

Data Governance Is Quality Control

John Brower, MSCS
Manager Epic Data Analytics and Interfaces
&
John Vo, MSN, RN, LSSGB
Manager Clinical Data Management



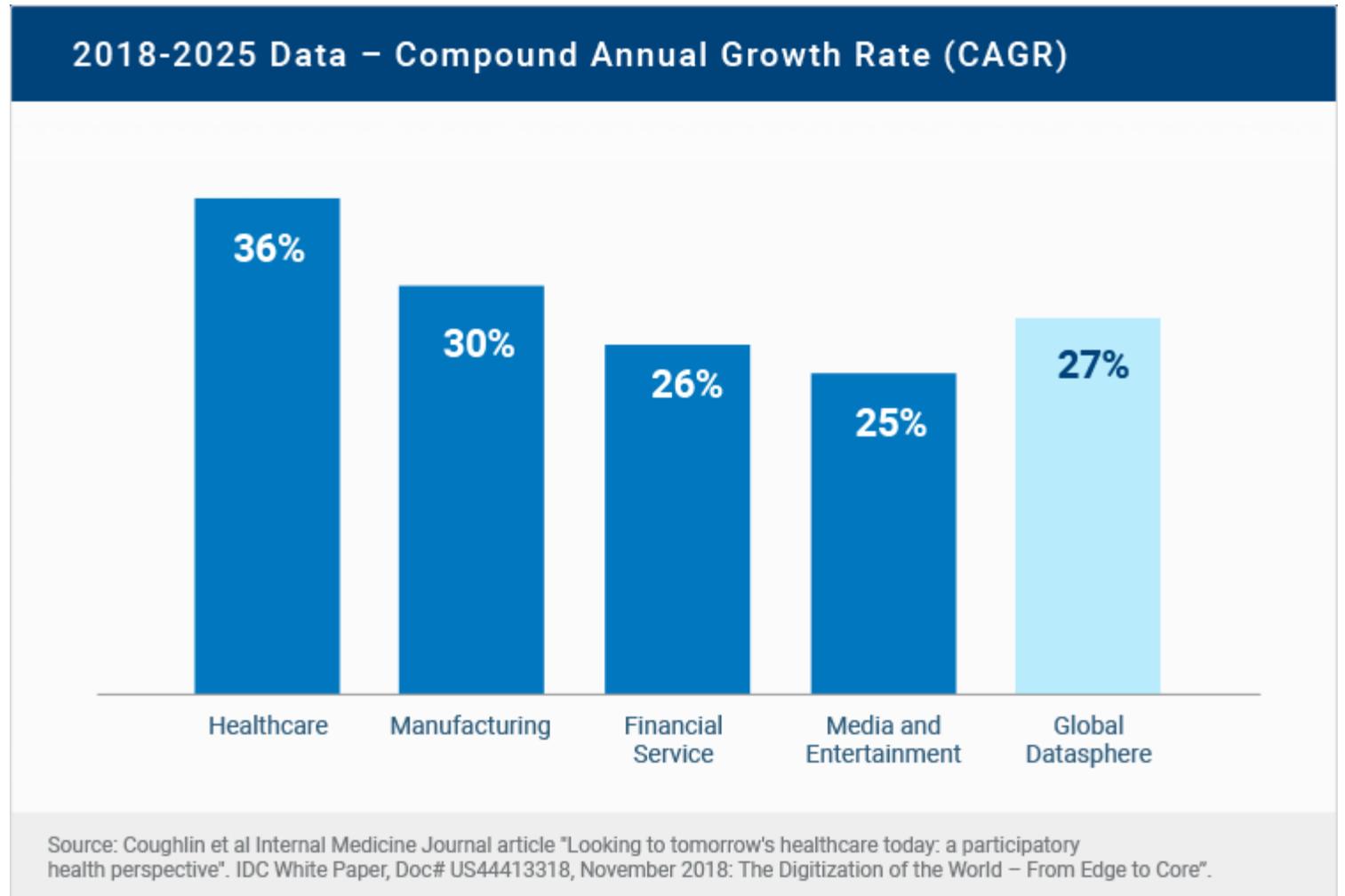
Trust. Methodist.

Objectives

- To educate and inform audience about best practices of Data Governance
- The audience will be able to learn data distribution and data analytics within a health system

The term “digital universe” quantitatively defines such massive amounts of data created, replicated, and consumed in a single year. International Data Corporation (IDC) estimated the approximate size of the digital universe in 2005 to be 130 exabytes (EB). The digital universe in 2017 expanded to about 16,000 EB or 16 zettabytes (ZB). IDC predicted that the digital universe would expand to 40,000 EB by the year 2020. To imagine this size, we would have to assign about 5200 gigabytes (GB) of data to all individuals (Dash et al., 2019).

Approximately 30% of the world's data volume is being generated by the healthcare industry (Callaway, 2021)



Importance of Data Governance

A 2021 survey of 100 healthcare executives showed 80% of the respondents identified high quality data is a top organizational priority. Only 8% of the respondents described their organizations data as “high quality” and 20% reporting full trust in the organizations data

(Glaikovskaya, 2021).



What is Data Quality?

- Data quality is “the planning, implementation, and control of activities that apply quality management techniques to data, in order to assure it is fit for consumption and meet the needs of data consumers.”
- This means that the quality of data is defined by the company’s context of quality. If the data process is fluid and consistent with the correct inputs, then the outputs of your data analysis system will yield the information you desire.



Data Completeness

Data completeness describes the comprehensiveness or wholeness of the information.

There should be no gaps or missing information for data to be truly complete. Sometimes incomplete data is unusable, but often it's still used even with missing information, which can lead to costly mistakes and false conclusions. (Tyagi, 2021)

Data Uniqueness

A discrete measure of duplication of identified data items within a data set or in comparison with its counterpart in another data set that complies with the same information specifications or business rules.



Data Timeliness

The degree to which the data is up-to-date and available within acceptable time frame, timeline and duration.

- Operational reporting focuses on the here and now
- Analytical reporting focuses on data over time

**OPERATIONAL DATA IS
IMPORTANT TO IT**



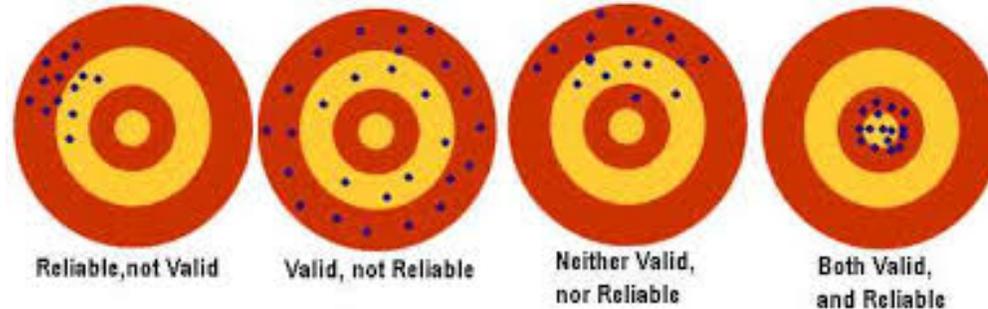
Data Validity

A measure of conformity to the defined business requirements and syntax of its definition. The scope of syntax may include the allowable type, range, format and other attributes of preference.



Data Accuracy

The degree to which the data item correctly describes the object in context of appropriate real-world context and attributes.

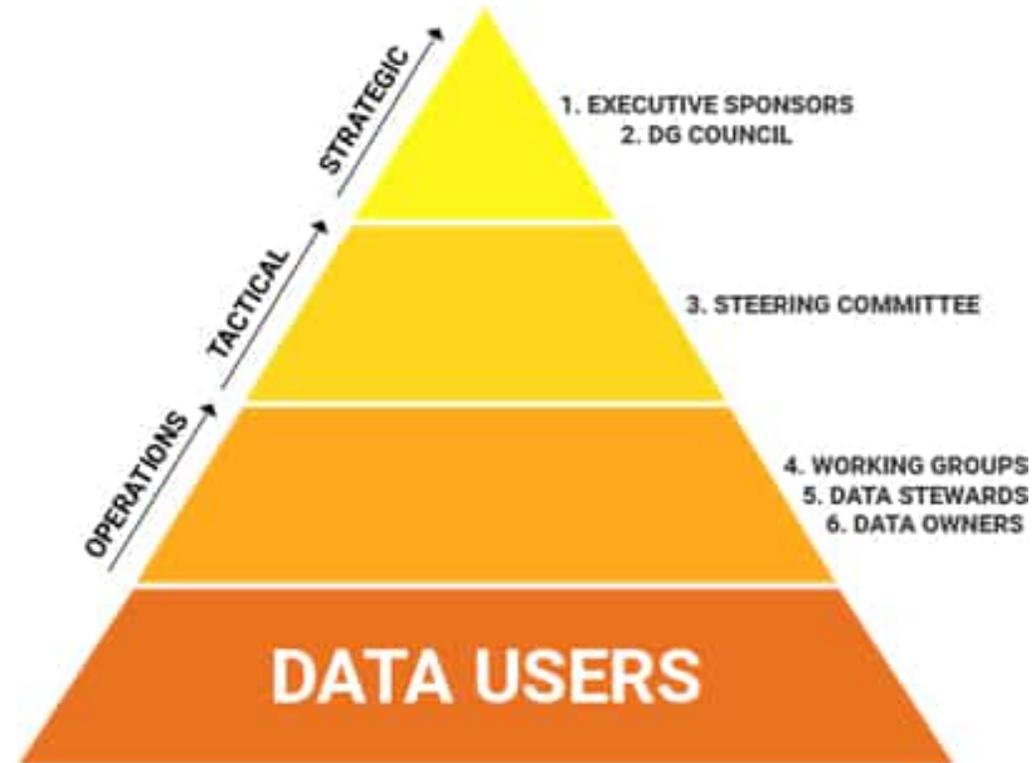


Data Consistency

Represents the absence of differences between the data items representing the same objects based on specific information requirements (Raza, 2018).



Data Governance Team



Data Governance



Data Quality

Data Creator

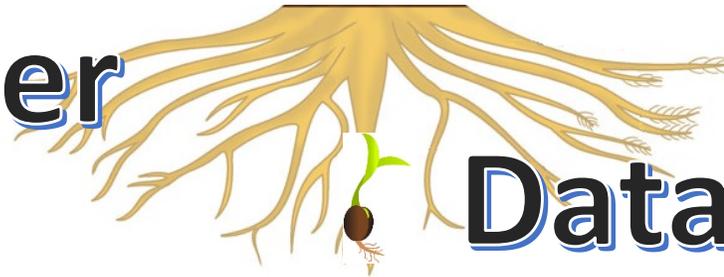


Data Governance



Data Quality

Data Owner



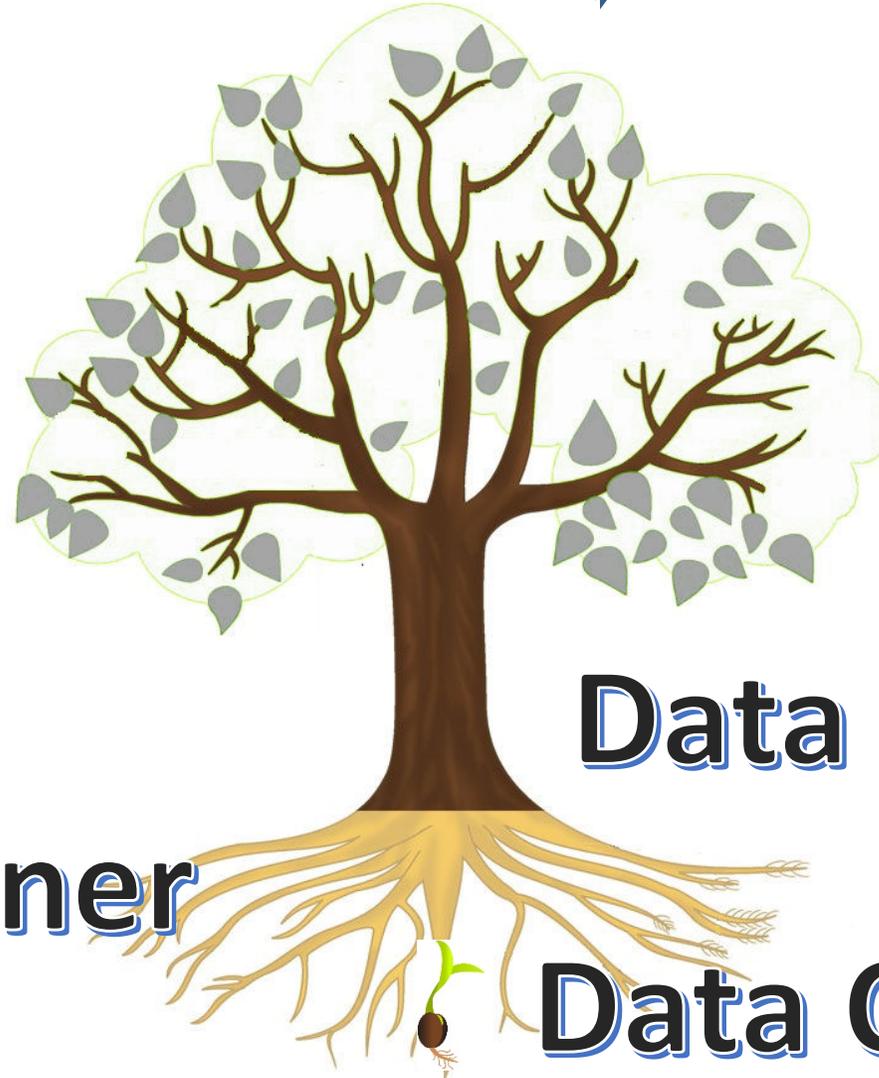
Data Creator



Data Governance



Data Quality



Data Steward

Data Owner

Data Creator

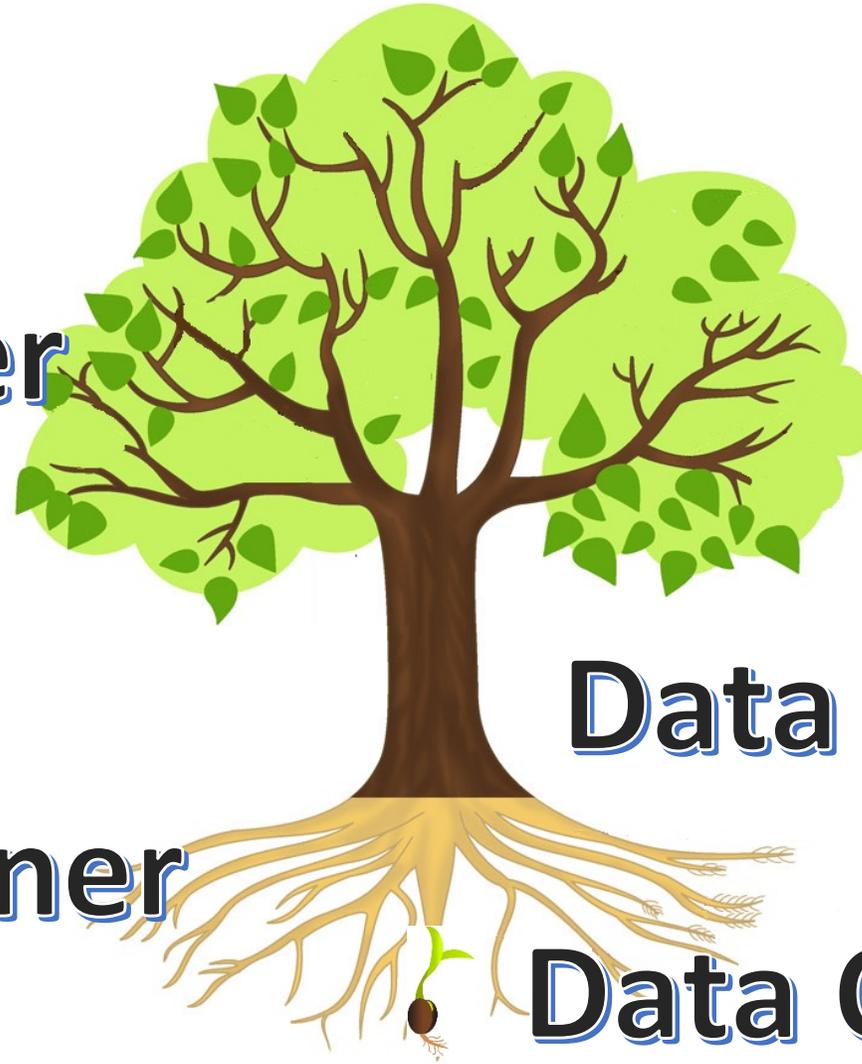


Data Governance



Data Quality

Data Developer



Data Steward

Data Owner

Data Creator



Data Governance



Data Quality

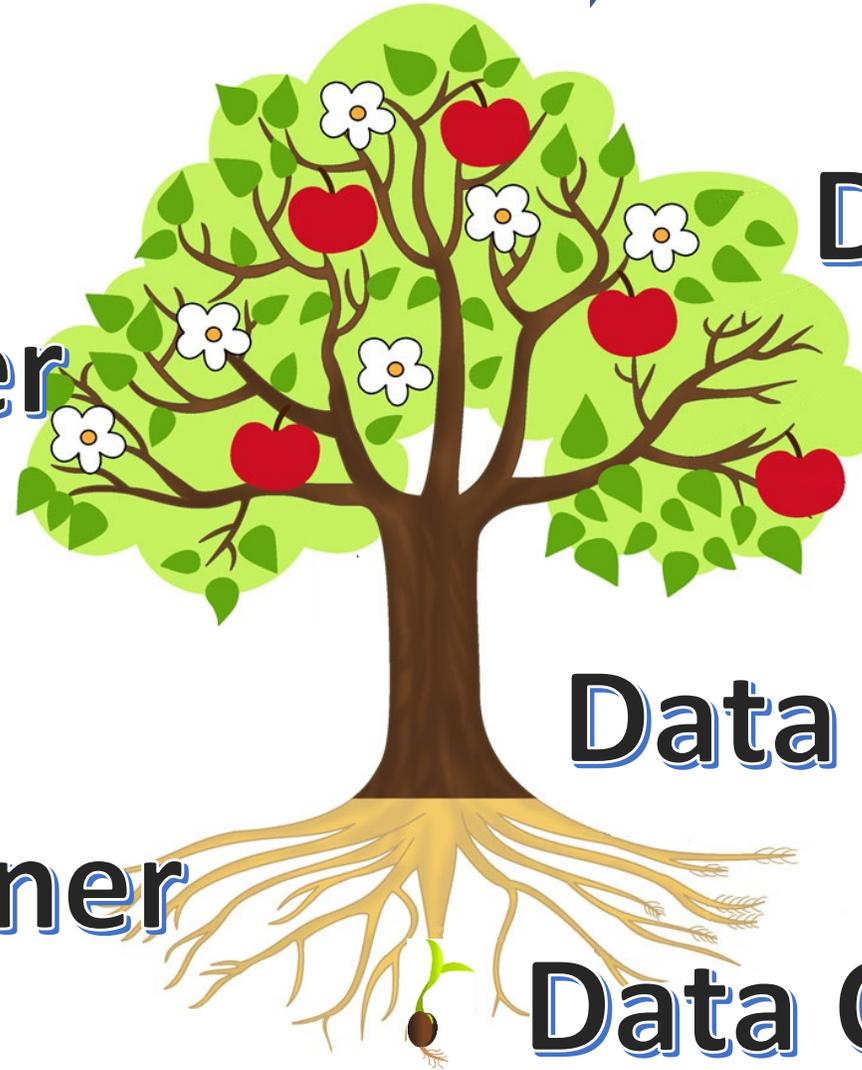
Data Developer

Data User

Data Steward

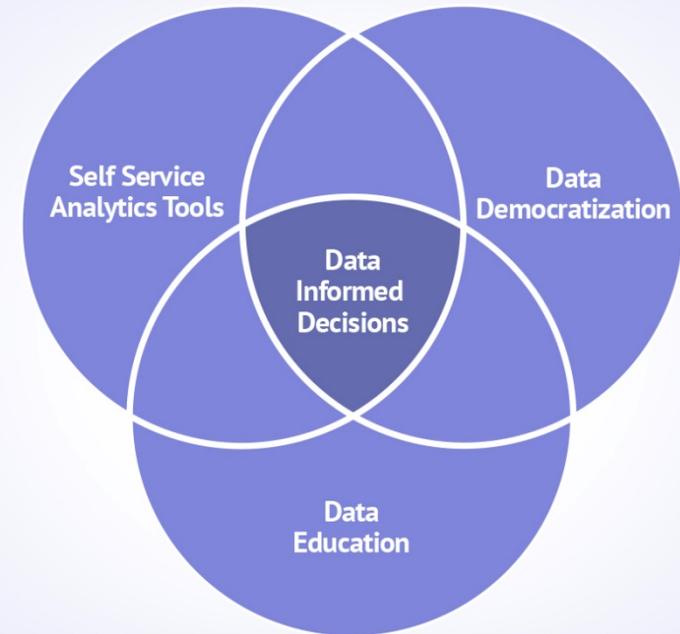
Data Owner

Data Creator



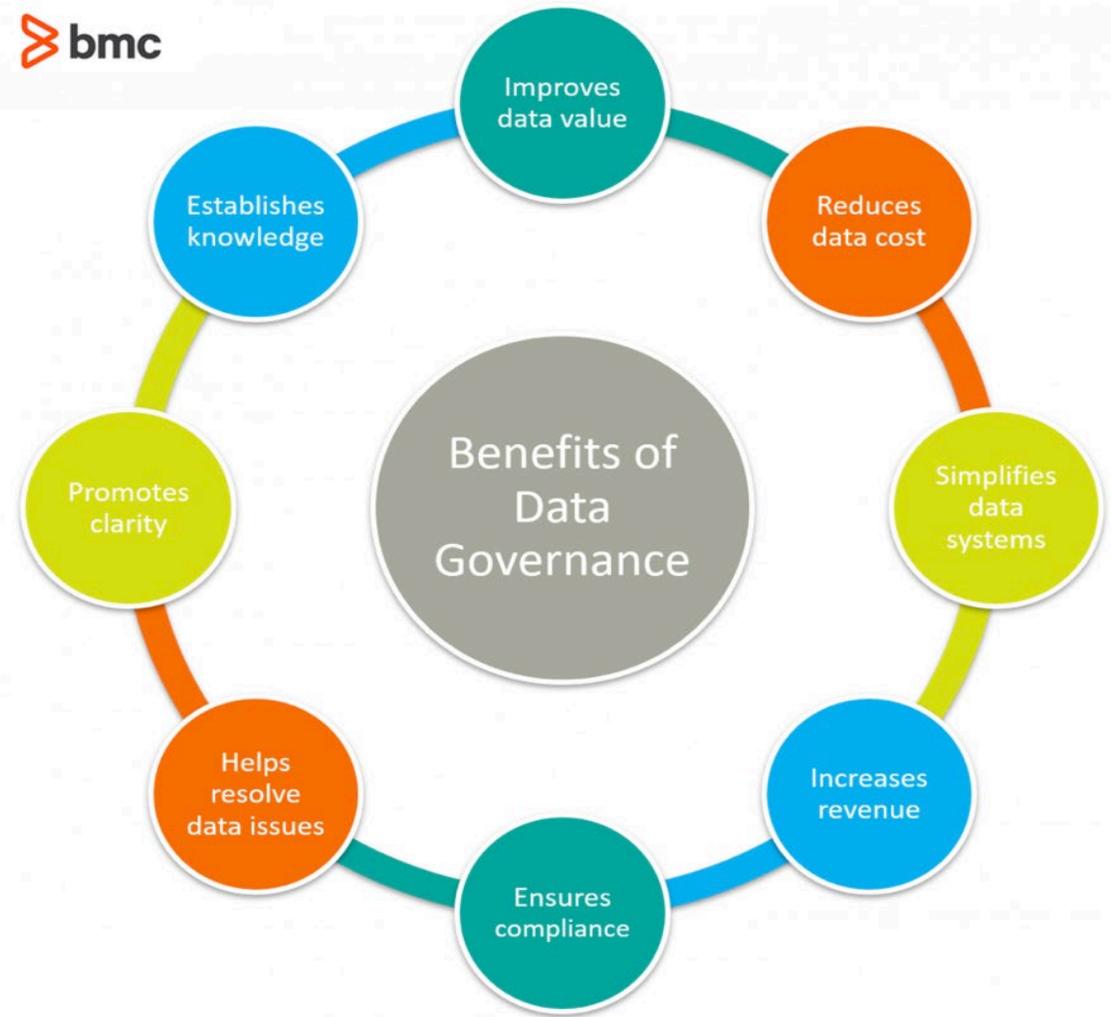
Data democratization

- Nearly all employees are expected to use data in their roles
- A legacy of a traditional approach to data (data in the hands of a few) persists in many organizations
- As a result, many employees don't have skills that would help them comfortably and confidently work with data
 - 21% of the global workforce is fully confident in their data literacy skills

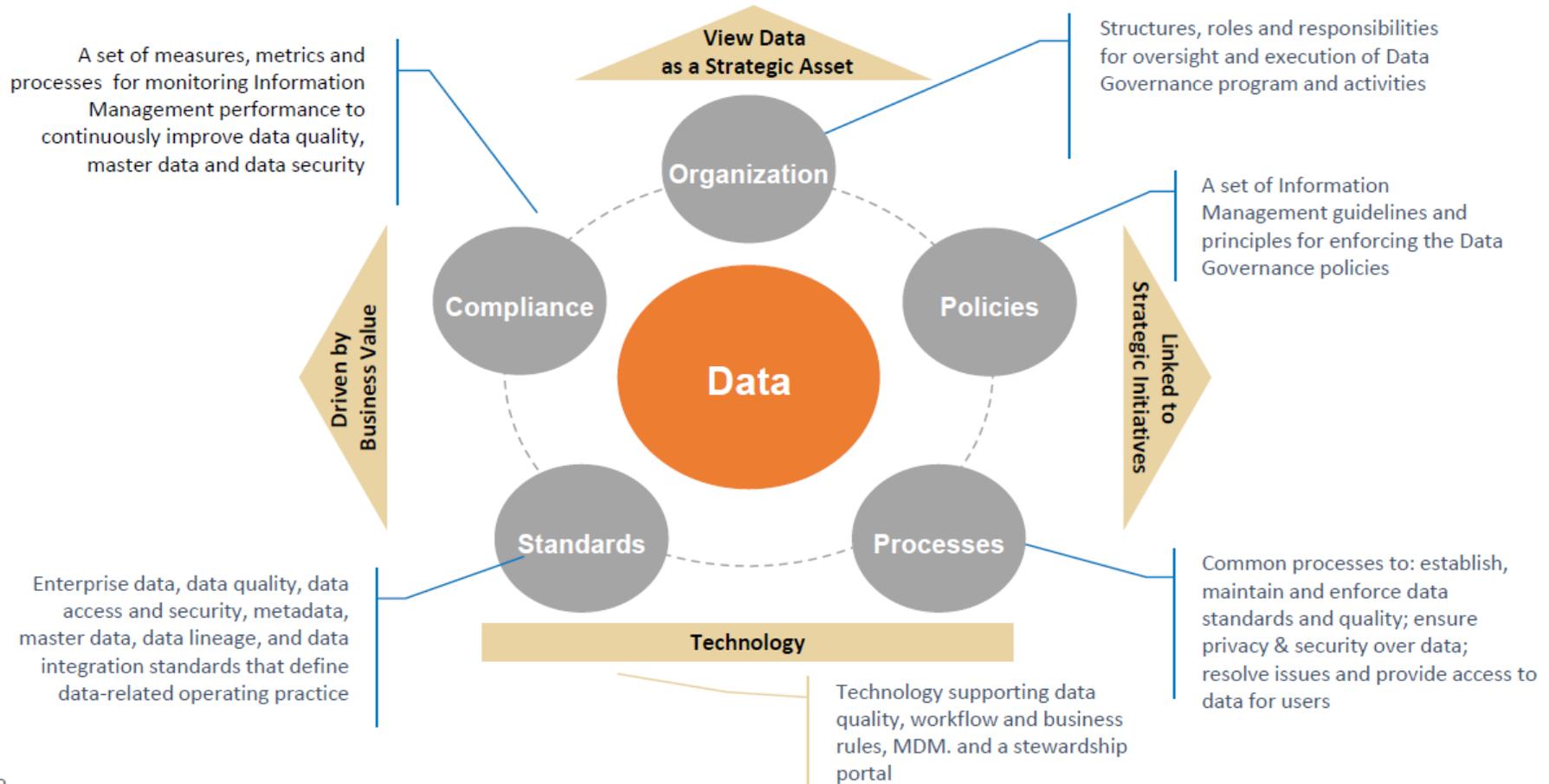


The Human Impact of Data Literacy

The Importance of Data Governance



Why Data Governance is Important



18

Data Governance at Methodist Health System

- Committee membership
 - Finance
 - IT
 - Clinical Informatics
 - Quality
 - Medical Staff
 - Accountable Care Organization
 - Revenue Cycle
 - Nursing

Data Definitions



Quality Scorecard - Operational Definition

Quality Scorecard - May 2021

Key Indicator	Unit of Measure	Calculation	Frequency	Time Period	Description	Data Owner	Source of Data
Overall Mortality	Ratio	Observed/Expected	Monthly (1 month data lag)	Current Value: Mar 2021 SPC Chart: Oct 2020-Mar 2021	Risk-adjusted ratio of observed mortality rate to expected mortality rate System aggregation: Equally-weighted aggregate of 4 hospitals (each hospital contributes 25% to the final score) (Feb '19 to Feb '20): Top Decile CareScience - 0.68, Top Quartile Mortality - 0.70	Hospital Administration and Quality	Premier
Sepsis Mortality	Ratio	Observed/Expected	Monthly (1 month data lag)	Current Value: Mar 2021 SPC Chart: Oct 2020-Mar 2021	Risk-adjusted ratio of observed mortality rate to expected mortality rate Sepsis cases are defined by any diagnosis of R65.21-SEVERE SEPSIS WITH SEPTIC SHOCK, R65.20-SEVERE SEPSIS WITHOUT SEPTIC SHOCK System aggregation: Equally-weighted aggregate of 4 hospitals (each hospital contributes 25% to the final score) (Feb '19 to Feb '20): Top Decile CareScience - 0.79, Top Quartile Mortality - 0.89	Hospital Administration and Quality	Premier
Sepsis 1 Compliance Metric Score	Percentage	Ranking	Monthly (1 month data lag)	Current Value: Mar 2021 SPC Chart: Oct 2020-Mar 2021	Overall Compliance with Sepsis Bundle - Benchmarking comes from Qnet's Benchmarks of Care System aggregation: Equally-weighted aggregate of 4 hospitals (each hospital contributes 25% to the final score)	Hospital Administration and Quality	Premier / QNet
Patient Falls w/ Injury per 1000 Pt Days (↓ is better)	Injury falls/ 1000 patient days	per 1000 patient days	Monthly (1 month data lag)	Current Value: Mar 2021 SPC Chart: N/A	Based on NDNQI unadjusted score for moderate or greater injury falls per 1000 patient days System aggregation: Equally-weighted aggregate of 4 hospitals (each hospital contributes 25% to the final score) YTD aggregation: Average of quarterly scores Benchmark : NDNQI best quartile over 2 years of data for all facilities	Risk Management	MHS Risk Managers
Hospital Acquired Pressure Injuries: Stage 2 and above	Pressure Injury	per 1000 patient days	Quarterly (3 month data lag)	Current Value: Q3-CY20 SPC Chart: N/A	NDNQI unadjusted score for pressure injury per 1000 patient days System aggregation: Equally-weighted aggregate of 4 hospitals (each hospital contributes 25% to the final score) YTD aggregation: Average of quarterly scores Benchmark : NDNQI best quartile over 2 years of data for all facilities	Nursing	NDNQI
HO-CDIFF	SIR	Observed/Predicted	Quarterly (3 month data lag)	Current Value: Mar 2021 SPC Chart: N/A	Hospital-Onset Clostridioides difficile infection: Those patients with laboratory testing positive for Clostridioides difficile on inpatient, calendar day four or later.	Infection Prevention	NHSN
CAUTI	SIR	Observed/Predicted	Monthly FYTD (1 month data lag)	Current Value: Mar 2021 SPC Chart: N/A	Catheter-Associated Urinary Tract Infection: A urinary tract infection where an indwelling urinary catheter was in place for more than 2 calendar days on the date of event, with day of device placement being day 1.	Infection Prevention	NHSN
Overall Geometric Length Of Stay	Days	Days/cases	Monthly (1 month data lag)	Current Value: Mar 2021 SPC Chart: Oct 2020-Mar 2021	Geometric length of stay for all inpatient cases System aggregation: System length of stay from Epic MHS Geometric Mean LOS System To get trending data for GMLoS data on accounts. This metric uses the mathematical equation of multiplying all the LOS then taking the root of the number accounts. MHS Geometric Mean LOS Hospital To get trending data for GMLoS data on accounts. This metric uses the mathematical equation of multiplying all the LOS then taking the root of the number accounts. Hospital targets for this year are reduction from FY20 baseline of: MDMC 0.10 MCMC 0.15 MMMC 0.15 MRMC 0.15 from baseline GMLoS Observed Data Source: LOS Epic Executive dashboards Benchmarks: Premier benchmark report for facilities with like CMTs	Leslie Pierce	Epic
Inpatient Experience – HCAHPS Rate the Hospital	Percentile	Ranking	Monthly (1 month data lag)	Current Value: July 2021 SPC Chart: Oct 2020-July 2021 *from dashboard as of 9/6/21	Percentile Ranking for Top Box % on the Rate the Hospital 0-10 question. Benchmark: All PG Database System Aggregate: Rank for aggregate of all HCAHPS surveys for MCMC, MDMC, MMMC, MLMC, MRMC	O/E	Press Ganey
ED Patient Experience - Overall Rating of ER Care	Percentile	Ranking	Monthly (1 month data lag)	Current Value: July 2021 SPC Chart: Oct 2020-July 2021 *from dashboard as of 9/6/21	Percentile Ranking for Top Box % on the Overall Rating of ER Care question. Benchmark: MCMC/MDMC: 75k+ Annual Visits Group; MMMC/MRMC: 50k-75k Annual Visits Group; MLMC: 10k-40k Annual Visits Group System Aggregate: Average of volume-based ranks for MCMC, MDMC, MMMC, MLMC, MRMC	O/E	Press Ganey
Advanced Care Planning	Percentage	Internal goal	Monthly (1 month data lag)	Current Value: Mar 2021 SPC Chart: N/A	Advance Care Planning of MMG patients	MMG -Dr. Marshall	Epic

Quality Scorecard

Quality Scorecard



Building an Enterprise Data Governance Program

- Foundation: Data Dictionary
- Building Blocks: Cogito Metric Glossary
- Online Shopping: Analytics Catalog (Available Epic 2018)

Cogito  for Data Governance
ergo sum

Catalog
Reports » Data Citizens



Glossary
Metrics » Data Stewards

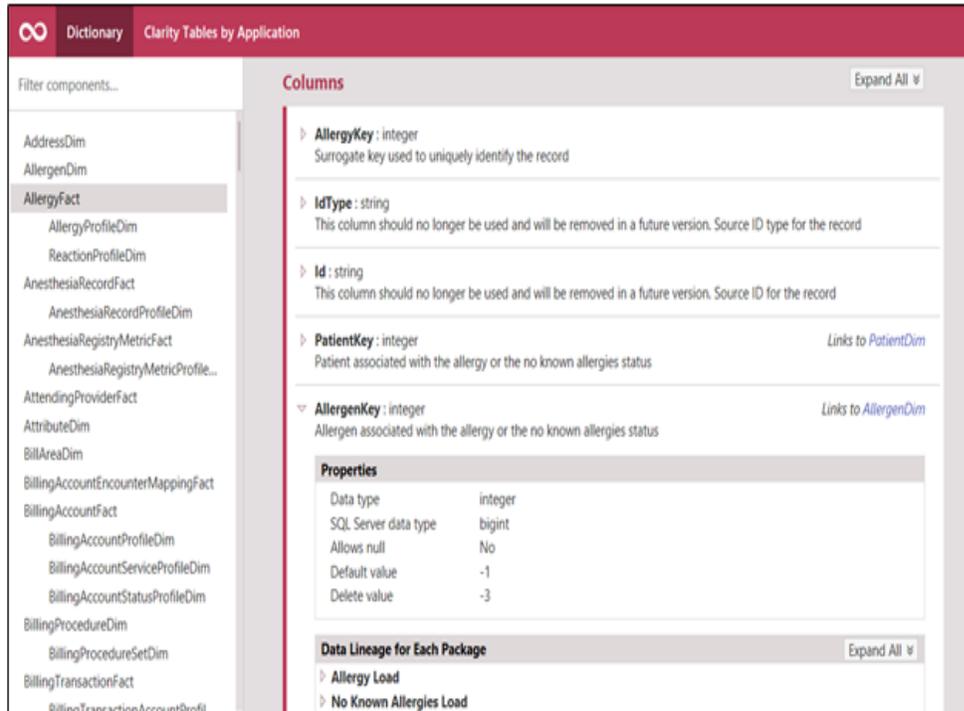


Dictionary
Items » Data Architects



Foundation: Data Dictionary

At the foundation of every data governance strategy is your enterprise data. Cogito uses the Caboodle Data Dictionary to provide a seamless way to create and update your Caboodle enterprise data, as well as record the associated metadata for things like data lineage, ontologies, and definitions. The Data Dictionary is intended for BI developers, ETL developers, and business analysts. It provides an easy one-stop-shop to both search and maintain your data definitions in a web-based console.



The screenshot displays the Caboodle Data Dictionary web console viewer. The interface is divided into a left sidebar and a main content area. The sidebar, titled "Dictionary", lists various components under the heading "Clarity Tables by Application". The "AllergyFact" component is selected and highlighted. The main content area, titled "Columns", shows the metadata for the "AllergyFact" table. It lists several columns with their data types and descriptions:

- AllergyKey**: integer. Surrogate key used to uniquely identify the record.
- IdType**: string. This column should no longer be used and will be removed in a future version. Source ID type for the record.
- Id**: string. This column should no longer be used and will be removed in a future version. Source ID for the record.
- PatientKey**: integer. Patient associated with the allergy or the no known allergies status. [Links to PatientDim](#)
- AllergenKey**: integer. Allergen associated with the allergy or the no known allergies status. [Links to AllergenDim](#)

Below the column list, there is a "Properties" section with the following details:

Property	Value
Data type	integer
SQL Server data type	bigint
Allows null	No
Default value	-1
Delete value	-3

At the bottom of the main content area, there is a "Data Lineage for Each Package" section with an "Expand All" button. It lists two packages: "Allergy Load" and "No Known Allergies Load".

Caboodle Data Dictionary web console viewer showing allergy metadata

Data Definitions



COVID-19 Hospital Trends
This dashboard shows COVID-19 specific metrics for hospital trends regarding patient volumes (ED, I...)

Infectious Disease +2 tags



Covid-19 Information

Dashboard 143673

Description

Dashboard for System wide Covid-19 data. Information about Covid-19 test, results, and Admissions.

Build Team
Cogito

Number of Favorites
64 favorites

Content Review

Overall Review Status
Approved

Last Reviewed On
5/22/2020

Reviewer

Review Status

HOHERTZ, MARIE D.

Approved

PERKINS-GROTH, ERIKA

Approved

VALERO, COURTNEY

Approved

Administrator Information

Metadata ID

HCM 1403221 - IDM 1405032 COVID-19 INFORMATION

Content Owner

PERKINS-GROTH, ERIKA

Notes

Dashboard is owned by CDM and Cogito. Please review with these teams for any changes.

Distribution

Report Groups

MHS Infection Control

MHS Hospital Operations - Past Week by Dynamic Location
Workbench Report 157408

Description
This report contains operational data for the MHS Hospital. Data displays for admissions, discharges, average length of stay, occupancy percentage, and observation hours, as well as separate sections for inpatient days by unit, financial class, and service for each day from the past week.

Add a tag
Cogito - Tested X

Build Team
Select a build team

Content Review

Overall Review Status: Set a review status
Last Reviewed On: Set the review date

Reviewer: [] Review Status: []

Administrator Information

Content Owner: JACKSON, CONNIE

Restore Save Cancel

MHS Hospital Operations - Past Week by Dynamic Location
Workbench Report 157408

Add a tag
Cogito - Tested X

Build Team
Select a build team

Content Review

Overall Review Status: Set a review status
Last Reviewed On: Set the review date

Reviewer: [] Review Status: []

Administrator Information

Content Owner: JACKSON, CONNIE

Notes
Other notes intended for administrators

Restore Save Cancel



METADATA FOR DATA GOVERNANCE

Building Blocks: Cogito Metric Glossary (Available August 2018)

Epic maintains over 700 healthcare-focused metrics from across the continuum of care in a business glossary, which includes the ability for your organization to add your own new metrics. The Cogito Metric Glossary allows you to document the business logic (calculations) and the metric definition for the most important KPIs and metrics in your organization.

Your metrics can appear in the glossary with a full human readable definition and owner.

Cogito Metric Glossary [3028869] as of Fri 1/11/2019 1:45 PM

Filters Options Metric Editor Metadata

Metric ID	Metric Name	Active?	Owning Application	Description	Metadata External ID	Overall Review Status	Owners	Last Review Date	Next Review Date	Reviewers
24007	ADT - Number of IP Discharges	Yes	GRAND CENTRAL	This is an ADT Metric designed to analyze ADT events. It calculates the number of inpatient discharges that have taken place. It identifies inpatient discharges by looking to the ADT event interpretations (ADT-80: Outgoing Event Interpretation and ADT-83: From Base Class) for the event, and not the ADT event type (ADT-30).						
24010	ADT - Number of ED Admissions	Yes	GRAND CENTRAL	This is an ADT Metric designed to analyze ADT events. It calculates the number of emergency department admissions that have taken place. It identifies emergency department admissions by looking to the ADT event interpretations (ADT-81: Incoming Event Interpretation and ADT-83: From Base Class) for the event, and not the ADT event type (ADT-30).	HCM 1352762 - IDN 90024010 ADT - NUMBER OF ED ADMISSIONS	Approved		01/11/19		ADT, ADMIN [ADTA
24012	ADT - Number of Obs Admissions	Yes	GRAND CENTRAL	This is an ADT Metric designed to analyze ADT events. It calculates the number of observation admissions that have taken place. It identifies observation admissions by looking to the ADT event interpretations (ADT-81: Incoming Event Interpretation and ADT-83: From Base Class) for the event, and not the ADT event type (ADT-30).						

Dashboards that contain ADT - Number of ED Admissions

Ready for Use	Contained in Components	Contained in Resources
IP Case Manager Reporting Homepage [3040000013]	ADT Census Metrics [31010000098]	ADT - Census Metrics - Number of ED Admissions [1001921]
Provider Executive Dashboard [3301100852]	Exec DB Daily Stats [3303350001]	ADT - Census Metrics - Number of ED Admissions [1001921]
CEO Dashboard (Email) [33010000009]	Exec DB Daily Stats (Email) [3303350002]	ADT - Census Metrics - Number of ED Admissions [1001921]

Glossary viewer showing metric metadata

Online Shopping: Analytics Catalog (Available Epic 2018)

You've carefully defined your enterprise data in Caboodle, and you have codified your KPIs and metrics into the Cogito Metric Glossary. Now, you can reach the third layer of Cogito data governance: the Analytics Catalog. The Analytics Catalog gives a comprehensive browsing experience for content in Reporting Workbench, Crystal, and Radar. From a single user interface, you can browse, favorite, and search for governed analytics content.

The catalog is embedded directly into the home page of Cogito so users don't need to go to a separate portal. Using integrated security and role definitions, users see only the content that is available to them. To help clarify the intended use of the report, the catalog supports storing and searching on metadata descriptions, tags, business owners, and administrative comments.

The screenshot displays the Analytics Catalog interface with a search bar at the top containing the text "bed planning". The interface is organized into a sidebar on the left and a main content area on the right. The sidebar includes sections for "Personal" (with checkboxes for "My Favorites" and "My Content"), "Content Type" (with checkboxes for "Dashboards", "Components", "Workbench Reports", "Crystal Reports", and "Web Reports"), "Tags" (with a "Select a tag" dropdown), and "Advanced" (with a "Clear Filters" button). The main content area displays six report cards in a grid. Each card features a purple hammer icon on a grid background. The first card is titled "ADT Bed Planning Productivity: Unit-to-Unit Turnaround Time" and includes a description: "The purpose of this report model is to report on the turnaround time in bed planning to move a...". The second card is titled "ADT Bed Planning Productivity: ED Admissions by Service (with Approval/Rejection Process)" and includes a description: "The purpose of this report model is to report on data about bed planning productivity in moving...". The third card is titled "Beds in Login Department" and includes a description: "This report covers beds in your current login department. It can be further refined to return only a...". The fourth card is titled "ADT Bed Planning Productivity: ED Admissions by Service" and includes a description: "The purpose of this report model is to report on data about bed planning productivity in moving...". The fifth card is titled "ED Bed Utilization (Crystal)" and features an orange diamond icon on a grid background. The sixth card is titled "Hourly Capacity Projections as Table" and includes a description: "Component that shows the expected future capacity of different types of bed based on type and...". Each report card has a "Bed Planning" button at the bottom.

Analytics Catalog

References

Callaway, A. (2021). *Navigating the Changing Face of Healthcare Episode*.
<https://www.rbccm.com/en/gib/healthcare/story.page>

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019, 2019/06/19). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, 6(1), 54. <https://doi.org/10.1186/s40537-019-0217-0>

Glaikovskaya, A. (2021). Most healthcare executives don't trust their organization's data, survey finds. *Fierce Healthcare*.
<https://www.fiercehealthcare.com/tech/majority-healthcare-executives-don-t-trust-their-organization-s-data-survey-finds>

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems*, 2, 3-3. <https://doi.org/10.1186/2047-2501-2-3>

Raza, M. (2018). Data Integrity vs Data Quality: An Introduction. *BMC Blogs*. <https://www.bmc.com/blogs/data-integrity-vs-data-quality/>

Tyagi, B. (2021). *Data consistency, completeness and cleaning*
<https://www.yumpu.com/en/document/read/50116664/data-consistency-completeness-and-cleaning-the-inclen-trust>



Questions?



Methodist
HEALTH SYSTEM

MethodistHealthSystem.org