Overcoming Barriers to the Adoption and Implementation of Predictive Modeling and Machine Learning in Clinical Care: Lessons from US Academic Medical Centers & Duke Health's Way Forward

Eric Poon, MD MPH FACMI
Chief Health Information Officer
Duke Health
Professor of Medicine
Duke University School of Medicine

Unlocking The Power of BIG DATA in Health

DukeHealth

Data Science Challenges in the Real World

Overcoming Barriers to the Adoption and Implementation of Predictive Modeling and Machine Learning in Clinical Care: Lessons from US Academic Medical Centers

Joshua Watson, MD MMCi; Carolyn Hutrya, MMCi; Shayna Clancy, MMCi; Anisha Chandiramani, MD MMCi; Armando Bedoya, MD MMCi; Kumar Ilangovan, MD MPH MMCi; Nancy Nderitu, MMCi; Eric Poon, MDM MPH

J Am Med Inf Assoc Open 2020 (In Press)

Unlocking The Power of BIG DATA in Health

DukeHealth

Study Goals

Given intense interest in Machine Learning and Predictive Modeling in healthcare:

- Define and disseminate the common challenges in developing and/or implementing machine learning models in healthcare organizations
- Understand and share current best-practices for overcoming these challenges

Study Methods (1)

- Semi-structured Interviews (~45 minutes) with Experts at Academic Medical Centers (AMCs)
- Targeted Experts: individuals who participate in the creation and/or implementation of Machine Learning/Predictive Modelling (ML/PM) algorithms
  - Recruitment through list of accredited Clinical Informatics Fellowship Directors
  - Assumption: AMCs are more likely to be ahead of non-academic institutions in encountering challenges and developing solutions

Unlocking The Power of BIG DATA in Health

DukeHealth

Study Methods (2):

- Discussion Guide Development
  - Definition of ML/PM in the Clinical Context
  - Literature review and Research Team Discussion
  - Life-Cycle Framework. Challenges & Best Practices during:
    - Model Development
    - Model Implementation
    - Model Maintenance

- Content Analysis of Interview Transcripts using Nvivo:
  - Grounded Theory approach
  - Preliminary → Final Code List
  - 3 pairs of Transcript Coders, each pair Coding 1/3 of Interview Transcripts

Unlocking The Power of BIG DATA in Health

DukeHealth

Characteristics of Informants:
33 Informants from 19 AMCs

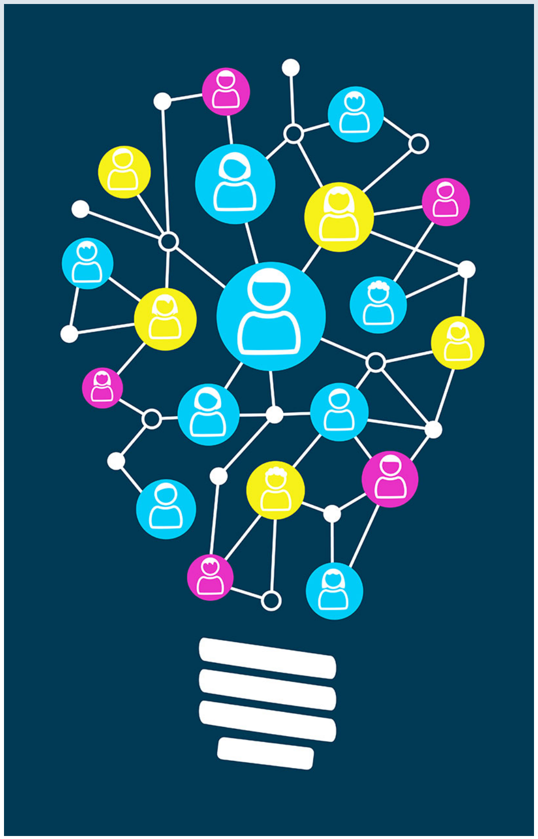
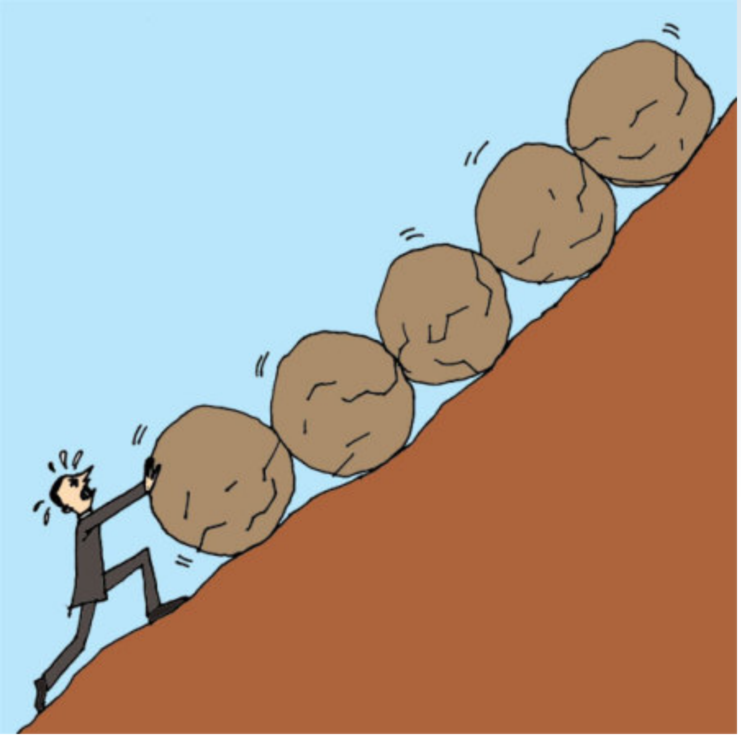


Table 1. Characteristics of informants and their home institutions

Location of institution by region ($n = 19$)	
Northeast	8
South	3
Midwest	3
West	5
Number of models in production ($n = 19$)	
<3 models in production	13
>3 models in production	6
Educational background of informants ($n = 33$)	
MDs	58%
PhDs	9%
Data Science	27%
MD/PhDs	6%
Seniority of informants ($n = 33$)	
Executive/senior role	64%
Non-executive	36%

Challenges & Facilitators: 5 Key Themes

- Culture and Personnel
- Clinical Utility
- Financing
- Technology
- Data

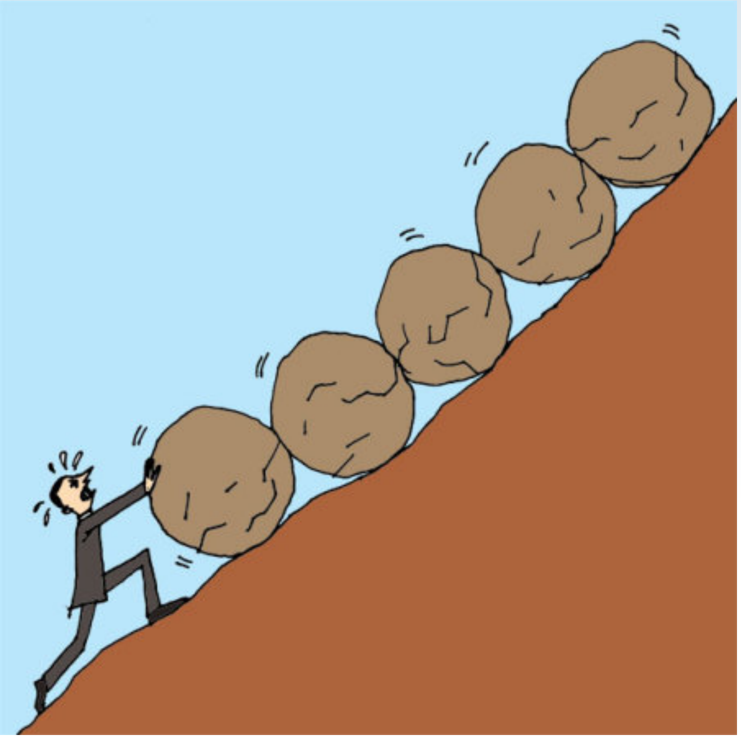


Unlocking The Power of BIG DATA in Health

DukeHealth

Challenges & Facilitators: 5 Key Themes

- Culture and Personnel
- Clinical Utility
- Financing
- Technology
- Data



Unlocking The Power of BIG DATA in Health

DukeHealth

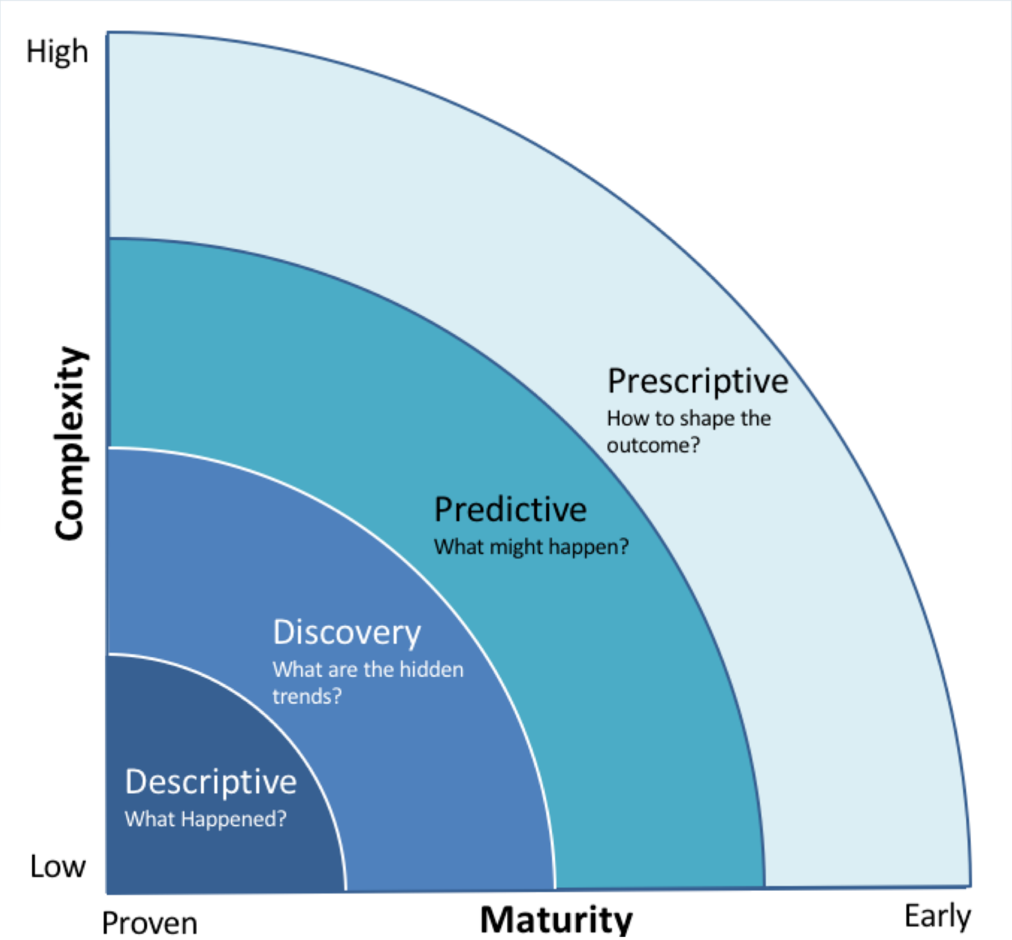
Culture and Personnel (1)

- Consensus Building & Direction Setting
  - Myriad of Stakeholders Needed in Collaboration:
    - Clinical and Data Experts need to Work Together From Early Stages of Model Development
    - "..more of the work is actually going to be focused on the intervention and the program to support that intervention in a sustainable way. The tech[nical] ..and analytics parts are getting easier and easier.."
  - Expectation Management: hype vs reality
  - Time Consuming

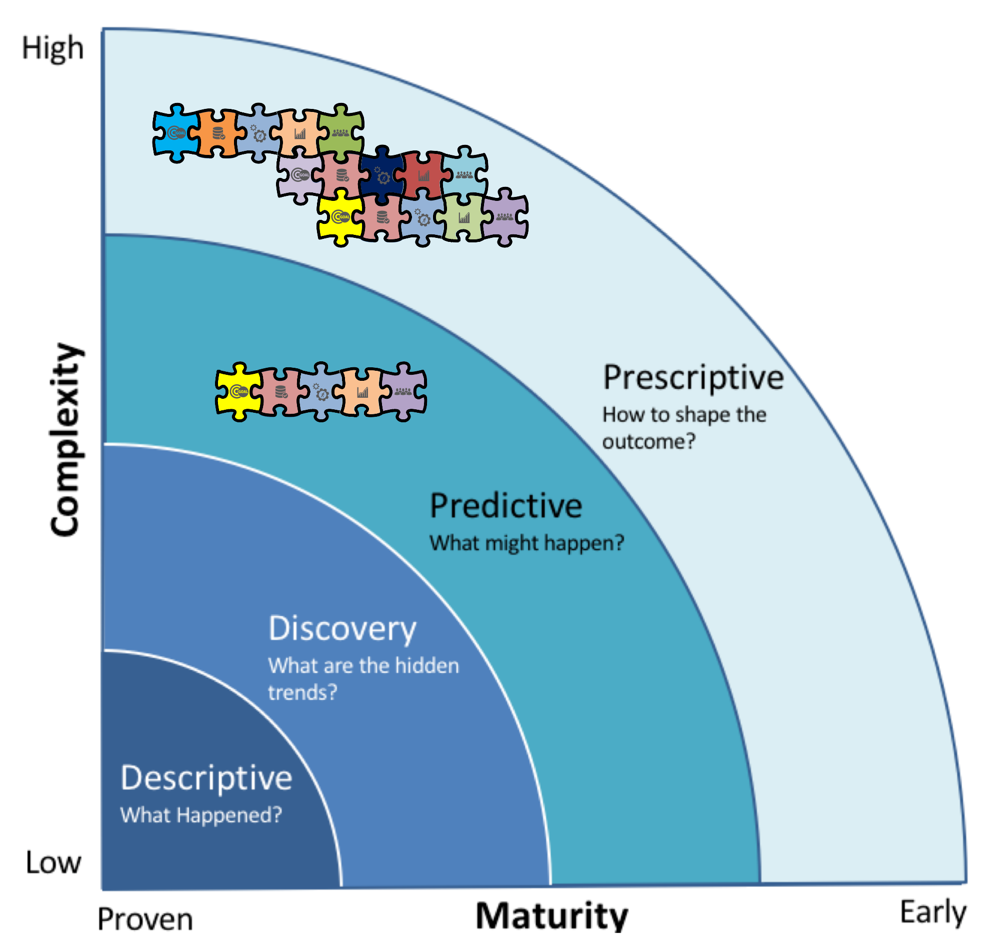
Unlocking The Power of BIG DATA in Health

DukeHealth

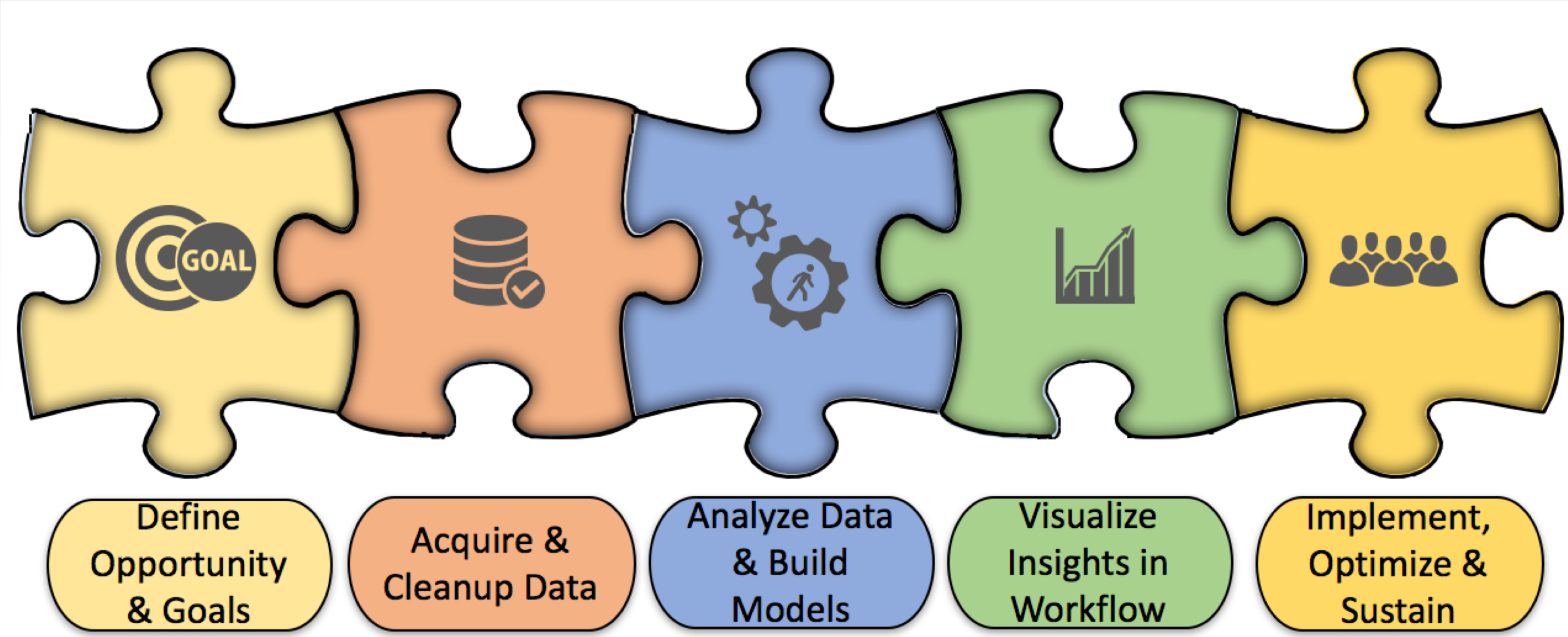
Spectrum of Analytics: Hindsight to Foresight



Spectrum of Analytics: Hindsight to Foresight



Anatomy of Impactful Advanced Analytics Projects

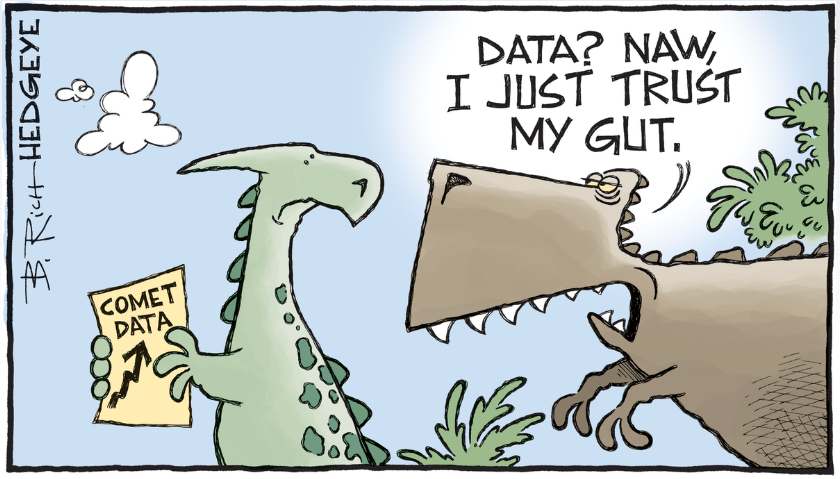


Unlocking The Power of BIG DATA in Health

DukeHealth

Culture and Personnel (2)

- Lack of Clinicians' Trust in ML/PM Model
  - Lack of Clinician Comfort Level with ML/PM Performance Metrics
  - 'Novel' algorithms often lack evaluation rigor/gravitas
  - Many high-performing models are 'black box' algorithms



Unlocking The Power of BIG DATA in Health

DukeHealth

Building Clinicians' Trust in ML/PM

- Transparency by Vendors & Model Developers
  - Reveal 'Inner Workings'
  - Interpretable AI
  - Retraining/Retuning of Model using local data
  - Evaluation Framework

Culture and Personnel (3)

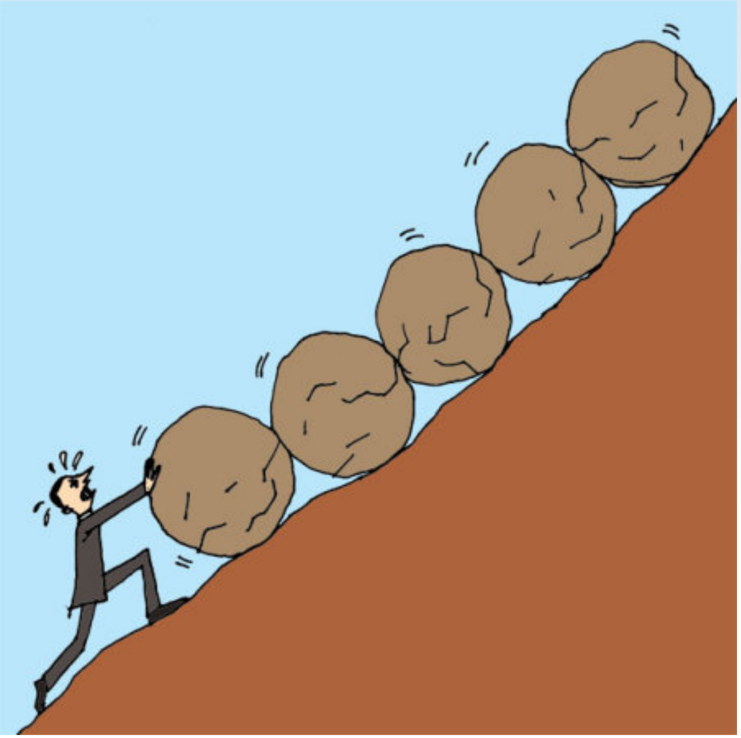
- Demand for staff skilled in development <u>and maintenance</u> of models far outstrip supply
- Vulnerability with Staff Turnover:
  – "When people with institutional knowledge [and] knowledge of models in production…move on to other institutions,.. It creates a knowledge gap"

Unlocking The Power of BIG DATA in Health

DukeHealth

Challenges & Facilitators: 5 Key Themes

- ✓ Culture and Personnel
- Clinical Utility
- Financing
- Technology
- Data



Unlocking The Power of BIG DATA in Health

DukeHealth

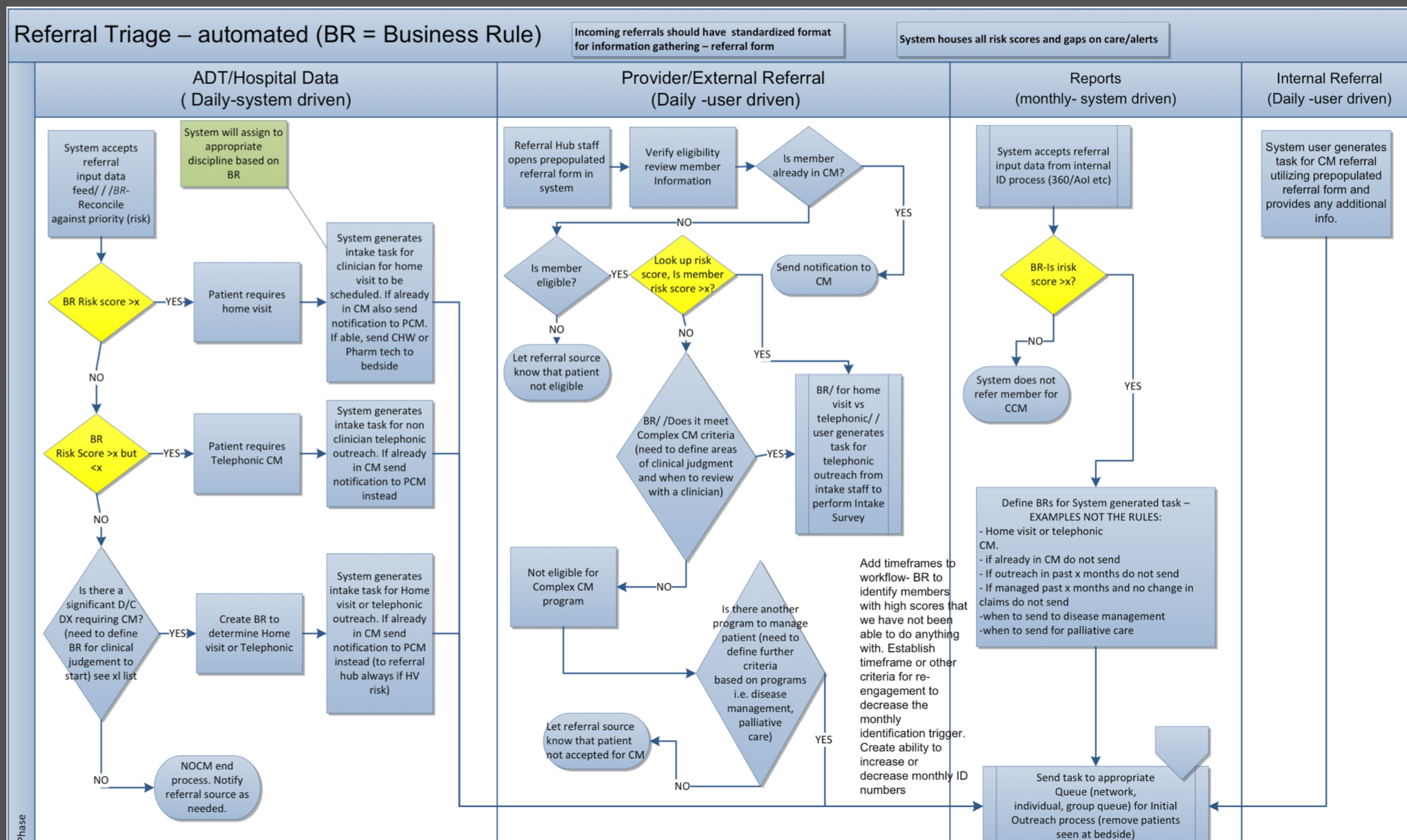
Clinical Utility

- Models need to be created with the end-goal in mind
  - Risk Stratification → Actionable advice → clinical intervention
  - "They validated that higher scores [indicate] a high risk of readmission, but what we don't know for sure whether you can do something about it"
- Will clinicians pay attention?
  - Alarm/Alert fatigue
  - "What signals should you send, who(m) to send them to, when, and how."

Unlocking The Power of BIG DATA in Health

DukeHealth

Workflow, Workflow, Workflow…



Predictive Model Evaluation Framework: v1.0 @ Duke

Decision to Pursue Silent Evaluation

Decision to Start Clinical Pilot

Decision to Scale Up/Disseminate

Decision to Retire

Model Development

- Project Goal & Outcome Definition
- Retrospective Model Evaluation to Determine Predictive Potential

Silent Evaluation

- Prospective Model Performance Evaluation
- Clinical Validation of Model Performance

Clinical Pilot

- Clinical Adoption Evaluation
- Clinical Impact Evaluation
- Iterative Refinement of Intervention

Ongoing Monitoring

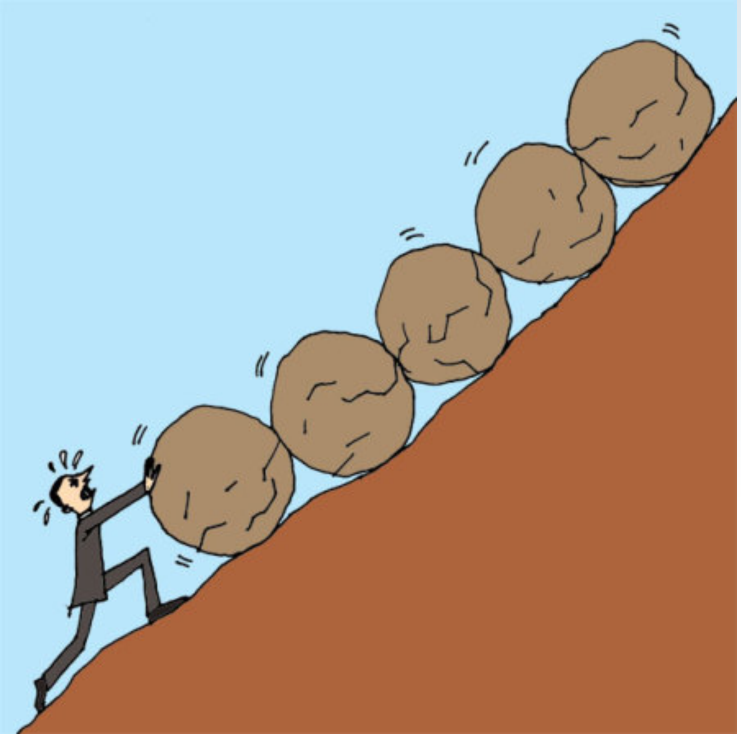
- Data Drift Monitoring
- Workflow & Clinical Impact Monitoring

Unlocking The Power of BIG DATA in Health

DukeHealth

Challenges & Facilitators: 5 Key Themes

- ✓ Culture and Personnel
- ✓ Clinical Utility
- Financing
- Technology
- Data



Financing

- Not hard to get started, but challenging to complete a project
  - Data Wranglers
  - Data Scientists
  - Integration
  - Change Management
  - Model Maintenance
- Mishmash of funding without sustainability plan
- Vendors may have a role particularly for smaller institutions, but no standard integration/interoperability approaches

Unlocking The Power of BIG DATA in Health

DukeHealth

Local Governance and Oversight

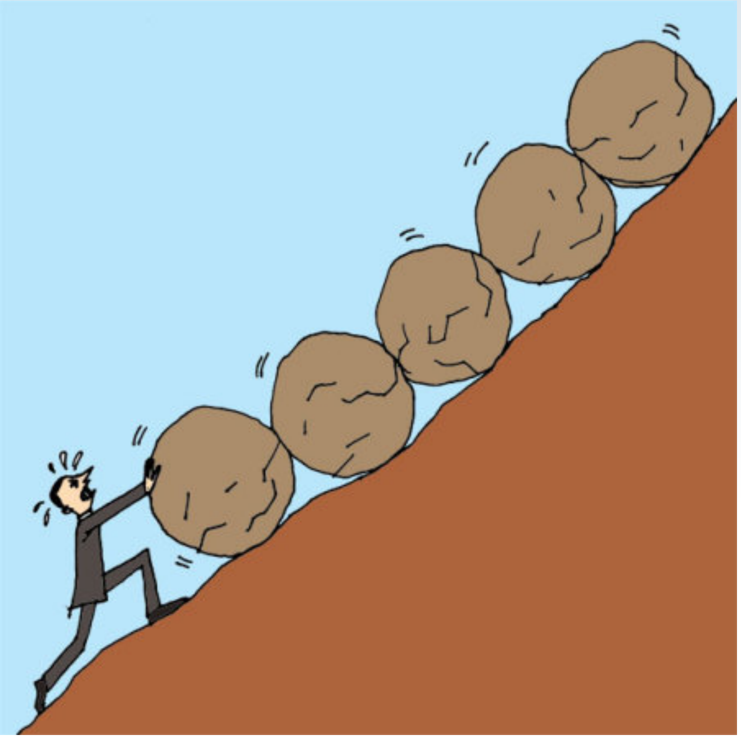
- Building Predictive Modelling on top of Standard of Care and Established Best Practices
- 5 Rights of Predictive Modeling
- Manage Prediction Fatigue Proactively
- Ensure initial and ongoing accuracy of predictive models through evaluation framework
- Oversee communication and user education

Unlocking The Power of BIG DATA in Health

DukeHealth

Challenges & Facilitators: 5 Key Themes

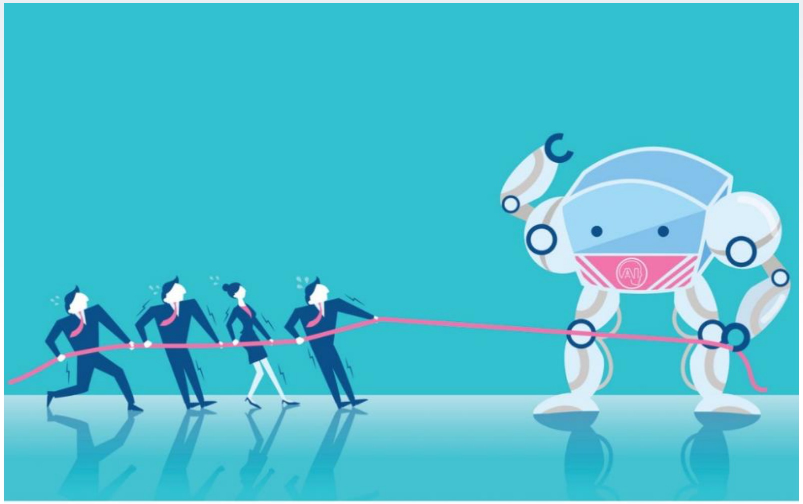
- ✓ Culture and Personnel
- ✓ Clinical Utility
- ✓ Financing
- Technology
- Data



Unlocking The Power of BIG DATA in Health

DukeHealth

Technology Challenges

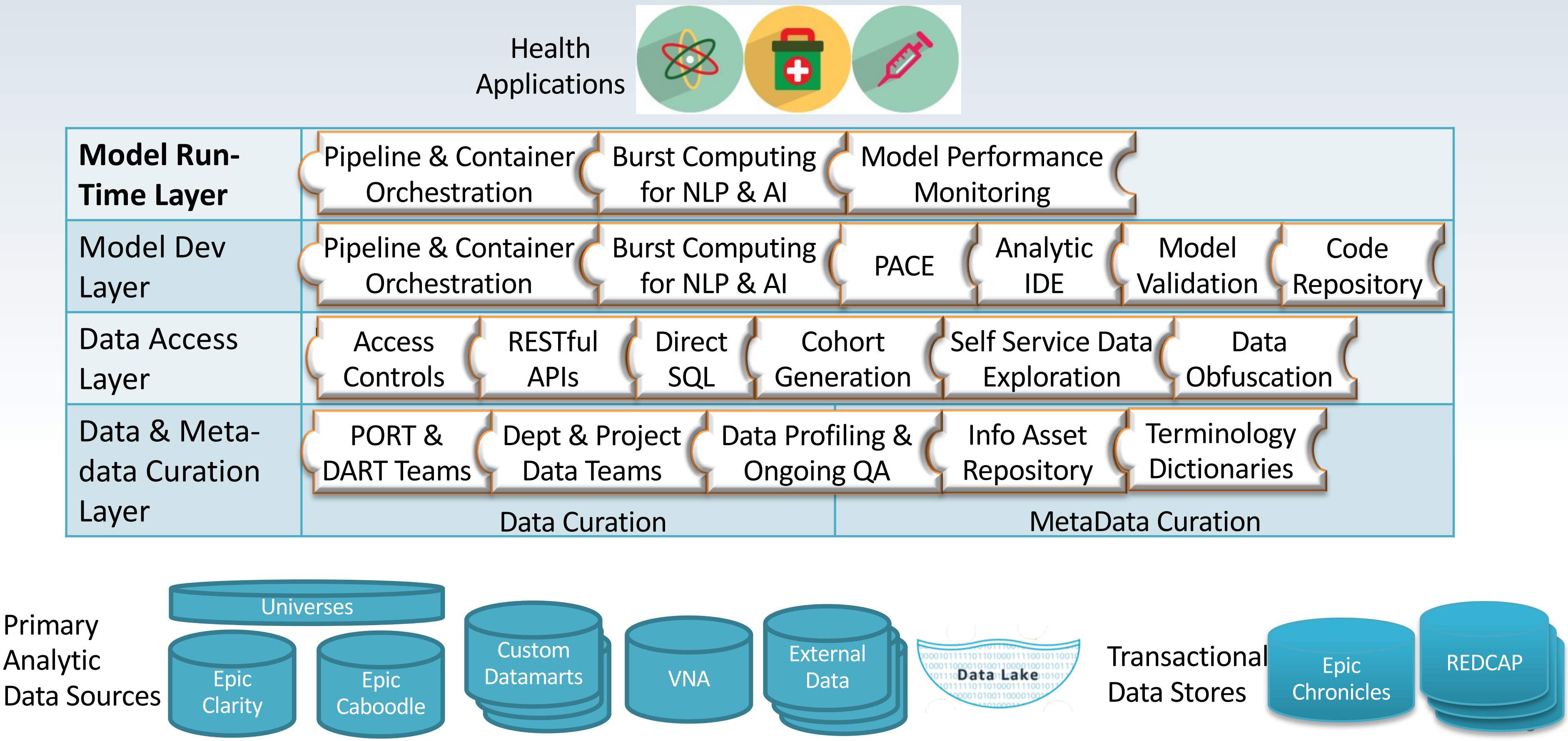


- Integration with EHR workflow
  - "How do you [implement ML/PM] in a standardized way [so] that a third party can build one tool and deploy it across different EHRs?.. Right now, none of the vendors have really sophisticated ways. The process is extremely naïve right now"
- Immature technology platform for PM and ML
  - "A key challenge is integration with legacy technology… Machine learning leverages the most modern tools, .. And integrating these modern tools with existing frameworks can pose interesting challenges"

Unlocking The Power of BIG DATA in Health

DukeHealth

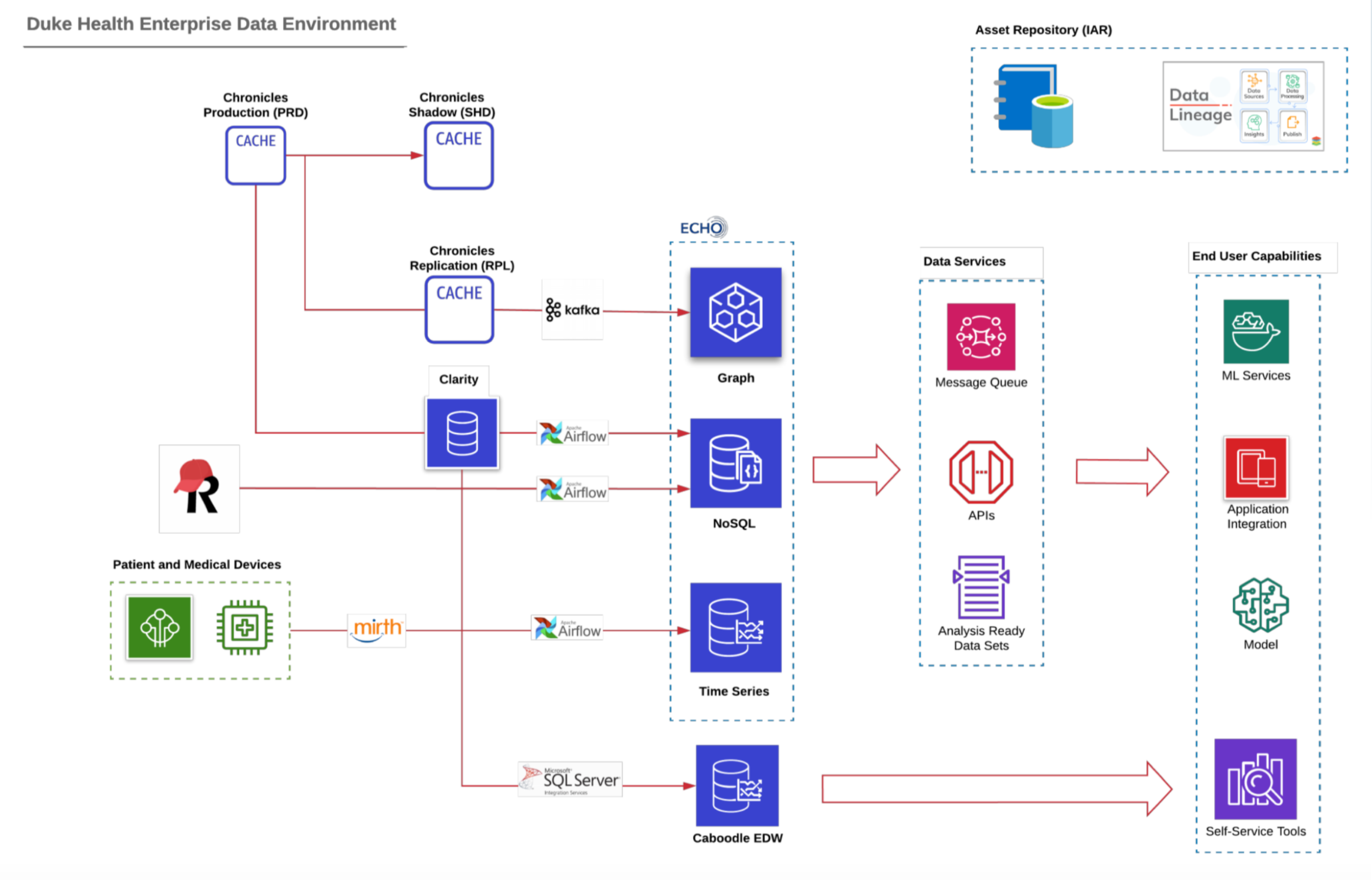
Data Science Technology Stack – Reference Architecture



Unlocking The Power of BIG DATA in Health

DukeHealth

Data Architecture for Now & Future



Unlocking The Power of BIG DATA in Health

DukeHealth

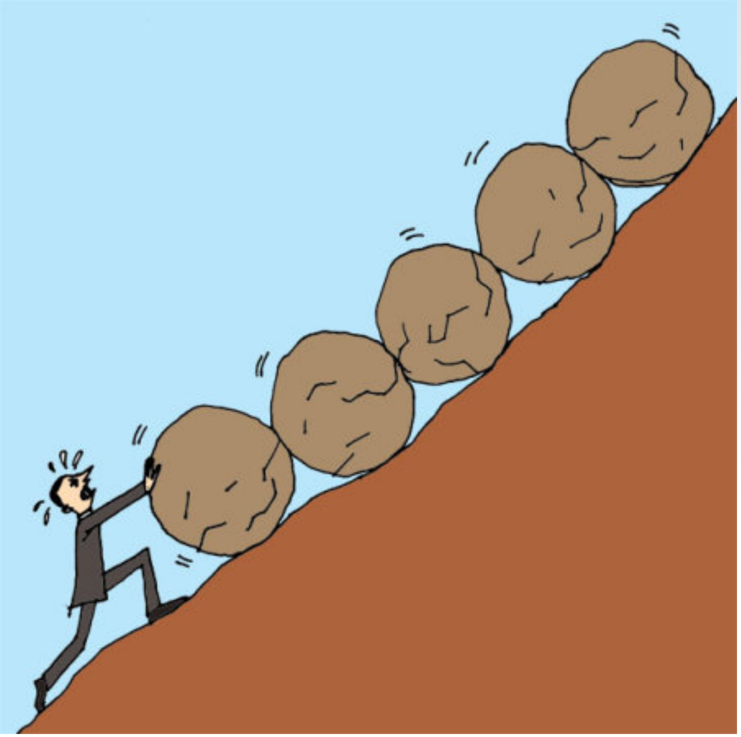
Progress Towards Reference Architecture

07/19 — 10/19 — 01/20 — 04/20 — 07/20

	Q1	Q2	Q3	Q4
Programmatic Access	✓ MongoDB & Sequelize Templates ✓ MongoDB from POC to PRD ✓ Airflow + RabbitMQ Install ✓ Mirth Interface Engine Upgrade ✓ Read (GET) API against IAR	✓ 45% of USCDI as APIs ✓ Update (POST) API against IAR ✓ Foundational Pipeline in Prod ✓ Clinical Notes in MongoDB	• 80% of USCDI as APIs • Philips Monitors Pipeline • ODS Replication Service (Pipeline)	• 100% of USCDI as APIs • API Environment Awareness (Data Sockets) • Data Extract Service (Pipeline)
Enterprise Content	✓ Epic CCD Pipeline w/ Prometheus	✓ Observations (Vitals) Data Mart ✓ Git Hosted Repository of Microservices	• Chronicles Data Provenance • Analytics DS-a-a-Service (BERD) • Redcap Pipeline (Surgery Data Project)	• Pub/Sub Exchange • Patient Generated Data (PGHD) • Chronicles "Echo"
Self-Service & Custom Analytics	✓ Datastage Retirement (ongoing 7/2021) ✓ Caboodle Server Upgrade ✓ Geocoding in Caboodle ✓ Caboodle access within PACE	✓ Data Prospector on LMS ✓ User Agreements and Use Cases in Phases ✓ Slicer-Dicer/DEDUCE integration	• Data Prospector Curriculum for Performance Services • Expanded Financial Data Model	• DEDUCE as Insight Pilot (Platform Agnostic) • Slicer Dicer Rollout (contd.)

Challenges & Facilitators: 5 Key Themes

- ✓ Culture and Personnel
- ✓ Clinical Utility
- ✓ Financing
- ✓ Technology
- Data



Unlocking The Power of BIG DATA in Health

DukeHealth

Data Challenges

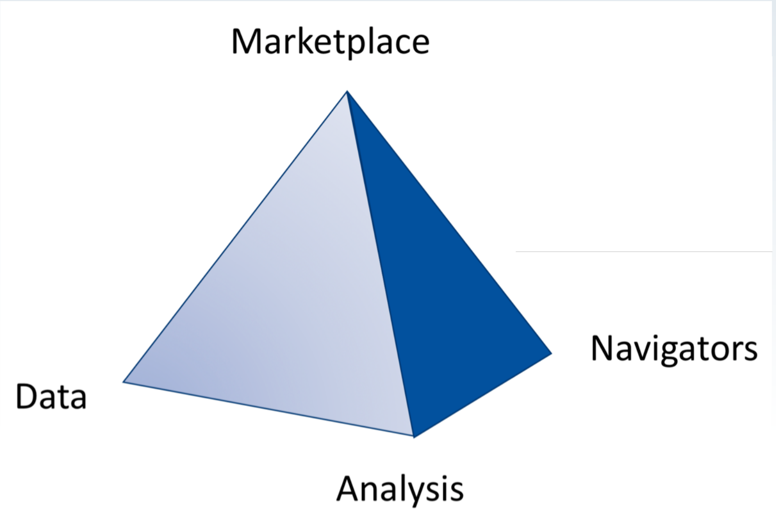
- Variable Data Quality from EHRs that May not reflect clinical reality
  - "There are many times when we pull data that look meaningful and carries a lot of signal, and when we talk to the clinicians, that data filed is laughed at. They don't use it, or it's just something that they have to enter for billing reasons, and they are horrified that we would actually use that in a model."
- Unstructured Data & Associated Methodological Challenges
- Variable Workflows across Health Systems

Unlocking The Power of BIG DATA in Health

DukeHealth

Data and Analytics Navigation:

Duke Data & Analysis Resource Center

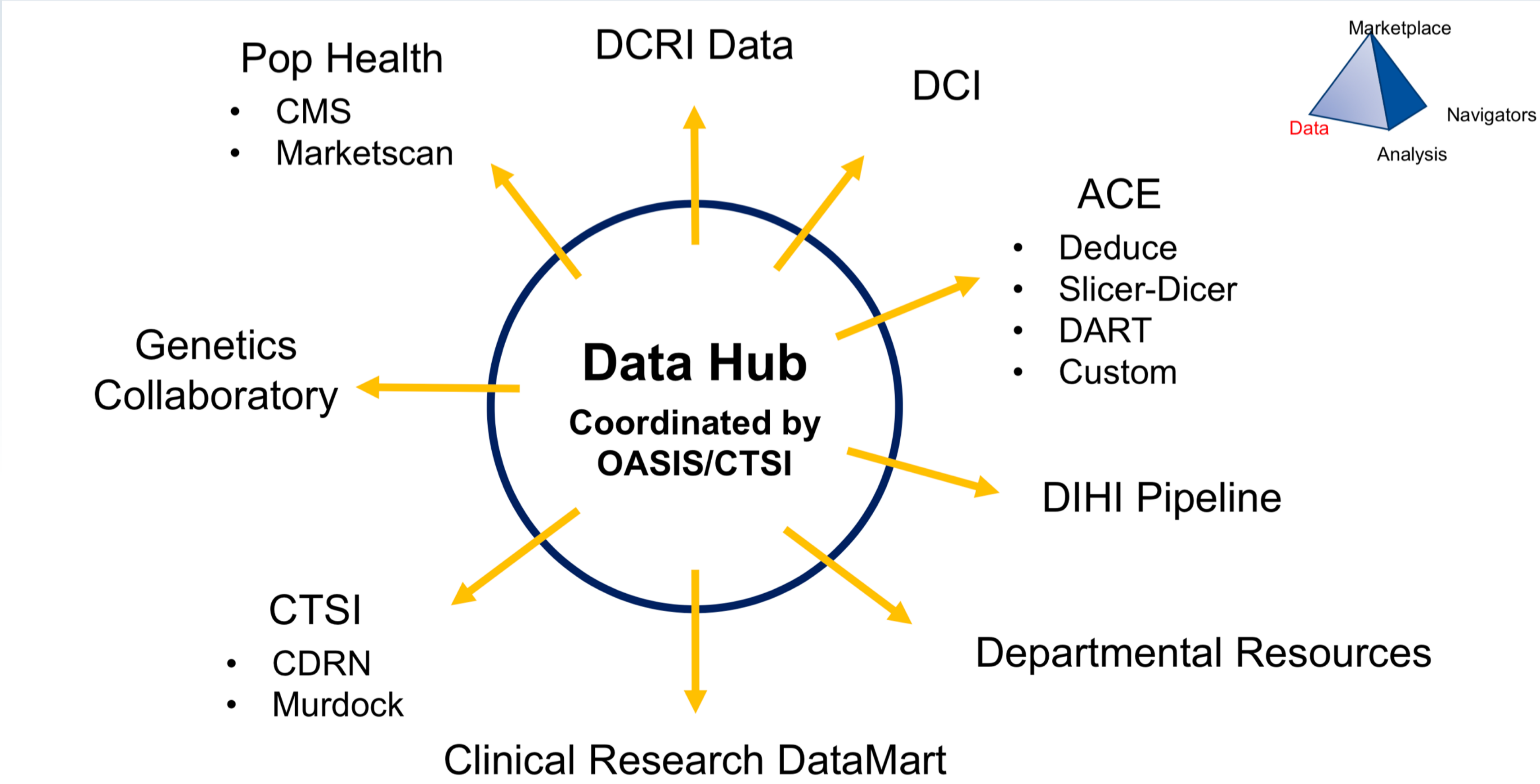


- A Front Door (one-stop shopping) to data resource for research at Duke (and beyond)
- Self-serve or facilitated (warm hand-off) access to data assets
- Connection to Appropriate Methodological Experts

Unlocking The Power of BIG DATA in Health

DukeHealth

Data Hub



Unlocking The Power of BIG DATA in Health

DukeHealth

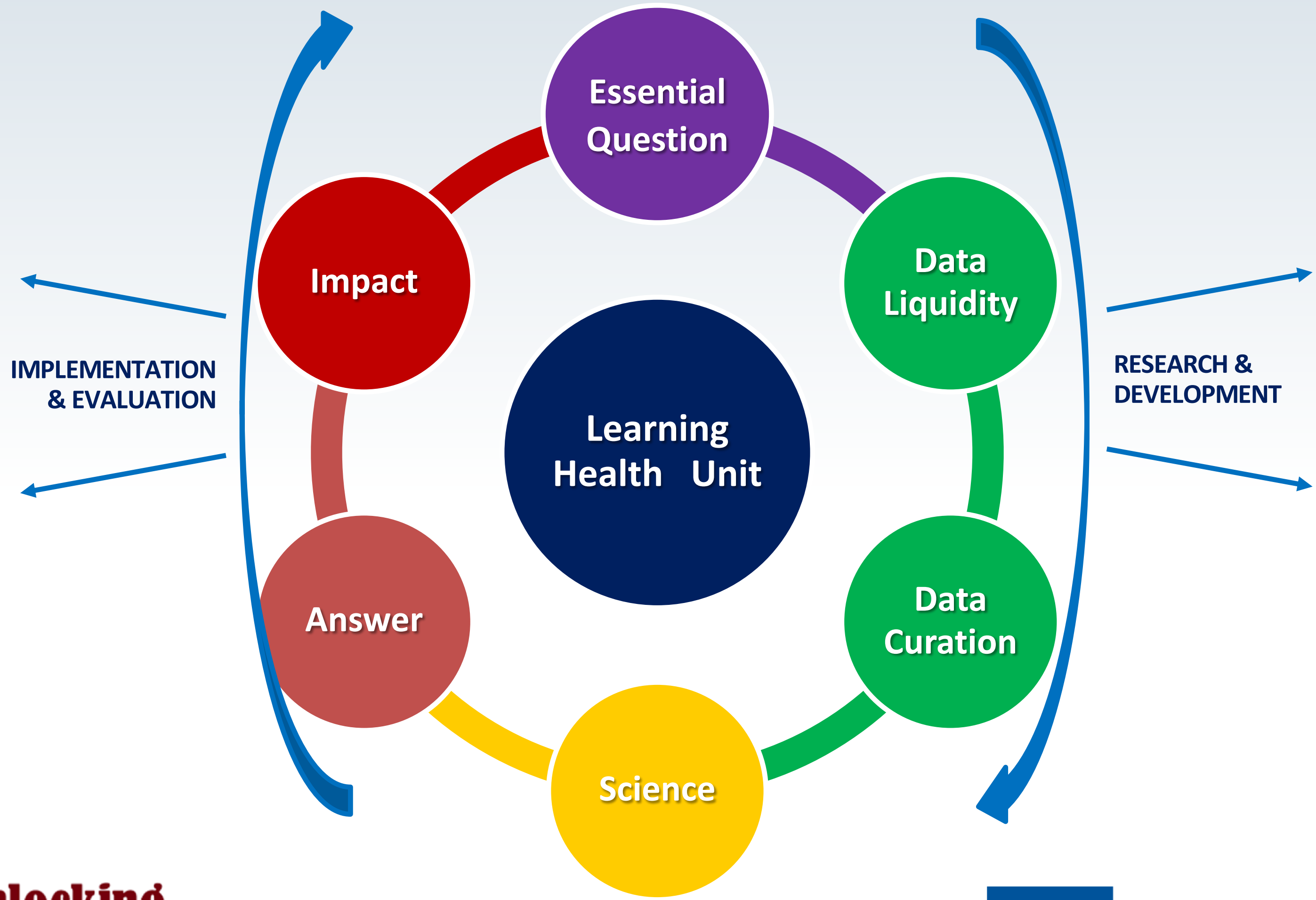
Approaches for Servicing the Data Hub

	Self-Service: Data Discovery Tools	Self Service: Data Science Pipeline	Custom Data Set Preparation	Shared Services: Dedicated FTEs	Shared Services: Dedicated Team
Offering					
Examples	DEDUCE Slicer Dicer	FHIR; Pythia; Data Sockets	Ad Hoc Fee-for-Service Data Extracts	Ortho; OB/Gyn; Transplant	PORT; PHMO Analytics
Cost	Free	$ to build Free to Reuse	$	$$$	$$$$$

Learning Health Units:
Bridging Research, Quality Improve and Data Science

Derivatives:

- Embedded data science
- Implementation science
- Clinical innovation



Derivatives:

- Curated cohorts for research
- Workforce development
- Academic advances

Key Insights from Duke and Other AMCs

- Big Data Needs to be Played as a Team Sport
  - Multi-disciplinary approach
  - 'End-to-End' Thinking
- Key Facilitators
  - Data Navigators
  - Clinical Workflow and Mindflow Expertise
  - Novel yet Scalable Technology Deployed on Reliable Platform
  - Comprehensive Evaluation and Sustainability Plan
  - Governance and Coordination

Thank You!

- Eric Poon, MD MPH FACMI
- eric.poon@duke.edu