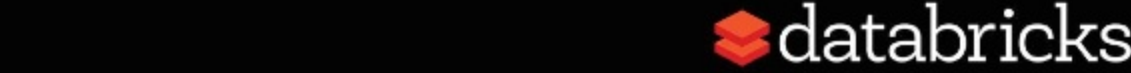
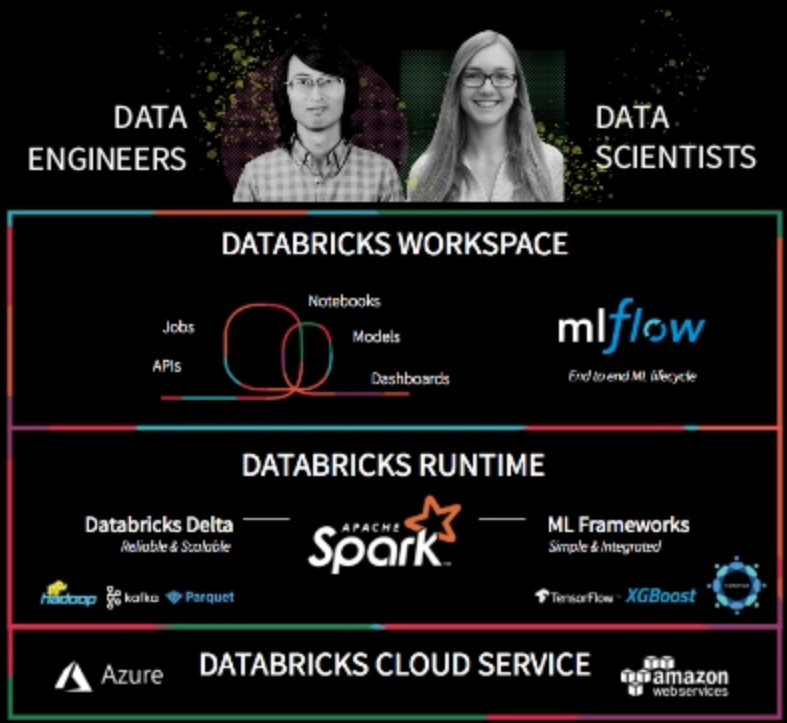
The Future of Healthcare with Big Data and AI




DATA ENGINEERS
DATA SCIENTISTS
UNIFIED ANALYTICS PLATFORM
DATABRICKS WORKSPACE
Notebooks
Jobs
Models
APIs
Dashboards
mlflow
End to end ML lifecycle
DATABRICKS RUNTIME
Databricks Delta
Reliable & Scalable
APACHE Spark
ML Frameworks
Simple & Integrated
Hadoop
kafka
Parquet
TensorFlow
XGBoost
Azure
DATABRICKS CLOUD SERVICE
amazon web services

databricks

DOMAIN EXPERT

EXPERTISE GAP

UNIFIED ANALYTICS PLATFORM

DATABRICKS WORKSPACE

Jobs
APIs
Notebooks
Models
Dashboards

mlflow
End to end ML lifecycle

DATABRICKS RUNTIME

Databricks Delta
Reliable & Scalable

APACHE Spark

ML Frameworks
Simple & Integrated

Hadoop kafka Parquet

TensorFlow XGBoost

Azure DATABRICKS CLOUD SERVICE amazon web services

databricks

DOMAIN EXPERT

UNIFIED ANALYTICS PLATFORM

INDUSTRY-SPECIFIC TOOLS

DATABRICKS WORKSPACE

Notebooks
Jobs
Models
APIs
Dashboards

mlflow
End to end ML lifecycle

DATABRICKS RUNTIME

Databricks Delta
Reliable & Scalable

APACHE Spark

ML Frameworks
Simple & Integrated

Hadoop kafka Parquet

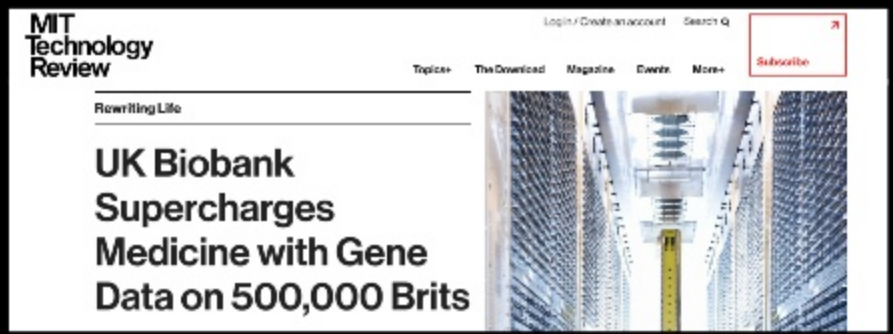
TensorFlow XGBoost

Azure DATABRICKS CLOUD SERVICE amazon web services

databricks


Unified Analytics Platform
for Genomics
databricks

Massive Investments in Genomic Data



databricks

Potential to Transform the Industry

Faster Drug Discovery

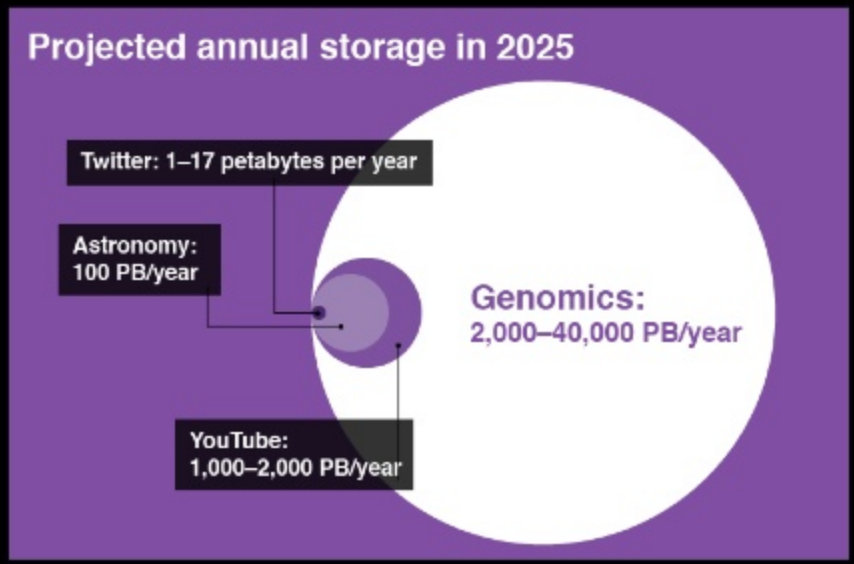
Reduced Health Claims

Better Patient Outcomes

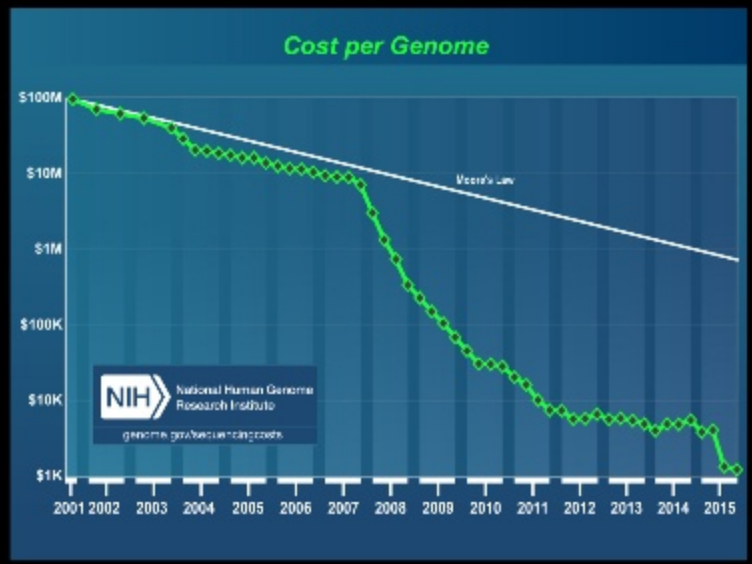
Genomic Data Volumes are Exploding

40,000 Petabytes / year by 2025

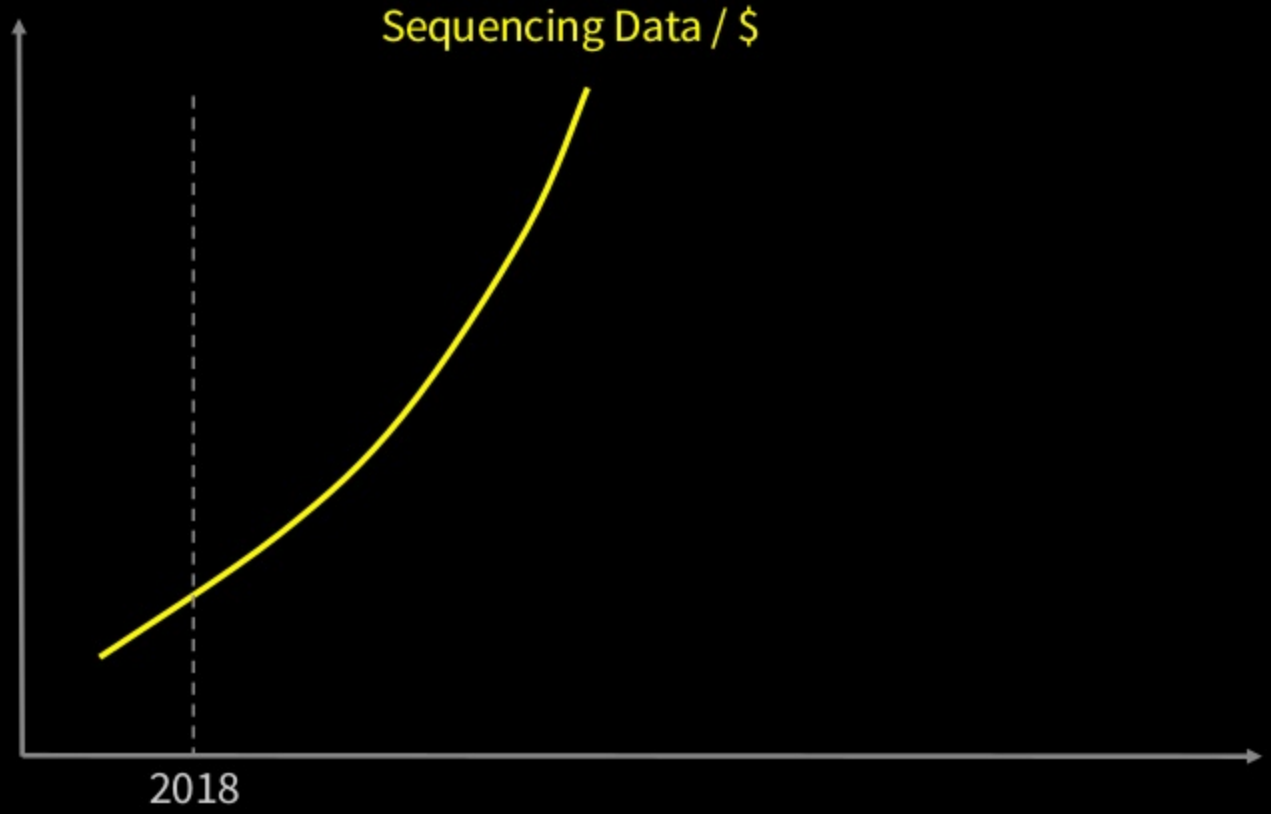
From $2.7B to <$1,000



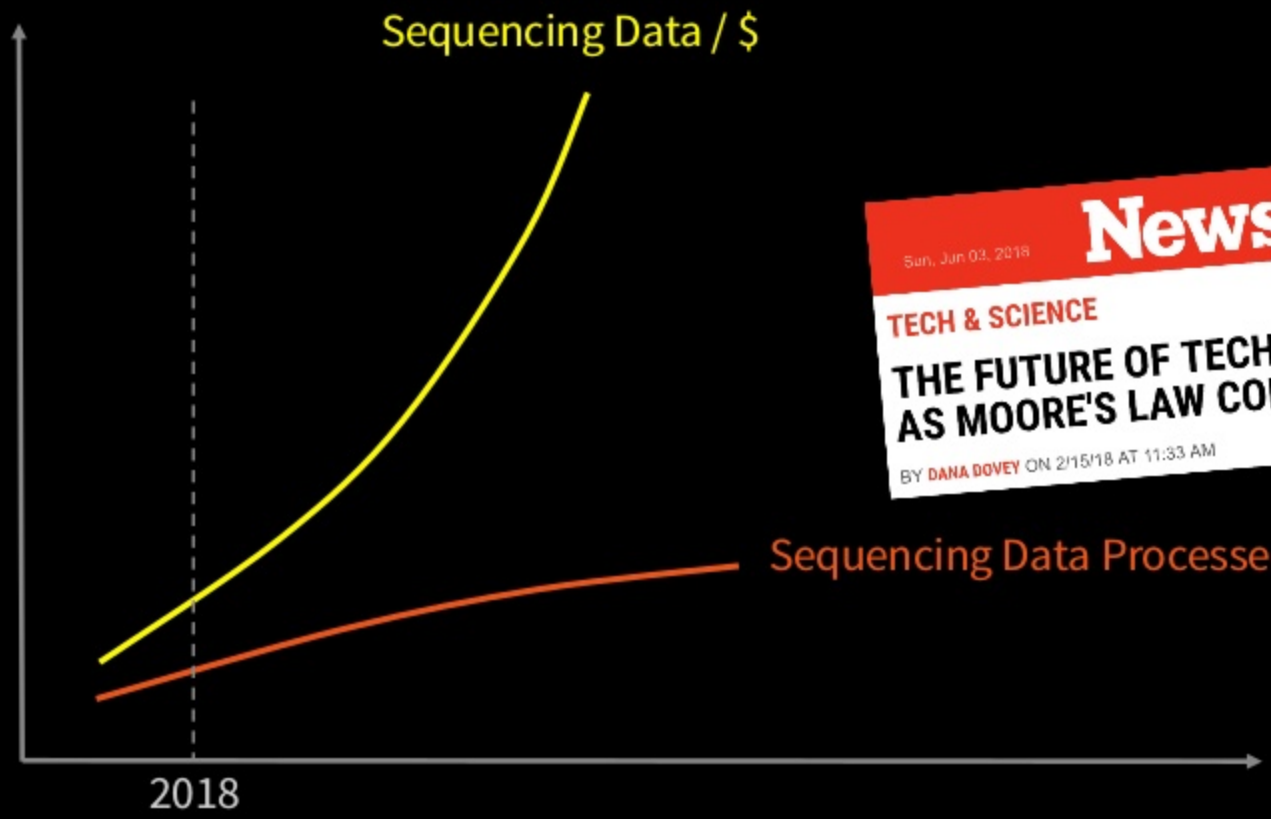
Source: "Big Data: Astronomical or Genomical? PLoS Biology, 7 July 2015



databricks

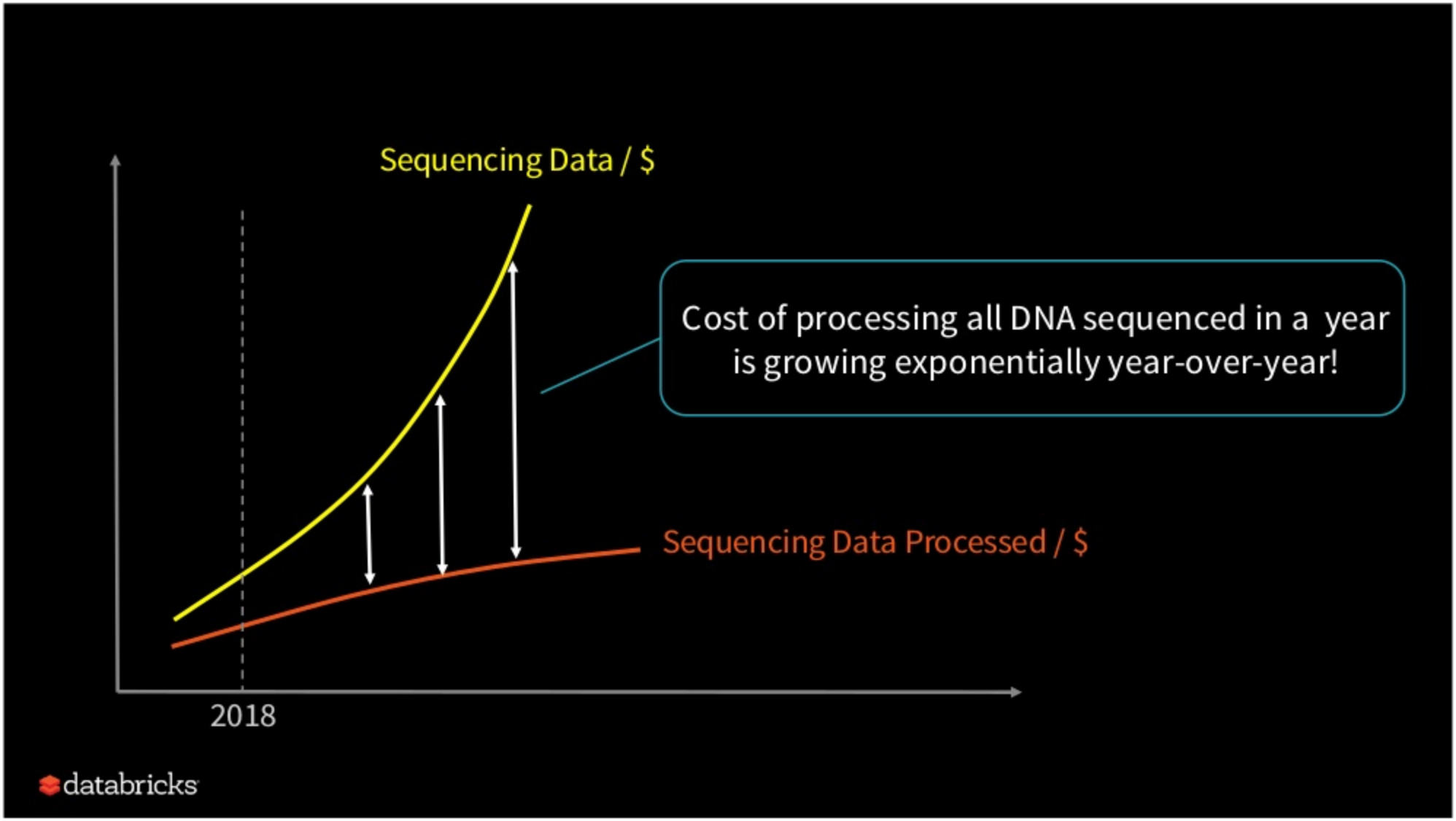


databricks


Sequencing Data / $
Sequencing Data Processed / $
2018


Sun, Jun 03, 2018
Newsweek
TECH & SCIENCE
THE FUTURE OF TECHNOLOGY IS UNCERTAIN AS MOORE'S LAW COMES TO AN END
BY DANA DOVEY ON 2/15/18 AT 11:33 AM

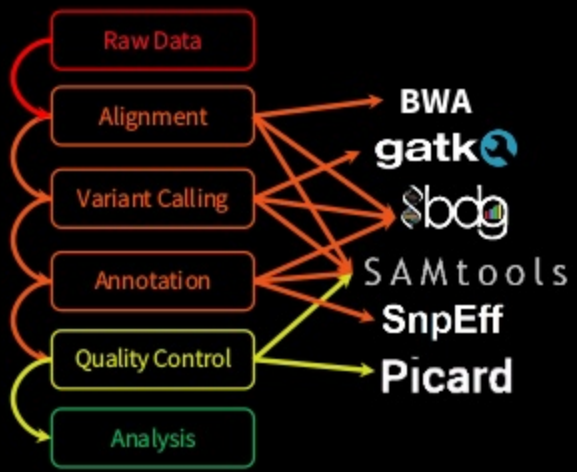
databricks


Sequencing Data / $
Cost of processing all DNA sequenced in a year is growing exponentially year-over-year!
Sequencing Data Processed / $
2018
databricks

Challenge #1: Complex Pipelines

Complex Genomic Pipelines

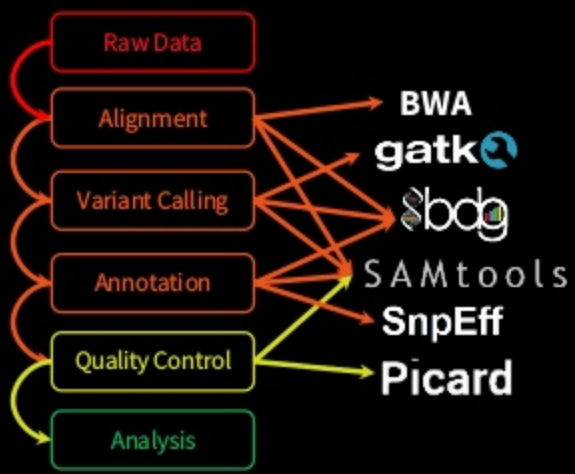
Costly and time consuming



Challenge #2: Rigid Analytics

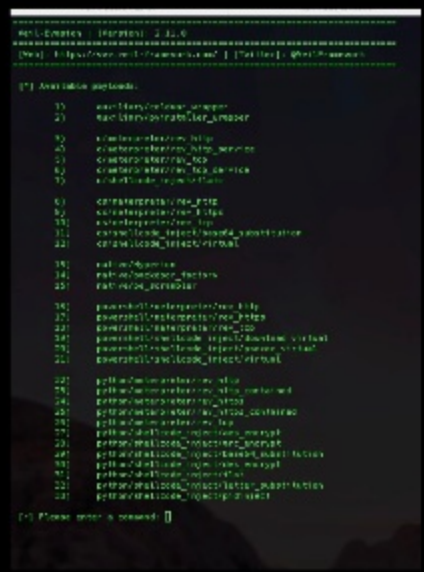
Complex Genomic Pipelines

Costly and time consuming



Rigid Analytics

Reduced Scope of Research

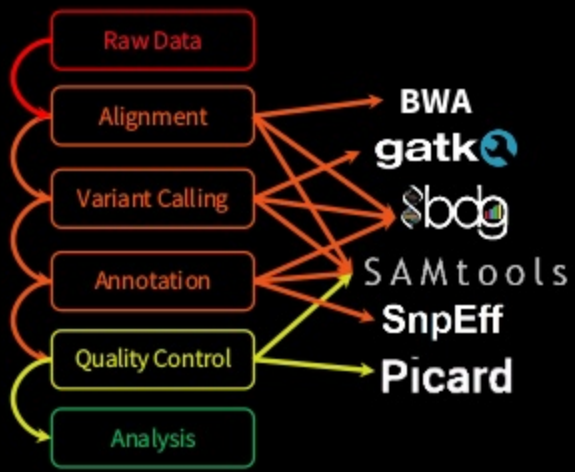


databricks

Challenge #3: Siloed Teams

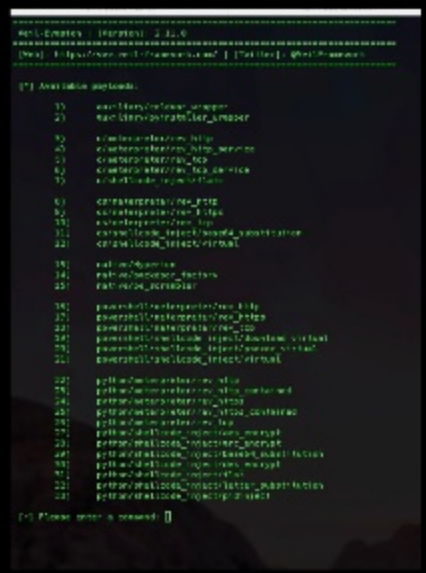
Complex Genomic Pipelines

Costly and time consuming



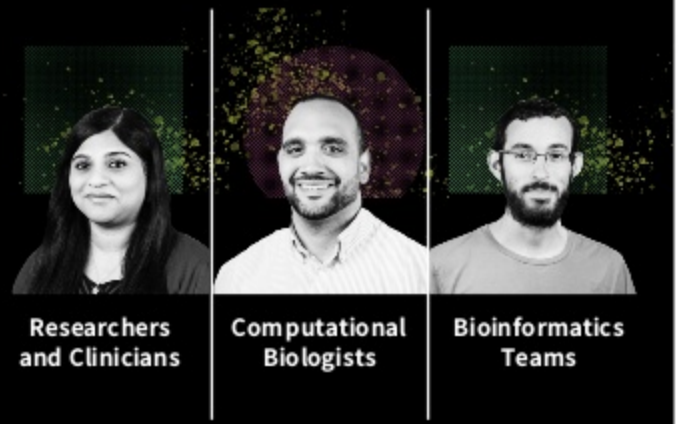
Rigid Analytics

Reduced Scope of Research



Siloed Teams

Lack of Productivity



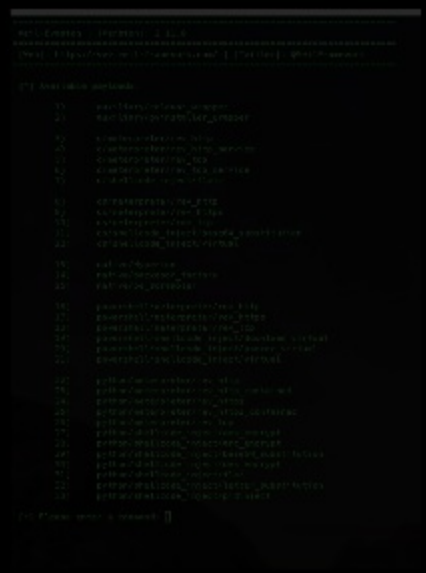
Researchers and Clinicians | Computational Biologists | Bioinformatics Teams

databricks

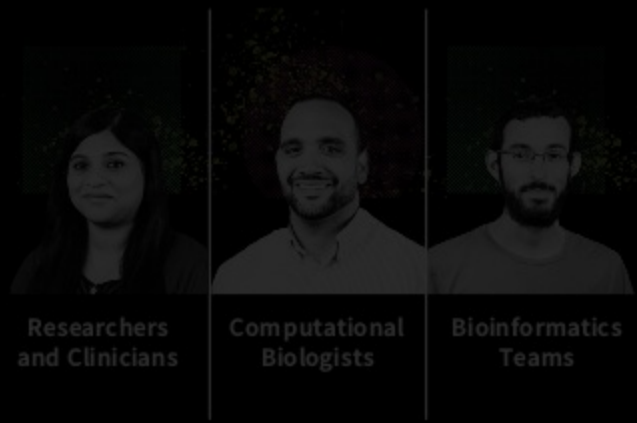
Solution #1: Prebuilt Pipelines

Best Practice Pipelines
"One click" execution

Rigid Analytics
Reduced Scope of Research



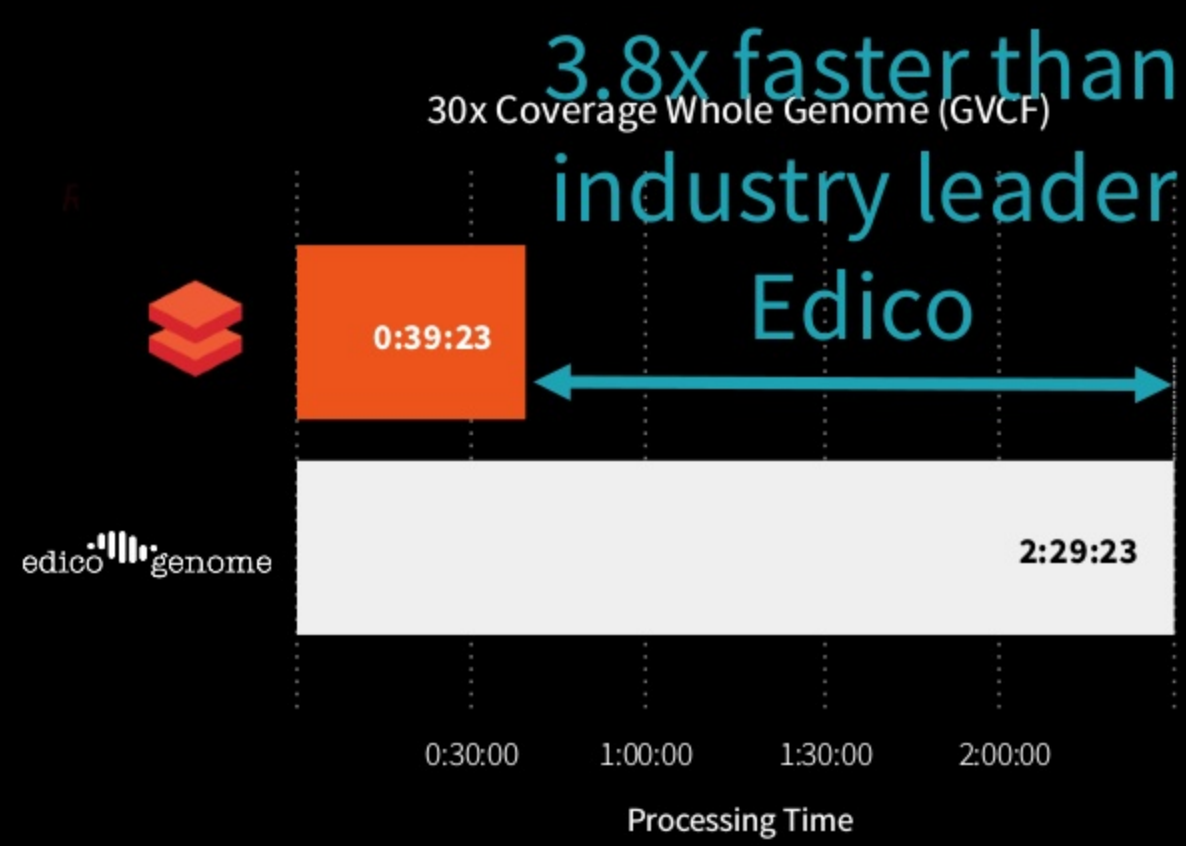
Siloed Teams
Lack of Productivity



databricks

Solution #1: Prebuilt Pipelines

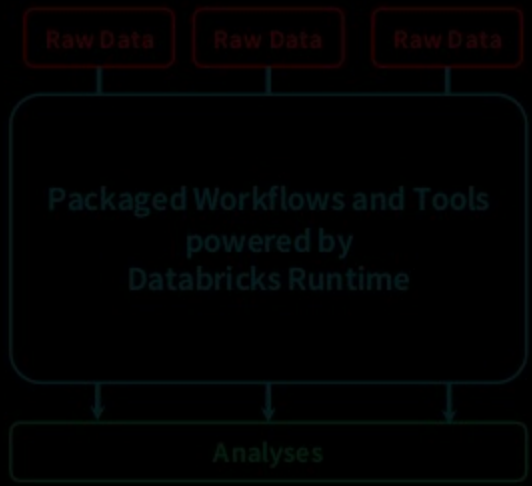
Best Practice Pipelines
"One click" execution



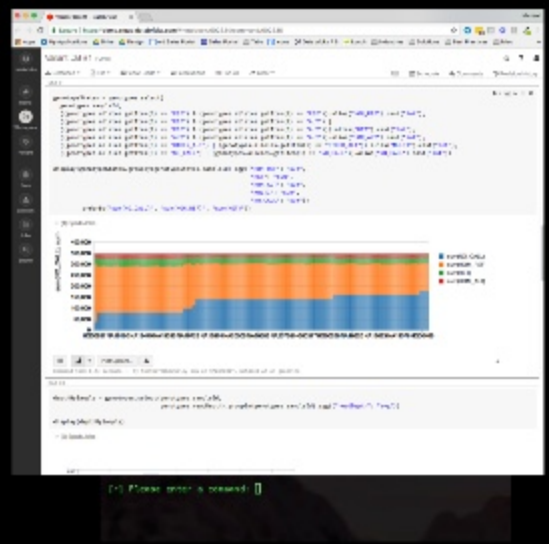
databricks

Solution #2: Powerful Analytics

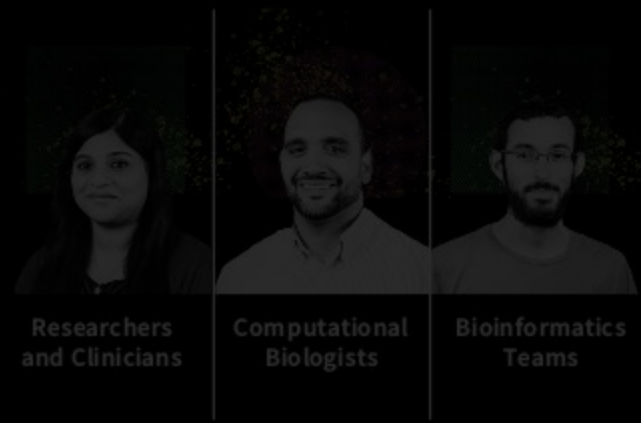
Best Practice Pipelines
"One click" execution



Powerful Analytics
From interactive queries to AI



Siloed Teams
Lack of Productivity

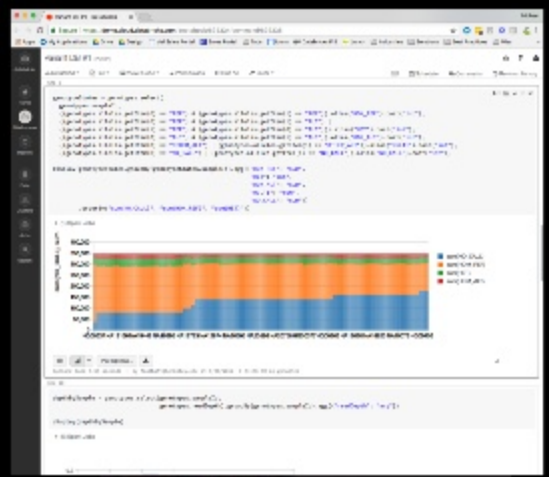


databricks

Solution #2: Powerful Analytics

"Queries on 60B+ genome associations in **3 seconds** vs. 30 minutes"

Powerful Analytics
From interactive queries to AI



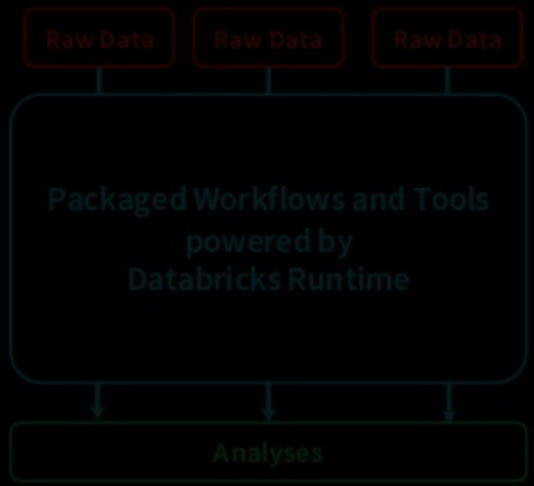
"Having the data is the first step, enabling drug development teams to answer questions with the data is how we are building the future of drug discovery."

Dr. Jeff Reid, Exec Dir at Regeneron

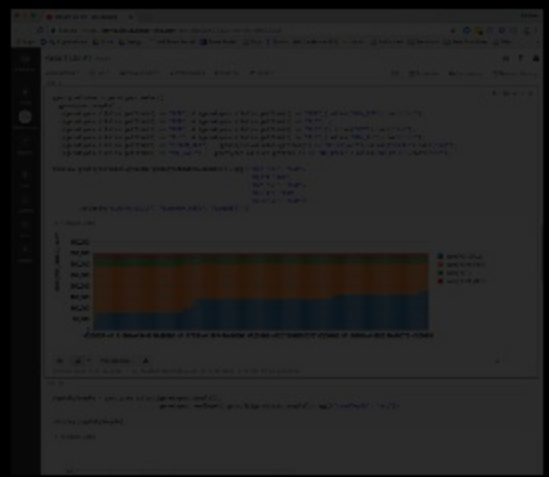
databricks

Solution #3: Collaborative Workspaces

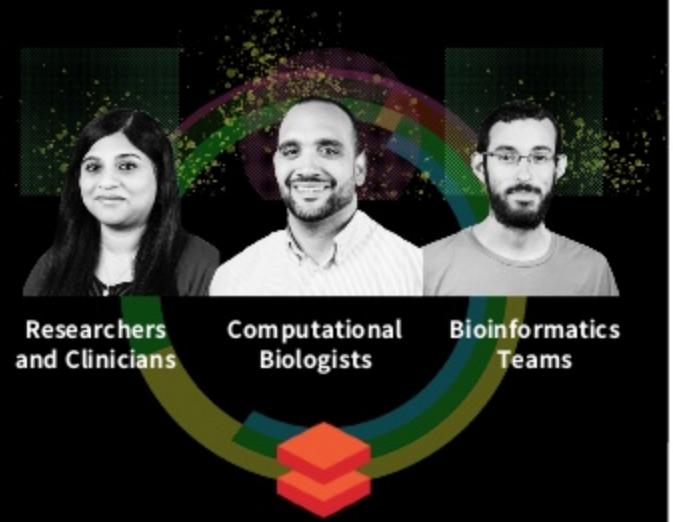
Best Practice Pipelines
"One click" execution



Powerful Analytics
From interactive queries to AI



Collaborative Workspaces
Dramatically Improve Productivity



databricks

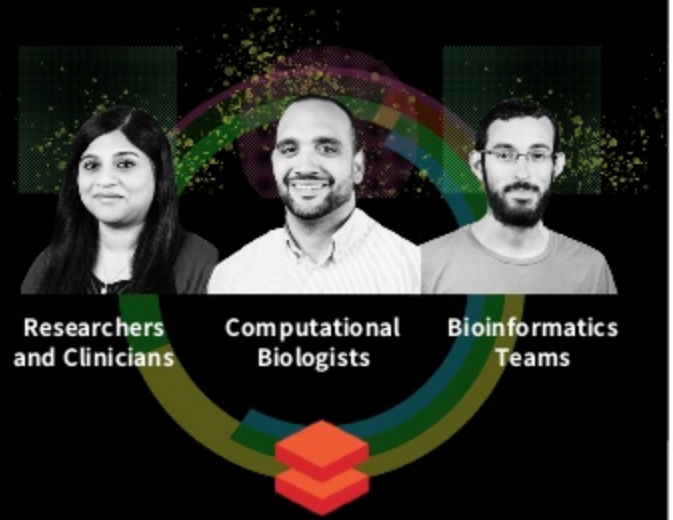
Solution #3: Collaborative Workspaces

"Databricks allows us to take clinical research and turn it into a clinically validated screen in far less time."

Sr. Director of Computational Bioinformatics, Lynn Carmichael

SANFORD HEALTH

Collaborative Workspaces
Dramatically Improve Productivity



Unified Analytics Platform for Genomics

All Your Genomic Data

databricks

Best Practice Pipelines

Tertiary Analytics and AI at Scale

Collaborative Workspaces

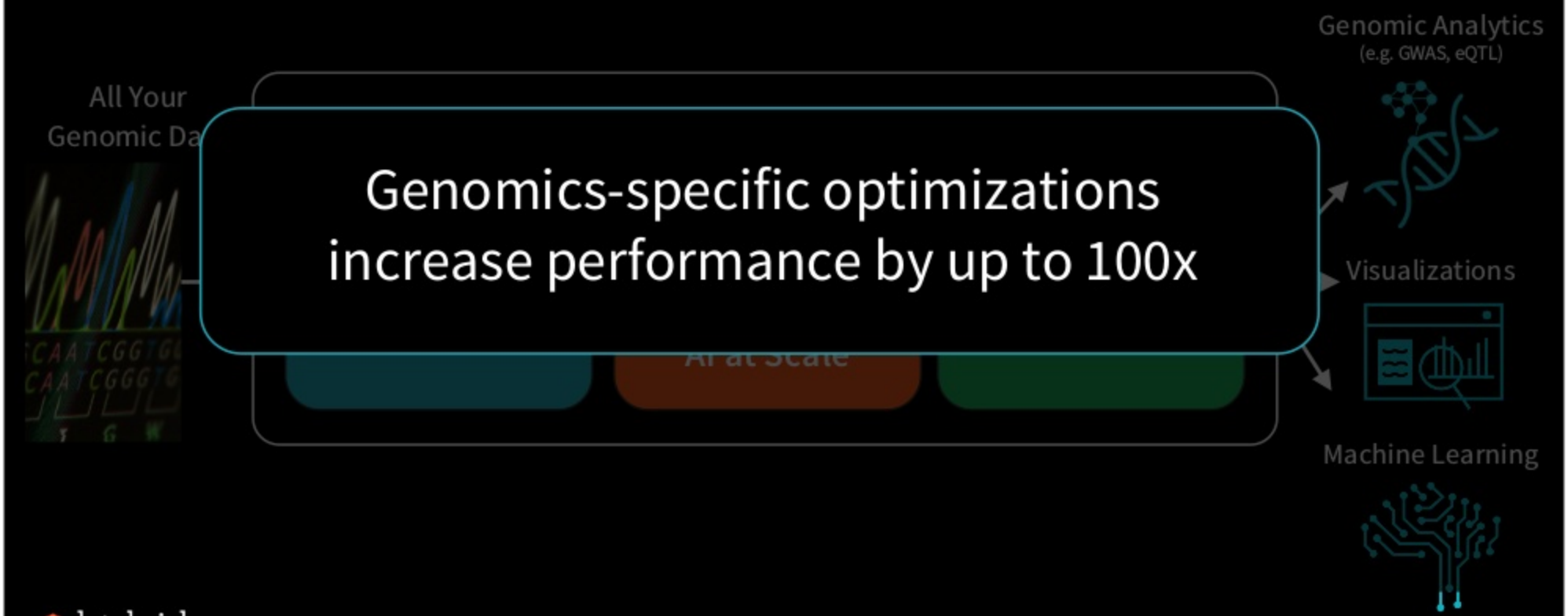
Genomic Analytics
(e.g. GWAS, eQTL)

Visualizations

Machine Learning

databricks

Unified Analytics Platform for Genomics



databricks


Sign-up for the preview
databricks.com/genomics
databricks

Accelerate Discovery

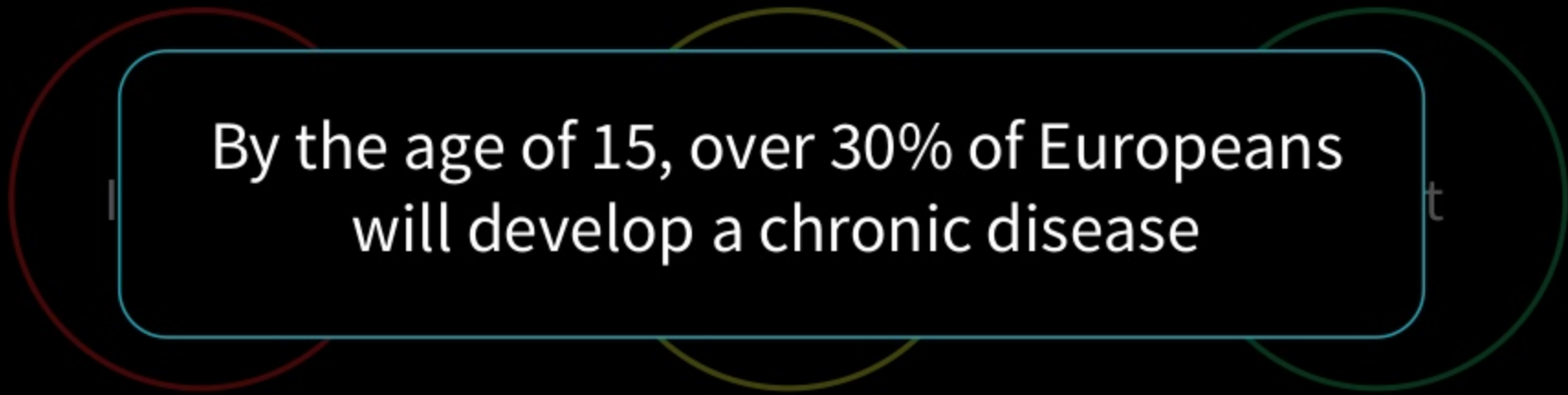
Demo: Preventing Disease with Genomics at Scale

Typical patient intake and treatment

Typical patient intake and treatment

...but this is very reactive and costly.

Typical patient intake and treatment



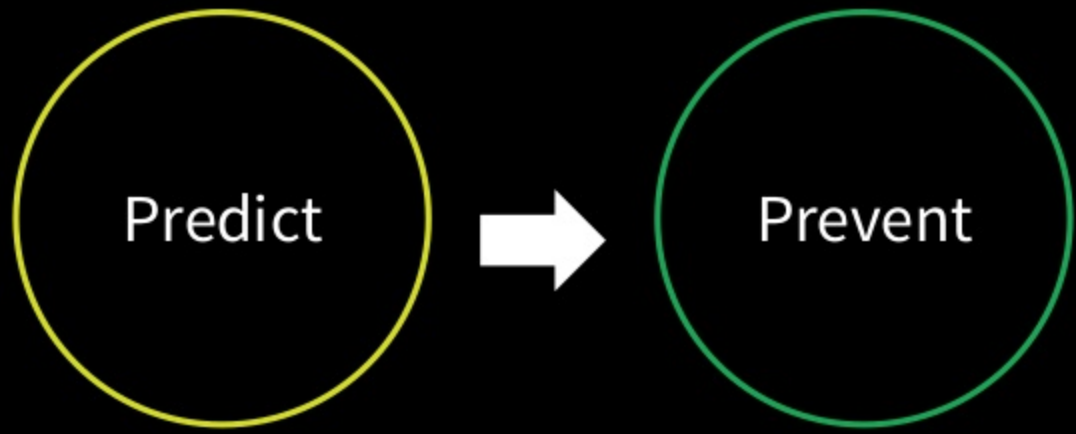
...but this is very reactive and costly.

Let’s shift our thinking

What if we could identify an individual’s risk for developing a disease and prevent that disease before it ever occurs?

The preventative care process

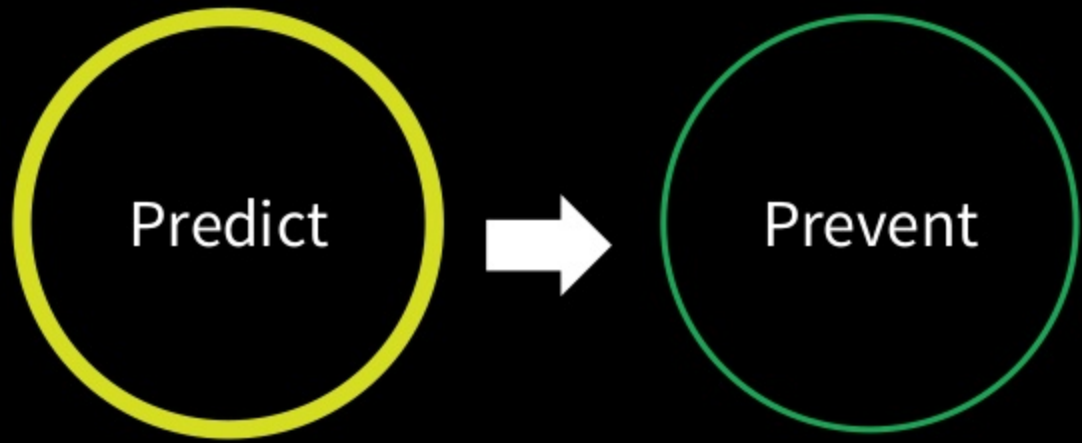
The preventative care process



Accelerated treatment improves outcomes

The preventative care process

Huge opportunity for genomics



Accelerated treatment improves outcomes

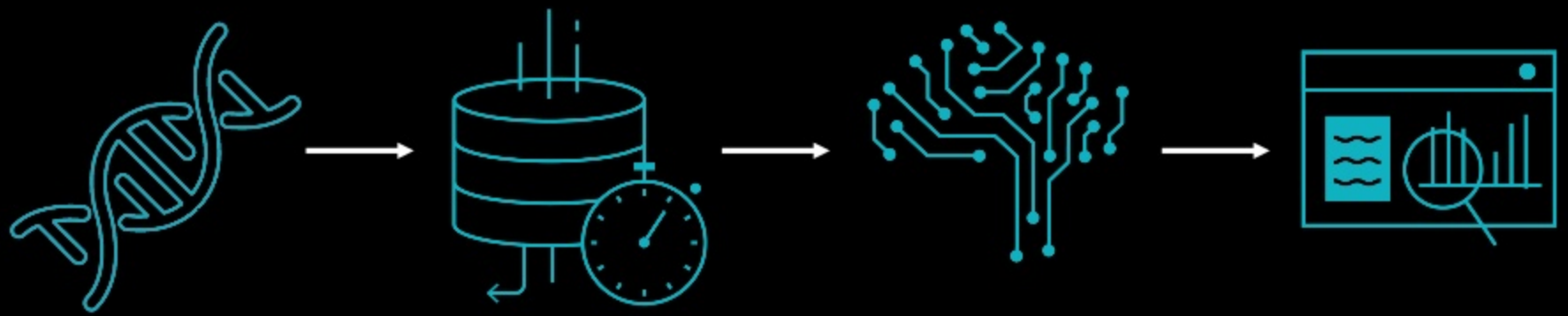
But genomic analysis is really hard

Population Scale Data Arrives (e.g. Biobank)

Process for Analysis

Export Model and Apply to Individual

Generate Dashboard for Clinician



databricks

Let’s try this with the Databricks Unified Analytics Platform for Genomics…


Sign-up for the preview
databricks.com/genomics
databricks