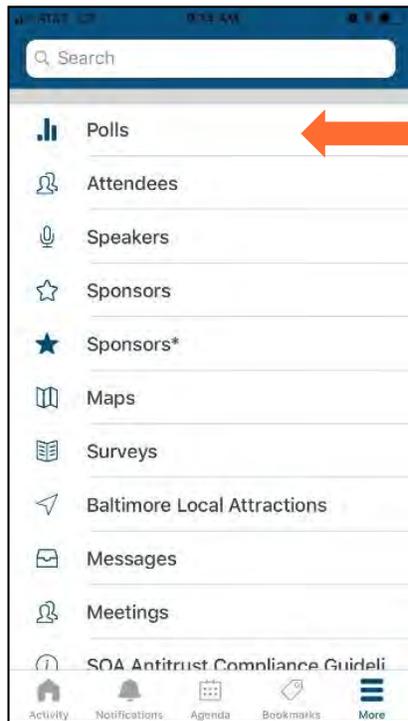
David Moore, FSA, MAAA

Wednesday, October 30th, 2019



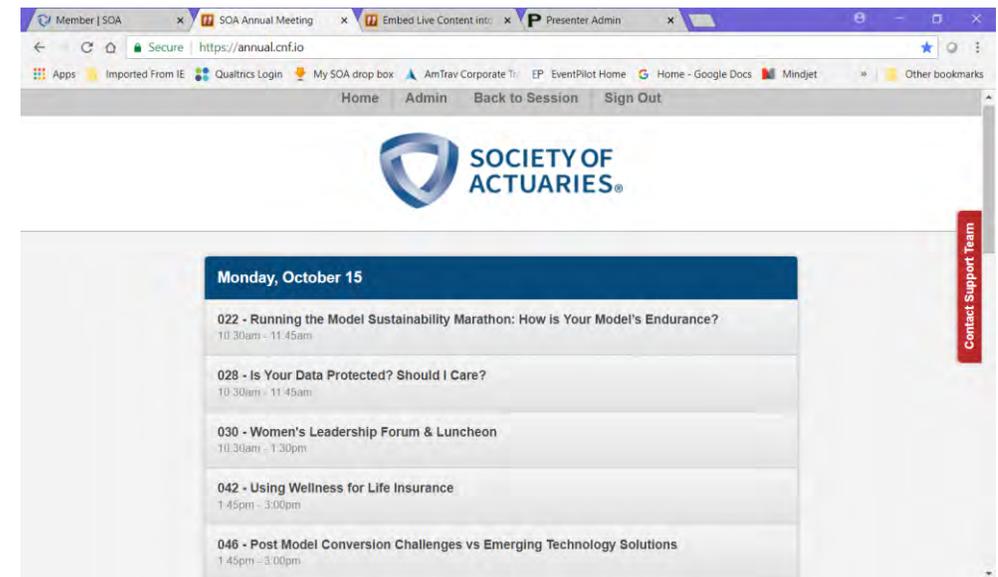
To Participate, look for Polls in the SOA Event App or visit annual.cnf.io in your browser

Find The Polls Feature Under **More**
In The Event App or Under This
Session in the Agenda



or

Type annual.cnf.io In Your Browser



“If your data sucks, you can’t make good decisions”

- My Boss

The current and future state of EHRs – Direct carrier perspective

Challenges to current underwriting process:

- Too slow
- Not enough data available at time of application
- Data required to make proper risk decisions is a pain to gather

Goal of this presentation:

- Share learnings from two pilot projects comparing EHRs to traditional sources of underwriting

Why Electronic Health Records?

- Current state of underwriting is (still) a long inefficient process
 - Electronic medical records can be leveraged to improve this process, but there are many issues with the data and structure
- Recent lab work or other measurements such as BMI can be used in lieu of ordering new requirements
- Reduce non-disclosure or misrepresentation
- Historical data from medical records can help us better understand the progression of disease, and the impact to future mortality
- **Goal: enable faster & more accurate underwriting decisions**

Live Content Slide

When playing as a slideshow, this slide will display live content

**Poll: Is your company currently
investigating the use of Electronic Health
Records for underwriting?**

Live Content Slide

When playing as a slideshow, this slide will display live content

Poll: Have you attempted to digitize prior medical requirements to better understand mortality risk and the underwriting process?

Case Study: Two EHR pilot projects

1. Patient driven EHR

- Use of a patient portal to allow applicant to capture EHR data and send to insurance company

2. Third party aggregation

- Obtain EHR through Health Information Exchange
- Use authorization provided on insurance application to obtain medical records

Patient Portal

A patient portal is a website for your personal health care. The online tool helps you to keep track of your health care provider visits, test results, billing, prescriptions, etc.

Benefits of a Patient Portal:

- You can access your secure personal health information and be in touch with your provider's office 24 hours a day. You do not need to wait for office hours or returned phone calls to have basic issues resolved.
- You can access all of your personal health information from all of your providers in one place. If you have a team of providers, or see specialists regularly, they can all post results and reminders in a portal. Providers can see what other treatments and advice you are getting. This can lead to better care and better management of your medicines.

Health Information Exchange (HIE)

HIE provides the capability to electronically move clinical information among different health care information systems. The goal of HIE is to facilitate access to and retrieval of clinical data to provide safer and more timely, efficient, effective, and equitable patient-centered care.

HIE systems facilitate the efforts of physicians and clinicians to meet high standards of patient care through electronic participation in a patient's continuity of care with multiple providers.

Secondary health care provider benefits include reduced expenses associated with:

- the manual printing, scanning and faxing of documents, including paper and ink costs, as well as the maintenance of associated office machinery
- the physical mailing of patient charts and records, and phone communication to verify delivery of traditional communications, referrals, and test results
- the time and effort involved in recovering missing patient information, including any duplicate tests required to recover such information

Patient Portal EHR Pilot

- Work with third party who provided a portal for the applicant/insured to login in to their personal health records and transfer them to the insurer
- To start the pilot, we identified recent applicants where we gathered the full UW requirements (application, Rx, labs, APS) in order have a complete underwriting record to compare against the EHR.
- These insureds were contacted by their agents and both given an incentive to participate in the study, and we ensured there would be no adverse action if EHR findings were not consistent with the risk class they just received on their policy

Results...

First, a disclaimer!

- This comparison is of traditional underwriting for some of our less healthy applicants. To test the EHR's, we wanted to use applicants with significant medical history to see what the EHR would provide
- As such, these results should not be generalized to the overall population, but instead provide us with insight in to whether the process is feasible

Note: The pilot results were only shared with the live conference audience

When will this be ready to use?

- You don't need 100% availability of medical records to successfully use EHR in your UW process – value can be achieved now
 - Post UW audits
 - Post claim audit
 - Proof of Concept for future state UW
 - Automate UW when data is available immediately; pivot others to full UW
 - Study mortality impact of longitudinal medical data (BMI, BP, Cholesterol, etc.)

Conclusions from Pilot Projects

- There is great potential – but we are not there yet

HIE

- High quality data, limited number of locations where it can be used
- The number of cases where valuable data is returned is relatively low, but can still provide considerable value to the underwriting process

Patient Portal

- High quality of data, limited to certain medical providers
- Limited buy in from agents/applicants – lack of awareness of medical records / unwillingness to take extra steps to provide insurer with data

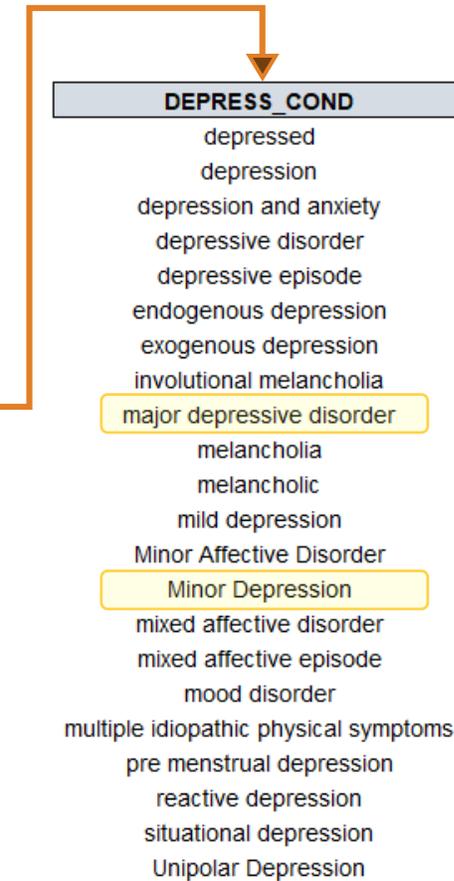
Future use case for EHR

- In these pilots, we focused on if the data available from EHRs was a suitable replacement for insurance labs and the APS
- Looking ahead, we are not trying to replicate the process with a different source of data, we are trying to change the process by leveraging new sources of data. This research is just the start of that effort

Transforming UW data with EHR/DHD

Underwriting Medical Condition Interpretation (current state)

NW's Condition Code	NW's Interpretation	Occurrences	Different Answers
HYPHBP_COND	High Blood Pressure (HBP)	7193	33
ELCHOL_COND	Elevated Cholesterol	4557	41
COLD_COND	Common Cold	3707	14
ANXIETY_COND	Anxiety	3438	23
ALLERG_COND	Allergies	2162	15
ESOPHAGITRFLX_COND	Reflux Esophagitis	2122	4
ALLERGS_COND	Seasonal Allergies	2033	6
DEPRESS_COND	Depression	1997	22
SINUSINFECT_COND	Sinus Infection	1874	5
INSOM_COND	Insomnia	1755	5



Challenges:

- Single condition code, varying answers
- Multiple condition codes, redundant answers
- Medical language interpretation
- Varying degrees of severity and granularity
- Unrecognized, rarely seen, misspelled
- Requires self-maintenance, -auditing, -governance

Codification & Standardization of Digital Health Data

The **Unified Medical Language System (UMLS)** integrates and distributes key terminology, classification and coding standards, and associated resources to promote creation of more effective and interoperable biomedical information systems and services, including electronic health records.

The Unified Medical Language System (UMLS)

1. SNOMED CT – Systemized Nomenclature Of MEDicine, Clinical Terms
2. ICD – International Classification of Diseases
3. RxNorm

You can use the UMLS to:

- Link terms and codes between doctor, pharmacy, and insurance company records
- Process texts to extract concepts, relationships, or knowledge
- Facilitate mapping between terminologies

Thank you!



Machine Learning for Medical Records

Munich Re

Dae Won Kim

October 2019

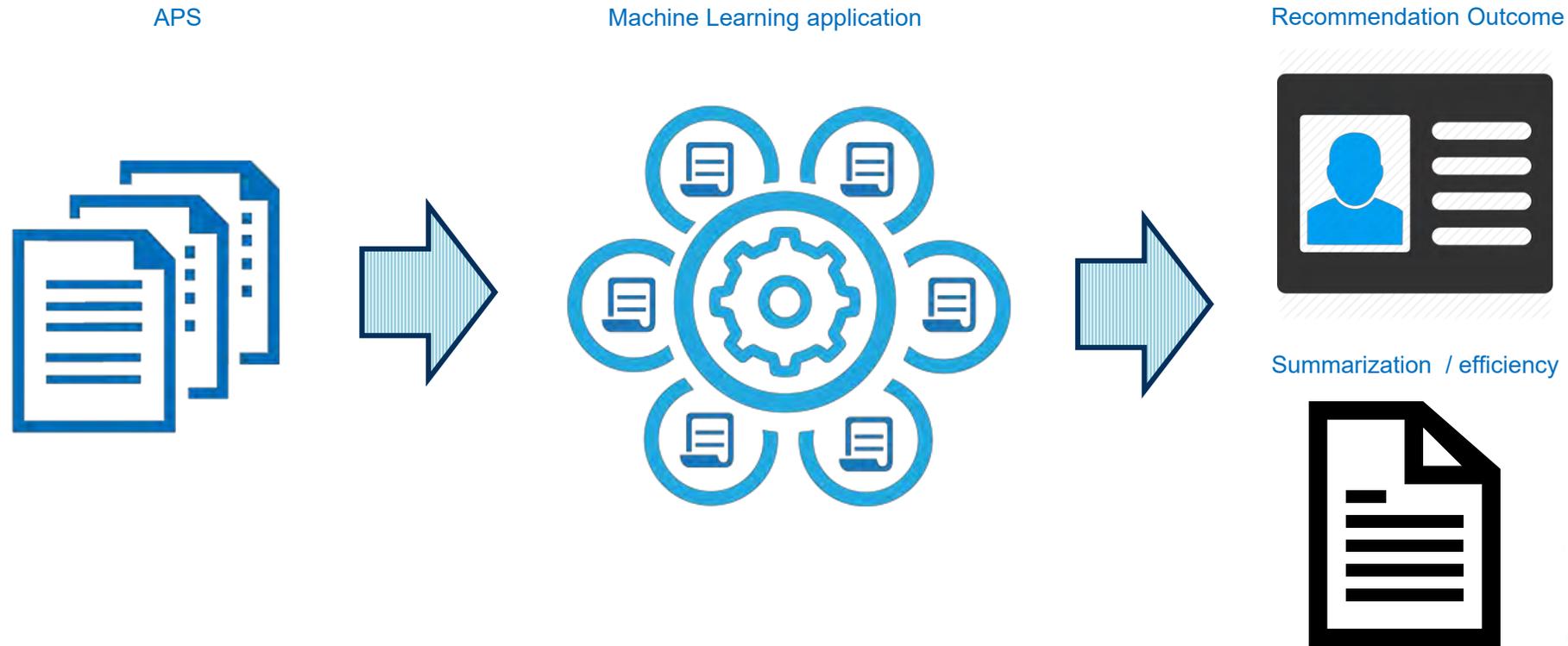
Agenda

- How ML and data engineering helps underwriters
- Understanding the different types of challenges
- Future strategies and directions for AI for APS'

Where we need technology for APS'

- APS' considered the “golden standard” of medical evidence
- Can be very long (hundreds to thousands of pages) and repetitive
- Characteristics vary greatly for different types of cases
- APS' require extra security
- Justification of decisions / analysis of bias

Integrating ML for APS'



Example: Decline recommendation

Metadata from underwriting file

BMI	Gender	Smoking Status	Issue Age	Face Amount
35	Male	Smoker	43	\$150,000

Words extracted from APS

The **patient** has had **surgery** due to **heart problems**.

ML Application

Refer to Underwriter

Recommend Decline

Poll: which is easier for AI / ML?

1) Newspaper article



2) Lab analysis




Certificate of Analysis

Powered by Confident Cannabis 1 of 2

Trim Ready/CBD Hemp Direct Sample: 1809NVC0923-4845

NV 89119 Strain: Juicy Fruit
 info@trimready.com Batch #: NA: Lot #: NA:
 (702) 338-6941 Sample Received: 09/05/2018; Report Created: 09/07/2018
 Lic. #

Juicy Fruit
 Plant, Flower - Cured
 Harvest Process Lot: METRC Batch: METRC Sample:



The photo on this report is of a sample collected by the lab and may vary from the final packaging

Safety

Pass	Pass	Pass
Pesticides	Microbials	Mycotoxins
Not Tested	Pass	Pass
Solvents	Heavy Metals	Foreign Matter

Cannabinoids

<LOQ	<LOQ	15.059%	10.5%
THCa	Total Potential THC	Total Potential CBD	Moisture
Analyte	%	%	mg/g
THCa	0.242	<0.242	<2.42
A9-THC	0.242	<0.242	<2.42
CBD	0.242	5.648	56.68
CBDa	0.242	10.708	107.08
CBC	0.242	0.689	6.89
CBG	0.242	<0.242	<2.42
CBN	0.242	<0.242	<2.42
THCV	0.242	<0.242	<2.42
A8-THC	0.242	<0.242	<2.42
CBGa	0.242	0.224	3.24
CBVD	0.242	<0.242	<2.42
Total		17.389	173.89

Terpenes

Analyte	LOQ	Mass	Mass
		mg/g	%
β-Caryophyllene	0.10	1.47	0.147
α-Humulene	0.10	0.64	0.064
(-)-Guaifol	0.10	0.56	0.056
β-Myrcene	0.10	0.50	0.050
α-Pinene	0.10	0.43	0.043
Caryophyllene Oxide	0.10	0.14	0.014
α-Bisabolol	0.10	<0.10	<0.010
β-Terpinene	0.10	<0.10	<0.010
Camphene	0.10	<0.10	<0.010
δ-3-Carene	0.10	<0.10	<0.010
β-Limonene	0.10	<0.10	<0.010
γ-Terpinene	0.10	<0.10	<0.010
Geraniol	0.10	<0.10	<0.010
Linalool	0.10	<0.10	<0.010
Nerolidol	0.10	<0.10	<0.010
Octenol	0.10	<0.10	<0.010
(-)-β-Pinene	0.10	<0.10	<0.010
(-)-Isopulegol	0.10	<0.10	<0.010
p-Cymene	0.10	<0.10	<0.010
Terpinolene	0.10	<0.10	<0.010

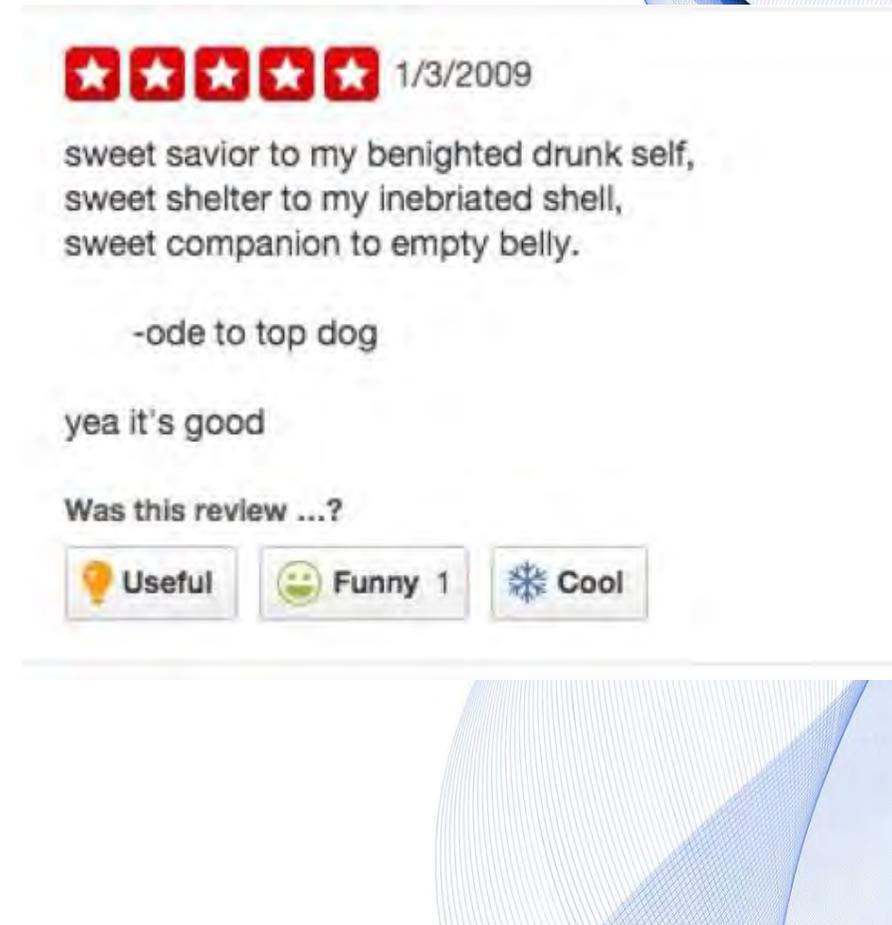
Notes:
 Total THC = THCa * 0.877 + a9-THC
 Total CBD = CBDa * 0.877 + CBD
 LOQ = Limit of Quantitation. The reported result is based on a sample weight with the applicable moisture content for that sample. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.
 LOQ = Limit of Quantitation. The reported result is based on a sample weight with the applicable moisture content for that sample. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
 Las Vegas, NV
 (702) 826-2700
 info@trimready.com

Hui Wang
 Scientific Director

Confident Cannabis
 All Rights Reserved
 support@confidentcannabis.com
 (866) 506-5866
 www.confidentcannabis.com

3) A Yelp review



Live Content Slide

When playing as a slideshow, this slide will display live content

**Poll: Which document is easiest to tackle
for AI/ML?**



A piece of text is easier to deal with when...

- 1) Information is relatively “flat” , as opposed to a hierarchy of concepts
- 2) Visual organization of text is less relevant and, if present, explicit
- 3) Text directly references any visuals provided
- 4) Information is localized, as opposed to distributed across different locations
- 5) Text has relatively little reliance on numbers and special characters/shapes
- 6) There are structured data associated with given block of text

So the easiest is....

 1/3/2009

sweet savior to my benighted drunk self,
sweet shelter to my inebriated shell,
sweet companion to empty belly.

-ode to top dog

yea it's good

Was this review ...?

 Useful  Funny 1  Cool

And the hardest is....




Certificate of Analysis

Powered by Confident Cannabis

1 of 2

Trim Ready/CBD Hemp Direct

NV 89119
info@trimready.com
(702) 338-6941
Lic. #

Sample: 1809NVC0923-4845

Strain: Juicy Fruit
Batch #: NA; Lot #: NA;
Sample Received: 09/05/2018; Report Created: 09/07/2018

Juicy Fruit
Plant, Flower - Cured
Harvest Process Lot: METRC Batch: METRC Sample:





The photo on this report is of a sample collected by the lab and may vary from the final packaging

Safety

Pass	Pass	Pass
Pesticides	Microbials	Mycotoxins
Not Tested	Pass	Pass
Solvents	Heavy Metals	Foreign Matter

Cannabinoids

Analyte	<LOQ	<LOQ	15.059%	10.5%
	THCa	Total Potential THC	Total Potential CBD	Moisture
THCa	<LOQ	<LOQ	15.059%	10.5%
Δ9-THC	0.242	<0.242	<2.42	
CBD	0.242	<0.242	5.668	56.68
CBDa	0.242	10.708	107.08	
CBC	0.242	0.689	6.89	
CBG	0.242	<0.242	<2.42	
CBN	0.242	<0.242	<2.42	
THCV	0.242	<0.242	<2.42	
Δ8-THC	0.242	<0.242	<2.42	
CBGa	0.242	0.324	3.24	
CBDV	0.242	<0.242	<2.42	
Total		17.389	173.89	

Total THC = THCa * 0.877 + Δ9-THC
Total CBD = CBDa * 0.877 + CBD
LOQ = Limit of Quantitation. The reported result is based on a sample weight with the applicable moisture content for that sample. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Cannabinoids analyzed by SOP-021.

Notes:

Terpenes

Analyte	LOQ	Mass	Mass
	mg/g	mg/g	%
β-Caryophyllene	0.10	1.47	0.147
α-Humulene	0.10	0.64	0.064
(-)-Guaiaol	0.10	0.56	0.056
β-Myrcene	0.10	0.50	0.050
α-Pinene	0.10	0.43	0.043
Caryophyllene Oxide	0.10	0.14	0.014
α-Bisabolol	0.10	<0.10	<0.010
α-Terpinene	0.10	<0.10	<0.010
Camphene	0.10	<0.10	<0.010
δ-3-Carene	0.10	<0.10	<0.010
δ-Limonene	0.10	<0.10	<0.010
γ-Terpinene	0.10	<0.10	<0.010
Geraniol	0.10	<0.10	<0.010
Linalool	0.10	<0.10	<0.010
Nerolidol	0.10	<0.10	<0.010
Ocimene	0.10	<0.10	<0.010
(-)-β-Pinene	0.10	<0.10	<0.010
(-)-Isopulegol	0.10	<0.10	<0.010
p-Cymene	0.10	<0.10	<0.010
Terpinolene	0.10	<0.10	<0.010

LOQ = Limit of Quantitation. The reported result is based on a sample weight with the applicable moisture content for that sample. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. Terpenes analyzed by SOP-022.

6631 Schuster Street
Las Vegas, NV
(702) 826-2700
http://www.nvcann.com



Hui Wang
Scientific Director

Confident Cannabis
All Rights Reserved
support@confidentcannabis.com
(866) 506-5866
www.confidentcannabis.com



All pass limits are as specified in MAC 453.A and DPB4 Policies. Unless otherwise stated all quality control samples performed within specifications established by the Laboratory. This product has been tested by NV Cann Labs using valid testing methodologies and a quality system as required by Nevada state law. Values reported relate only to the product tested. NV Cann Labs makes no claims as to the efficacy, safety or other risks associated with any detected or non-detected levels of any compounds reported herein. This Certificate shall not be reproduced except in full, without the written approval of NV Cann Labs. Uncertainty information is available upon request. pH/Water activity is NOT ISO 17025 accredited.

How about APS'?

Place of Service

[REDACTED]

Diagnoses

	Codes	Comments
Celiac disease - Primary	K90.0	
Osteoporosis	M81.0	
Weight gain	R63.5	
Hot flashes	N95.1	
Screening for breast cancer	Z12.39	
Need for pneumococcal vaccination	Z23	
Medicare annual wellness visit, subsequent	Z00.00	

Reason for Visit

Complete Physical Exam
Reason for Visit History

Vitals Recorded in This Encounter

[REDACTED]

BP:	124/70
Pulse:	64
Resp:	20
Weight:	154 lb (69.9 kg)
Height:	5' 2.25" (1.581 m)

Orthostatic Vitals

[REDACTED]
0929

Orthostatic Vitals

BP	124/70
Pulse	64
Positioning	
Site	
Cuff Size	
Release	

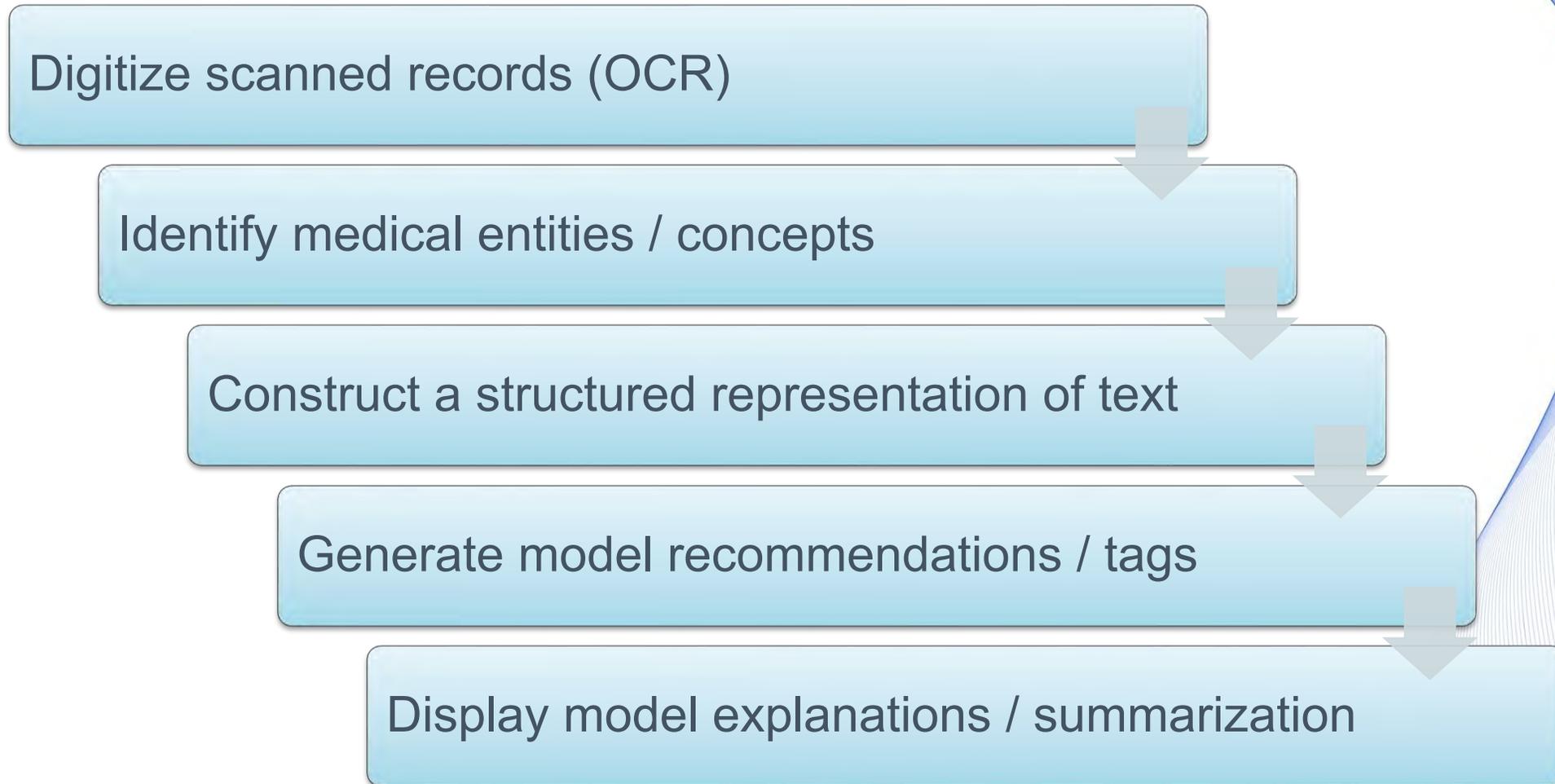
Past Medical History

Diagnosis	Date
• Bronchitis, acute	[REDACTED]
• Tobacco dependence	
• Lung disease, chronic obstructive	
• Osteoporosis	<i>Improved on Rx--osteopenia</i>
• Fatigue	
• Pneumonia due to other virus not elsewhere classified	<i>LLL 1996, cleared</i>
• Abnormal exam/test finding	<i>Elevated SGOT back to 1995, w/u neg</i>
• Colonic polyp	<i>Hyperplastic, CLS 2002</i>
• Celiac sprue	[REDACTED]
• Urinary, incontinence, stress female	
• Sigmoid diverticulosis	[REDACTED] <i>CLS repeat 7-10 years</i>
• Internal hemorrhoid	[REDACTED]
• Bronchiectasis	[REDACTED] <i>mild, chest CT</i>
• Basal cell carcinoma	[REDACTED] <i>chest, Dr. Ferzoco</i>
• Squamous cell cancer of skin of hand	

Current Outpatient Prescriptions:

raloxifene (EVISTA) 60 mg tablet	Take 1 tablet by mouth daily
fluticasone (FLONASE) 50 mcg/actuation Spray, Suspension	1-2 puffs in each nostril daily
montelukast (SINGULAIR) 10 mg Oral tablet	Take 1 tablet by mouth every evening
fluticasone (FLOVENT HFA) 220 mcg/actuation Inhalation HFA Aerosol Inhaler	Use 2 puffs twice daily rinse your mouth thoroughly afterward
fluocinonide 0.05 % Topical Cream	Apply to the affected area twice daily sparingly to affected area; avoid face.

APS processing pipeline



Engineering challenges

- Objective Character Recognition(OCR)
 - Slow, error-prone and non-deterministic
- Occurrence of typos that are hard to fix

Diabates → Diabetes vs 50mg → 60mg

- Checkboxes and lists used for selecting relevant conditions from a wide variety

Modeling challenges

- Numerical representation difficult
 - Word counts?
 - Sequences
 - Priority of concepts tricky
- Not enough data / documents
- Negation detection

$$\begin{array}{l} \text{Vocab} = \begin{bmatrix} \text{I} \\ \text{love} \\ \text{like} \\ \text{cats} \\ \text{dogs} \end{bmatrix} \\ \\ v(\text{I}) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v(\text{like}) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, v(\text{dogs}) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

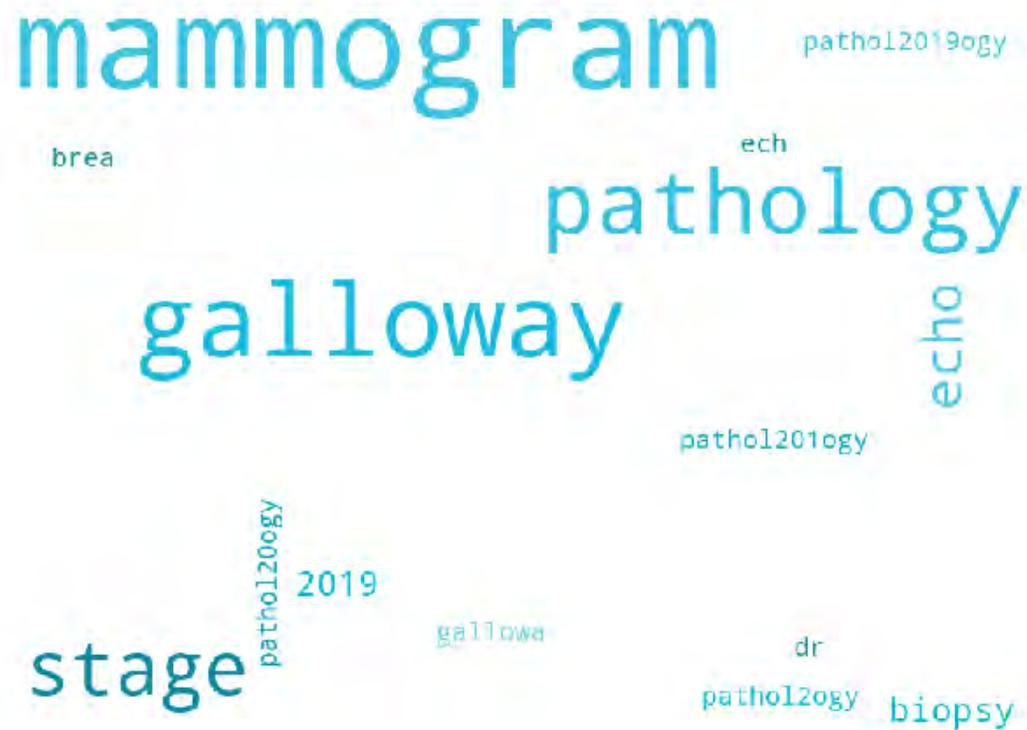
Future strategies

- Rule-based approach:
 - Bottom-up approach
 - Build rules/heuristics to resolve information within paragraphs and sections
- Modeling approach
 - Top-down approach
 - Collect telemetry on underwriter's interactions with PDF
 - Constructing models most receptive to hierarchical presentation of APS'

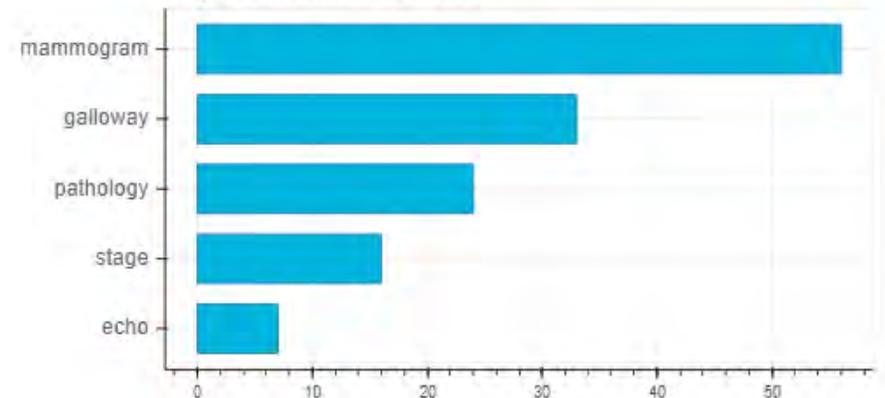
Telemetry analytics

Primary Impairment

Words Searched



Top 5 Searched Words



Thanks!

Q & A