DATA DRIVEN APPROACH IN CLINICAL TRIALS / DRUG DISCOVERY

Sep 28 2023 | 9:30-11 AM PST

In conversation with...

Jinu Jose
IQVIA
VP, R&D Solutions

Ken McLaren
Frazier Healthcare Partners
Partner for Data & AI

Abhishek Ray
Finarb Analytics
Chief of Data Science

Debashis Banerjee
Finarb Analytics
Vice President of Technology

This session is being recorded

Audience is on mute

Audience Q&A

IMPACT OF AI ON PATIENT LIVES AND CLINICAL TRIALS

Jinu Jose
VP, R&D Solutions, IQVIA

Current responsibility

Jinu heads clinical operations, customer relationships and sales at IQVIA R&D Solutions India. He is part of IQVIA's India Leadership council and represents IQVIA in various industry forums (ISCR, OPPI, DIA)

Profile overview

Jinu is a global life sciences executive with 22+ years of experience in the life sciences industry

11+ years at IQVIA heading various clinical functions and 9 years at Cognizant as part of Cognizant's life sciences practice

Areas of expertise

Clinical Research and Clinical Biometrics

Clinical Technology & Analytics

Clinical Global Capability Centers

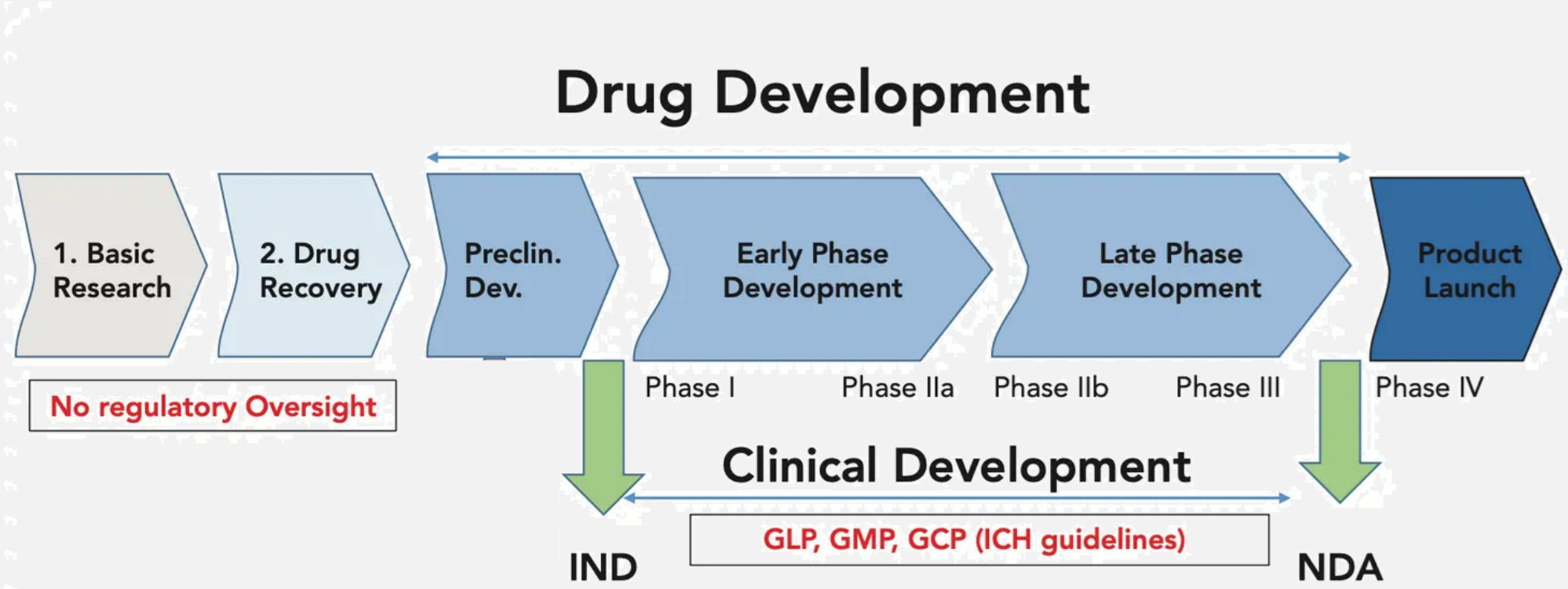
Sales and Business Development

Education

B.Tech, CET Trivandrum

PGDBM (XLRI, Jamshedpur)

CLINICAL RESEARCH & DEVELOPMENT – AN OVERVIEW



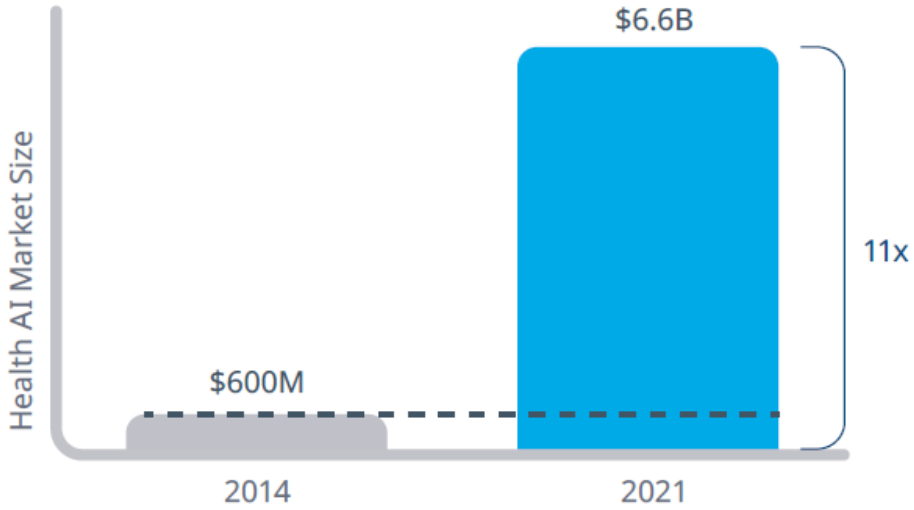
Key Considerations

- **Highly regulated process** – Global in nature - FDA, EMA, MHRA, PMDA, CDSCO
- **Data Intensive** – Collected Patient Data (Safety, Efficacy etc), Device Data, Real World Data, Public Registries, Trial Operations Data, Hospital EMR Data, Lab Data
- **Key Operational Challenges**– Development timelines, Costs and Productivity, Patient Recruitment, Selecting the right sites, Study Start-up timelines
- **Innovation (Risk)** - ONLY 10% of drug candidate entering clinical trials end up being approved drugs

AI ENABLED TECHNOLOGY IN CLINICAL R&D

- AI and ML tools have the potential to transform how clinical development occurs
- Top Pharma are adopting or experimenting with AI/ML at an increasing rate across the product lifecycle
- While the emphasis has been more concentrated on Discovery/ Pre-Clinical applications, increasing activity in Trial Planning and Start-up, as well as in Phase IV.
- Potential to deliver significant time and cost efficiencies while providing better faster insights to inform decision making
- As these tools evolve, new opportunities will continue to emerge that drive further benefits to the clinical research landscape

Significant growth in the AI health market



HEALTH AI CONSIDERATIONS

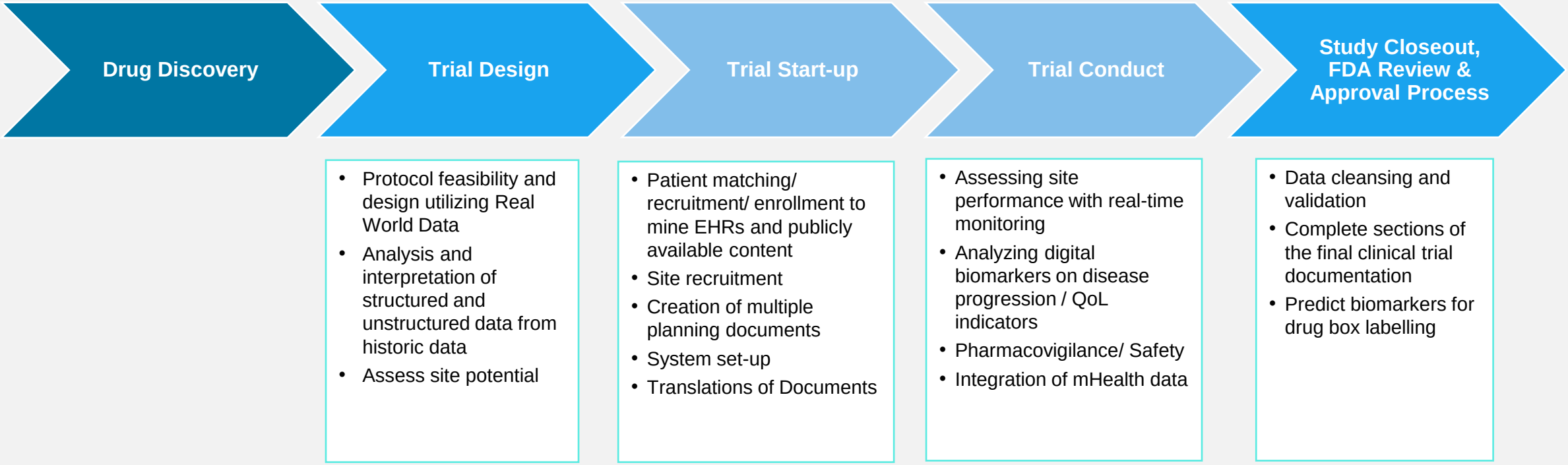
Ethical

- Regulation
- Privacy
- Bias
- Transparency
- Relevance

Legal

- Governance
- Confidentiality
- Liability
- Accuracy
- Decision Making

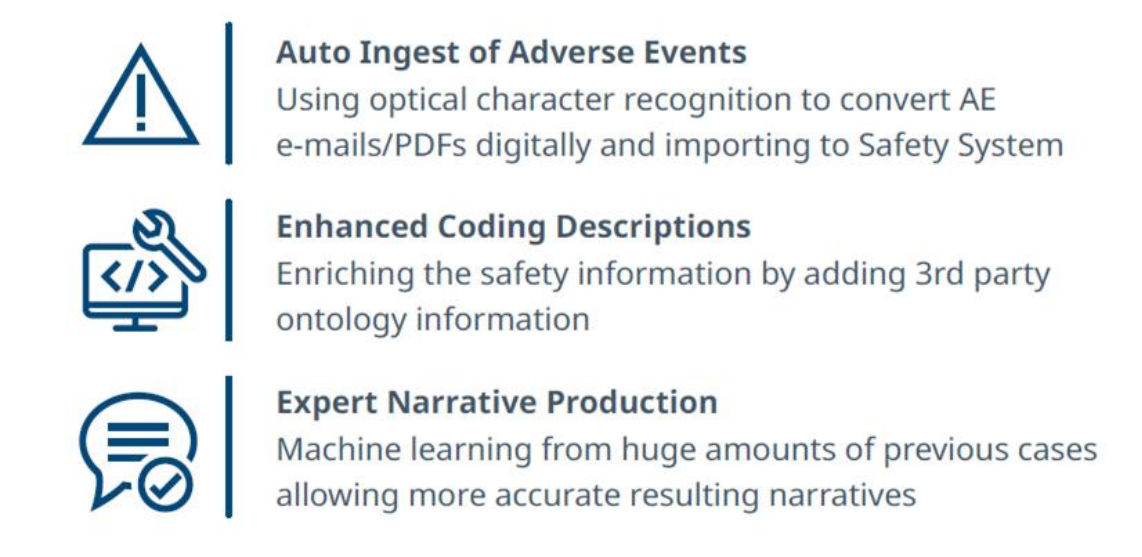
AI ENABLED USE CASES (REPRESENTATIVE)



PHARMACOVIGILANCE (PV)

PHARMACOVIGILANCE

- In PV, massive amounts of structured and unstructured data must be integrated and reviewed to ensure quality and oversight
- AI/ ML tools can be used to automate highly manual processing tasks and translate and digitize safety case processing and adverse drug reaction (ADR) documents to make them more usable.
- They can also perform data listening tasks to monitor conversations on social media and other platforms, ensuring adverse events are promptly identified.

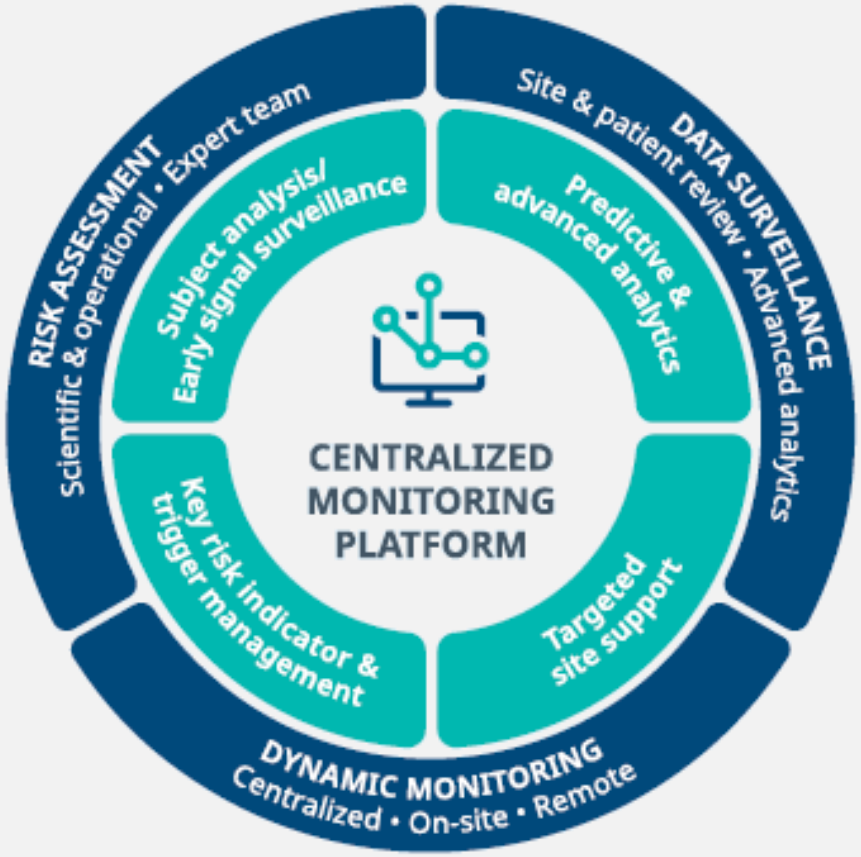


- 20 years of historical data
- Converting major adverse event reporting media from various formats (often paper and some “dead” image scans) into a digitized and standard format.
- Full auditability
- Improving efficiency and productivity
- Leveraging semantic ontology data to enrich the information
- Producing a new XML digital library for learning

TRIAL MONITORING

TRIAL MONITORING

Managing Risks

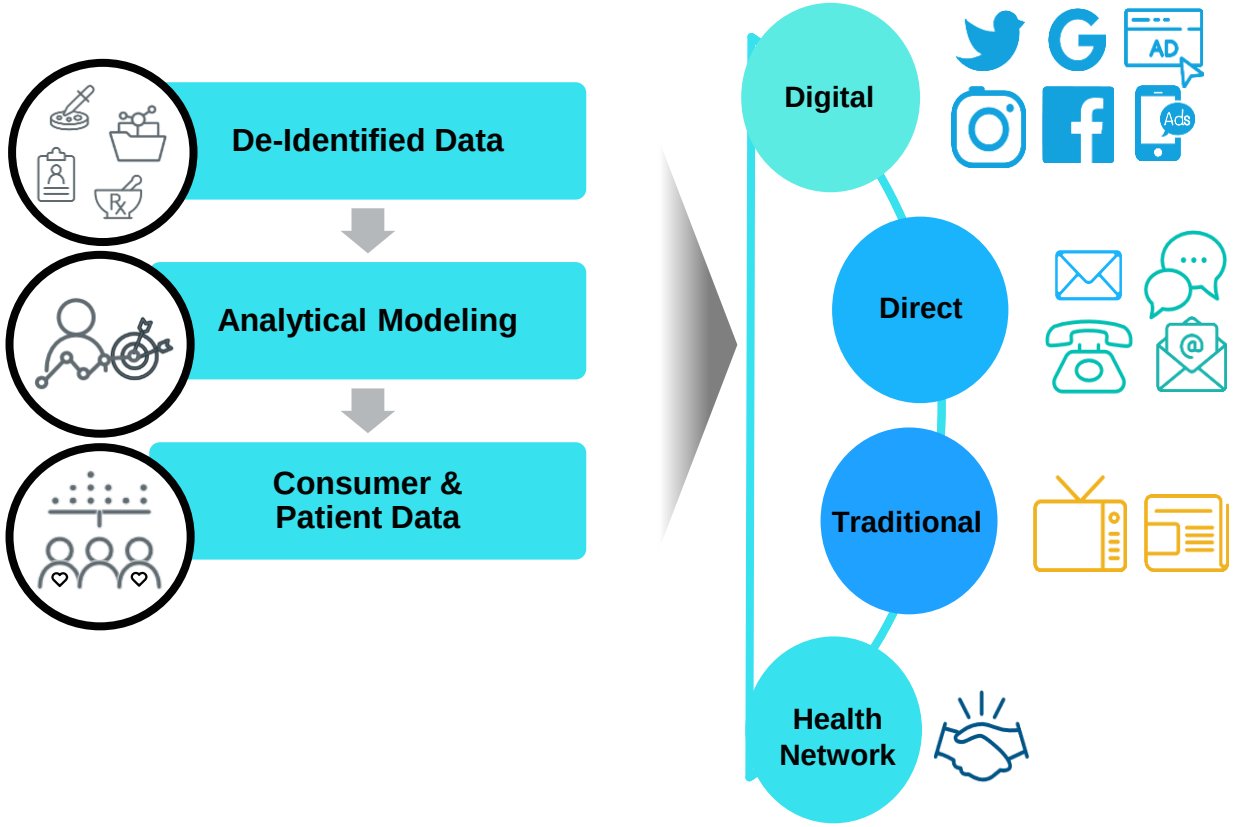


- Tremendous manual effort is spent analyzing site risks and generating “action items” to mitigate those risks.
- AL and ML concepts can alleviate these pressures by manually assessing the risk environment and delivering predictive analytics to generate more effective clinical monitoring insights.
- Advanced analytics provide composite site rankings for holistic risk assessments of sites, allowing for more specific identification of risks and removal of false positive
- They can also be used to proactively identify which sites are more likely to have recruitment and performance issues, or which patients are at higher risk for potential AEs.
- Some use case include detection of duplicate subject registrations, to auto-identify subjects with abnormal lab test results, to detect any vital sign outliers
- With these insights, we can take actions faster and avoid potential issues

SITE IDENTIFICATION AND PATIENT RECRUITMENT

SITE IDENTIFICATION AND PATIENT RECRUITMENT

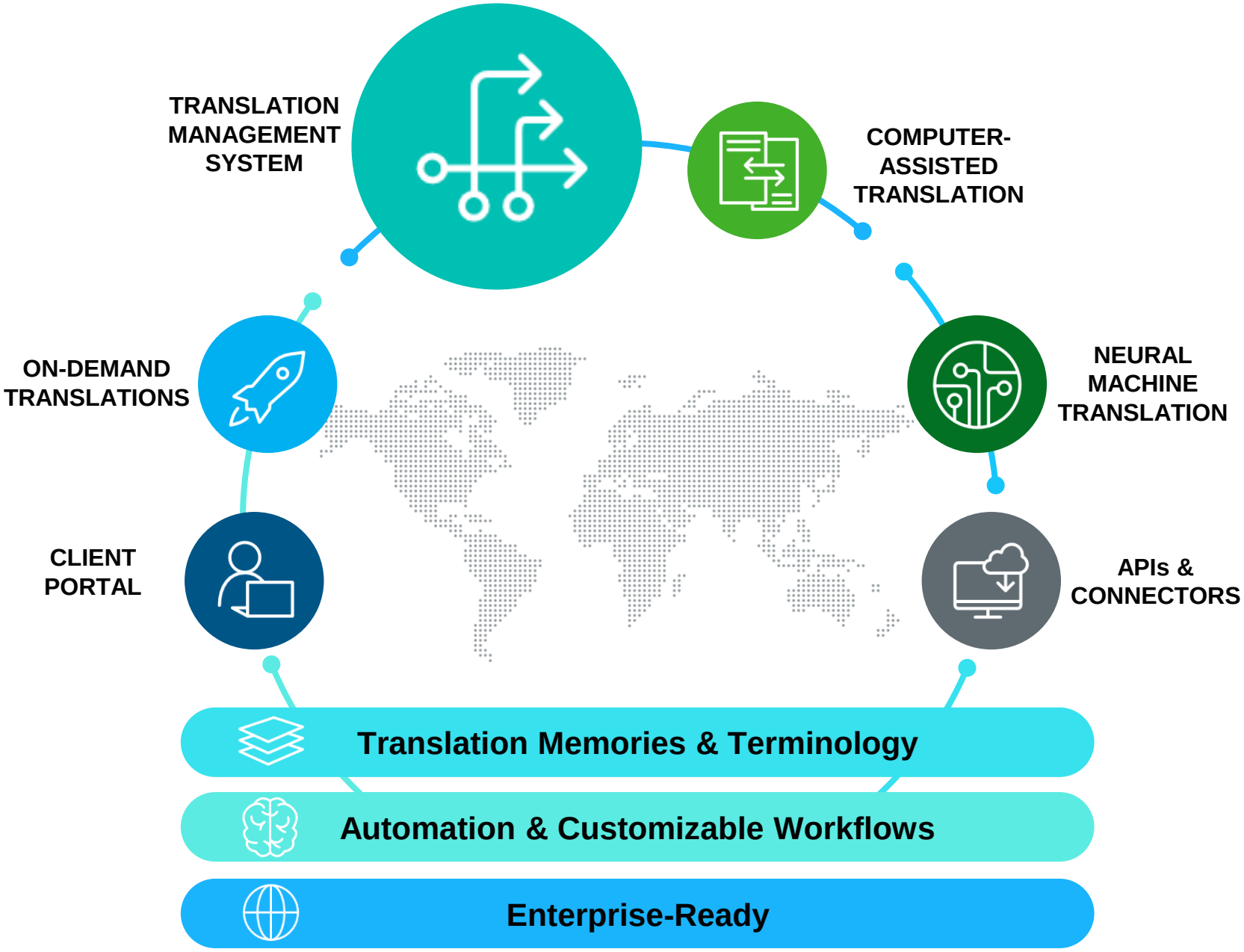
- Identifying trial sites to successfully perform clinical research with access to enough patients who meet inclusion/exclusion criteria is an ongoing challenge.
- As studies target more specific populations, recruiting goals become even harder to achieve, driving costs up, increasing timelines, and raising the risk of failure.
- AI and ML can mitigate these risks is by identifying the sites with the highest recruitment potential and suggesting appropriate recruitment strategies.
- This involves mapping patient populations and proactively targeting sites with high predicted potential to deliver the most patients – before a single site is opened – and identifying the best avenues to recruit them.
- Analyze huge amounts of disease and region-specific healthcare data to locate targeted patient pools, as well as the research sites, healthcare facilities and physicians who treat them



TRANSLATION SERVICES

Translation Services Technology Suite

Robust technology powering end-to-end processes for high quality, fast, secure, and cost-effective translation & localization services



TRANSLATIONS TECHNOLOGY SUITE TO SERVE NEEDS EFFICIENTLY

Translations Technology Suite to Serve Needs Efficiently

CLIENT PORTAL

Submit translations requests

- Access point for customers to submit translations requests & pick up translated files
- Business management functionality, including reporting & quoting
- Provide feedback to improve future translations

ON-DEMAND TRANSLATIONS INTERFACE

Generate quick automated translations

- Easy-to-use, secure Machine Translation Portal for generating quick machine translations
- Unreviewed MT output for internal communications and non-patient/RA-facing documentation

TRANSLATION MANAGEMENT SYSTEM

Manage workflows & resource assignment

- Project management system to apply translation workflows
- Monitor & manage ongoing project progress
- Align best-fit resources for translations & quality review activities

COMPUTER ASSISTED TRANSLATION TOOL

Enhance and improve machine translations

- Efficient tool for post-editing and translation workflows
- Leverage translation memories, term bases & glossaries
- Perform quality checks & reviews

NEURAL MACHINE TRANSLATION ENGINES

Use specialized fit-for-purpose engines

- Suite of machine translation engines with secure access
- 24 custom engines built on IQVIA life sciences & clinical trial contents
- Industry-leading engines (Google, Amazon, DeepL)

APIs & CONNECTORS

Access files and contents

- Embed machine translations and translation services workflows in software applications
- Additional APIs & connectors to be developed to integrate with your technology

Client-Facing

Internal

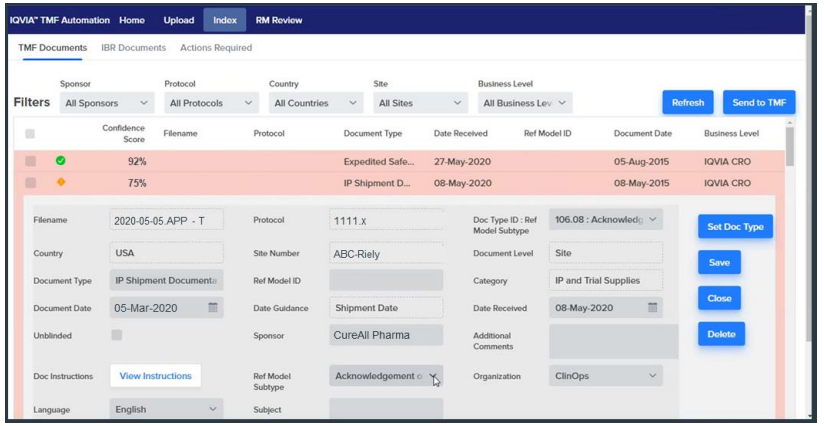
DOCUMENT MANAGEMENT

Intelligent eTMF™

Insight

- What is the quality of the scanned document?
- What type of document is it?
- What is the country, site, protocol number, and subject that is included in the document?

Action



Outcome

- 75% reduction in TMF filing time

I-eTMF (AUTOMATOR)

AI IN CLINICAL TRIALS

Suggested Readings

- AI in Clinical Development – IQVIA White Paper
- Recruiting Rare Disease Patients Just Got Easier – IQVIA White Paper
- AI Chatbot Spontaneously Develops A Theory of Mind – Discover Magazine

FINARB ANALYTICS CONSULTING

CREATING IMPACT THROUGH AI

80% of our clients are the Forbes 1000 enterprises With **78%** being long term partners

Named **'Business Spotlight Leader of the Year 2023'** by Outlook Magazine

Recognized by our clients with high client satisfaction score of **98%**

90% growth

Since our inception, we have been growing in terms of #new clients, repeat business and top line

4 Patents

Finarb's market leading AI models are patent, and IP protected

Industries & Clients

Our projects span 7 major industries and sub domains and 30+ clients

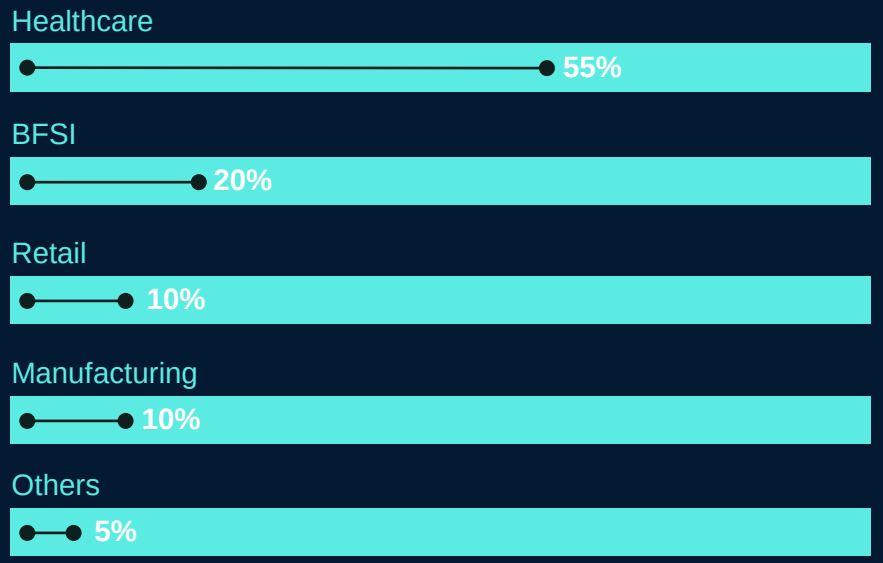
115+ Projects

Our AI models have consistently shown accuracy higher than industry benchmark

OUR CLIENTS

Parata	teva	FRAZIER HEALTHCARE PARTNERS	cps
PCCI	Apollo Intelligence	SOLIS Mammography	Synergy Medical A Parata Company
South Side Specialty Pharmacy	veinclinics of America	P&G	BMW
Wharton University of Pennsylvania	Swissquote	Black Diamond Risk	GETCAPITAL
Bluewater Analytics LLC	darwinbox	PFITR	GATED LIST

OUR TOP VERTICALS



AI USE CASES IN ADHERENCE, COHORT IDENTIFICATION, PREDICTING DROP OUTS & IMAGE AI

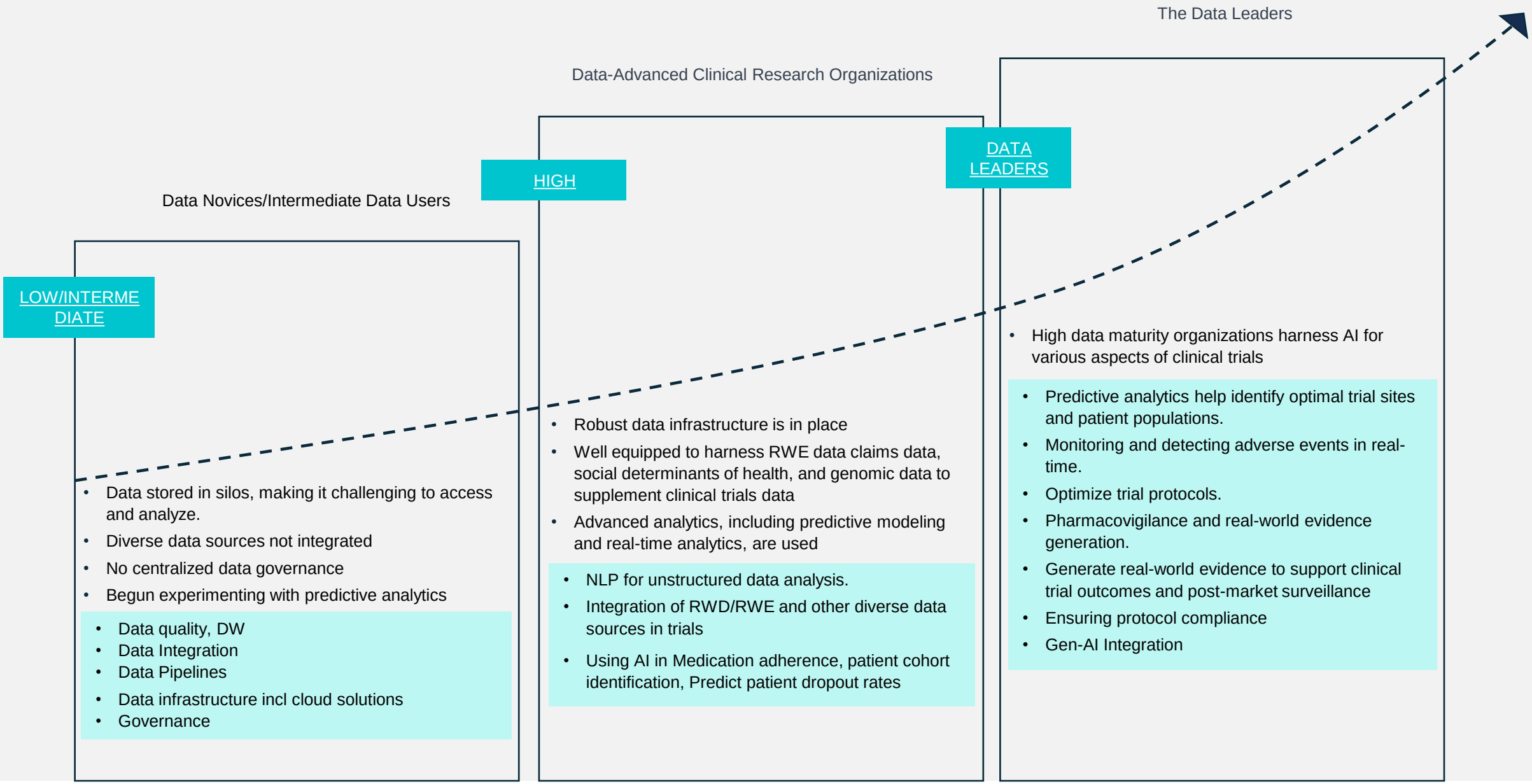
Abhishek Ray
Chief of Data Science,
Finarb Analytics Consulting

Debashis Banerjee
VP of Technology,
Finarb Analytics Consulting

Abhishek Ray - he is the CEO/Chief of Data Science at Finarb Analytics Consulting, a consult-to-operate advanced analytics & AI firm. He brings over 15 years of experience in predictive analytics, modeling, model validation. He has worked with some of the biggest names in the corporate world, Parata (BD Group), PCCI, P&G, and BMW, amongst others.

Debashis Banerjee - Debashis is VP-Technology at Finarb, specializing in AI, Blockchain, and Healthcare Big Data. With over 15 years of experience in the tech industry, he oversees developing and deploying secure cloud-based applications leveraging MLOps, DevOps, and Machine Learning

DATA MATURITY STAGES FOR CRO'S



IMPROVING MEDICATION ADHERENCE IN TRIALS

CLINICAL DATA POINTS

Patient Identifiers

- Name/Gender
- Age
- Health records
- SDOH data

EHR/public datasets

Drug Dispense Data

- Drug Dispensed
- Quantity and Date Dispensed
- Days Supply

CRO/Trial Site Records

Drug Identifiers

- Drug NDC
- Drug Strength
- Dosage form
- Batch/Lot No.

Sponsors

RWD/Patient Survey Data

- Adherence to Refills
- Symptoms/Condition Flare
- Drug Adverse Events
- Barriers to care
- Patient reported outcome/effectiveness
- Patients' satisfaction with medical care
- Medication/dosing related
- Lifestyle

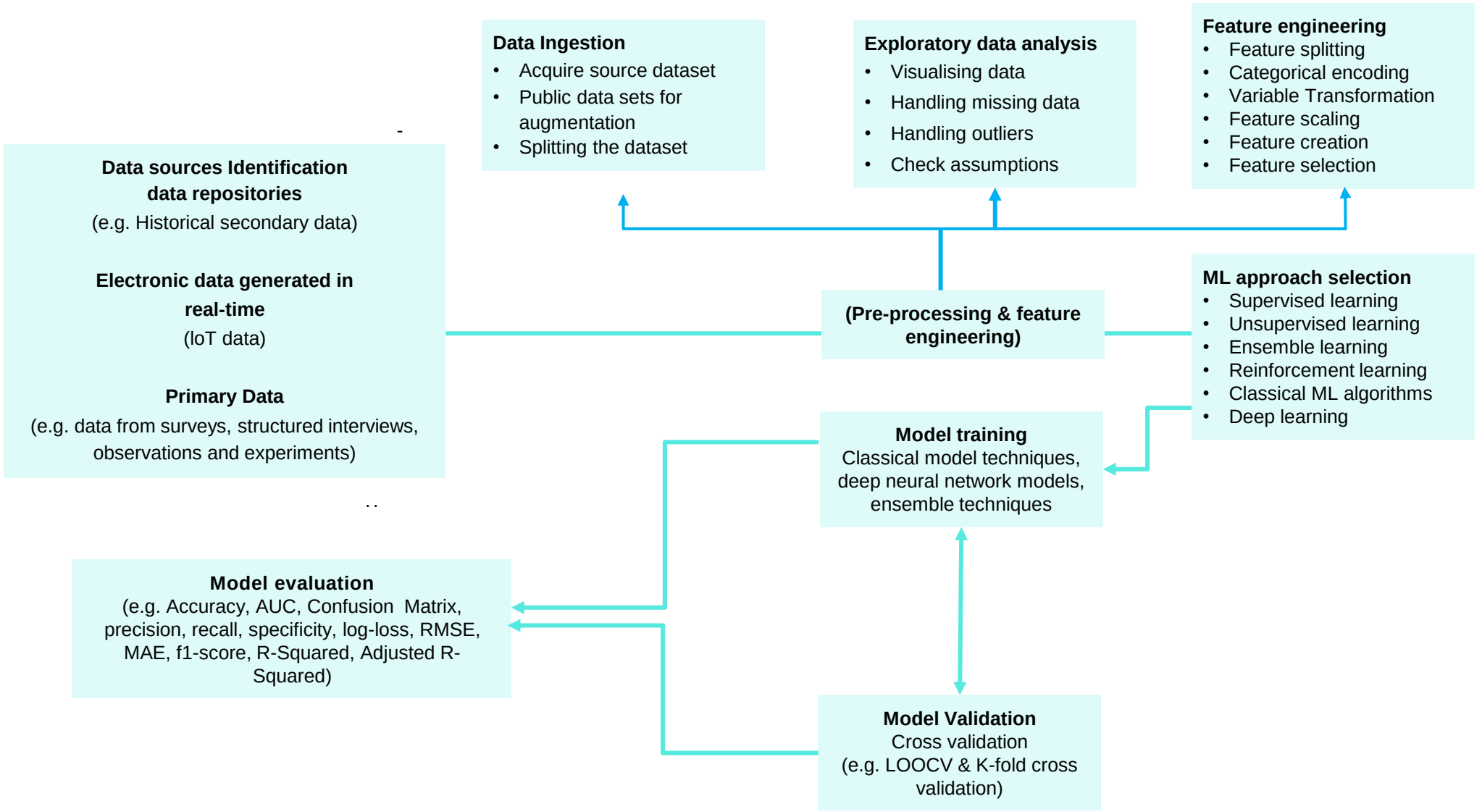
Patient Survey/Assessment

Clinical Assessment

- Data on disease progression or regression based on biomarkers, lab tests, imaging data, clinical notes

Lab Information Systems

CREATION OF DATA PIPELINES, EDA, FEATURE ENGINEERING



APPROACH

- We determine MA (medication adherence) in terms of Proportion of Days Covered (PDC) or Medication Possession Ratio (MPR). We are conversant with both approaches

$$PDC = \left(\frac{\text{Number of days in period "covered"}}{\text{Number of days in period}} \right) \times 100\%$$

$$MPR = \left(\frac{\text{Sum of days' supply for all fills in period}}{\text{Number of days in period}} \right) \times 100\%$$

- A patient PDC score **≥0.9** is conventionally considered adherent, and below is considered non-adherent.
- The model predicts probability of a patient at high risk of being non-adherent
- High risk patients & key drivers of weak adherence are identified to mitigate areas of concern.

Development Strategy

- Data pre-processing, extraction from unstructured text data
- Data augmentation
- Development of ETL pipelines for predictive models
- Building Patient Adherence Predictive Models
- Feature Engineering and Augmenting Model Performance
- Risk Stratification of Patients
- Identification of Drivers of non-adherence
- Developed Interventions for High-risk patients
- Continuous Refinement of Predictive Models and Associated Drivers

PATIENT COHORT IDENTIFICATION – ASTHMA, US

PATIENT COHORT IDENTIFICATION

Problem

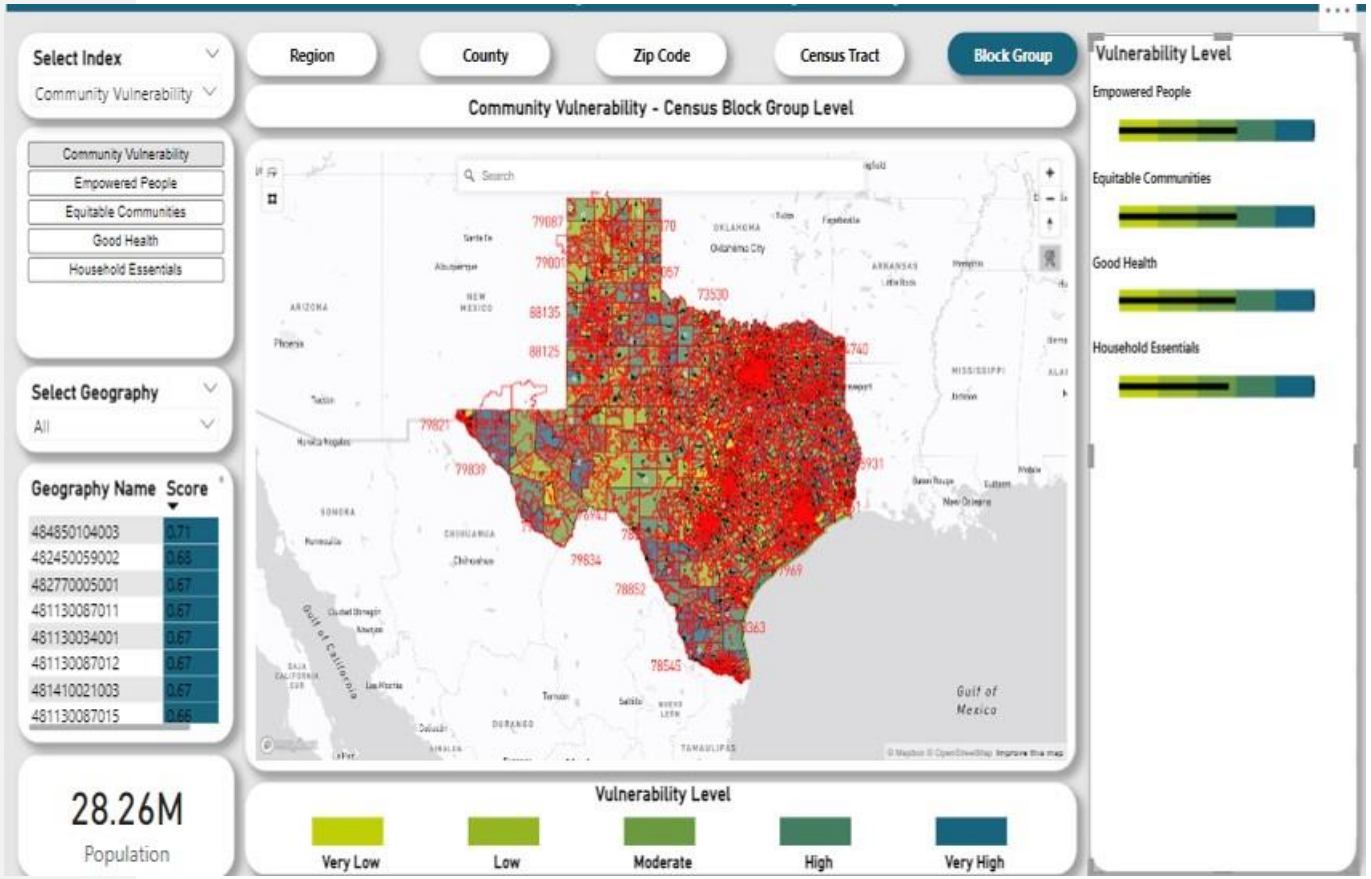
Cohort Identification is time-consuming and highly prone to errors. Unsuccessful cohort determination can lead to worst monetary burdens, high dropout rates.

Solution Impact

- Based on the following data sources we identified the cohorts and segmented them into various categories. The result is depicted in the heatmap

Data Sources

- Asthma Prevalence Data
- Demographic Data
- Environmental Data
- Geographic Data
- Socioeconomic Data
- Behavioural and Lifestyle Data
- Occupational Data
- Historical Data



PREDICTING LIKELIHOOD OF PATIENT DROP- OUTS DURING TRIAL

PREDICTING LIKELIHOOD OF PATIENT DROP-OUTS DURING TRIAL

We determine **Probability of Drop-Out** by assessing various predictor variables such as patient/trial/medication outcome related features and influence on probability of drop out.

Clinical Data Inputs: Structured - SDOH, EHR, CTMS for current/previous trial data, or unstructured via patient surveys

Key predictor variables

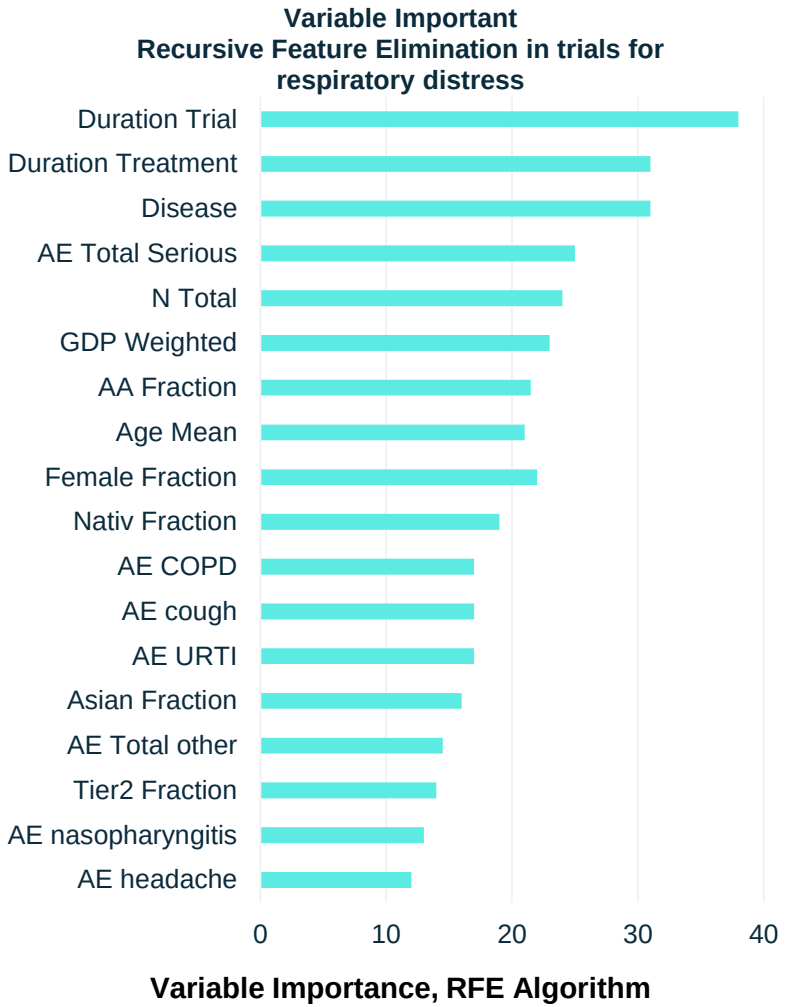
- Trial specific data (Duration, trial phase, treatment type, and study site)
- Adverse events experienced by the patient
- Adherence history (e.g., previous missed appointments)
- Treatment-related factors (e.g., type of treatment, dosage, duration)
- Medication efficacy/disease progression
- Engagement metrics (e.g., response to communications, attendance at follow-up visits).
- Medical history and comorbidities.
- Demographic information (e.g., age, gender).
- Additional predictive variables can also be created

Analyze importance of these features in predicting drop-out probability, and calculate patient risk scores using ML classifier models

Segment patients into High, Medium, Low risk of discontinuation

Identify risk factors associated with high drop-outs

Target customized interventions



MEDICAL IMAGING AI IN TRIALS

MEDICAL IMAGING AI IN TRIALS

~10% of clinical trials are using medical imaging AI for managing eligibility decisions or final imaging endpoints

Beneficial For

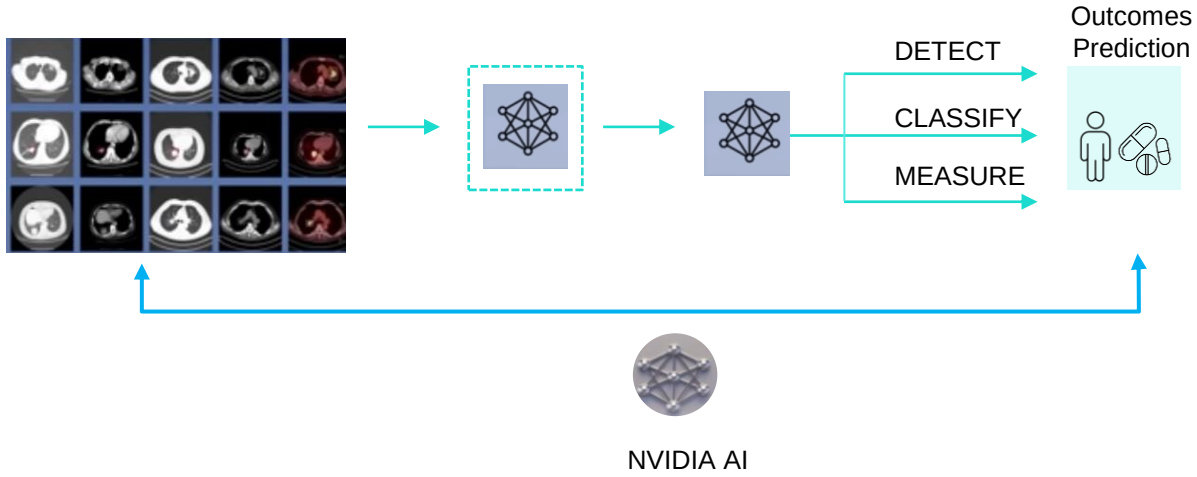
- Expedited review of images
- Reduced turnaround time from days to milliseconds
- Assess even poor quality medical images, helping radiologists avoid missed lesions, misclassifications
- Sponsors see ~40-50% variability on different radiologists' readings of scans. Variability with ML can be virtually eliminated
- Redact PHI from DICOM tags and pixel data before transmission to other users.

Critical success factors

- Requirement of large image banks reviewed by experts, for training the model
- Integrated image analysis pipeline directly built into the radiologist & clinical trial workflows

Regulations: FDA: existing guidelines & framework for integration- AI/ML)-Based Software as a Medical Device (SaMD)

USING MACHINE VISION AI

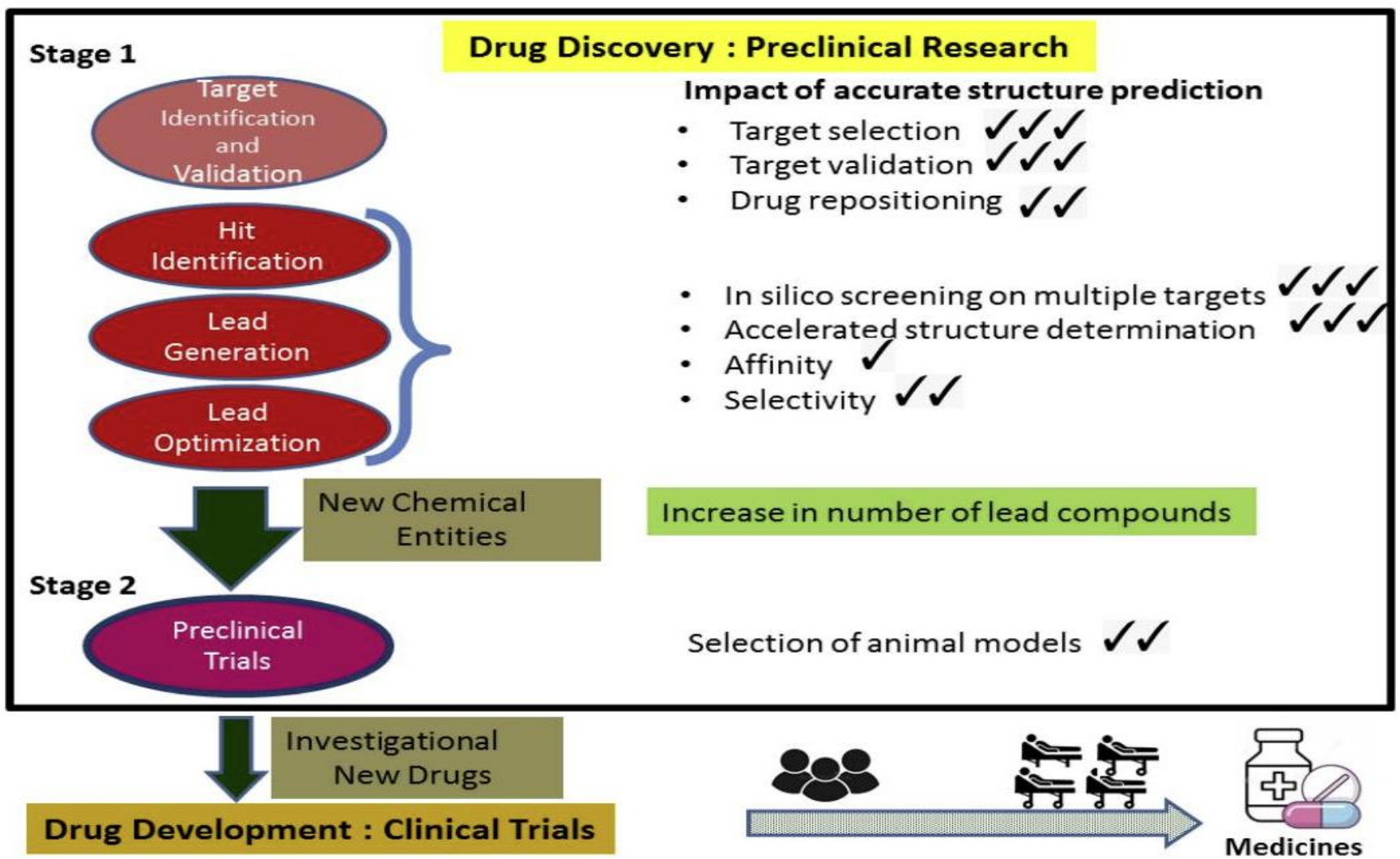


GENERATIVE AI IN DRUG DISCOVERY

Ken McLaren
Partner for Data & AI, Frazier Healthcare Partners

Ken McLaren is a Partner for the Data & AI CoE at Frazier Healthcare, one of the leading mid-market healthcare investment firms. His experience spans 20 years in technology and analytics led life sciences transformation, and has held global leadership positions at Clarivate Analytics and DRG prior to joining Frazier

DRUG DISCOVERY PROCESS



Sources: https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

GENERATIVE AI FOR PREDICTING PROTEIN STRUCTURES

FRAZIER HEALTHCARE PARTNERS

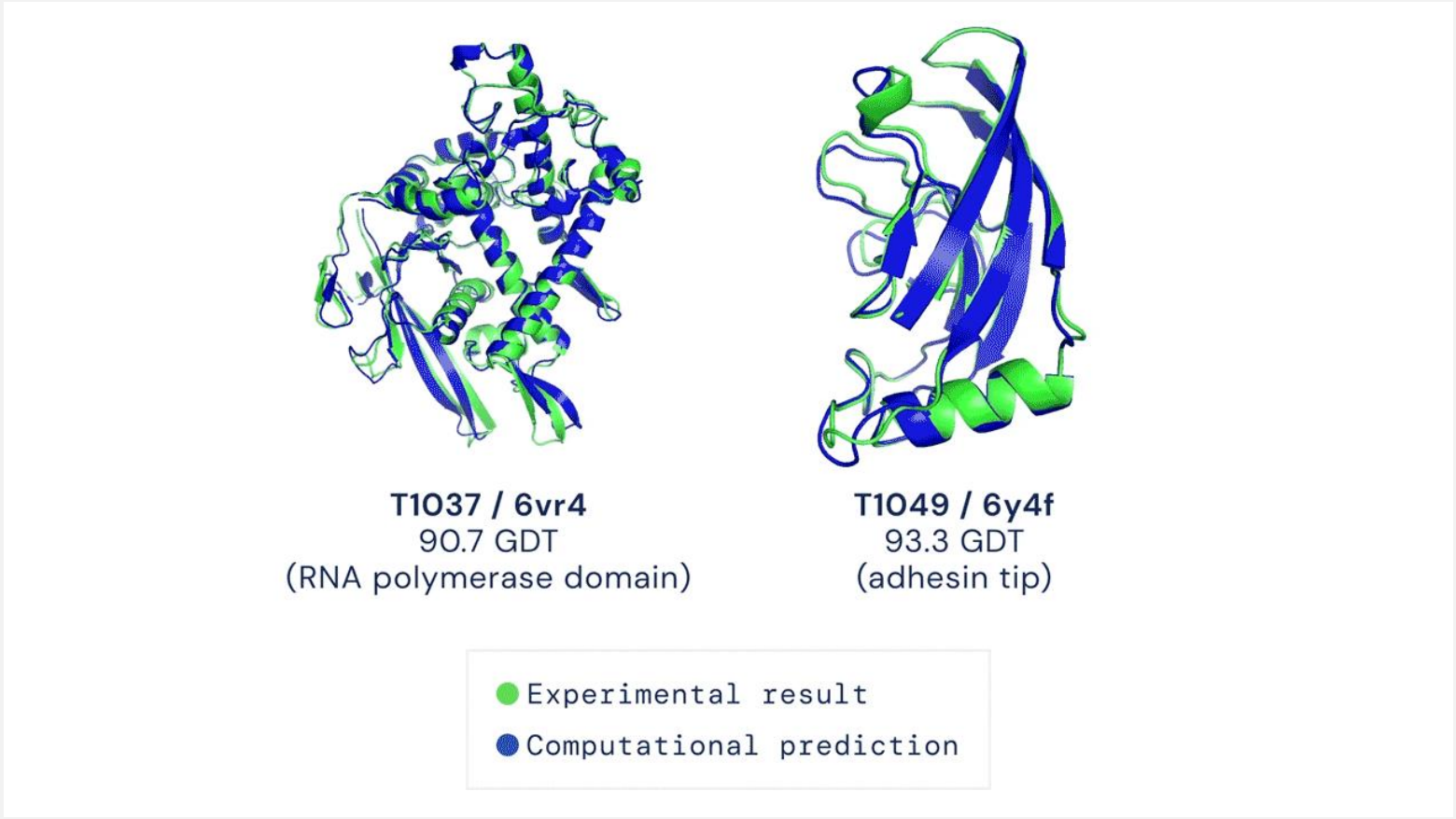
Google DeepMind

Research

AlphaFold: a solution to a 50-year-old grand challenge in biology

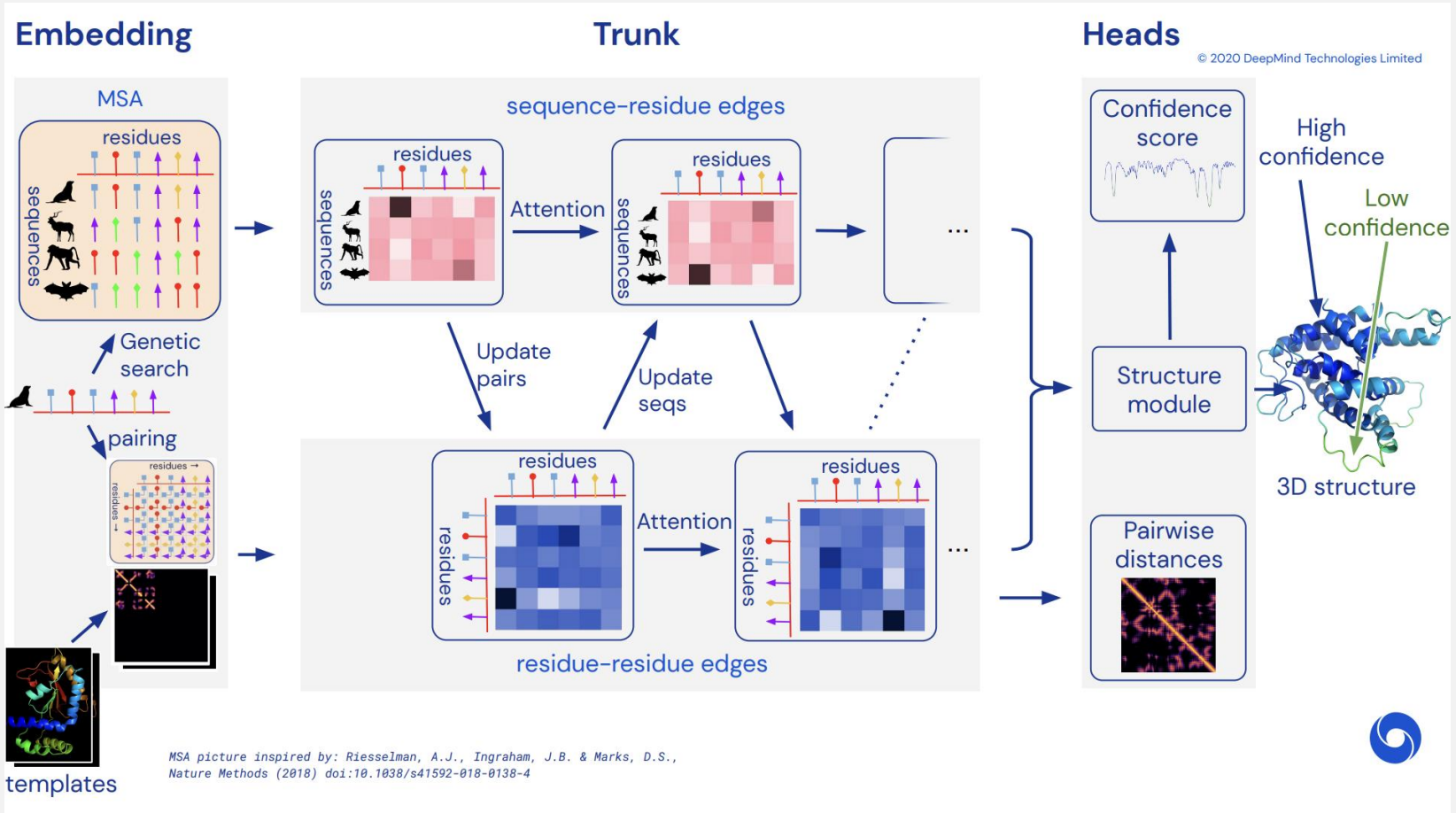
November 30, 2020

AlphaFold is a deep learning system that can predict the 3D structure of a protein from its amino acid sequence with high accuracy, including proteins that are difficult to study experimentally. In the CASP14 protein structure prediction competition, AlphaFold was the top-ranked method by a large margin, with its predictions being more accurate than those of any other method, including experimental methods.



Sources: https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

ALPHAFOLD MODELLING PROCESS



DOWNSTREAM ADVANTAGES

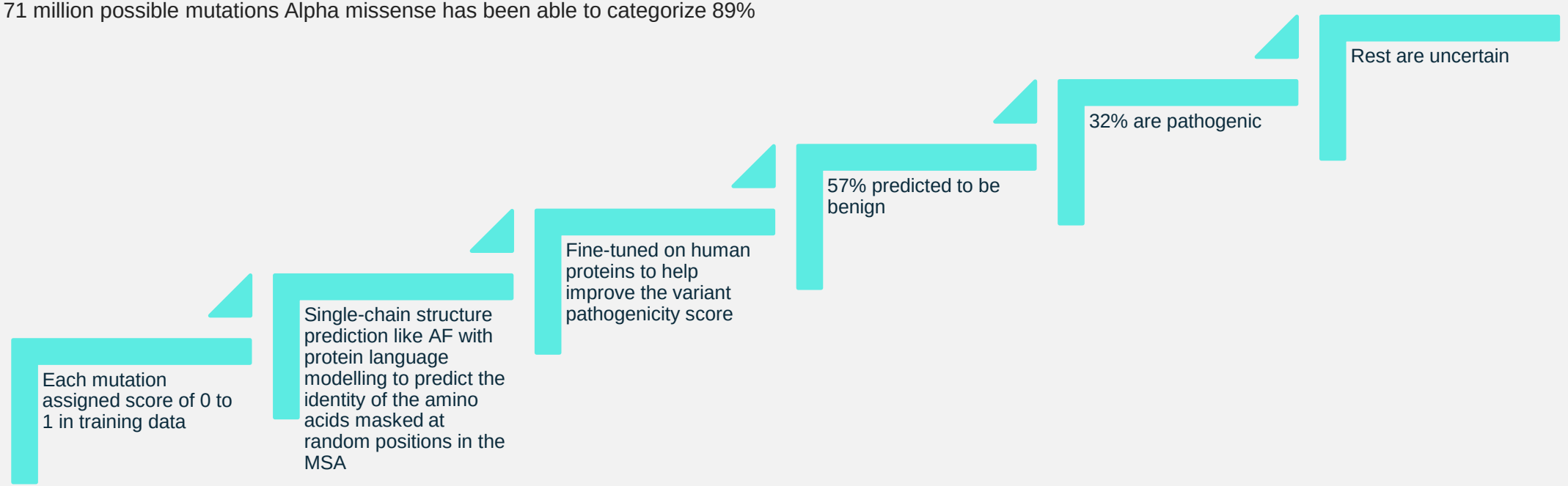
AI is being used for discovering and predicting protein folds and that opens up new avenues for the drug discovery process. Here are some of the downstream advantages of accurately predicting the protein fold structure -

- Accelerated Drug discovery to reduce the go-to-market time
- Treatment of auto immune diseases
- Rational Drug Design leading improved efficacies, reduced adverse effects and reduced trial and error
- Personalised medication improvements
- Speeding up the process of drug discovery for rare & genetic diseases
- Reduce environmental impact by reducing the quantity and intensity of human/ animal trials
- Drug Repurposing

ALPHA MISSENSE

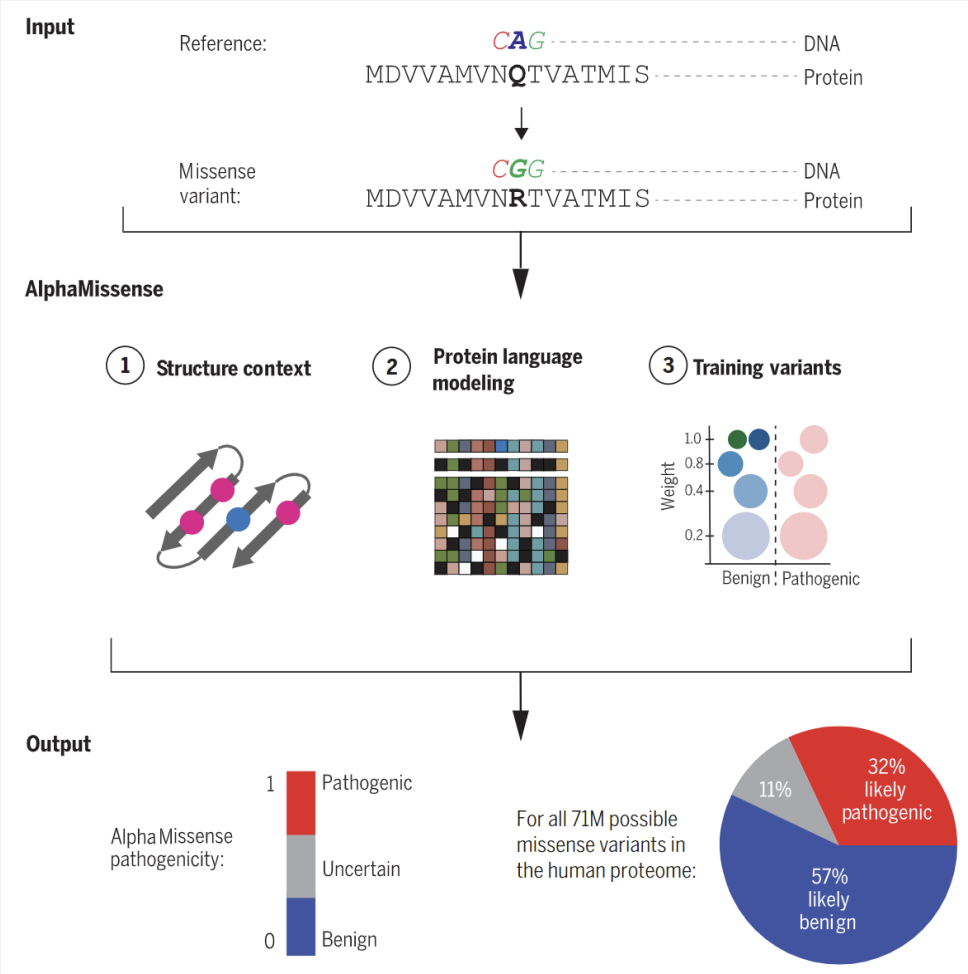
Alphafold missense – focusses on missense mutation where we have a single nucleotide change

- Individual has 9000 mutations approximately
- Around total of 4 million mutations have been observed
- 2 percent of them have been classified as pathogenic or benign
- Out of 71 million possible mutations Alpha missense has been able to categorize 89%



Sources: https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

Google DeepMind



Sources: https://www.science.org/stoken/author-tokens/ST-1429/full

RESEARCH

RESEARCH ARTICLE SUMMARY

MACHINE LEARNING

Accurate proteome-wide missense variant effect prediction with AlphaMissense

Jun Cheng*, Guido Novati, Joshua Pan†, Clare Bycroft†, Akvilė Žemgulytė†, Taylor Applebaum†, Alexander Pritzel, Lai Hong Wong, Michal Zielinski, Tobias Sargeant, Rosalia G. Schneider, Andrew W. Senior, John Jumper, Demis Hassabis, Pushmeet Kohli*, Žiga Avsec*

ALPHA MISSENSE

DOWNSTREAM ADVANTAGES

- **Disease Diagnosis**: Identify pathogenic missense mutations for genetic disorders.
- **Personalized Therapies**: Tailor treatments based on mutation-driven protein function changes.
- **Drug Target Identification**: Pinpoint mutation-induced protein vulnerabilities.
- **Guided Functional Studies**: Prioritize mutations with predicted significant impact for lab research.
- **Evolutionary Insights**: Discern protein evolutionary pressures via mutation effect analysis.
- **Population Genetics**: Detect variants under positive or purifying selection in populations.
- **Protein Engineering**: Design proteins with desired functionalities via mutation predictions.
- **Enhanced Genetic Testing**: Make genetic test results more actionable.
- **Risk Stratification**: Assess genetic risk profiles for hereditary conditions.
- **Genetic Counselling**: Inform decision-making with predictive mutation insights.

FINARB ANALYTICS CONSULTING

Finarb is a leading provider of Advanced AI ML and Data Analytics solutions. We partner with our clients on their business transformation journey enabled by AI starting from problem discovery – strategy & consulting – rapid prototyping & MVP – build – deploy – manage – operate. Our vision is to democratize AI and enable large, mid-sized and SMB organizations take advantage of the power of AI for sustained growth, competitive advantage and be resilient enterprise of future!

80% of our clients are the Forbes 1000 enterprises with 78% of them being our long term partners. We are recognized for our high client satisfaction score of 4.9/5. Our AI models are pioneering in their own right with 4 patents filed so far.

Creating Impact Through Data & AI

https://finarbconsulting.com/ || sales@finarbconsulting.com