CMS 2024 Quality Conference

Resilient and Ready Together

Creating an Optimal Environment for Quality Healthcare for Individuals, Families, and Communities

Infusing Technology and Innovation to Reduce Burden and Improve Services for Medicare Beneficiaries

April 2024

CENTERS FOR MEDICARE & MEDICAID SERVICES

Leaning Objectives

- Strategies and impact of using technology to reduce provider burden
- Strategies and impact of using technology to improve delivery of beneficiary and family-centered care
- Strategies and impact of using technology to protect the Medicare Trust Fund

CMS 2024 Quality Conference

Resilient and Ready Together

Creating an Optimal Environment for Quality Healthcare for Individuals, Families, and Communities

Wendy Gary (moderator)

Avar Consulting

ED BFCC-QIO Operations

Brian Salzer

Avar Consulting

Data Science Team Lead

Jed Shakarji

Avar Consulting

Data Analyst II

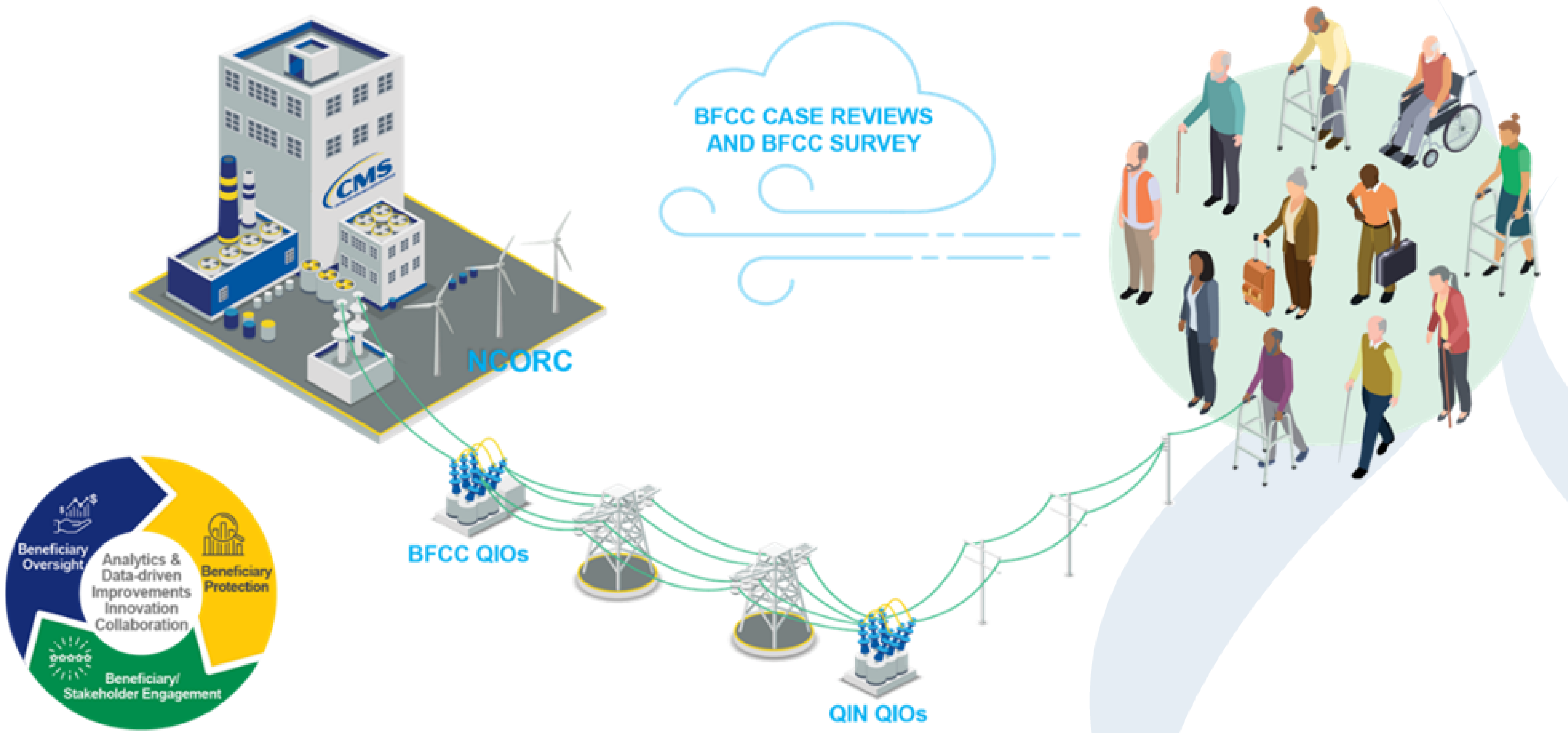
Josh Dominick

Kepro

Director, Business Intelligence

Core Functions of the QIO

Beneficiary and Family Support through the QIO Program



CMS 2024 Quality Conference
Resilient and Ready Together
Creating an Optimal Environment for Quality Healthcare for Individuals, Families, and Communities

Background

Brian Salzer

Analyses & Reviews

- Avar Consulting performs several types of analyses and reviews on behalf of the BFCC NCORC to ensure Medicare beneficiaries are receiving the care that they deserve

- It is not feasible to review everything, so random sampling is used to identify which data to manually investigate

- Data can generally be divided into 2 categories:
  - Structured Data
    - Well defined and easily searchable
    - Examples: Databases, Spreadsheets
  - Unstructured Data
    - Difficult to categorize and search
    - Examples: Text, Images

Medical Record Reviews

- Avar physicians and nurses manually review samples of medical records to determine the prevalence of patient safety events (PSEs) during hospital stays

- Medical records are a good example of unstructured data
  - Medical records in this context are medical documentation from the date of admission through the date of discharge that includes all pertinent information contributed to by the entire inter-disciplinary team.

- Medical records of a beneficiary's stay are received in the form of PDF files from providers
  - They come in a variety of different formats and layouts
  - They can contain text, image, and images of text
  - They can be thousands of pages long

Hurdles to Automating the Review of Unstructured Medical Records

Need for Human Expertise

- Computers are not medical professionals
- The developers are not medical professionals either
- Clinical team has years of experience and Physicians are currently practicing

Properly handling PHI/PII

- Shouldn't be leveraging third party tools/APIs
- All data must be processed in house in a secure environment
- The tools may also not be useful due to our use case

Pertinent data

- Only 5% of a medical record contains information relevant for review objectives
- No knowledge of contents without manual review
- Medical records also contain significant amounts of similar data but non-pertinent

EMR Vendors

- Medical records are generated by a variety of electronic medical record systems/vendors
- Medical records formats vary greatly
- Even within the same vendor, formats are different

Data Quantity / Noise

- Medical records could require OCR which has known issues with whitespace
- Large quantity of irrelevant information/noise
- Counting on human input into the EMR systems

CMS 2024 Quality Conference
Resilient and Ready Together

How to increase efficiency of the review process

- Structured data approach:
  - Use data from CMS systems combined with publicly available data to assign likelihoods of review outcomes before the review takes place
    - Focus more reviewer resources on cases that are predicted to be more likely to have a negative outcome
    - Internal quality control: Select cases in which a reviewer's conclusions don't match the predicted outcome
- Unstructured data approach:
  - Use data from the medical record PDF's themselves
    - Identify pertinent areas of the records so staff won't need to find them
    - Once pertinent areas are identified additional techniques can be applied to achieve similar results to the structured approach
- These technologies will not replace clinician reviews, rather it will allow them to:
  - Be conducted with greater efficiency
  - Produce more accurate reviews of beneficiary data

CMS 2024 Quality Conference

Resilient and Ready Together

Creating an Optimal Environment for Quality Healthcare for Individuals, Families, and Communities

Predictive Model using Structured data

Brian Salzer

CMS

CENTERS FOR MEDICARE & MEDICAID SERVICES

Approach

- Through our work, we have built datasets of manually reviewed cases and outcomes
- Using those datasets we successfully trained models on CMS data in combination with publicly available data to predict outcomes before a human review even takes place
- This allows us to prioritize the selection of cases consisting of specific scenarios

Data Sources

Neighborhood

- U.S. Census
- Area Deprivation Index
- Social Vulnerability Index
- Health Literacy

Hospital

- Medicare Hospital Quality Chartbook
- Medicare Hospital Care Compare Measures
- CMS Hospital Provider Cost Report

Patient

- Claim Data
- HCUP Nationwide Readmissions Database
- Case Mix Index

Variable Selection

- 3,076 variables to select from
  - 648 individual variables
  - 2,428 interaction terms of any two variables

Variable Selection – Lasso Regression (1 of 2)

- LASSO Regression
  - Least Absolute Shrinkage and Selection Operator
  - Type of machine learning for building models
- Purpose
  - Removes unimportant variables from our selection
  - Prioritizes less variables to reduce over fitting of a model

Variable Selection – Lasso Regression (2 of 2)

- When analyzing PSEs in medical records, we sample 4,000 medical records
- For the same level of effort we can shift the sampling methodology to identify 104% more hospital stays with PSEs

Random Sampling

Selecting 4000 cases randomly, 1,610 cases with PSE's will be found (40%)

Risk-Based Sampling

Selecting 4000 cases with the highest risk score, 3,283 cases with PSE's will be found (82%)

CMS 2024 Quality Conference
Resilient and Ready Together

How will this help improve review efficiency?

- Can automatically screen medical records containing specific healthcare outcomes to conduct analysis
  - Can be used to automatically categorize records by specific healthcare outcomes
  - Reviewers are only given records deemed to have a relevant outcome
- Can be used to quality control purposes for reviews
  - Example: Flag reviews where the predicted healthcare outcome does not match with the reviewer's outcome

Approach Limitations

- Limitations/Hurdles
  - Variables we have public access to are limited
  - Variables were selected by an algorithm and not a trained expert
    - Selected variables may not be meaningful clinically
  - The predictive model may change when adding more data or variables

CMS 2024 Quality Conference

Resilient and Ready Together

Creating an Optimal Environment for Quality Healthcare for Individuals, Families, and Communities

Medical Record Content Analysis

Jed Shakarji

CMS CENTERS FOR MEDICARE & MEDICAID SERVICES

Medical Records

- The unstructured nature of medical records creates a considerable amount of potentially unproductive load on clinicians
    - To make the jobs of clinician reviewers more efficient, we must first:
        - Be informed of the structure and contents of data
        - Know where to look for pertinent information in unstructured data (text)

- The records come from a number of medical providers and can be sent from the provider to CMS in a variety of ways which makes traditional methods of analyzing forms unsuitable

Methods to apply structure to medical records

- Keyword Search
  - Regular Expressions
- EMR Vendor Identification Model
  - LGBM Classifier
- Medical Record Section Identification Model
  - LGBM Classifier
  - LSTM

Pertinent data

- Only 5% of a medical record contains information relevant for review objectives
- No knowledge of contents without manual review
- Medical records also contain significant amounts of similar data but non-pertinent

Concepts: LSTM (Long Short Term Memory Network), RNN (Recurrent Neural Network), LGBM (Light Gradient-Boosting Machine)

Using Keywords to Search for Pertinent Data

- Logical/Simple approach to locating pertinent information within medical records is to perform a search for keywords
  - Example: searching for the word "discharge" when trying to locate the discharge summary section
- Results
  - Data Quality Problems (OCR)
  - Keywords indicating sections appearing in other sections
  - Very poor

Data Quantity / Noise

- Medical records could require OCR which has known issues with whitespace
- Large quantity of irrelevant information/noise
- Counting on human input into the EMR systems

EMR Vendor Classification

- Medical records are semi-automatically generated using Electronic Medical Record (EMR) processing tools
  - There are many vendors who have developed such tools
  - Each vendor structures their medical records differently from one another
- Results
  - Developed a supervised model to classify medical records
  - Provided little predictive power to other models

EMR Vendor

- Medical records are generated by a variety of electronic medical record systems/vendors
- Medical records formats vary greatly
- Even within the same vendor, formats are different

CMS 2024 Quality Conference
Resilient and Ready Together

Medical Record Section Classification Model

- As a part of our process, we have a dataset of pages containing pertinent information
- There are 4 sections that we target and have information on which pages they are located on
  - Those sections account for ~5% of a medical record

Pertinent data

- Only 5% of a medical record contains information relevant for review objectives
- No knowledge of contents without manual review
- Medical records also contain significant amounts of similar data but non-pertinent

Medical Record Class Imbalance

- While many machine learning classification models may be trained with imbalanced datasets, the imbalance usually has a negative effect on learning outcomes
  - Depending on the model, they would quit without results

- To address the imbalanced classification, we sampled the oversized class of non-pertinent information to reduce it to the same size of other classes
  - Data manipulation techniques called Random Undersampling

Results

- We’ve successfully built a classification model to identify pertinent sections which are supplied to clinicians
  - EMR Vendor Identification Model
    - Accuracy = 87.5%
  - Medical Record Section Identification Model
    - LGBM Accuracy = 93.8%
    - LSTM Accuracy = 91.5%

Concepts: LSTM (Long Short Term Memory Network), RNN (Recurrent Neural Network), LGBM (Light Gradient-Boosting Machine)

CMS 2024 Quality Conference

Resilient and Ready Together

Creating an Optimal Environment for Quality Healthcare for Individuals, Families, and Communities

Data and Innovation

Josh Dominick

CMS

CENTERS FOR MEDICARE & MEDICAID SERVICES

Key Objectives

Strategies and impact of using technology to…

Reduce provider burden

Improve delivery of beneficiary and family-centered care

Protect the Medicare Trust Fund

Analytics Intelligent Document Processing (AIDP) Pipeline

What is an AIDP Pipeline?

- AIDP is technology that uses artificial intelligence (AI) algorithms and techniques to extract data and insights from documents
- AIDP typically consists of several stages:
  - **Document ingestion**
    - Involves scanning and uploading of documents through a pipeline
  - **Preprocessing**
    - Involves various techniques to prepare documents for analysis
  - **Data validation**
    - Process of checking extracted data for accuracy and completeness
  - **Data extraction**
    - Stage where relevant data is identified and extracted from documents
  - **Data exportation**
    - Storage and presentation of data for gained value and insights

CMS 2024 Quality Conference
Resilient and Ready Together

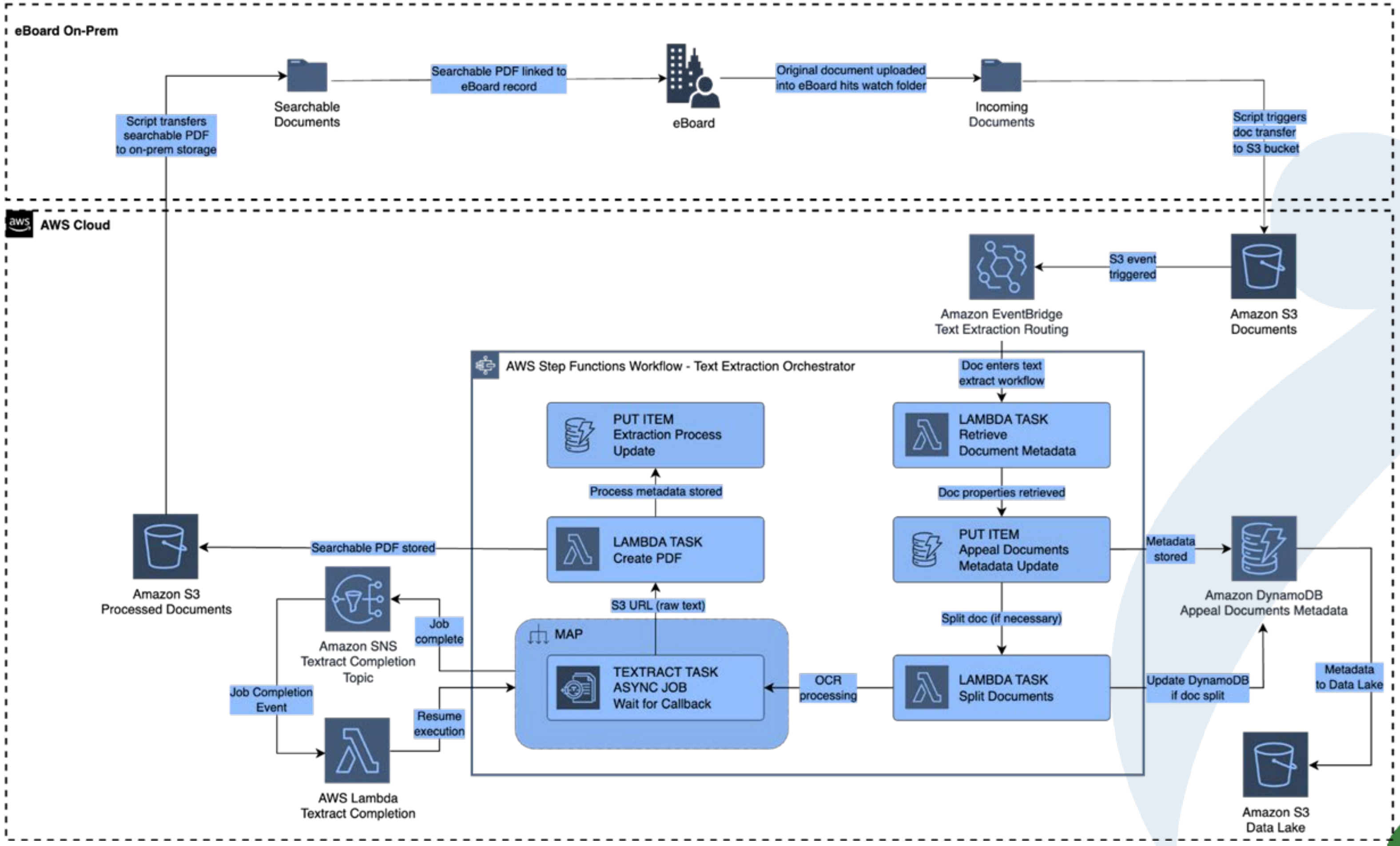
Analytics Intelligent Document Processing (AIDP) Pipeline – (continued)

What is the added value of an AIDP Pipeline?

- Improved efficiency and accuracy in processing, saving time and reducing errors
- Improved data quality, improving outcomes and supporting decision making
- Reveal key insights and value that can otherwise be difficult to identify
- Integration with existing systems & workflows

AIDP Architecture

A robust, modern architecture designed for current and future capabilities.

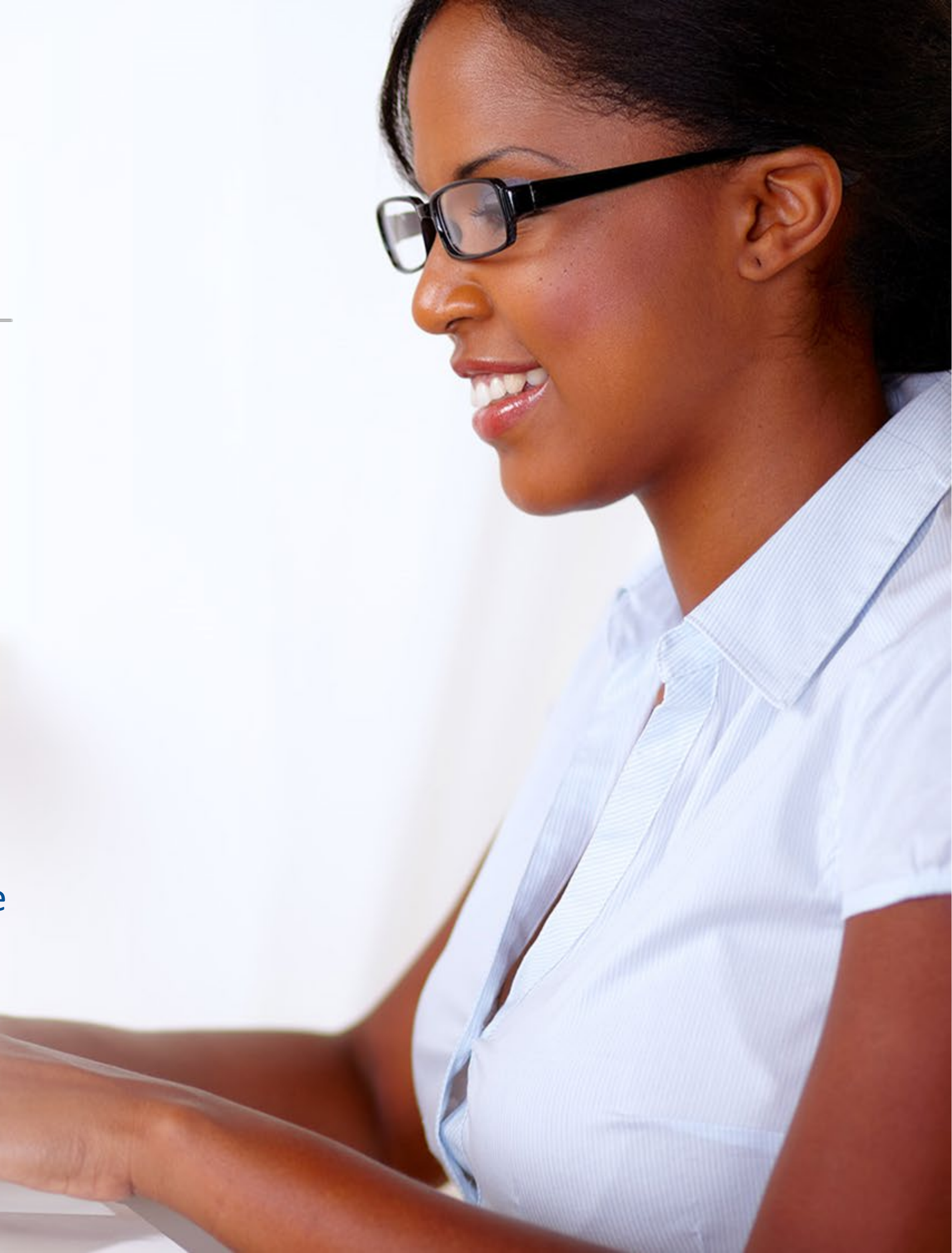


Reducing Provider Burden

Document Ingestion | Preprocessing | Data Validation | Data Extraction | Data Exportation

- Barcode scanning and medical record renaming
  - Ingest incoming faxes and submissions to rename files for quicker identification and workflow routing
  - Benefit to providers is more streamlined submission of medical records, increasing turnaround time (TAT)

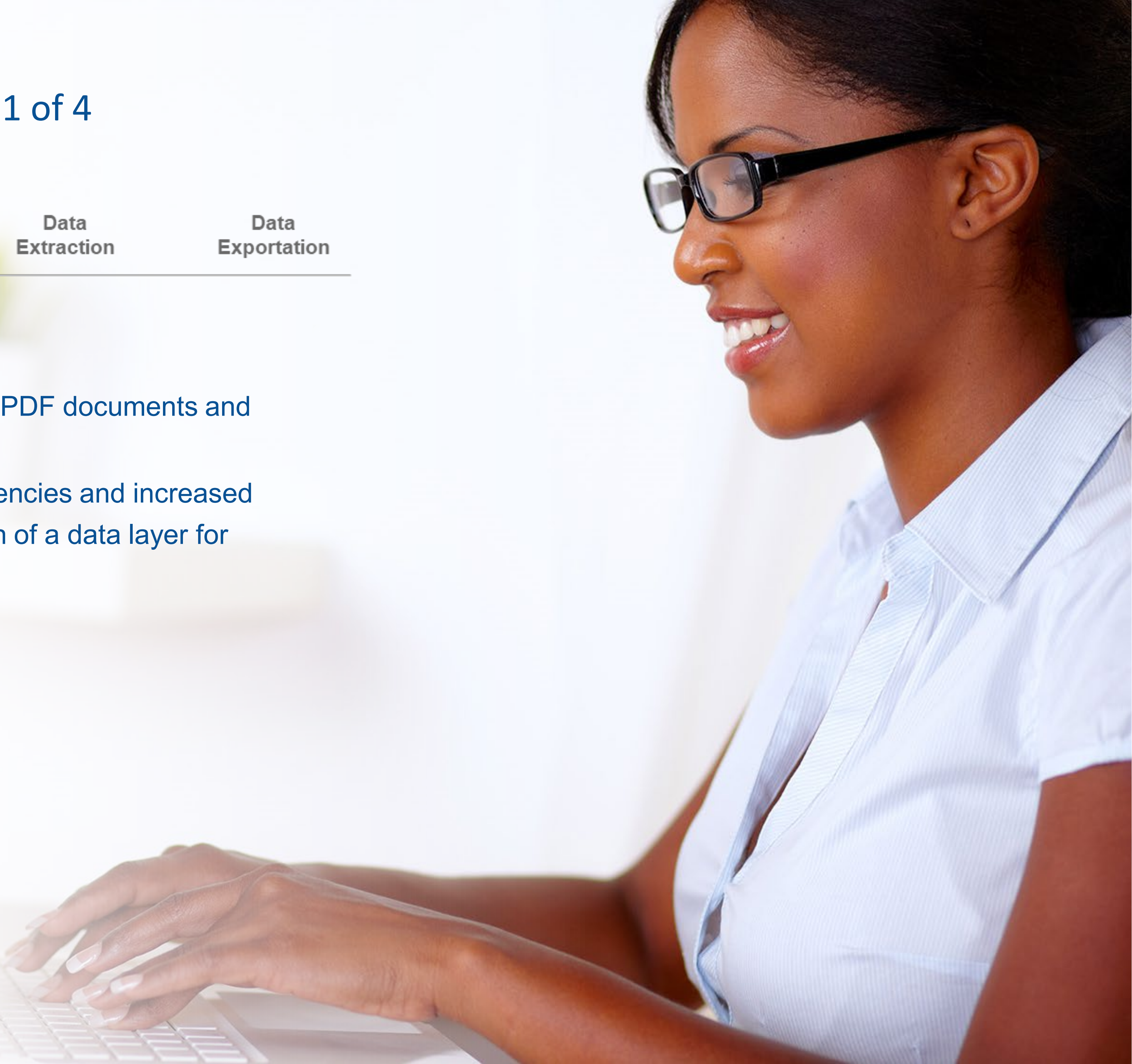
- Metadata extraction and tagging
  - Extract metadata from faxes and submissions, including page counts, submitting IDs, quality of images, etc.
  - Benefit to providers is less back and forth when identifying missing records and submissions



Improve Delivery of BFCC 1 of 4

Document Ingestion | **Preprocessing** | Data Validation | Data Extraction | Data Exportation

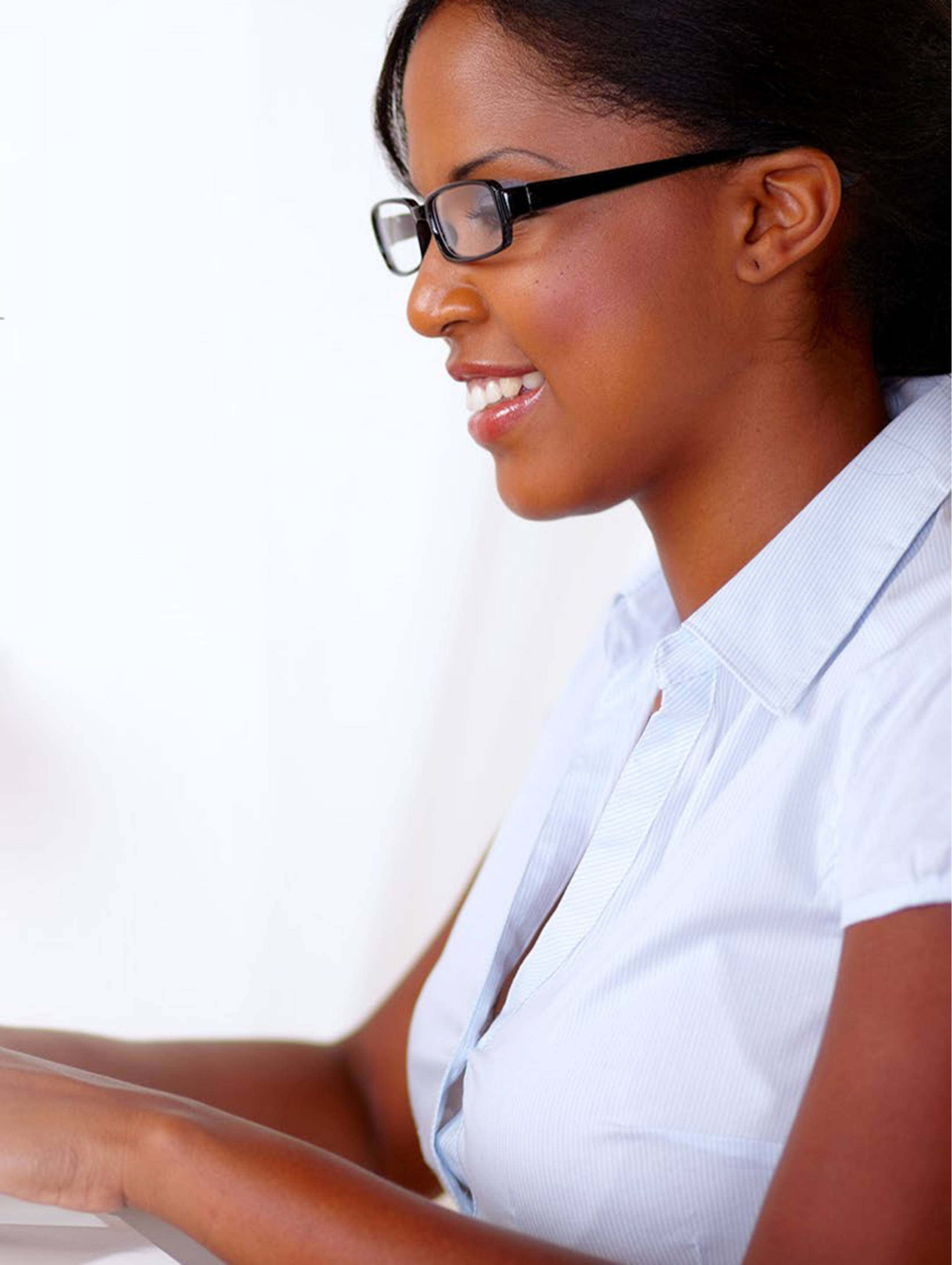
- Optical character recognition (OCR)
  - ➢ Converts image files to searchable/copyable PDF documents and creates a data layer for analysis
  - ➢ Benefit to the BFCC-QIO is operational efficiencies and increased accuracy of transcriptions, as well as creation of a data layer for analysis and quality assurance



Improve Delivery of BFCC 2 of 4

Document Ingestion | Preprocessing | **Data Validation** | Data Extraction | Data Exportation

- Accuracy and logic analysis
  - Programmatically check accuracy of OCR to determine if it meets threshold standard
  - Page number comparison
  - Key value pair relationship and pattern matching analysis
  - Benefit is more consistent data and higher quality extraction

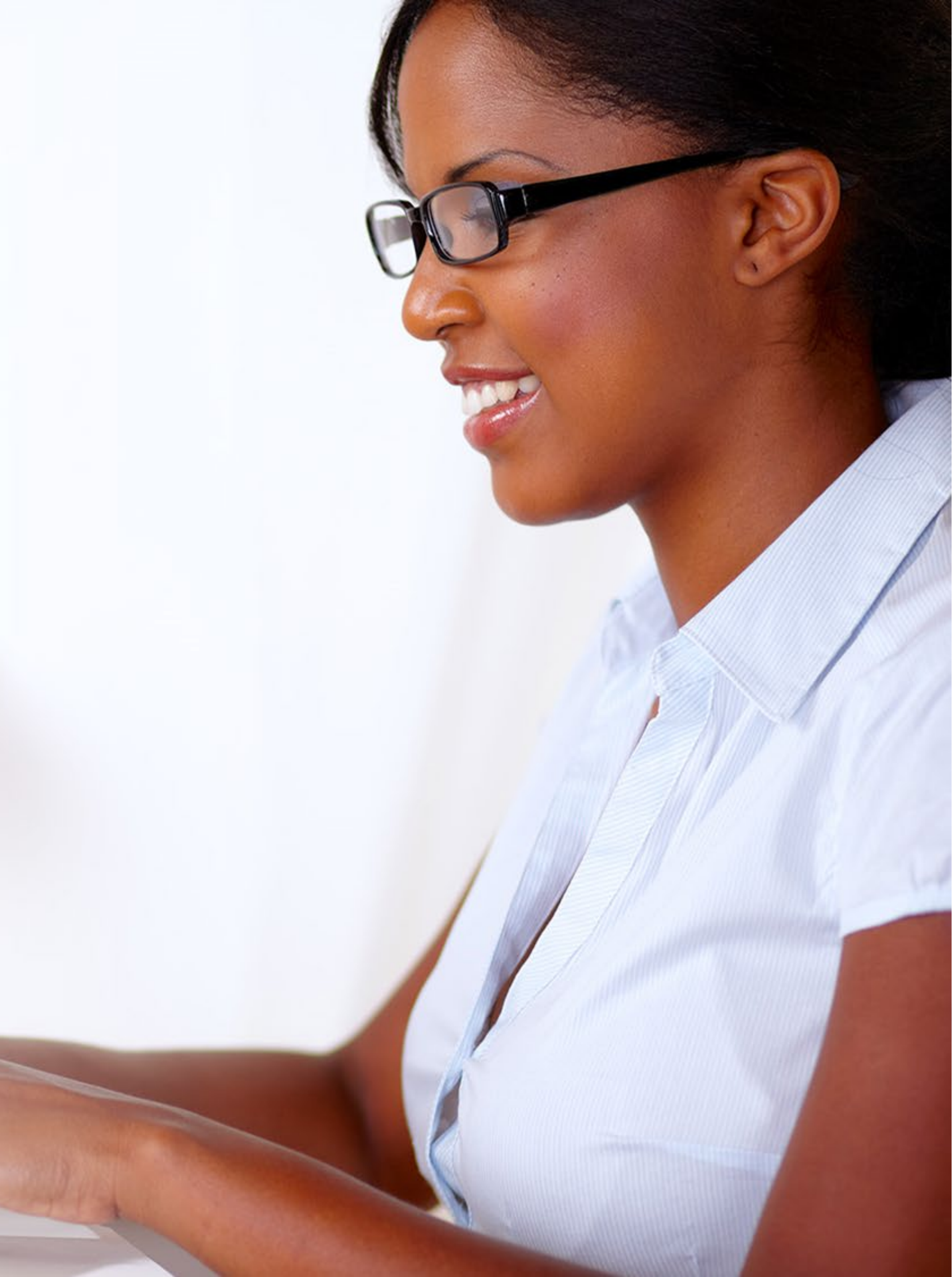


Improve Delivery of BFCC

Document Ingestion | Preprocessing | Data Validation | **Data Extraction** | Data Exportation

- Intelligent bookmarking
  - Leveraging rule-based as well as large language model (LLM) techniques to intelligently bookmark key sections of medical records
  - Benefit to the BFCC-QIO is operational efficiencies through decreased manual work and higher coverage of automation

- Physician review assessment form (PRAF) automation
  - Leveraging rule-based as well as large language model (LLM) techniques to intelligently search, classify and summarize key findings within records
  - Benefit to the BFCC-QIO is operational efficiencies through standardization of criterion selection and identification



Intelligent Bookmarking: *A Closer Look*

Document Ingestion | Preprocessing | Data Validation | **Data Extraction** | Data Exportation

- Semantic search vs keyword search
  - A keyword search looks for *exact* word matches in the record
  - A semantic search *improves* accuracy by understanding intent through contextual meaning
  - Difference between *"give me what I said"* versus *"give me what I want"*
- Key-value pairs
  - Consist of a key and its corresponding values
  - Keys and values are stored together
  - Key serves as unique identifier for data, while the value represents the data itself

Example:

2

"The patient didn't take the pills because **they** were large"

1

2

"The patient didn't take the pills because **she** was scared"

1

Improve Delivery of BFCC 4 of 4

Document Ingestion | Preprocessing | Data Validation | Data Extraction | **Data Exportation**

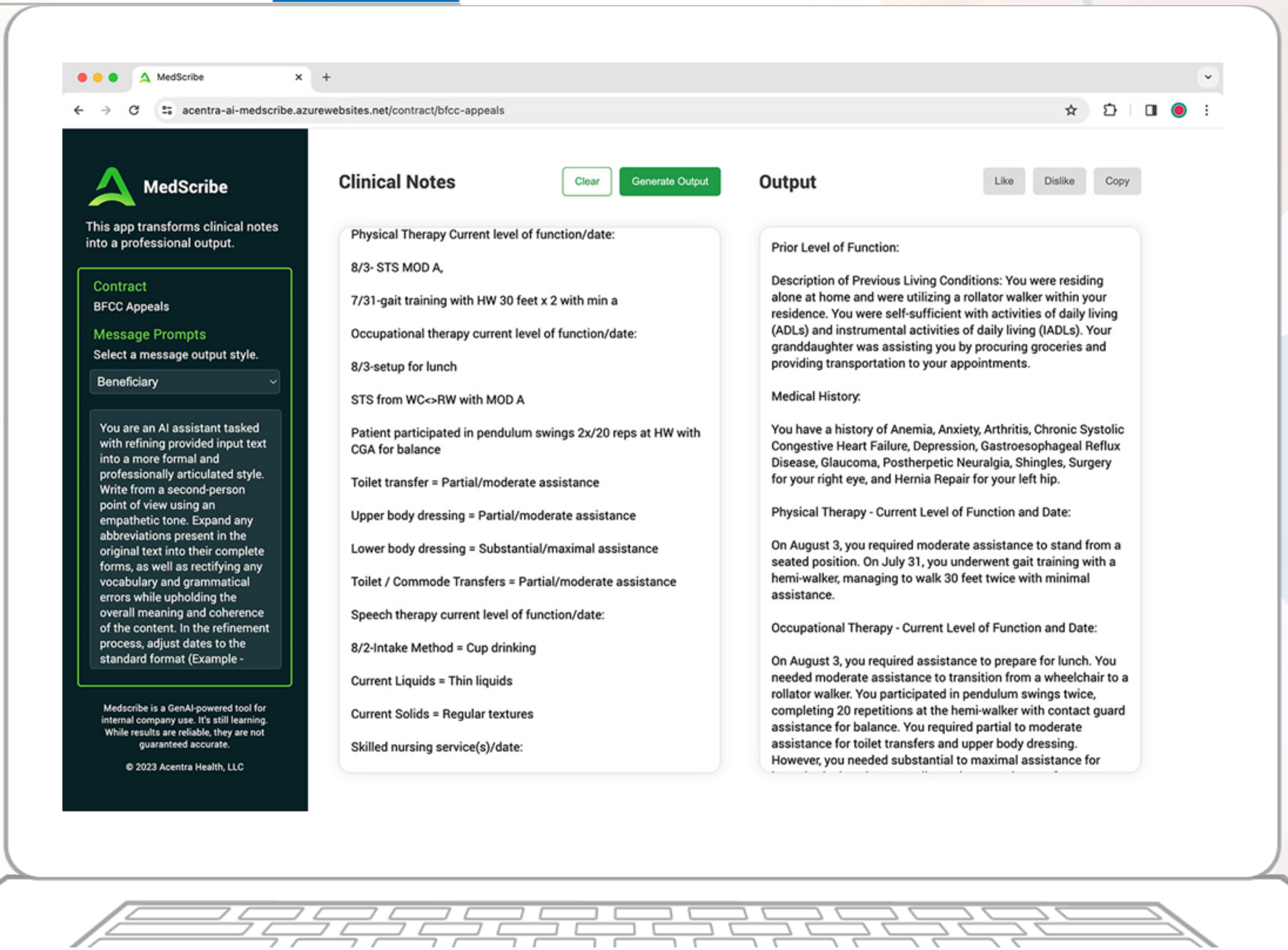
- Automatic routing of supporting documentation to clinicians and Physicians
  - ➢ Seamlessly transmit data through the pipeline to users who need it most
  - ➢ Benefit to the BFCC-QIO is increased automation in workflows, leading to quicker turnaround times (TAT)

- Kepro MedScribe
  - ➢ LLM to transcribe physician rationale into empathetic, easily understandable output that is still unique to each case
  - ➢ Benefit to the BFCC-QIO is the standardization of approach to technical writing

MedScribe: *A Closer Look*

Document Ingestion | Preprocessing | Data Validation | Data Extraction | **Data Exportation**

- MedScribe leverages the power of LLM (GPT4) to transform clinical rationale into concise empathetic output
- Through prompt engineering, the tool can succinctly generate content that meets business requirements
- Easily adapted to meet changing business requirements and use cases



Responsible Use

Human in the Loop (HTL)

HTL is a process where a human is involved in all facets of the decision-making, validation and feedback processes.

- Algorithms and technology are used to augment and assist human decision-making, rather than **replace** decision-making.
- Benefits:
  - Improves the accuracy and effectiveness of systems
  - Helps to build trust and transparency in systems and supports explanations for why decisions are made
  - Helps to improve the performance of human decision-making by leveraging data-driven recommendations

Kepro AI Council

- Kepro is committed to the responsible and ethical use of technology
- The AI Council is a 14-member cross-functional team driving consistent and persistent use of AI

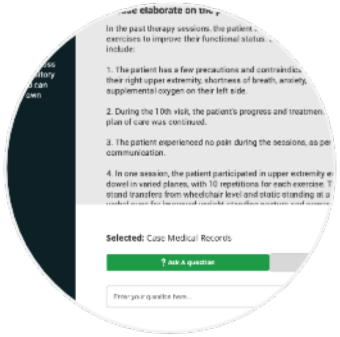
The council is tasked with overseeing:

- Governance
- Legal and Compliance Alignment
- Strategy Alignment
- Tool Standardization
- Promotion
- ROI and Measurement
- Education and Advocacy
- Monitor
- Partnership

Our Vision

Where Do We See the Future Taking Us?

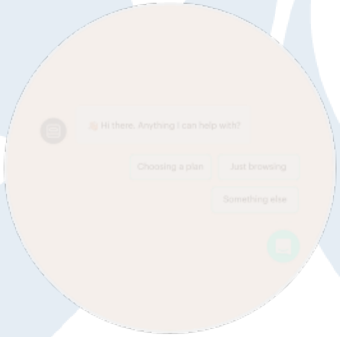
GenAI is so powerful, that completely new capabilities have emerged <u>organically</u>.



Search, Q & A

Writing and Summarization

Software Programming



AI Assistants

CMS 2024 Quality Conference
Resilient and Ready Together

Elevate AI - automating knowledge-based tasks with GenAI

- The ***Elevate AI*** initiative will advocate and promote the meaningful use of GenAI throughout our organization
- It will establish the AI "standard" for Kepro
- Process
  - Discovery sessions with functional areas
  - Build *AI Playbook* for each (use cases, training, best practices, measurement)
  - Deploy an AI Ambassador across functional areas to drive adoption
  - Identify goals for each function and measurement results

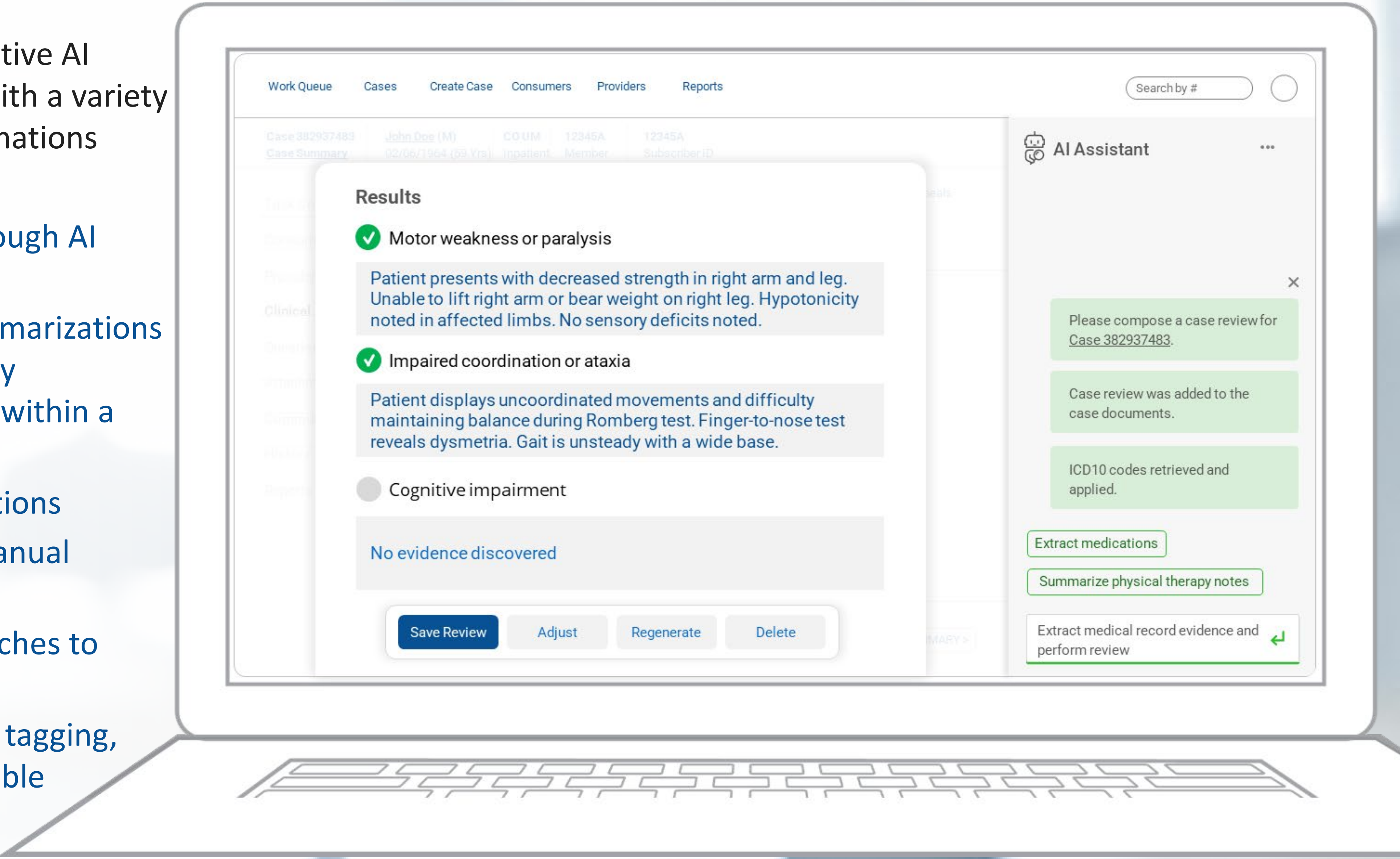
Generated by ChatGPT DALL-E

CMS 2024 Quality Conference
Resilient and Ready Together

Kepro Copilot

Leveraging LLMs and generative AI (GenAI) to assist clinicians with a variety of workflow tasks and automations

- Medical record Q & A through AI assisted chatbot
- Intelligent document summarizations and analysis to identify key components and findings within a medical record
- Rule-based criteria validations
- Automation of existing manual workflows
- More streamlined approaches to inter-rater reliability
- Better document storage, tagging, and cataloging for actionable analytics



Closing

Bottom Line

Innovations in technology, AI and Machine Learning can deliver value to the Beneficiary and Family Centered Care program

Streamlines review processes and workflows	Makes it faster (and easier) to discover key information	Saves review time and money	Allows clinician resources to focus on their most important tasks	Delivers operational excellence

Ethical use of AI/ML

Oversight

The Kepro AI Council will approve tools, use cases, and appropriate safeguards.

Responsible AI

We are utilizing approved datasets and ensuring proper levels of testing. Additionally, our solution designs mitigate the risks associated with AI.

Human in the loop

Our solutions will be designed to support clinicians in their decision-making process, rather than independently making determinations.

Quality

Model updates and quality checks will be carried out on a scheduled basis. All use cases will undergo approved levels of quality and reliability testing.

CMS 2024 Quality Conference

Resilient and Ready Together

Creating an Optimal Environment for Quality Healthcare for Individuals, Families, and Communities

Questions?

Thank You!

- Wendy Gary
wgary@avarconsulting.com

- Brian Salzer
bslazer@avarconsulting.com

- Jed Shakarji
jshakarji@avarconsulting.com

- Josh Dominick
Joshua.Dominick@acentra.com