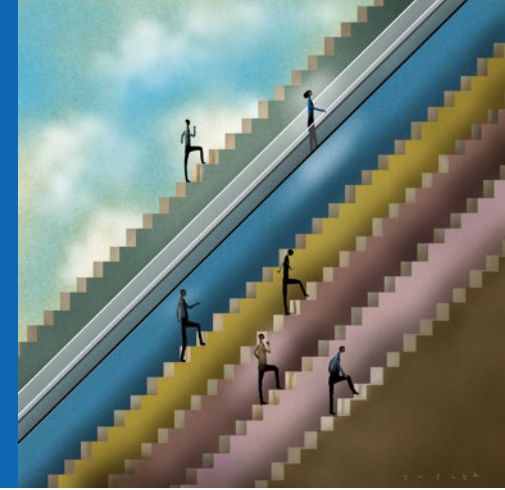


2019

CAQH ProView Symposium

Achieving Provider Data Excellence



Machine Learning and the Future of Provider Data Quality

Amanda Clifton, Senior Director, Provider Data Solutions, Centene

Hassaan Sohail, CAQH Product Manager



Provider Data Goals

Easy to do business

Simplify provider data acquisition and processing from the provider perspective

Compliant

Increase directory accuracy & member transparency

Timely services

Increase speed and accuracy of provider configuration for member assignment, claims payment, call centers, and directories

We are all solving the same problem...

- An “accurate” provider setup varies based on product, contract, state, health plan, and provider type.
- Rules are not universally defined...meaning “standards” are hard to follow or implement.

Claims are processed with multiple inputs and through varying systems/methodologies

- Varies by provider contracts
- Product (Medicaid, Medicare, Commercial, Marketplace) introduces additional variations
- Provider billing practices

Medicaid offerings introduce unique state-by-state variation

- FQHCs and FQHC look-alikes drive us to multiple provider set-ups that differ by state and product
- Medicaid eligibility drives more than just whether we can pay a claim, but also what provider type, billing practices, and locations are eligible
- Member assignment methodologies can also vary by state and product
- Credentialing requirements

What about alternate provider types for complex care products?

- No NPI
- Degrees vary broadly and specialties are less specific
- How do we identify them?

Provider Directory Challenges

49%

Directory locations
are inaccurate

30%-60%

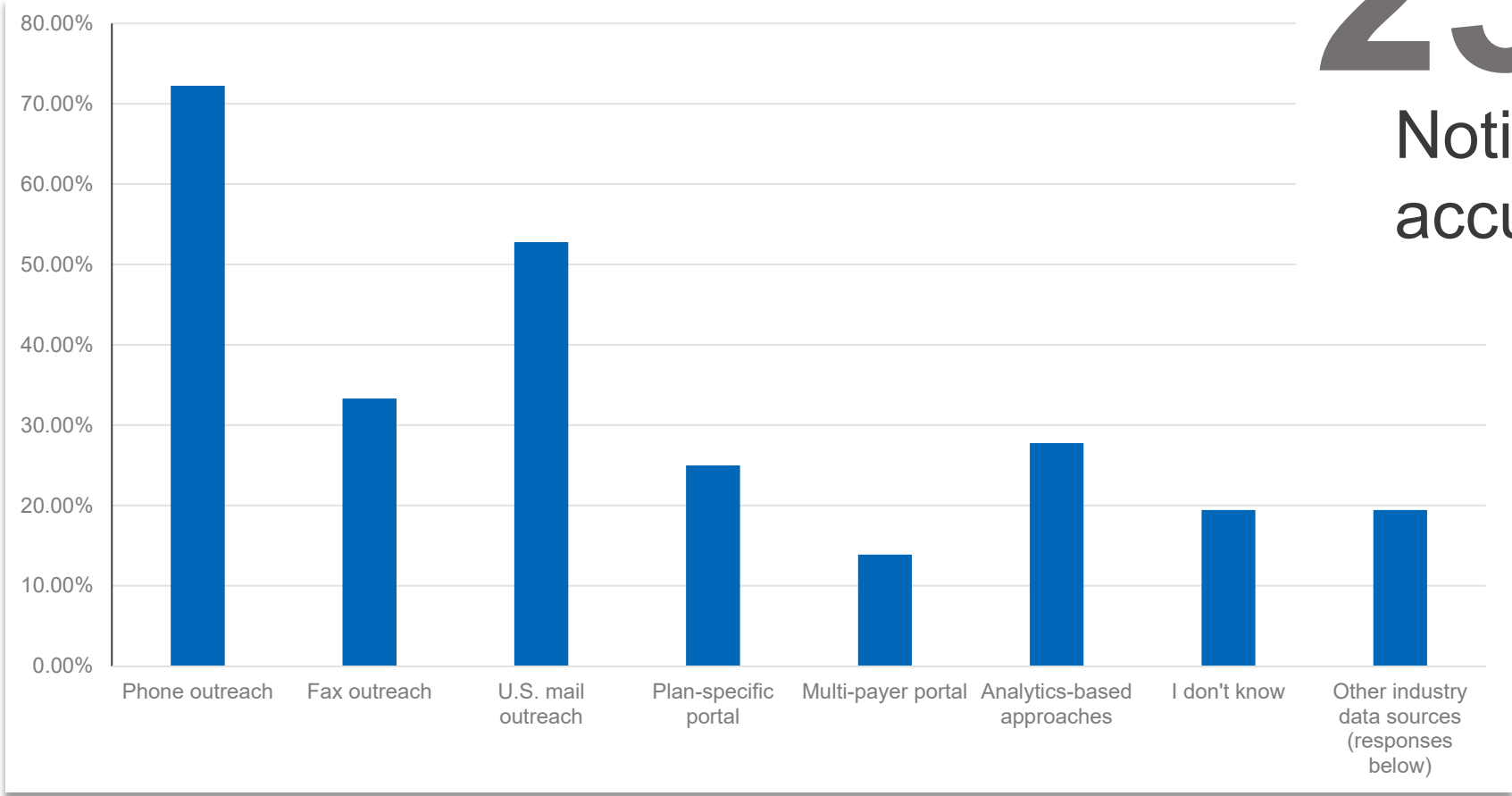
Average inaccuracy rate for
Medicare Advantage plans

66%

Errors identified are where
the 'Provider is not practicing at location.'

* Online Provider Directory Review Report.
Centers for Medicare and Medicaid Services, November 28, 2018.

Health Plans Rely on Phone and Mail Outreach for Data Quality Improvement



25%
Noticed data
accuracy increases

Tackling Top Directory Accuracy Issues

<i>Issues</i>	<i>Tactics & Success/Fail</i>
Practitioner does not practice at the address listed	<i>Lexis Nexis, CAQH, Manual audits:</i> <i>Improvement</i> , but no long-lasting solution yet
Practitioner office hours are incorrect	<i>Lexis Nexis, CAQH, Manual audits, Google API: Unsuccessful</i>
Location phone accuracy	<i>LexisNexis, Audits, Google: Successful</i> , improvement to >90% phone accuracy
Provider group name accuracy	<i>Provider outreach, rosters, Google:</i> Solutions offer <i>improvement</i> but not scale

Challenges with Telephonic Outreach



Practices are receiving **an increased number of outreaches** from health plans



Practices receive **~100 calls or faxes a week**



Call campaigns to providers **negatively impact** patient access and care delivery



Provider Impact

Cost to Health Plans



Annual cost ~\$10-12 to outreach providers



National health **plans are spending millions of dollars** on outreach



50% of plans will reduce phone and fax outreach for directory updates

Providers are Annoyed

“We are constantly inundated with requests for directory updates. Sometimes they won’t even identify WHICH provider they want – its just a generic notification to report our directory data to them.”

“Payers seem to be calling us all the time. I would think if they looked at CAQH, they wouldn’t need to call.”

“Even though my CAQH ProView profile is up to date, insurance companies keep asking to verify directory information.”

“We would much prefer to support CAQH as the one best source of directory info but it seems that is possibly being derailed..... providers feel coerced into reporting provider info to multiple sources.”

Top directory issues shouldn't surprise us

Our problem has layers. A physician or practitioner profile is simple, but their context changes based on patient, product, group/facility and location. A singular profile does not exist.

Physician Profile

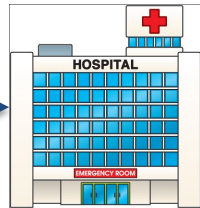


Dr. Jones
Gender: Female
DOB: 5/1/1973
SS#: xxx-xx-xxxx
Specialties: Family Medicine,
Emergency Medicine

Provider Associations



Provider Name: Best Primary Care,
LLC
Provider Specialty: Single Specialty
Group Physicians
Provider Tax ID: xx-xxxxxxx
Provider NPI: 1xxxxxxxxxx



Provider Name: SaveMe Hospital
Provider Specialty: General Acute Care
Hospital
Provider Tax ID: xx-xxxxxxx
Provider NPI: 1xxxxxxxxxx



Provider Name: SaveMe UrgentCare
Provider Specialty: Urgent Care
Provider Tax ID: xx-xxxxxxx
Provider NPI: 1xxxxxxxxxx

Locations /Office Hours

123 Sesame Street
St. Louis, MO

Monday – Friday
Practices Family Med, Primary Care

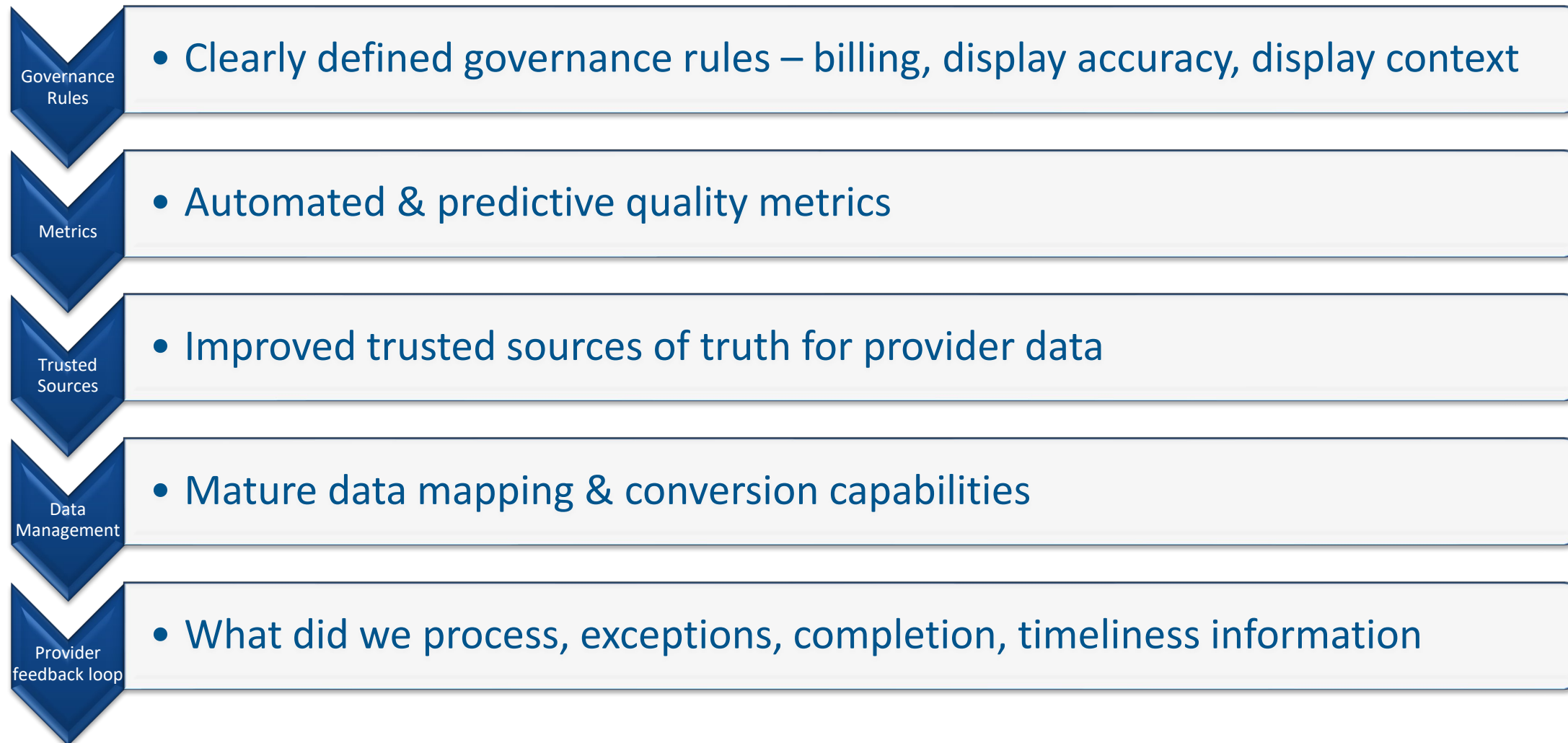
101 Main Street
St. Louis, MO

Once a month, weekends
Practices Emergency Med

201 Main Street
St. Louis, MO

One Friday per month
Practices Emergency Med

Solution Needs



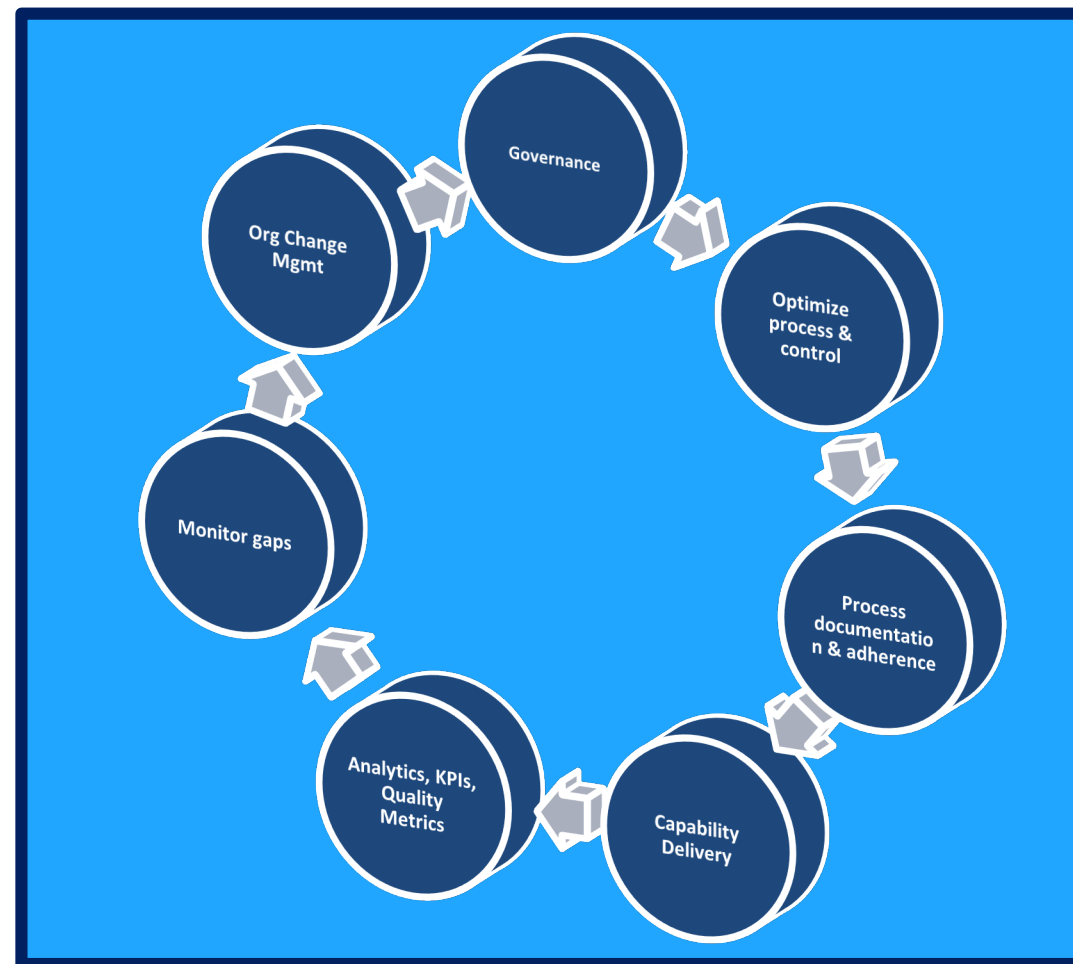
Continuous Improvement Approach

Establish governance to drive policy and process adherence monitoring

Weave quality control into data processing, and enforce process standards

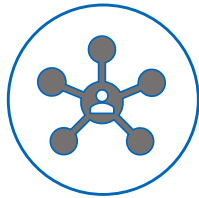
Deliver capabilities to drive efficiency and automate where possible

Measure baseline, continuously improve, re-measure



Machine Learning + Provider Engagement

The CAQH Advantage



- Large repository of provider attested data in CAQH ProView.
- 1.5 million providers.
- Up to 800 data elements per provider.
- Nearly 1M providers attesting on a frequent basis.
- Health plan data across organizations via DirectAssure.

Case for a Machine Learning Solution.



- Can predict probability of a provider at a given location.
- Does not rely on a single source of truth.
- High development cost but low run cost in comparison to conventional solutions.
- Eliminates provider abrasion.

Using a Probabilistic Model to Predict a Deterministic Outcome

Real World

Dr. Joe Chen 123 Main Street



- In the real world, either the provider is at the location or they are not.
- Results are deterministic.
- Making a phone call to the office can provide these results.

Probabilistic Machine Learning Model

96%

Dr. Joe Chen, 123 Main Street



70%

Dr. Sally Patel, 345 Maple Street

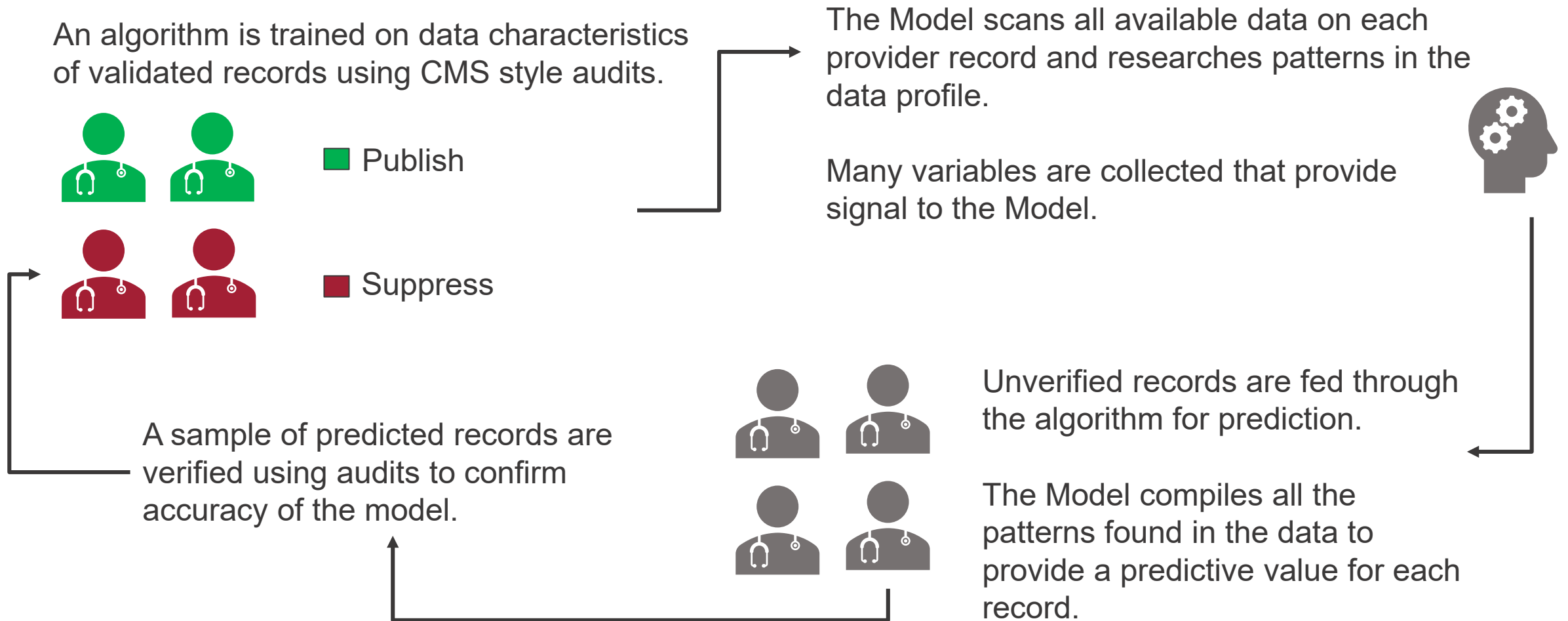
15%

Dr. Veera Davis, 932 Lee Highway



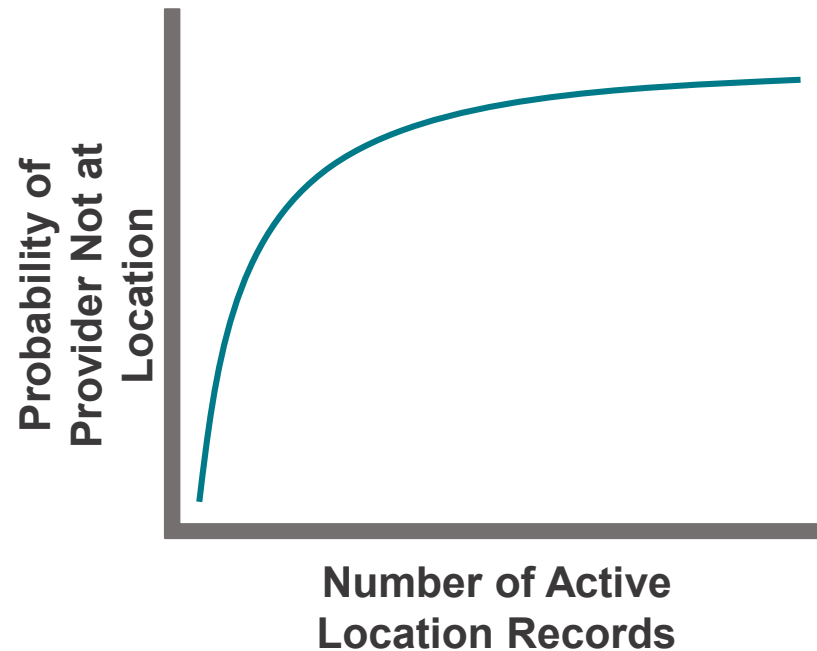
- In a probabilistic model, results are on a spectrum of probabilities.
- Records with very high or very low probabilities can be used to confidently make business decision.

How Does CAQH's Machine Learning Model Predict Provider at Location?



Various Data Elements Can Become Variables that Have Predictive Value

Hypothesis: The presence of other locations decreases the probability that the location in question is incorrect.



Increase in the number of active locations is the variable that is created

Over 100 Variables to Predict the Accuracy of Directory Records

1

Number of Practice Locations

Number of practice locations in a provider's profile where longer list associated with greater DNPL.

2

Number of Providers Phone that Share Phone Number

Takes the union of associated phone networks and computes the total number of providers.

3

Correlated List Length

Number of records in lists, where the length is correlated to DNPL (e.g. credentialing contacts).

4

Number of Providers that Share a TIN

Takes the union of associated Tax ID networks and computes the total number of providers.

5

Correspondence Address

Reviews if provider has listed this location as a place to receive correspondences.

CAQH Developed Scores that Assess the Quality of Provider-Locations Records

Provider Details

Name: Bob Patel
DOB: 08/10/1980
Gender: Male
Specialty: Cardiology
Board Certification: Cardiology
License ID: 123123123123
License State: VA

Last Attest Date: 05/11/2018
Currently Practicing: Y
Type 1 NPI: 139212392
DEA: AC1239182
CDS: M12322
Education: UPenn Medicine, MD
Hospital Affiliation: GW Hospital
Medicare ID: 1231212
Medicaid ID: 1232902

Location 1

123 Main Street
Suite A
Washington DC 20016
(202)123-1231

Address Phone Align:
Accepting Patients:
Provider at Location: **TRUE**
Specialty at Location: Cardiology



Location 2

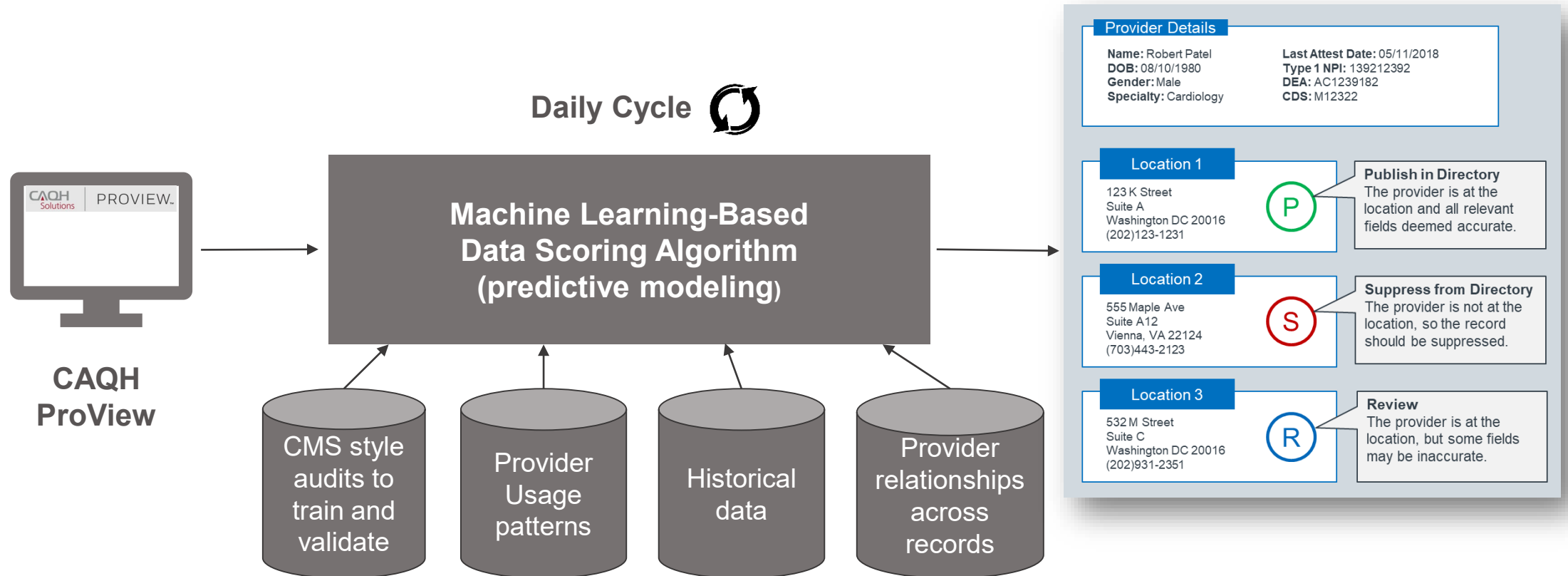
421 Apple Street
Suite 231
Washington DC 20018
(202)432-1232

Address Phone Align:
Accepting Patients:
Provider at Location: **FALSE**
Specialty at Location: Cardiology

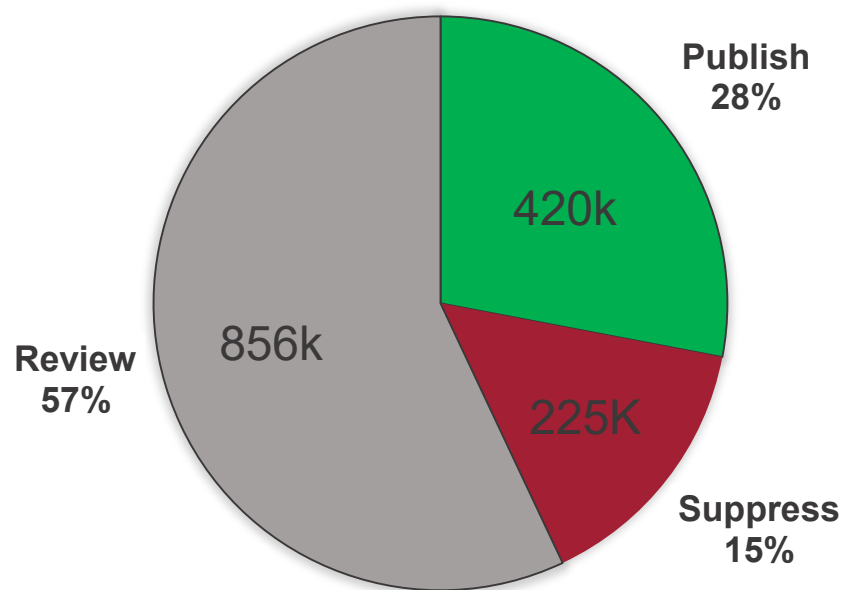


Score	Description	Confidence
Publish	The provider is at the location and should be published in a directory.	90% probability of accuracy
Suppress	The provider is not at the location, so the record should be suppressed from directories.	80% probability of accuracy
Review	We were unable to make a confident assessment using the data available.	

How the Scoring Model is Designed



Machine Learning Scores for Provider Location Records in CAQH ProView



- Using only the machine learning model over 600k records in CAQH ProView can be validated with a high degree of confidence.
- Scoring is assessed for nearly 200 provider specialties, focusing on those listed in health plan directories.
- Improvements to coverage of the model are currently in progress.

Health Plans ‘Publish’ Scores at Greater Than 90% Precision

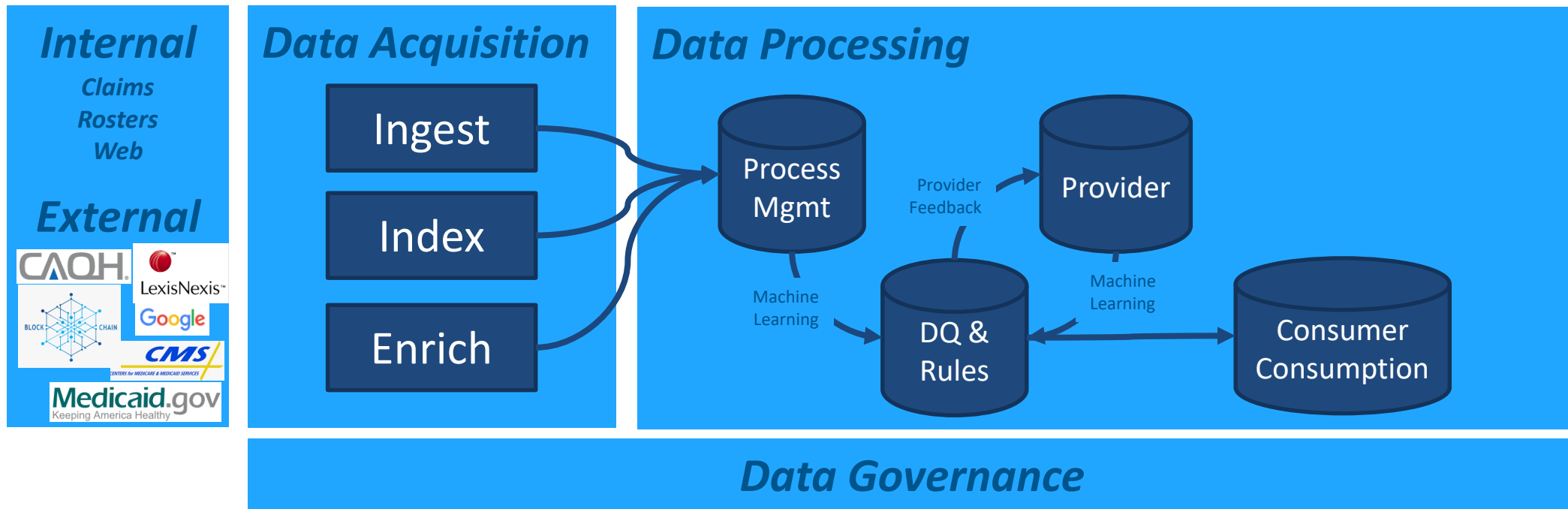
- Accurate records as determined by the algorithm (with 90% confidence) were randomly sampled and shared with some CAQH member plans.
 - “Accurate” is currently defined as attested provider locations that are predicted to pass a CMS-style audit for provider at location data element. The definition will be expanded as the model is expanded to cover additional data elements.
- These member plans were asked to validate the scoring algorithm.
- 6,685 calls were completed by these health plans. The validation rate was 94% (i.e., the score was correct in 94% of the time), compared to the goal of 90%.

Plan	Completed Calls*	Data Confirmed	Percentage Confirmed
A	962	937	97%
B	742	691	93%
C	515	512	99%
D	973	878	90%
E	994	903	91%
F	1,500	1,402	93%
G	550	536	97%
H	448	413	92%
Total	6,685	6,272	94%

*Attempted but not completed calls were omitted from the analysis to be consistent with CMS methodology.

Building & Scaling Intelligent Data Capabilities

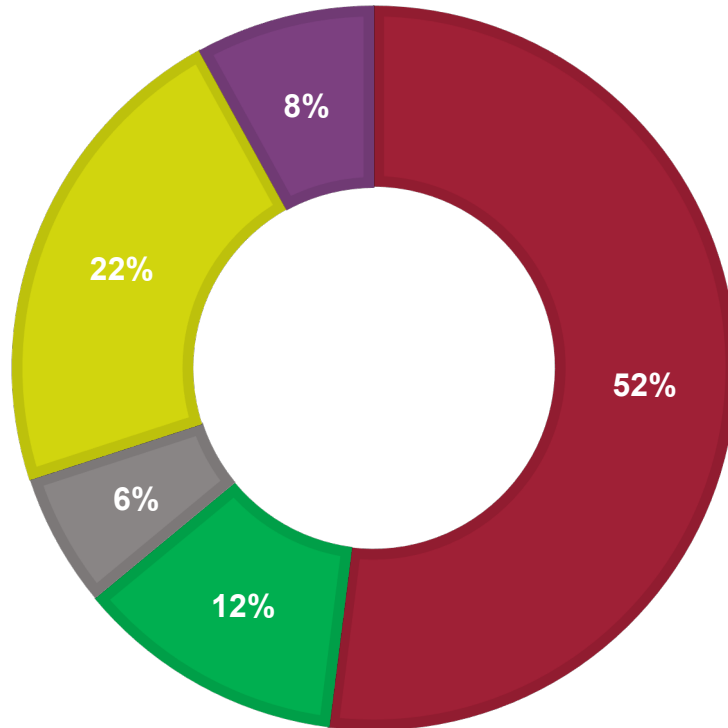
1. Define standards & exception processing
2. Establish approach & data governance **process**
3. Monitor the direction of the industry
4. Review capability needs. Capabilities delivered must follow process.



Use Case – Incorporate ML Scores Alongside Provider Responses to Validate Accuracy of Health Plan Networks at Meaningful Volumes

HEALTH PLAN NETWORK IN GEORGIA

■ Suppress ■ Publish ■ Non-Directory
■ Likely Publish ■ Likely Suppress



- CAQH reviewed one health plan's network in the state of Georgia and was able to provide high confidence indicators on 70% of records.
- Leveraged data from DirectAssure in combination with machine learning technique allows automated directory validation to cover a larger segment of the population.
- Automated validation allows cost savings/calls prevented on an on-going basis.

Use Case—Enhance Attestation Workflow and Clean Up Data Upstream

Make sure to enter all group/practice information in the Employment Information section of your profile.



Our research indicates that you do not practice at one or more of the locations listed below. Please confirm the status of your active practice locations in order to share accurate information with health plans.

PRACTICE LOCATIONS

[Import](#) [ADD](#)

Physician Group/Practice Name	Tax ID	Location	Actions
Johnstown Hospital	23-4918194	3903 HANGING MOSS RD JACKSON, MS 39206-4710 Phone: 738-229-4837	Primary Practice Edit Archive
Please Archive Albertson Family Medicine	03-3929103	2304 Seven Mile Lane Des Moines, IO 40292-1393 Phone: 530-239-2203	Archive Confirm Active I Don't Know
Please Archive Smith Family Medicine	03-3929789	6534 Stills Lane Des Moines, IO 40292-1393 Phone: 530-239-2203	Archive Confirm Active I Don't Know

[Save and Go Back](#) [Save and Continue](#)

- Introducing machine learning insights into CAQH ProView's attestation workflow is scheduled for Summer 2019.
- Enables upstream data scrub where scores indicate the provider is not practicing at a given location.
- New workflows will capture provider behavior to be used as training data for future iterations of our machine learning models.

Learn How Machine Learning Can Improve Your Provider Data Quality



Hassaan Sohail

Data Quality Product Manager

CAQH

HSohail@CAQH.org

- Multiple health plans are implementing machine learning scores as early adopters.
- Contact me or your Account Manager if you are interested in learning more about machine learning.
- Let's schedule an initial conversation to see how your organization would utilize this capability to improve your provider data quality.

Contact



Hassaan Sohail

Data Quality Product Manager
CAQH

HSohail@CAQH.org



DirectAssure.org
[@CAQH](https://twitter.com/CAQH)



Amanda Clifton

Senior Director, Provider Data Solutions
Centene

