

# Big Data Management and Advanced Analytics in Healthcare



*Jason Burke*  
*System VP and Chief Analytics Officer*  
*UNC Health Care*

# Agenda

---

**Define the evolving role of data management in the field of health care “big data”**

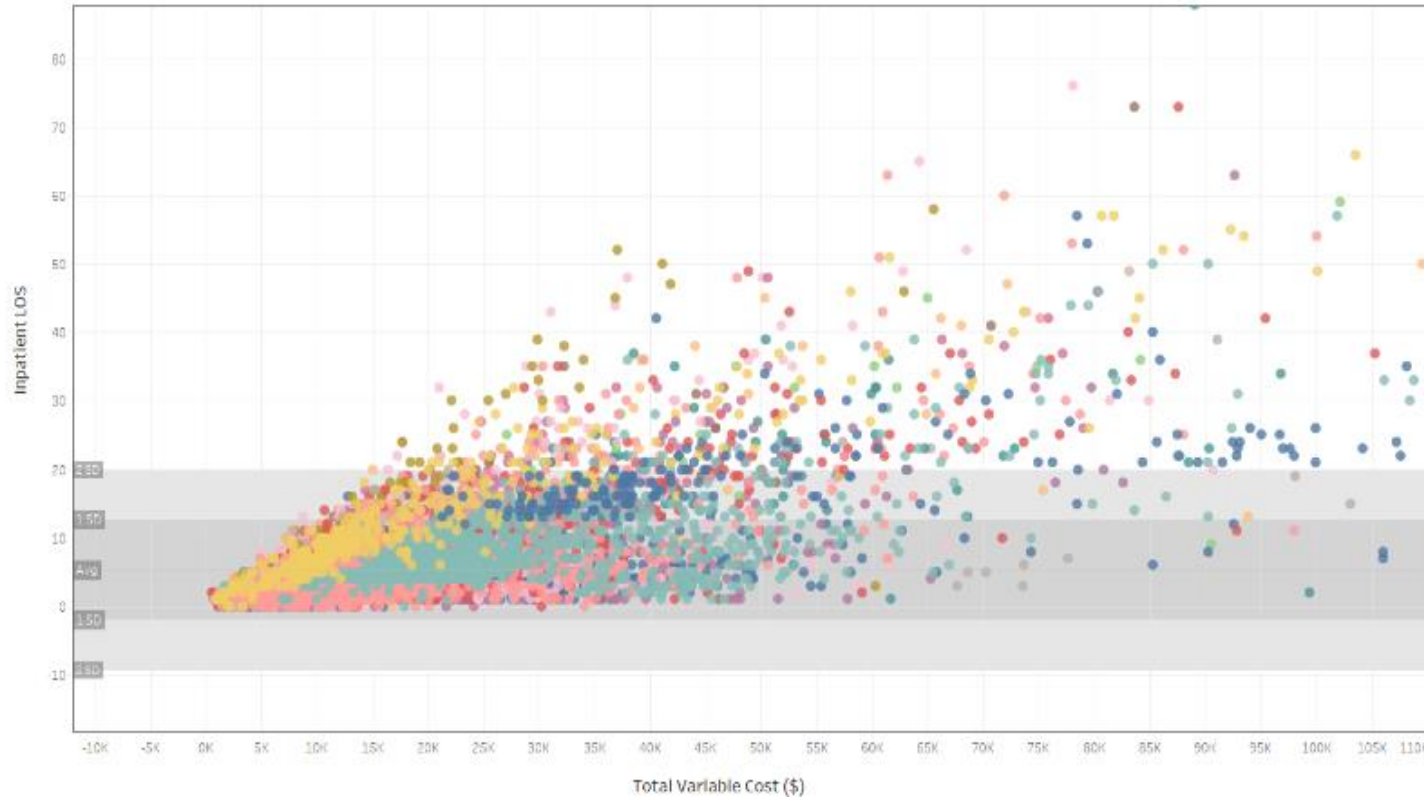
**Discuss some of the new competencies for effectively leveraging big data at scale**

**Share some examples of how this is being applied at a large health care system**

# Here's Your Report!

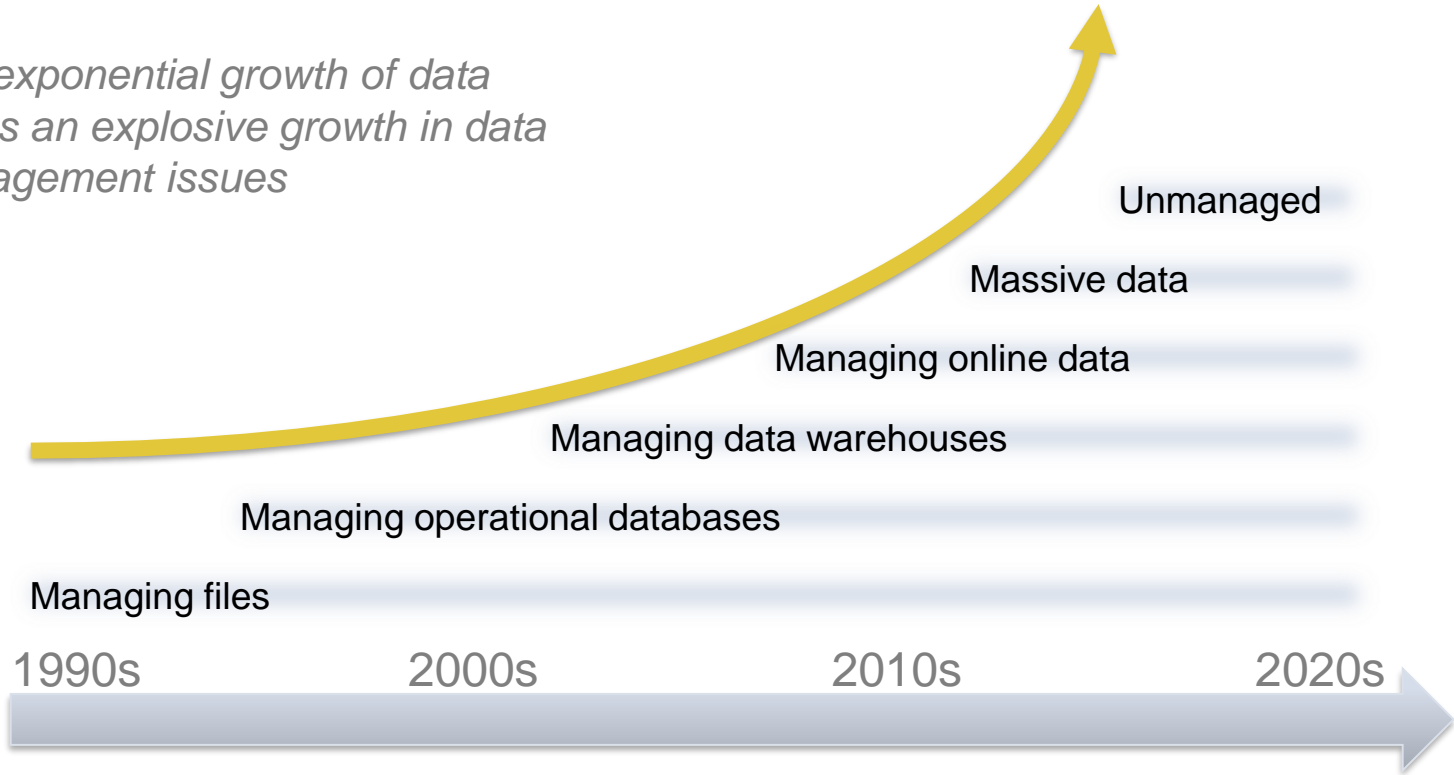
*Starting with the end in mind...*

Total Variable Cost vs. Inpatient LOS  
Attributed by Service Line (color)



# The Changing Role and Scope of “Data Management”

*The exponential growth of data drives an explosive growth in data management issues*



# Human Valuation and Big Data

---

## Valuation Area

**What is the source of the data?**

**What do we know about this data and what it means?**

**What's in the data; how representative is it?**

**How was the data obtained and managed?**

**Who else is using this data?**

**How consistent is the data with other data?**

**How good was the process used to create this data?**

**How current is the data?**

**How much is it changing?**

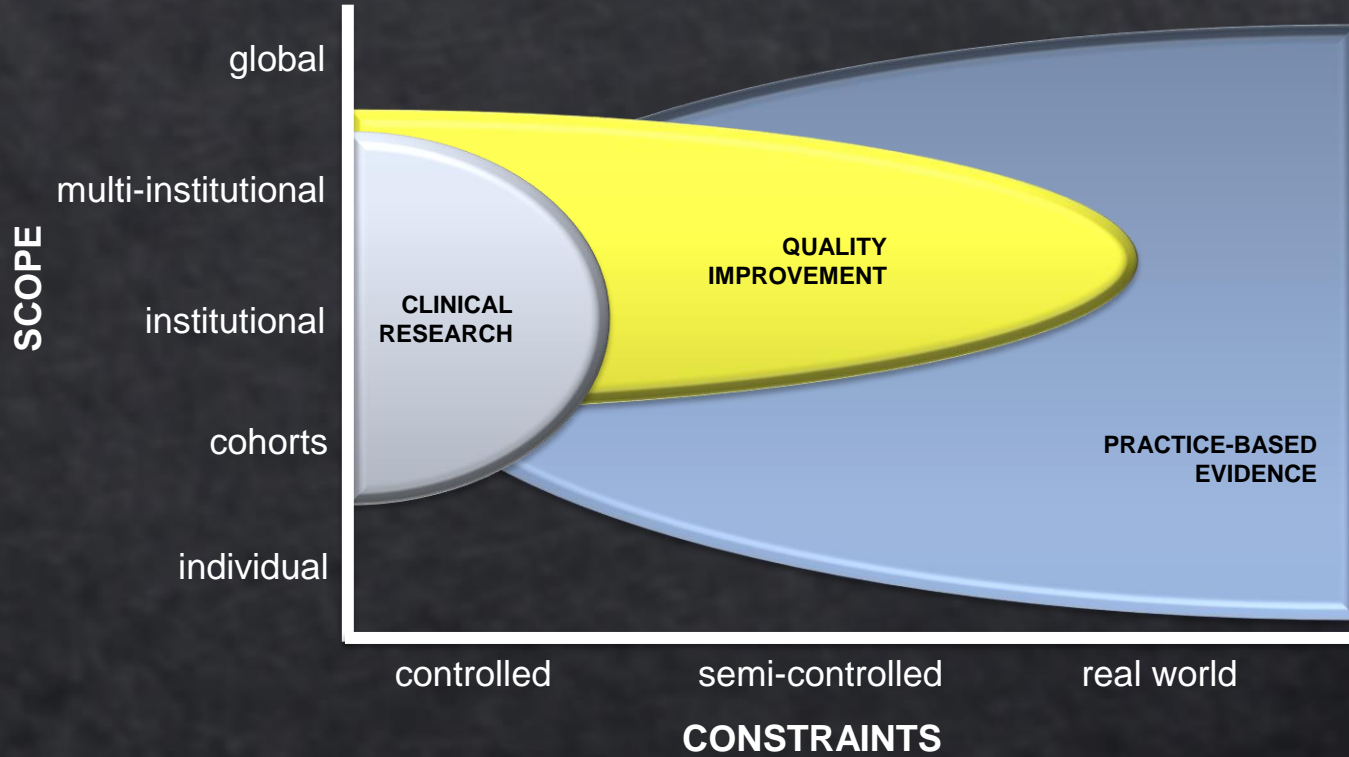
# Implications of Today's Big Data

*We have to know and manage a lot ABOUT our data,  
not just the data itself*

What is the source of the data?  
What do we know about this data and what it means?  
What's in the data; how representative is it?  
How was the data obtained and managed?  
Who else is using this data?  
How consistent is the data with other data?  
How good was the process used to create this data?  
How current is the data?  
How much is it changing?

Lineage / Pedigree  
Business Context  
Quality  
Reasonableness  
Consistency  
Pervasiveness  
Socialization  
Controls  
Currency  
Volatility

# Why is this so different?



# Dismantling the Hype

---

## The Buzz

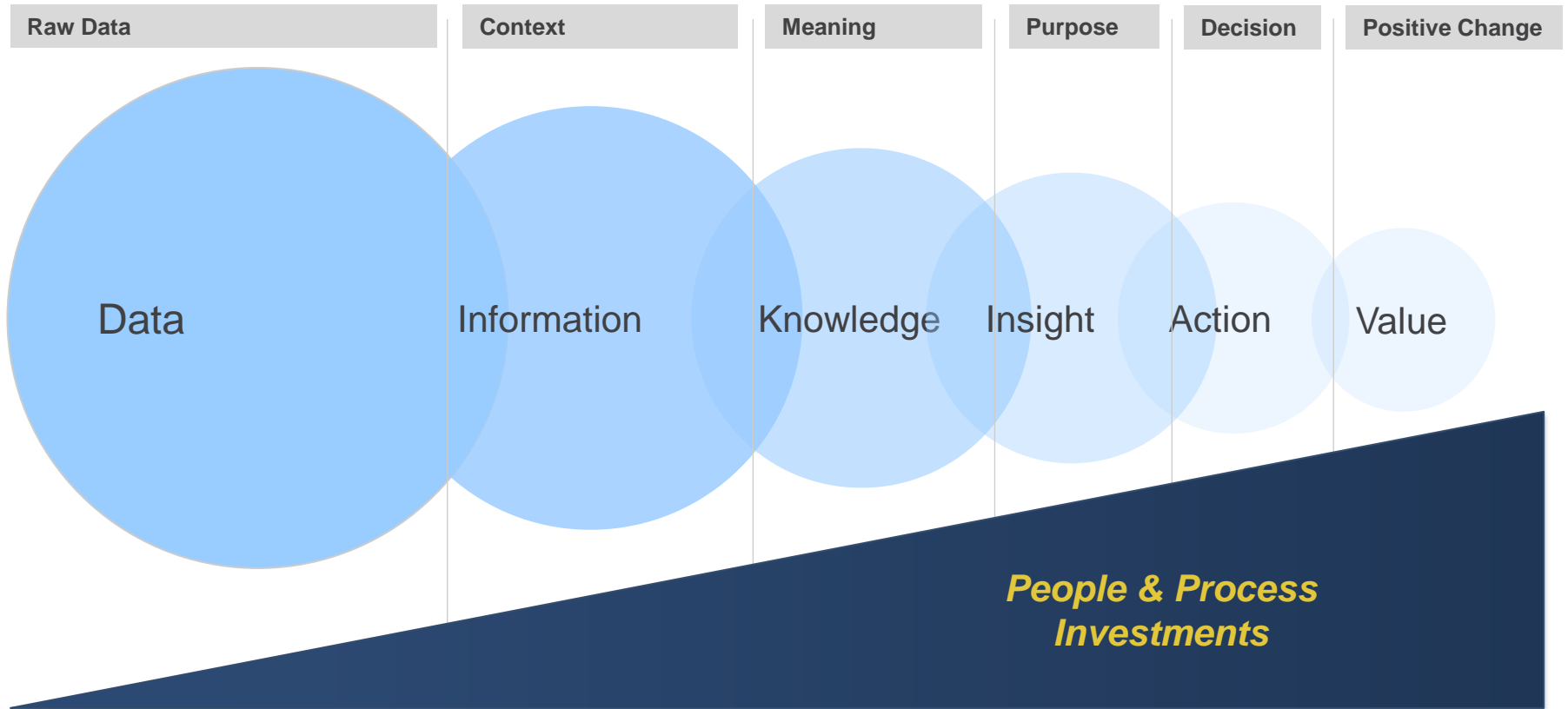
- A technology problem
- You can fix it with text mining, machine learning, AI, Hadoop, or another buzzword
- Standards is the fix
- Sticking to one vendor is a fix
- Outsourcing helps

## The Reality

- A people / process problem
- Garbage in, garbage out...but we can improve it even if we can't fix it
- Standards help a lot
- Sticking to one vendor is impossible, but does help
- No vendor knows your business better than you



# Harnessing Big Data is an Organizational Competency





*Designing for Reusability*

# Reusability in Analytics: Products!

## PROJECTS

- Stakeholders are project area experts
- Effort is focused on predefined questions
- Work is relevant to project team
- Timeline is project driven
- Data definitions are project specific
- Data structured for single use
- Little-to-no analytical code reuse
- Release available to project stakeholders



## PRODUCTS

- Stakeholders are functional experts
- Questions are not predefined
- Work must be relevant to multiple customers
- Timeline is engineering driven
- Data definitions are enterprise-wide
- Data is structured for broad re-use
- Analytical models are built for multiple projects
- Release available to entire enterprise



# Investing in Competencies

<b>DATA GOVERNANCE</b>	Data Policies and Standards	Data Roles & Stewardship	Governance & Decision Making	Master Data Management (MDM)	Asset Provisioning, Management & Certification	Data Quality Management
<b>BUSINESS OPERATIONS</b>	Data Strategy Formalism	Consulting and Guidance	Analytics Competency Development & Staffing	Data Operations Management and Controls	Knowledge Management	Community Engagement
<b>ENGINEERING OPERATIONS</b>	Engineering Management	Architecture Design and Management	Data Model Engineering & Management (Domains)	Analytical Model Engineering & Management	Lifecycle and Quality Management	Metadata Management

# Centralized vs. Federated Capability Development

Like many academic medical systems, UNC HCS has a diversified, empowered culture

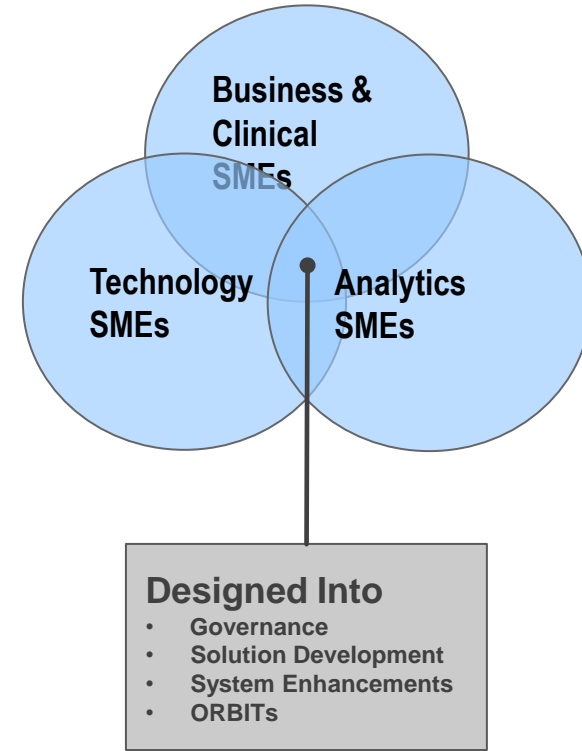
- Organizational units with deep subject matter experts (SMEs)
- Spirit of research, innovation, and entrepreneurship

Any journey with analytics must respect those cultural norms

- Fully centralized → “ivory tower”, bottlenecks, loss of institutional context and SME
- Fully federated → no economies of scale, impossible to establish a single source of truths

We opted to pursue a hybrid model

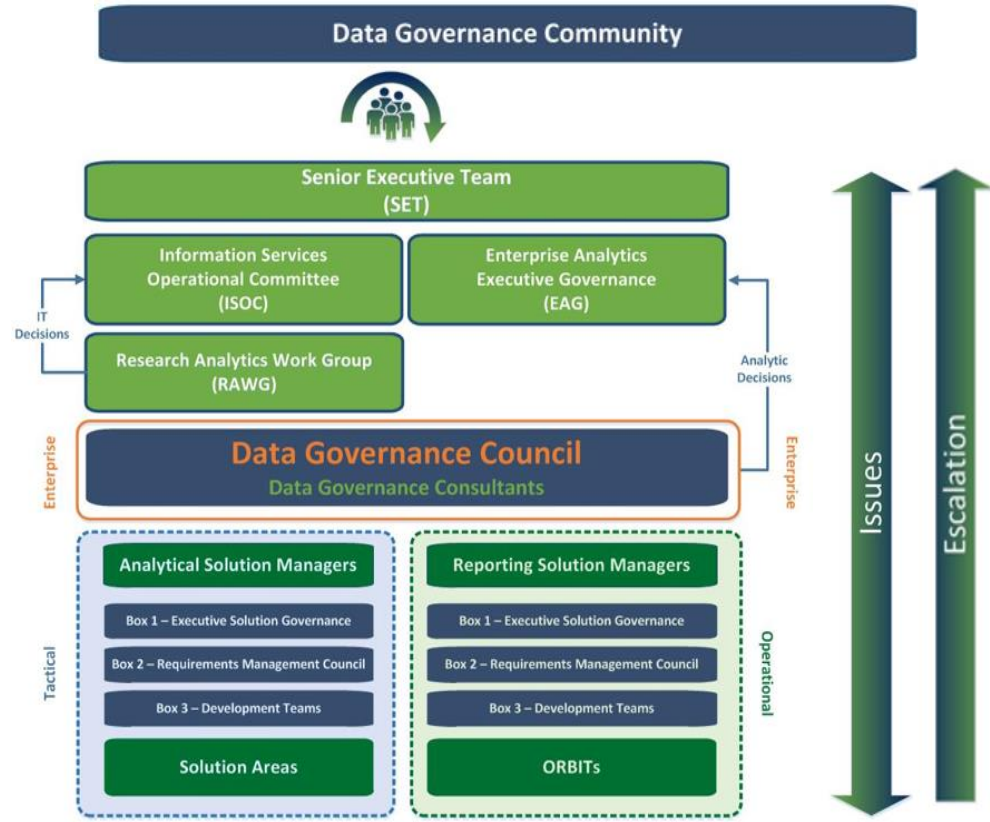
- Centralize building reusable assets
- Federate the use and extension of those assets
- Bring federated SMEs into all build-related work
- Help federated users be more effective with data / analytics





DATA GOVERNANCE	Data Policies and Standards	Data Roles & Stewardship	Governance & Decision Making	Master Data Management (MDM)	Asset Provisioning, Management & Certification	Data Quality Management
BUSINESS OPERATIONS	Data Strategy Formalism	Consulting and Guidance	Analytics Competency Development & Staffing	Data Operations Management and Controls	Knowledge Management	Community Engagement

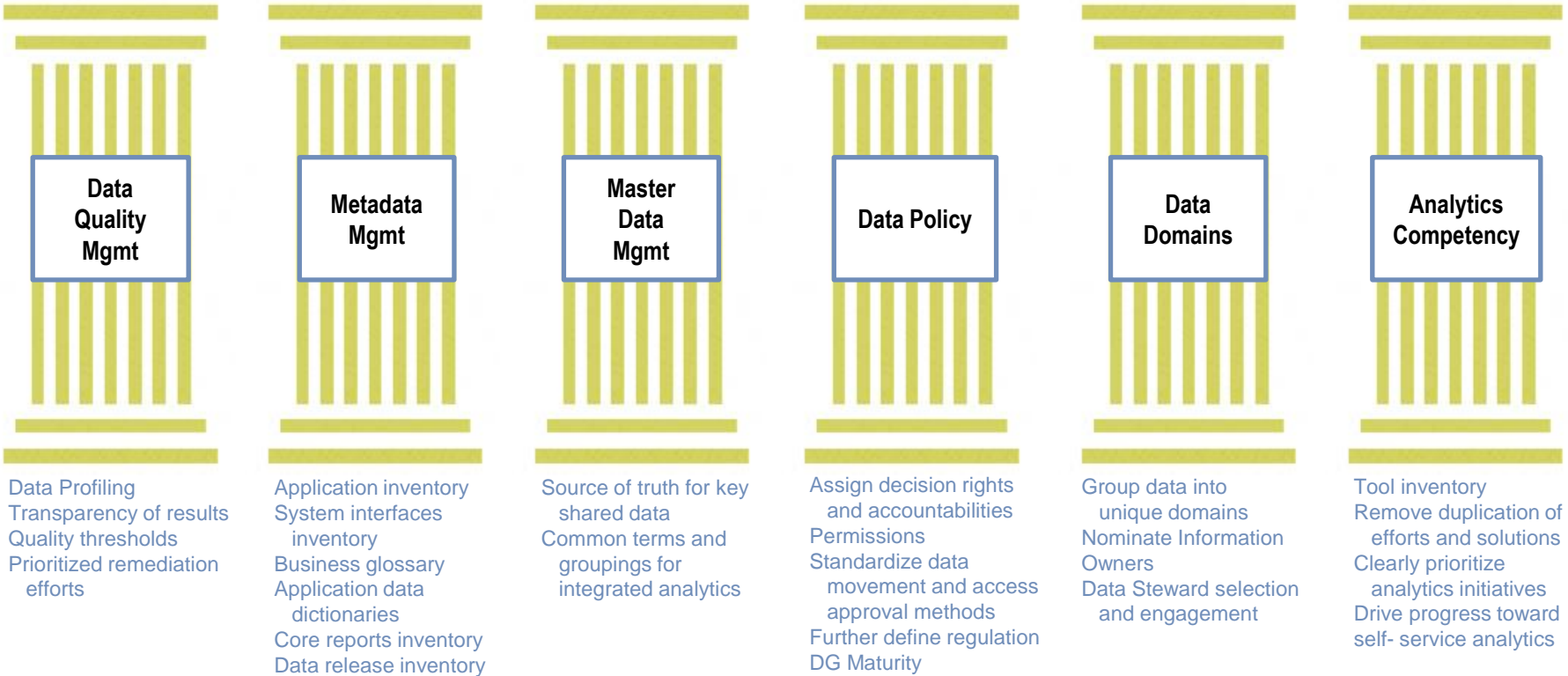
# Data Governance



DATA GOVERNANCE	Data Policies and Standards	Data Roles & Stewardship	Governance & Decision Making	Master Data Management (MDM)	Asset Provisioning, Management & Certification	Data Quality Management
BUSINESS OPERATIONS	Data Strategy Formalism	Consulting and Guidance	Analytics Competency Development & Staffing	Data Operations Management and Controls	Knowledge Management	Community Engagement

# Data Governance Program Pillars

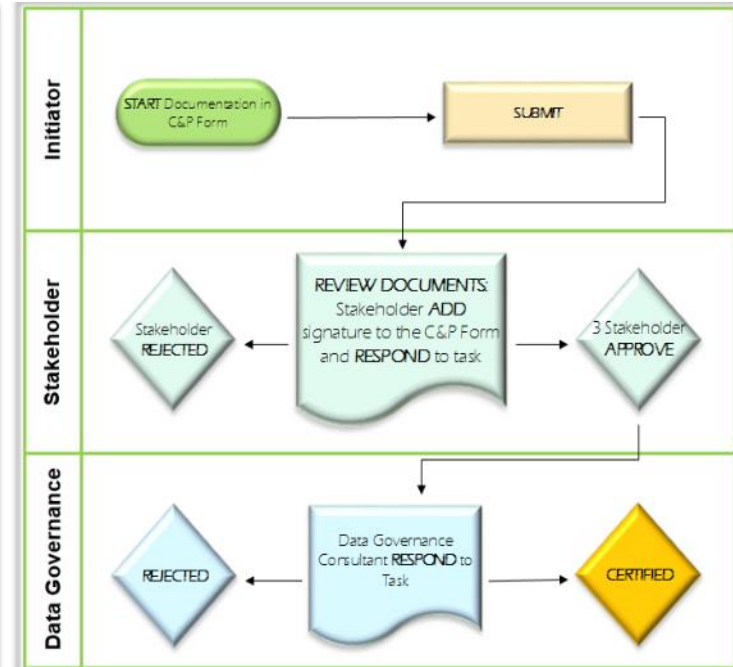
UNC HEALTH CARE SYSTEM



# Asset Certification

- Certification is a review process of key elements and tools used for decision making across the enterprise. Certification provides:
  - Reliability
  - Traceability
  - Standardization
  - Documentation
- Certification Levels
  - Gold
  - Silver
  - Bronze
- Certification Drivers
  - Initiator
  - Stakeholders
  - Data Governance

CERTIFICATION & PUBLICATION FORM	
Analytical Asset Name	A word or set of words by which the analytical asset being certified is known, referred to, or addressed.
Attachments	Click here to attach a file
Analytical Asset Type	Analytical assets are separate and distinct discrete units of information or decision-making tools. For example a data element and metric are distinct units of information whereas reports and dashboards are decision-making tools.
Analytical Asset Description	A detailed explanation highlighting the purpose and underlying concepts of the analytical asset so it is clear and comprehensible to the reader.
Analytical Asset Location	A detailed description or link to where the asset is located.
Analytical Asset Intended Usage	A detailed description of the planned insights drawn from the analytical asset.
Analytical Asset Calculation	A symbolic or written mathematical representation of the metric narrative.
Analytical Asset Inclusions	A detailed description of any data elements specifically included in the metric calculations. Used to inform others who may not be familiar with the data context.
Analytical Asset Exclusions	A detailed description of any data elements specifically left out of the metric calculation. Used to inform others who may not be familiar with the data context.





# Community Knowledge & Culture

Knowledge Management

Community Engagement



## Welcome to EADSpedia!

EADSpedia is intended to be a one-stop shop for curated information around data, reporting, & analytics at UNC Health Care. The success of EADSpedia relies on engagement from the entire analytics user community to provide content and share insights.   
 \*\*Please be patient with us and forgive our dust as we work diligently to build out this resource!

## Certified Analytical Assets

Review the analytical assets that have been submitted for Certification and their current status. If you have a request for an item to be certified, you can [submit it here](#).



Enterprise Analytics & Data Sciences | EADSpedia

### Analytical Asset Repository

+ new item

Status: **Certified (73)**

ID	Analytical Asset Name	Analytical Asset Type	Strategic Pillars	Solution Area	EADS Certification Level
57	Patient Mailing List Report (Professional Revenue Cycle)	Report	People	Finance & Revenue Cycle	Bronze
23	Hospital Revenue Cycle Denials Dashboard	Dashboard	Value	Finance & Revenue Cycle	Bronze
24	Denials Rate (Hospital Revenue Cycle)	Metric	Value	Finance & Revenue Cycle	Bronze
27	Preventable Denial Write Off Rate (Hospital Revenue Cycle)	Metric	Value	Finance & Revenue Cycle	Bronze
45	Hospital Revenue Cycle Preventable Denial Write Offs Dashboard	Dashboard	Value	Finance & Revenue Cycle	Bronze
105	Late Charges % (Hospital Revenue Cycle)	Metric	Value	Finance & Revenue Cycle	Bronze
48	228 (NQF 0028): Tobacco Use, Screening and Cessation Intervention	Metric	Quality & Service	Quality	Bronze
49	008 (NQF 0083): Heart Failure (H): Beta Blocker Therapy for LVSD	Metric	Quality & Service	Quality	Bronze
50	023 (NQF 0239): Perioperative Care: VTE Prophylaxis (When Indicated in ALL Patients)	Metric	Quality & Service	Quality	Bronze
51	032 Stroke and Stroke Rehabilitation: Discharged and Antithrombotic Therapy	Metric	Quality & Service	Quality	Bronze
52	143 (NQF 0384): Oncology: Medical and Radiation - Pain Intensity Quantified	Metric	Quality & Service	Quality	Bronze
53	131 (NQF 0420): Pain Assessment and Follow-Up	Metric	Quality & Service	Quality	Bronze



## Analytics Community

Be a part of the community that helps drive content for EADSpedia through knowledge sharing, networking, and discussion.

If you're an analyst or often wear the "report-building" hat for your area, come be a part of the community!

## Documentation

Search below or use this link to find information about a dashboard, report, WebI Universe, training, or something else.

Search...



Enterprise Analytics & Data Sciences | EADSpedia

### Hospital Revenue Cycle Denials Dashboard

ID: Analytical Asset Type: Business Point of Contact: Technical Point of Contact: EADS Solution Manager: Strategic Pillars: Solution Area: EADS Certification Level: EADS-Certification Effective Date: 23 Dashboard: Davis, Jennifer (Denial Mgt) | Dong, Liam | McDaniel, Rebecca | Value | Finance & Revenue Cycle | Bronze | 10/25/2017

#### Background & Purpose

This dashboard was built by members of the Hospital Revenue Cycle GRBIT to give end users a summarized view of how a particular area is performing with regards to their denials rate and give visibility into trends and payor information.

**Industry Standard Alignment:** The denials rate on this report relies on the HFMA guidelines around calculating a denials rate for hospital revenue cycle. For more information about this or other HFMA Map Keys: <https://www.hfma.org/Map/MapKeys/>

#### Dashboard Example:

**Denials Rate August 2017**

Denial Rate	1.7%
Aug 17 Denial Claims	1,753
Total Claims	42,205

**Past 12 Months Denial Rate**

Denial Rate	1.6%
Denial Claims	52,278
Total Claims	454,618

**REX HOSPITAL August 2017 Denial Rate**

**Denial Claims & Total Claims**

Notes: The Denial Rate is calculated based on the number of Denial Claims divided by the Total Claims. Denial Rate includes Denial Claims and Rejected Claims only. (Denial Rate = Denial Claims / Total Claims) (Denial Rate = 1,753 / 42,205 = 4.15%)

## A Word on Ethics

---

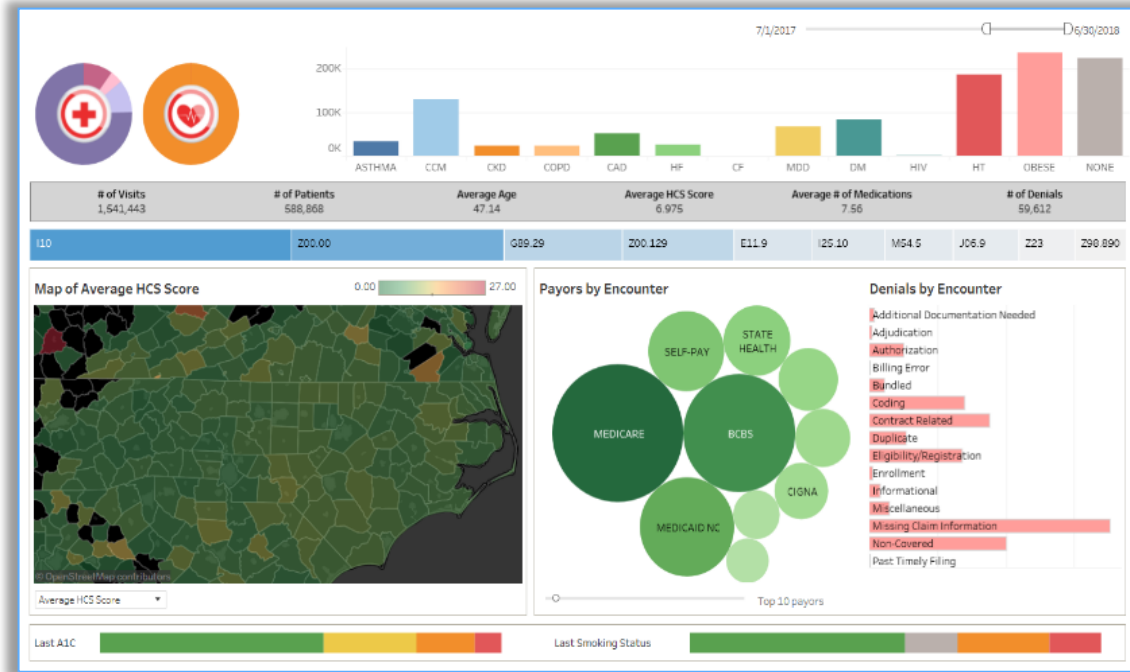
**Data availability velocity > policy development velocity**

**Obtaining multidisciplinary perspectives on a big data opportunity up front is critical to ensuring the right questions are being asked**

**Questions to ask as soon as possible:**

- Who are the stakeholders for this work?
- Who has granted consent for this work, when, and why?
- What are potentially negative outcomes of this work, and who would be impacted (e.g., sponsors vs. owners vs. stakeholders)?
- If we told patients and/or physicians we were doing this, what would they think?
- Where is the line between “improvement” and “research”?
- What controls can be used to mitigate potentially negative impacts of this?

# Example: Utilization Modeling

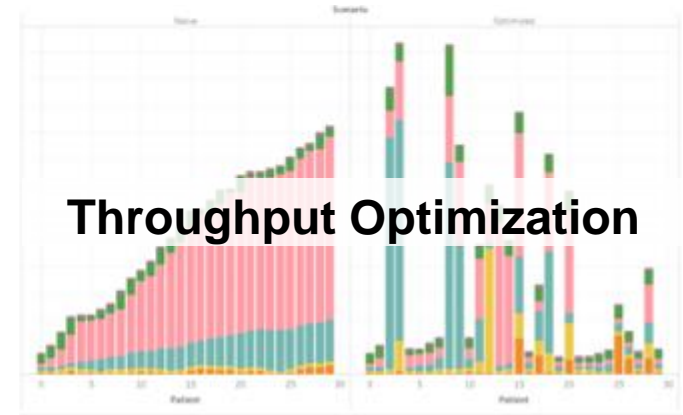
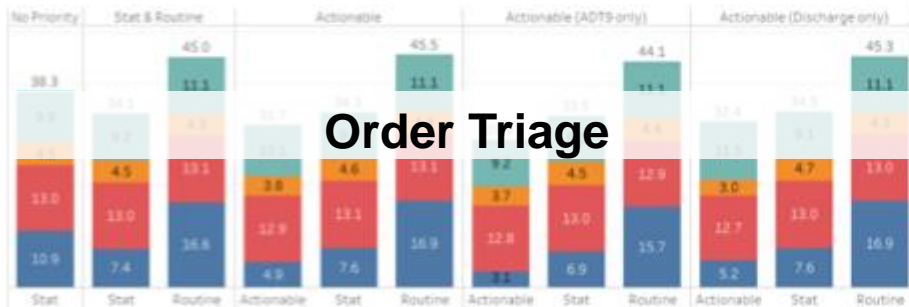
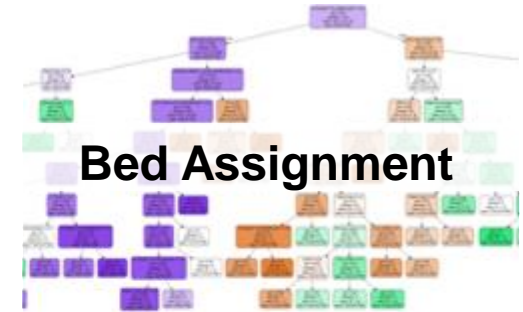


**Common views of health care system utilization through shared semantic representation**

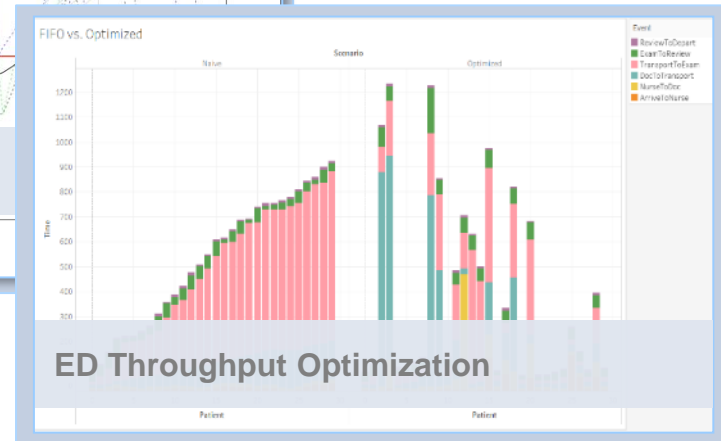
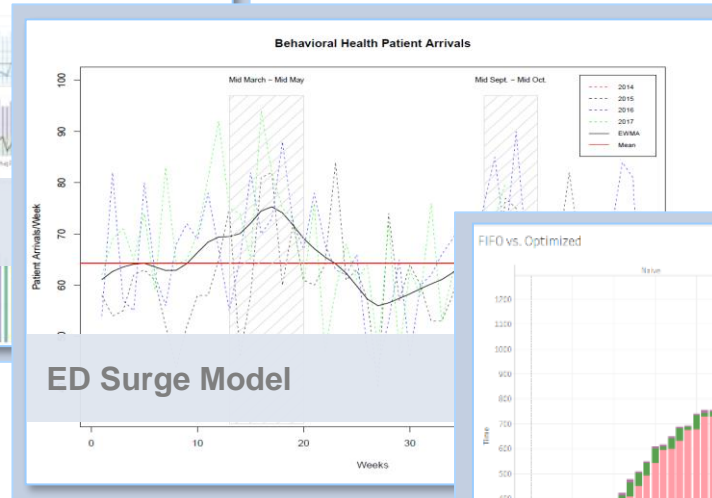
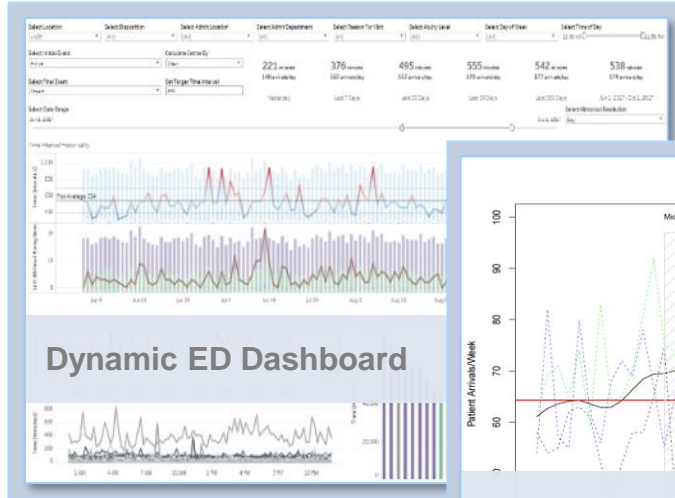
**Requires data domains through a more sophisticated data strategy linked to data governance**

# Example: Patient Throughput Optimization

Using discrete event simulation to drive operational efficiencies



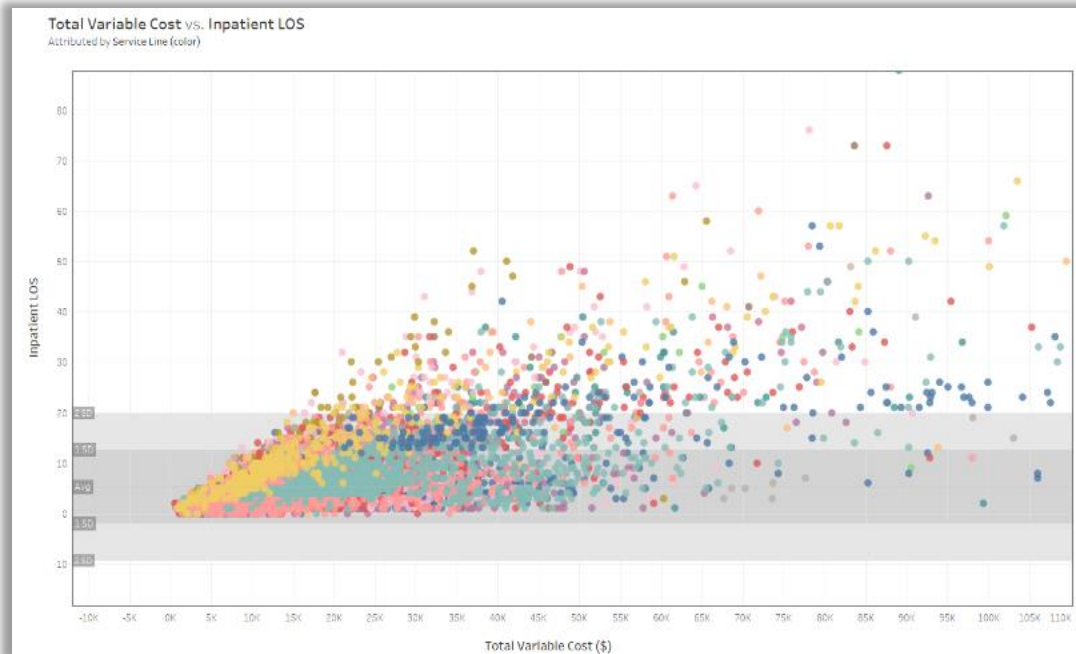
# Example: Emergency Operations



Models can be interpreted consistently despite different methods and focus

## Example: Care Variation

### *Mapping clinical context to big data*



- Individual diseases have disease-specific models
- Performance is defined against system-level standards
- Analytics are used to normalize comparisons



*Do we care about “big data” or “big insights”?*



**DATA Issues**

Storage  
Structure  
Timeliness  
Semantics & Language  
Validity  
Reliability  
Triage  
Pedigree

**INSIGHT Issues**

Innovation  
Health Outcomes  
Profitability  
Productivity  
Translational Science  
Customer Intimacy  
Risk  
Value



## Summary

---

- 1. Deriving value from big data is about a lot more than traditional “data management”**
  - What we know about our data →  
What we know from our data
- 2. Data governance is one of the key domains required to effectively operationalize big data**
- 3. Routine value creation from big data is dependent on growing and transitioning enterprise capabilities**
  - Reusable designs and assets
  - New process development
  - Business and clinical engagement
- 4. Governance programs need to be ever mindful of ethics considerations**
  - New data use is sometimes unplanned data use

# Questions

---

**Jason Burke**  
System VP and Chief Analytics Officer

**Email: [Jason.Burke@unchealth.unc.edu](mailto:Jason.Burke@unchealth.unc.edu)**

**Twitter: [@jaburke](https://twitter.com/jaburke)**

**LinkedIn: [jasonburke](https://www.linkedin.com/in/jasonburke)**