Day Zero Diagnostics (DZD)

The Future of Microbiology is Digital

Analog, culture-based diagnostics

Digital, sequencing-based diagnostics
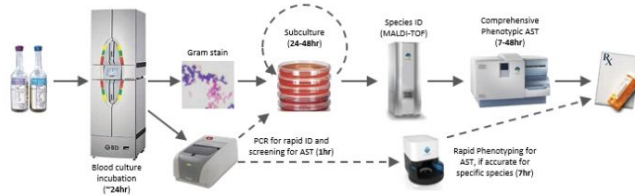
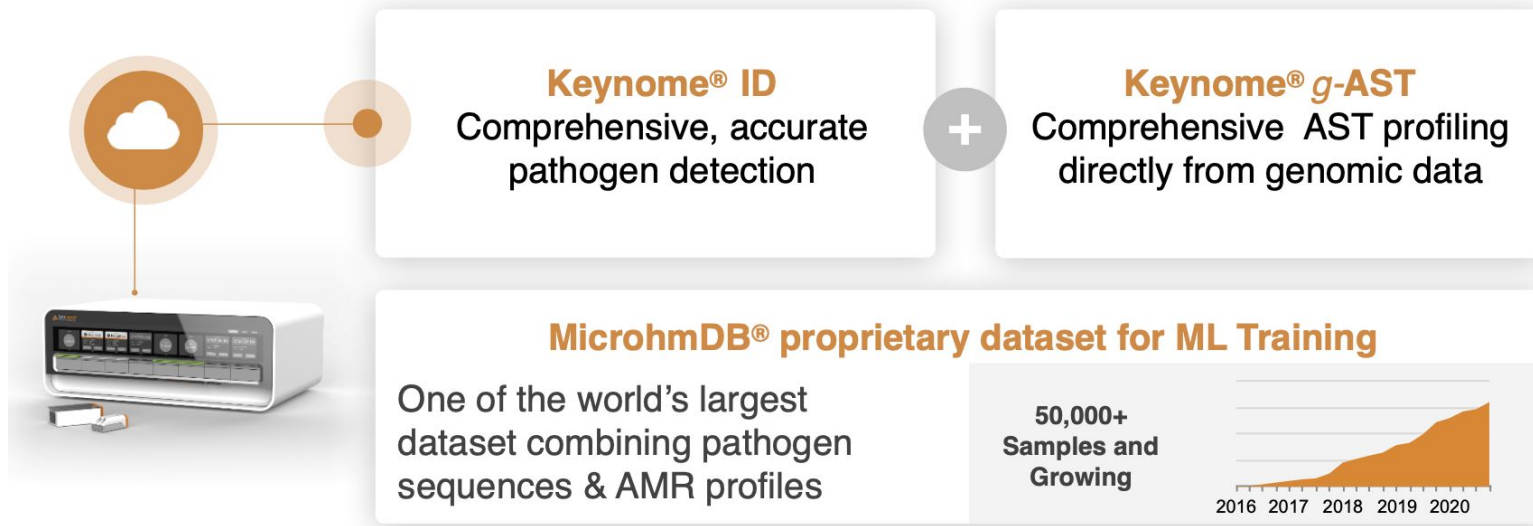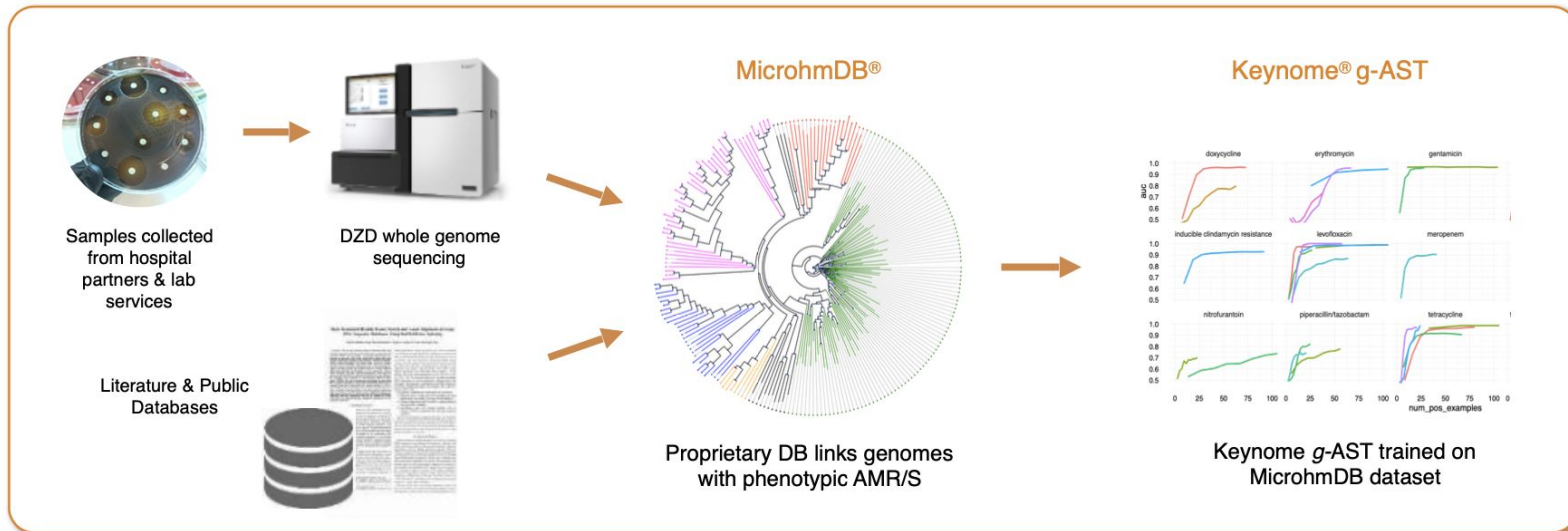|			
Slow	**Days to weeks** to diagnose	**Fast**	Answers within **hours**
Low Resolution	Fails to grow 40% of the time, requires multiple assays for full diagnosis	**High Resolution**	High sensitivity, full genomic interrogation in a single diagnostic
Limited	Optimized for core diagnostic use-cases	**Comprehensive**	Capable of comprehensive organisms, HAI & epidemiology, forecasting

Day Zero Diagnostics (DZD)

Advanced ML Algorithms
Complete diagnosis from genomic data

Keynome® ID
Comprehensive, accurate pathogen detection

+

Keynome® g-AST
Comprehensive AST profiling directly from genomic data

MicrohmDB® proprietary dataset for ML Training

One of the world's largest dataset combining pathogen sequences & AMR profiles

50,000+ Samples and Growing

2016 2017 2018 2019 2020

Day Zero Diagnostics (DZD)

Keynome® g-AST
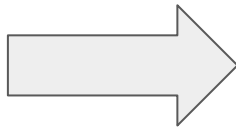## DZD Algorithms Are Trained on MicrohmDB for AST Calls

MicrohmDB: Cloud migration and versioned releases

When I took this over in late 2019, it was poorly designed and technical debt was slowing data integration and sourcing..
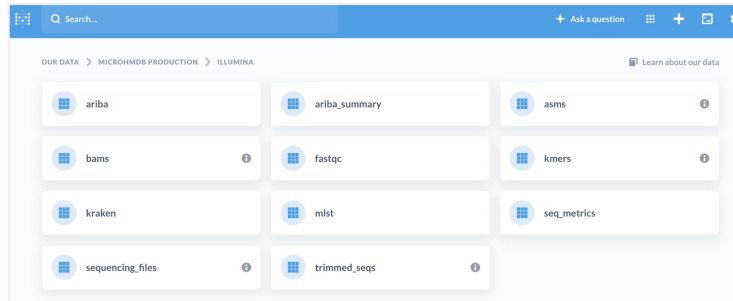
Legacy system

- Data not stored in cloud
- Compute not scalable
- Analysis not reproducible
- Data poorly modeled
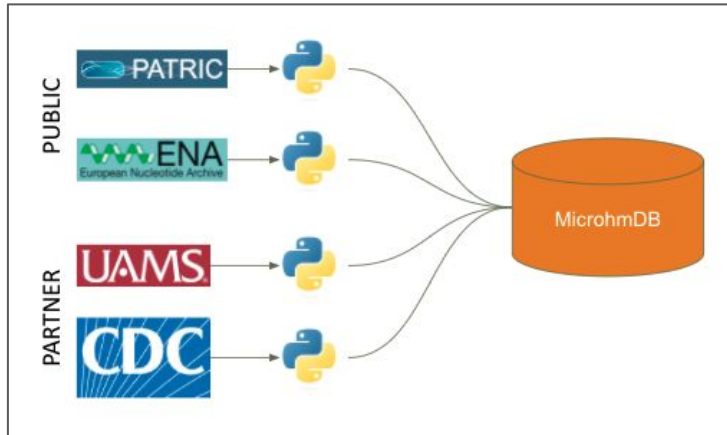- Data not versioned

MicrohmDB v1.0

- Data stored in cloud
- Scalable compute
- Analysis reproducible
- Simper, cleaner design
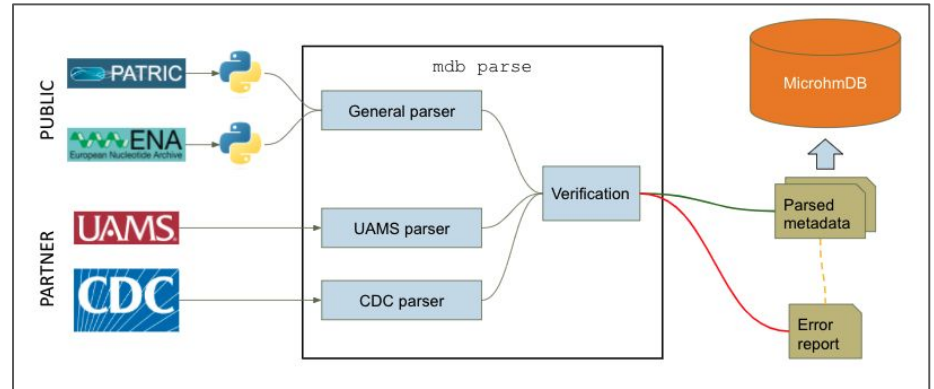- Data releases quarterly

MicrohmDB: Data integration and QC

.. as we rapidly integrated new data sources, we also wanted to put data quality checks/ verification in place to ensure we were only including high quality data
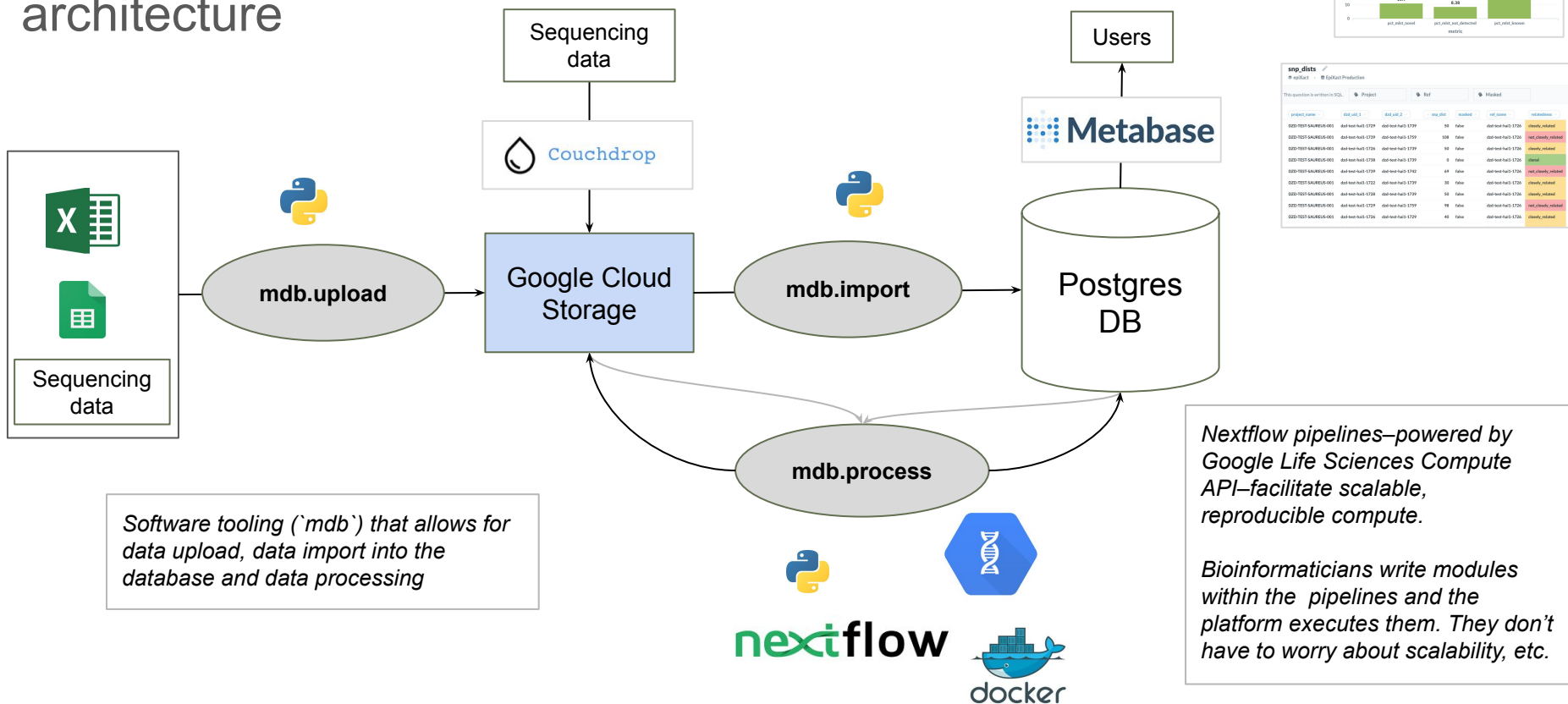
Legacy system

MicrohmDB v1.0

MicrohmDB:
Data platform architecture

Partners and vendors can transfer sequencing data directly over FTP

Users can interact with data through BI dashboards, making plots and saving common queries

Sequencing data

Couchdrop

Users

Metabase

Sequencing data

mdb.upload → Google Cloud Storage → **mdb.import** → Postgres DB

Sequencing data

mdb.process

Software tooling (`mdb`) that allows for data upload, data import into the database and data processing

nextflow docker

Nextflow pipelines–powered by Google Life Sciences Compute API–facilitate scalable, reproducible compute.

Bioinformaticians write modules within the pipelines and the platform executes them. They don't have to worry about scalability, etc.

MicrohmDB®
Large Scale Dataset of Pathogen Genomes and AMR Profiles

MicrohmDB: 50,000+ and growing

50,000

37,500

25,000

12,500

0

Non-Isolate Source
Hospital Partners & Biobanks

MDB v1.0

2016   2017   2018   2019   2020   2021
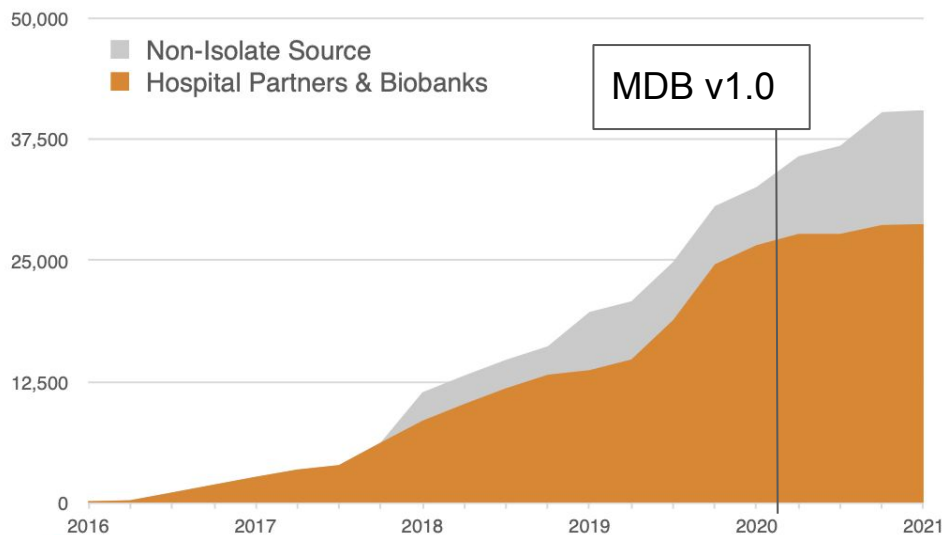
Customers, Contributors & Collaborators

MASSACHUSETTS GENERAL HOSPITAL
MGH 1811

UAMS®
University of Arkansas for Medical Sciences

UCIrvine

BWH BRIGHAM AND WOMEN'S HOSPITAL

NEW ENGLAND BAPTIST HOSPITAL

ihma

CANADIAN ANTIMICROBIAL RESISTANCE ALLIANCE CARA

ARLG
Antibacterial Resistance Leadership Group

bei RESOURCES
SUPPORTING INFECTIOUS DISEASE RESEARCH

New partners since 2019