Building Infrastructure for a Data Driven Organization: CHOP Data Blocks

Caroline Burlingame
David Maier
Christian Minich
Mary Rosman

Children's Hospital of Philadelphia®
Center for Healthcare Quality & Analytics

Presentation Overview

- Data Barriers at CHOP
- The CHOP Blocks Project
  - User Requirements
  - Content Development
  - Governance
- Technical Infrastructure
- Adoption and Next Steps

Children's Hospital of Philadelphia℠

CHOP is becoming increasingly data driven

- Analytics Leadership held an organization wide retreat
- Common themes were revealed by consumers

Children’s Hospital of Philadelphia℠

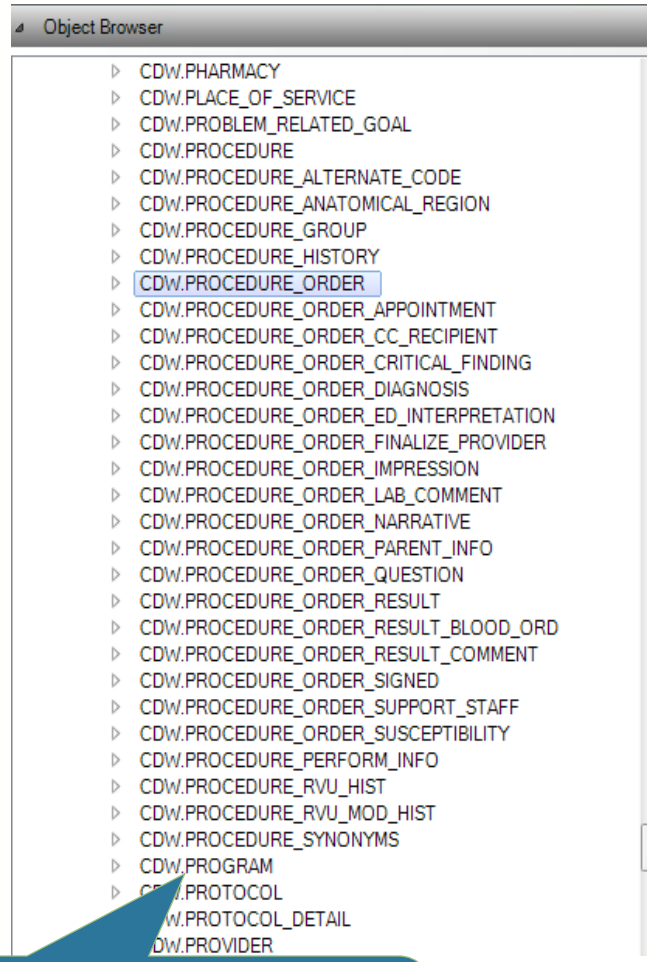
Many of these challenges were rooted in the analyst workflow

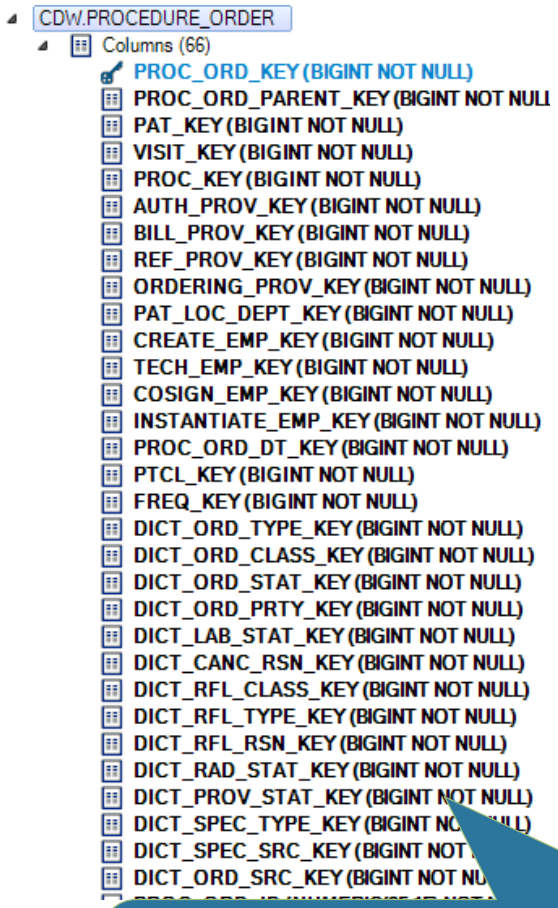
- A routine data request could take hours because a data analyst needs to

  - Find data tables and explore their contents
  - Join them together
  - Calculate metrics and fields

- Re-writing code that contains repetitive business logic, which could result in different answers to foundational questions
- These tasks, while not difficult, take away analyst capacity

Find data tables and explore their contents





- To make data easy to store and analyze, it is broken down into its most granular pieces
- However, that means there are also many different tables (1,000+)
- Data Analysts are specially trained to extract these data with queries and with data tools

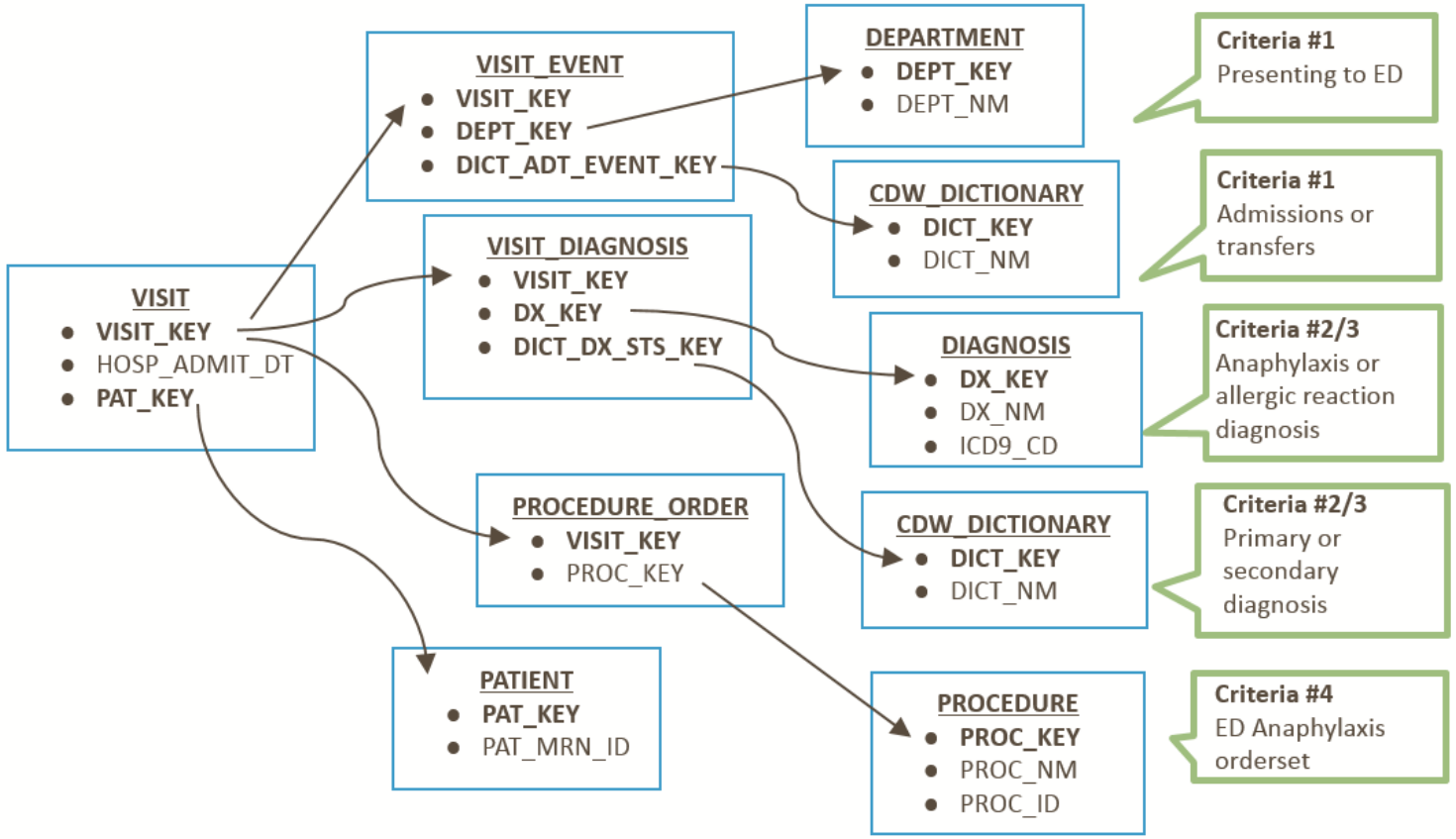
All these tables hold data from procedure orders!

Tables have many fields, often names are not intuitive and must be ‘decoded’

Join tables to create a cohort

I want to review data for patients:

- Presenting to the ED (including transfers)
- With a primary or secondary diagnosis of anaphylaxis
- *or* allergic reaction with the ED anaphylaxis orderset started



Write queries to calculate metrics

Example Request:

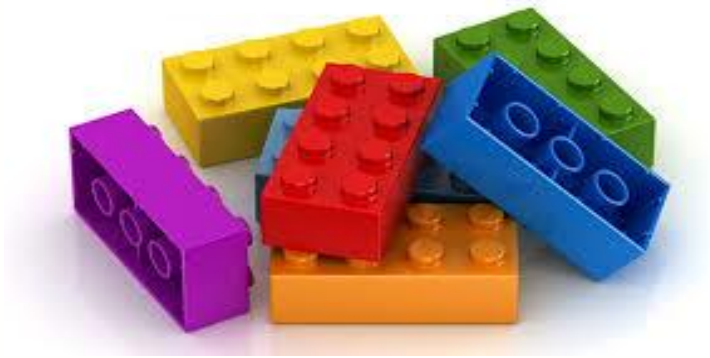
How many inpatients have we seen with a diagnosis of bronchiolitis under the age of one?

- 6 Joins
- 1 Calculation
- 30 Lines of Code

```
select
v.visit_key
, v.ENC_ID
, v.HOSP_ADMIT_DT
, v.HOSP_DISCHRG_DT
, extract(day from v.Hosp_admit_dt - p.dob) as age_days
, p.PAT_MRN_ID
, p.PAT_KEY
, p.DOB
, p.full_nm

from visit v
join cdw_dictionary DICT_ENC_TYPE_KEY on v.DICT_ENC_TYPE_KEY=DICT_ENC_TYPE_KEY.dict_key
join patient p on v.pat_key=p.pat_key

/*Primary diagnosis only*/
join visit_diagnosis vd on vd.visit_key=v.visit_key
join CDW_DICTIONARY dict_dx_sts_key on dict_dx_sts_key.dict_key=vd.dict_dx_sts_key
join CDW_DICTIONARY dict_dx_type_key on dict_dx_type_key.dict_key= vd.dict_dx_type_key
join diagnosis diag1 on diag1.dx_key=vd.dx_key

/*Hospital Encounter came through ED*/
join visit_ed_event vee on vee.visit_key=v.visit_key
join CDW.MASTER_EVENT_TYPE met on vee.event_type_key=met.event_type_key
join department dept1 on dept1.dept_key=v.eff_dept_key

where
/*Acute Bronchiolitis*/
(diag1.icd9_cd like '466.1%'or diag1.icd10_cd like 'J21%')

/*Diagnosis types, primary DX, or when ED doesnt mark one as primary, do the first one
listed*/
and dict_dx_sts_key.DICT_NM in ('HSP ACCT FINAL - PRIMARY')

/*Date Range*/
and date_trunc('day', v.HOSP_ADMIT_DT) >= to_date('2012-12-01', 'yyyy-mm-dd')

/*Hospital Encounter Came Thru the ED*/
and DICT_ENC_TYPE_KEY.dict_nm= 'HOSPITAL ENCOUNTER'
and met.event_id in (60, 50, 95)
and upper(dept1.dept_nm) not like '%URG%'

/*Age Range of Cohort between 29 days and 1 year*/
and extract(day from v.Hosp_admit_dt - p.dob) > 28
and extract(day from v.Hosp_admit_dt - p.dob) < 365.25
```

How can CHOP overcome these barriers?

- De-normalize the data to decrease number of joins
- Pre-calculate common measures
  - Length of stay
  - Revisit indicators
- Remove uncertainty surrounding definitions
- Provide self service tools based on trusted data



Children's Hospital of Philadelphia℠

Create a curated data layer in the data warehouse called CHOP Blocks

GOALS

- Make routine querying simpler
- Ensure data, metrics, definitions are accurate and consistent
- Create Blocks that are use case driven
- Build the foundation for a visual analytics tool

Children's Hospital of Philadelphia℠

What's in a name?

Branding is important for:

- Leadership buy-in
- Marketing
- Fun!

Why CHOP Blocks?

- Differentiate from existing fact tables/data marts
- Conveys the idea that users should build on this layer

Children's Hospital of Philadelphia℠

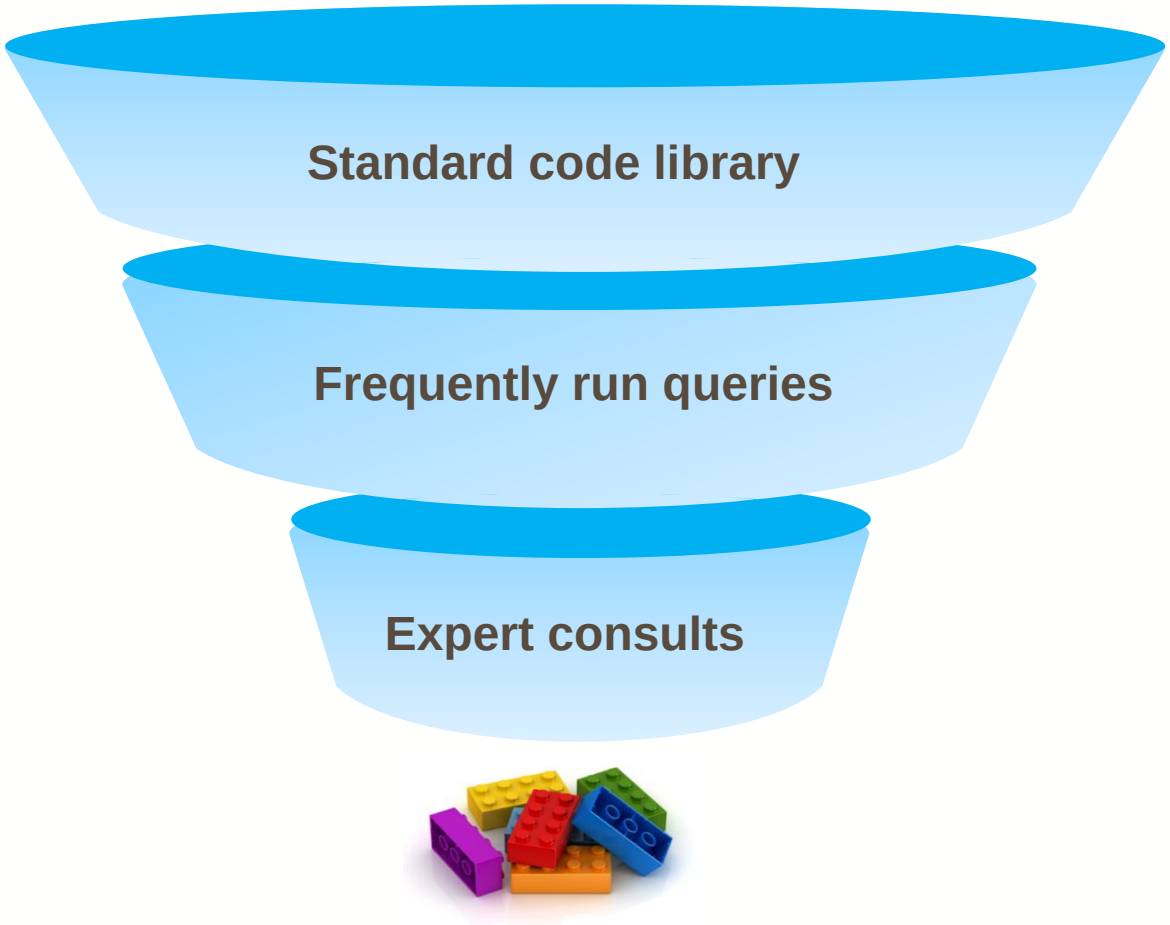
Scrum gets the work done!

- Cross-functional team
  - Analysts
  - Data Engineers
  - Data Governance
- Bi-weekly meetings with analytics leadership

- Agile Scrum Methodology
  - Clear product owner, scrum master roles
  - Groomed backlog
  - Sprint planning
  - True team mentality

Children's Hospital of Philadelphia℠

We used multiple methods to define block content



Children's Hospital of Philadelphia℠

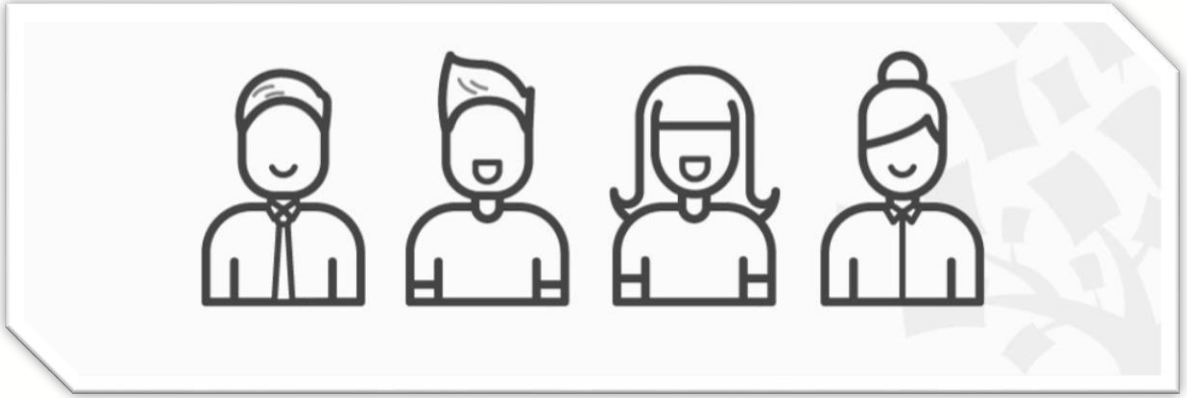
User profiles help drive requirements

Understand needs and experiences of users

Role + goal based

Who should the blocks serve?

- Analysts
- Consumers



Children's Hospital of Philadelphia℠

Analysts want to be more productive

Walter

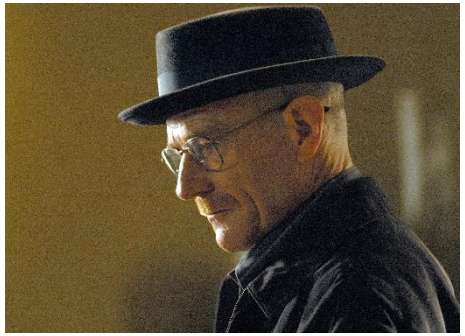
- Novice analyst
- New to the team

Pain points:

- Continuously answering the same types of questions from data requests
- Unsure of what is contained in every EDW table
- Too many joins!

"I want to see if patients are flagged for having a complex chronic condition"
"I don't want to repeatedly calculate BMI"

Block goal:
Reduce time it takes to query data



Heisenberg

- Advanced analyst
- Been at CHOP for several years

Pain points:

- Too much time spent creating and understanding my data sets
- Too many queries to maintain
- Unsure of how data elements are defined

"I want to build a predictive model using CDW data"
"I don't want to update/don't know when to update my query if definitions are changed"

Block goal:
Generate most features for my predictive model

Children's Hospital of Philadelphia℠

Consumers want access to data without an analyst

Pam

- Clinical Nurse Specialist

Pain points:

- Can't access any data on their own
- Waits too long to receive data requests

"I want to have use CDW data without having to write queries"
"I want to expand my data knowledge"
"I don't want to submit data requests for simple questions"

Block goal:
Have data easily accessible in a self service tool

Ruth

- Department Quality Improvement Lead

Pain points:

- Current ways to access data aren't well communicated
- Has to go through someone else to get data, and then wait for the answers

"I want to know that elements are defined consistently and that definitions are transparent"

"I want to build data sets on my own"

Block goal:
Use block data for trusted decision making

Children's Hospital of Philadelphia℠

Block category content defined with stakeholder input

- **Encounters**
- Flowsheet
- Patient
- Diagnosis
- Procedure Order
- Medication Order
- ADT

For ease of use, each table contains the same core columns and common naming conventions.

CLINICAL.ENCOUNTER_INPATIENT
- Columns (35)
  - VISIT_KEY (BIGINT)
  - PATIENT_NAME (CHARACTER VARYING(911))
  - MRN (CHARACTER VARYING(255))
  - DOB (TIMESTAMP)
  - ENCOUNTER_DATE (DATE)
  - CSN (BIGINT)

CLINICAL.PROCEDURE_ORDER_CLINICAL
- Columns (31)
  - **PROC_ORD_KEY (BIGINT NOT NULL)**
  - PATIENT_NAME (CHARACTER VARYING(911))
  - MRN (CHARACTER VARYING(255))
  - DOB (TIMESTAMP)
  - ENCOUNTER_DATE (DATE)
  - CSN (BIGINT)

Children's Hospital of Philadelphia℠

Each category consists of several tables with a specific focus

- **Encounters**
- Flowsheet
- Patient
- Diagnosis
- Procedure Order
- Medication Order
- ADT

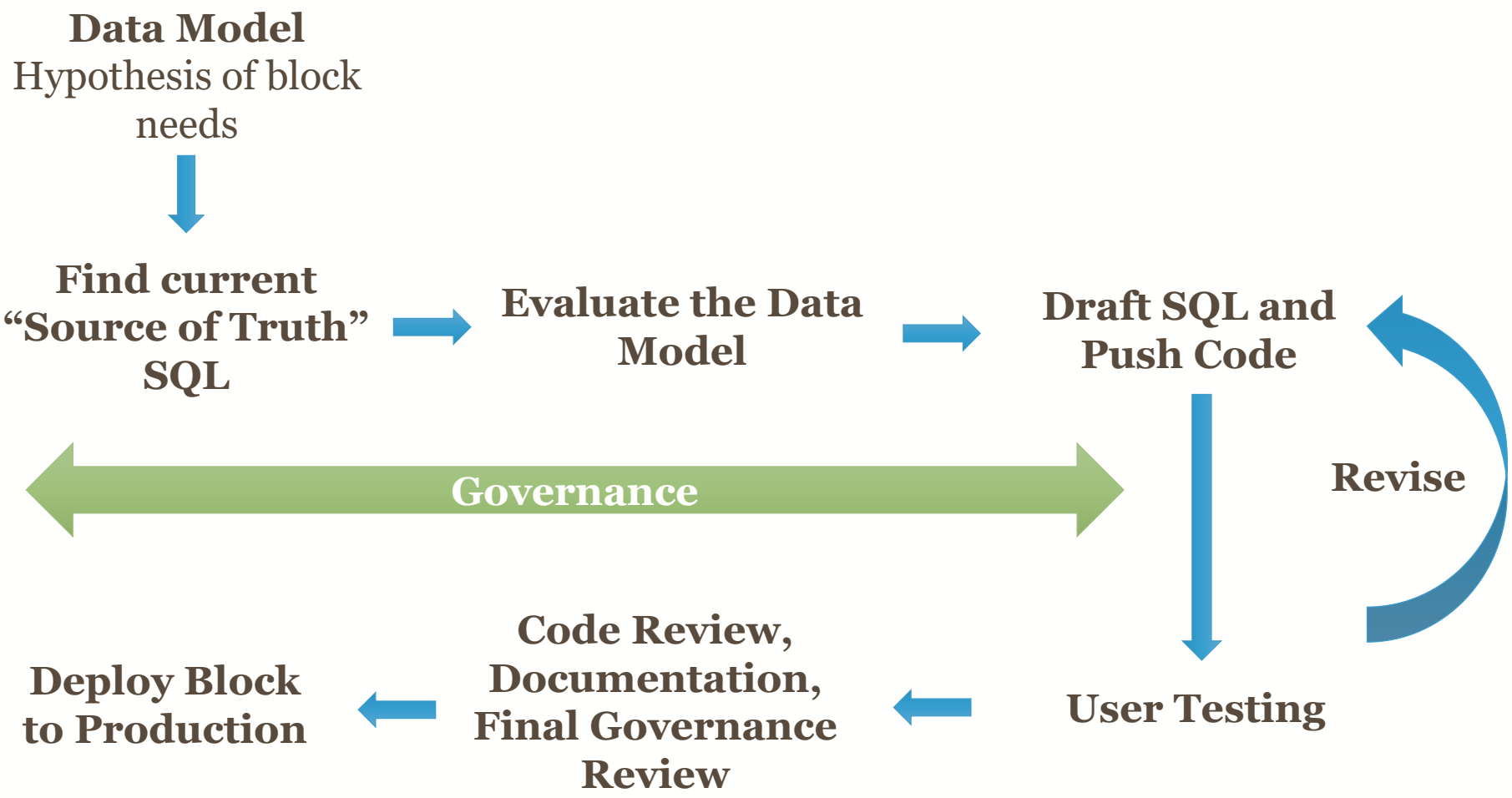
- Encounter_Inpatient
- Encounter_Specialty_Care
- Encounter_Primary_Care
- Encounter_ED
- Encounter_All

Inpatient Definition

- Visits with a hospital encounter visit type
- Visits that have a non-canceled ADT event in a specified list of inpatient departments
- Excludes visits that only touch the ED or ED observation unit

Children's Hospital of Philadelphia℠

Making blocks is an iterative process



Children's Hospital of Philadelphia

Governance is essential in the development process

- Engaged subject area experts to review fields and definitions
- Created action-oriented definitions to provide complete context
  - ~~"The time a patient was checked in"~~
  - "The time a CHOP employee marked in Epic that a patient checked in for an appointment"
- Maintained consistency between block and enterprise definitions

Baked in governance ensures accurate information

- Code changes cannot be put in production without validating that metadata exists for the change
- Definition changes are part of release notes
- Using blocks helps prevent discrepancies
- All metadata is published in an open dictionary tool for block users

Children's Hospital of Philadelphia℠

Infrastructure Goals for Sustainability

- Transparent code
- Open source contribution model
- Reusable architecture pieces
- Continuously validated data and metadata

Children's Hospital of Philadelphia℠

Transparent Code

- Help establish trust with analysts who want to see the logic behind a block
- All code is openly posted in a GitHub repository
- Release notes are published so analysts know “what’s new?”
- Code is formatted according to a standardized SQL style guide

Open Source Contribution Model

- Analysts are the content experts
- Keeps the blocks relevant as the business changes
- Minimizes ongoing engineering work to implement Blocks

User proposes change

Blocks team reviews

User codes change

Blocks team reviews, approves, & documents

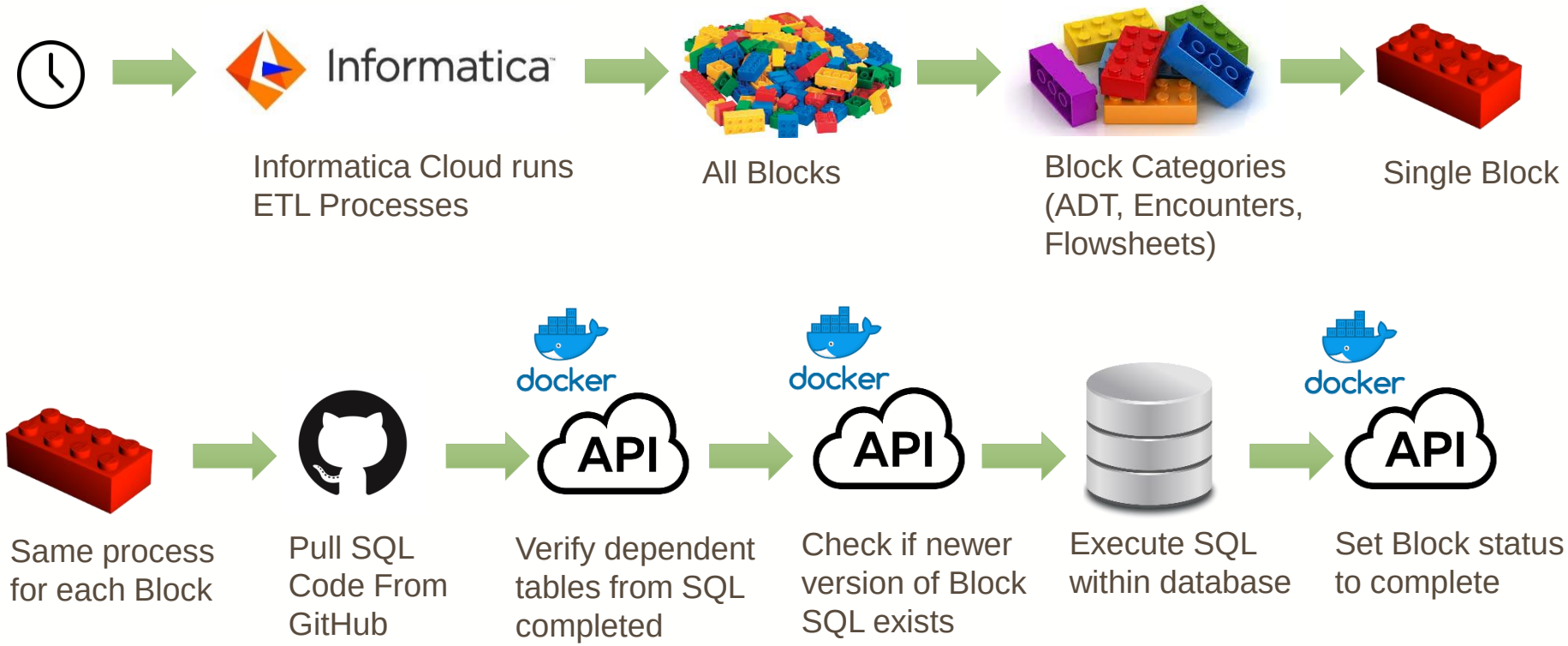
WAYS TO CONTRIBUTE

- Creating or commenting on GitHub issues
  - 19 issues created by 9 non-Core blocks team member
- Submitting a proposed code change (business logic updates, new block, etc.)
  - 5 code contributions from 4 external contributors

Children's Hospital of Philadelphia

Reusable Architecture

- Built pieces to be reused by data engineers outside of the blocks project



Children's Hospital of Philadelphia

Continuously Validated Data and Metadata

Matt proposes adding discharge disposition into the `encounter_inpatient` block

Pull Request

`EAR/Chop-Data-Blocks`

`devinem4/Chop-Data-Blocks`

GitHub Hook

Jenkins

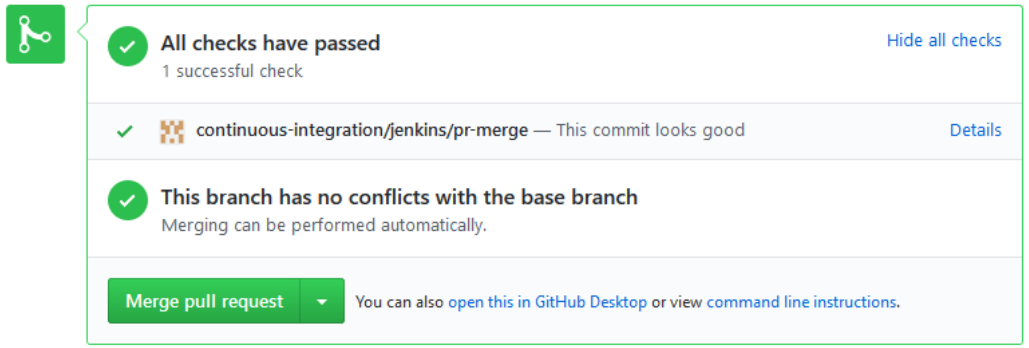
Jenkins checks out the version of the code at `devinem4/Chop-Data-Blocks`

Jenkinsfile

docker

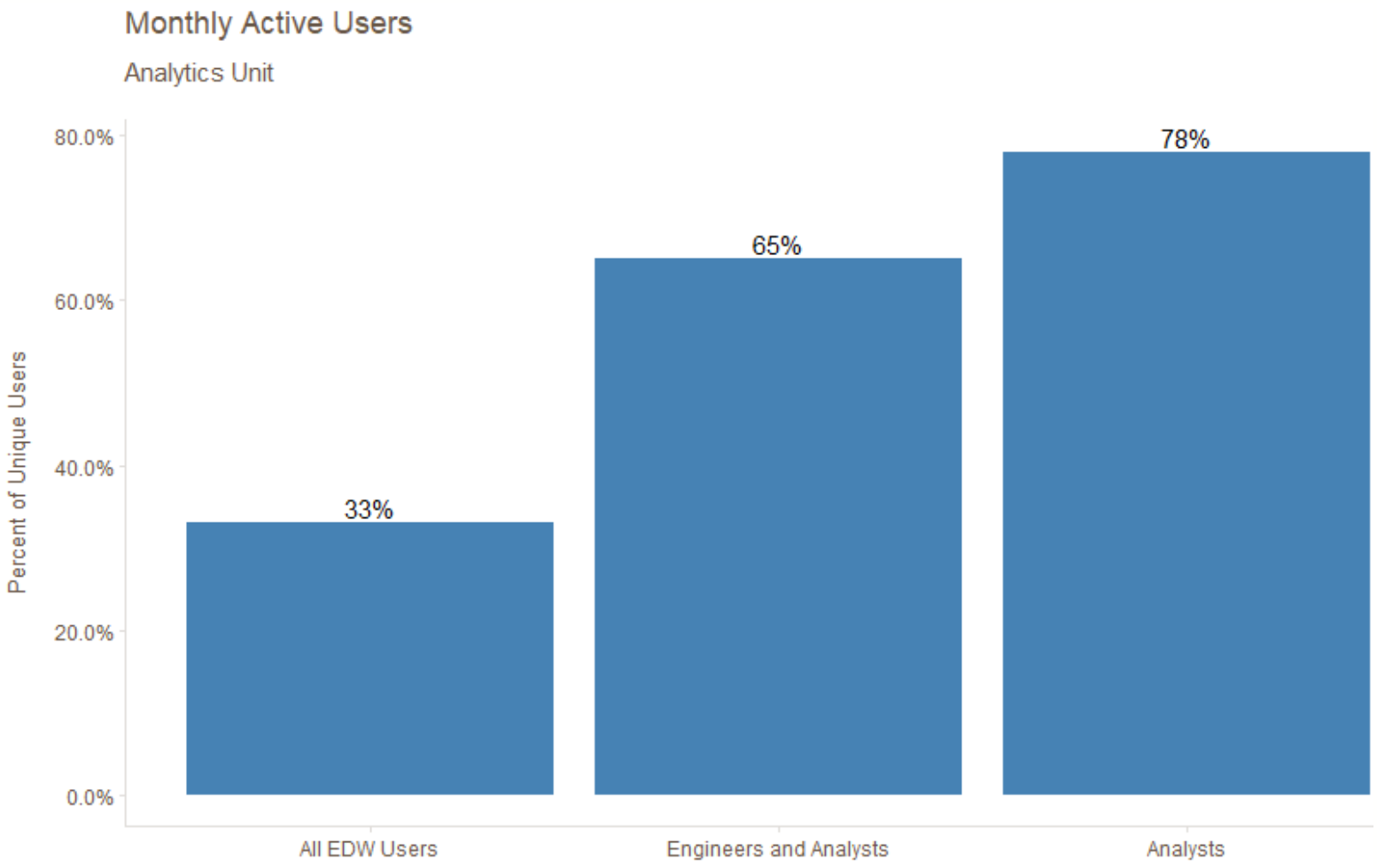
python

Jenkins sends a message to GitHub, saying that Matt's changes have passed all our tests.



Children's Hospital of Philadelphia

Most analysts use Blocks regularly

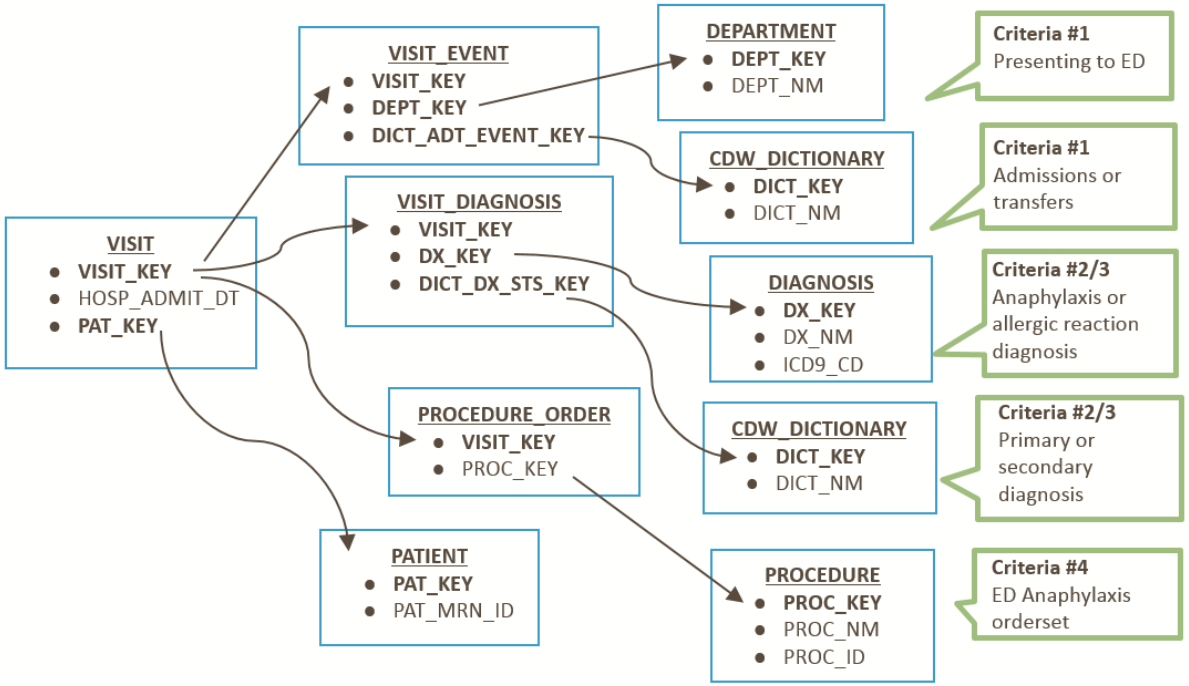


55 Users

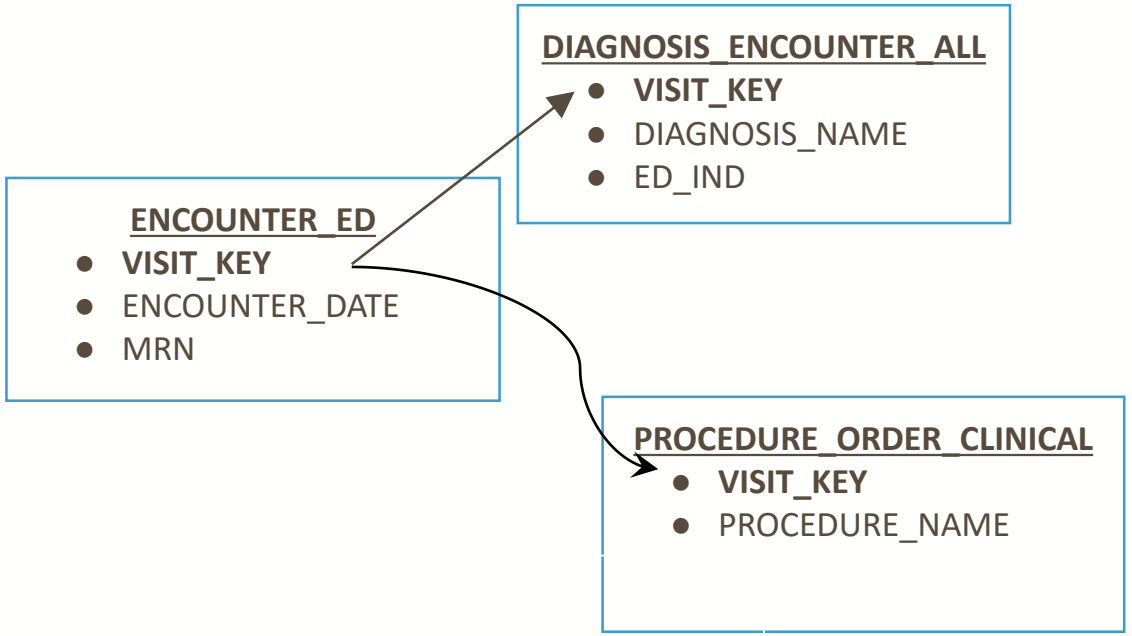
- Marketing
  - User testimonials
  - Presenting at forums
  - Release Notes
- Analyst ownership
- Stakeholder buy-in

Children's Hospital of Philadelphia

Remember Before…



Look Now!



Children's Hospital of Philadelphia

1000 + tables before blocks… To only 17!

- CDW.PPLS_PATIENT
- CDW.PPLS_STAFF_LOCATION_HISTORY
- CDW.PPLS_STAFF_STATUS_HISTORY
- CDW.PPLS_WORKFLOW_MILESTONE_HISTORY
- CDW.PPLS_WORKFLOW_PRIMARY_STAFF_HISTORY
- CDW.PPLS_WORKFLOW_TAGGABLE
- CDW.PRECISION_QUEUE
- CDW.PROBLEM_RELATED_GOAL
- CDW.PROCEDURE
- CDW.PROCEDURE_ALTERNATE_CODE
- CDW.PROCEDURE_ANATOMICAL_REGION
- CDW.PROCEDURE_GROUP
- CDW.PROCEDURE_HISTORY
- CDW.PROCEDURE_ORDER
- CDW.PROCEDURE_ORDER_APPOINTMENT
- CDW.PROCEDURE_ORDER_CC_RECIPIENT
- CDW.PROCEDURE_ORDER_COMMENT
- CDW.PROCEDURE_ORDER_CRITICAL_FINDING
- CDW.PROCEDURE_ORDER_DIAGNOSIS
- CDW.PROCEDURE_ORDER_ED_INTERPRETATION
- CDW.PROCEDURE_ORDER_FINALIZE_PROVIDER
- CDW.PROCEDURE_ORDER_IMPRESSION
- CDW.PROCEDURE_ORDER_LAB_COMMENT
- CDW.PROCEDURE_ORDER_NARRATIVE
- CDW.PROCEDURE_ORDER_PARENT_INFO
- CDW.PROCEDURE_ORDER_QUESTION
- CDW.PROCEDURE_ORDER_RESULT
- CDW.PROCEDURE_ORDER_RESULT_BLOOD_ORD
- CDW.PROCEDURE_ORDER_RESULT_COMMENT
- CDW.PROCEDURE_ORDER_SIGNED
- CDW.PROCEDURE_ORDER_SUPPORT_STAFF
- CDW.PROCEDURE_ORDER_SUSCEPTIBILITY
- CDW.PROCEDURE_PERFORM_INFO
- CDW.PROCEDURE_RVU_HIST
- CDW.PROCEDURE_RVU_MOD_HIST
- CDW.PROCEDURE_SYNONYMS
- CDW.PROCURE_VENDOR_AGREEMENT
- CDW.PROCURE_VENDOR_AGREEMENT_LINE
- CDW.PROGRAM
- CDW.PROTOCOL
- CDW.PROTOCOL_DETAIL

- Databases (8)
  - ADMIN.BLOCKS
    - Tables (17)
      - CLINICAL.ADMIN_BLOCK_ETL_INFO
      - CLINICAL.ADT_DEPARTMENT
      - CLINICAL.DIAGNOSIS_ENCOUNTER_ALL
      - CLINICAL.DIAGNOSIS_MEDICALLY_COMPLEX
      - CLINICAL.ENCOUNTER_ALL
      - CLINICAL.ENCOUNTER_ED
      - CLINICAL.ENCOUNTER_INPATIENT
      - CLINICAL.ENCOUNTER_PRIMARY_CARE
      - CLINICAL.ENCOUNTER_SPECIALTY_CARE
      - CLINICAL.FLOWSHEET_ALL
      - CLINICAL.FLOWSHEET_GROUP_LOOKUP
      - CLINICAL.FLOWSHEET_VITALS
      - CLINICAL.MEDICATION_ORDER_ADMINISTRATION
      - CLINICAL.PATIENT_ALL
      - CLINICAL.PROCEDURE_ORDER_ALL
      - CLINICAL.PROCEDURE_ORDER_CLINICAL
      - CLINICAL.PROCEDURE_ORDER_RESULT_CLINICAL

Children's Hospital of Philadelphia℠

Blocks give analysts more time for analysis

```
select
v.visit_key
, v.ENC_ID
, v.HOSP_ADMIT_DT
, v.HOSP_DISCHRG_DT
, extract(day from v.Hosp_admit_dt - p.dob) as age_days
, p.PAT_MRN_ID
, p.PAT_KEY
, p.DOB
, p.full_nm

from visit v
join cdw_dictionary DICT_ENC_TYPE_KEY on
v.DICT_ENC_TYPE_KEY=DICT_ENC_TYPE_KEY.dict_key
join patient p on v.pat_key=p.pat_key

/*Primary diagnosis only*/
join visit_diagnosis vd on vd.visit_key=v.visit_key
join CDW_DICTIONARY dict_dx_sts_key on dict_dx_sts_key.dict_key=vd.dict_dx_sts_key
join CDW_DICTIONARY dict_dx_type_key on dict_dx_type_key.dict_key= vd.dict_dx_type_key
join diagnosis diag1 on diag1.dx_key=vd.dx_key

/*Hospital Encounter came through ED*/
join visit_ed_event vee on vee.visit_key=v.visit_key
join CDW.MASTER_EVENT_TYPE met on vee.event_type_key=met.event_type_key
join department dept1 on dept1.dept_key=v.eff_dept_key

where
/*Acute Bronchiolitis*/
(diag1.icd9_cd like '466.1%'or diag1.icd10_cd like 'J21%')

/*Diagnosis types, primary DX, or when ED doesnt mark one as primary, do the first one listed*/
and dict_dx_sts_key.DICT_NM in ('HSP ACCT FINAL - PRIMARY')

/*Date Range*/
and date_trunc('day', v.HOSP_ADMIT_DT) >= to_date('2012-12-01', 'yyyy-mm-dd')

/*Hospital Encounter Came Thru the ED*/
and DICT_ENC_TYPE_KEY.dict_nm= 'HOSPITAL ENCOUNTER'
and met.event_id in (60, 50, 95)
and upper(dept1.dept_nm) not like '%URG%'

/*Age Range of Cohort between 29 days and 1 year*/
and extract(day from v.Hosp_admit_dt - p.dob) > 28
and extract(day from v.Hosp_admit_dt - p.dob) < 365.25
```

Old Code

- 6 joins
- 1 Calculation
- 30 Lines of Code

Example Request:

How many inpatients have we seen with a diagnosis of bronchiolitis under the age of 1

```
select
,visit_key,Encounter_ID
,Hospital_Admission_Date
,Hospital_Discharge_Date
,age_years
,mrn
,pat_key
,DOB
,Patient_name

from encounter_inpatient

where primary_dx_icd10 like 'J21%'
and AGE_DAYS > 28
and AGE_DAYS < 365
and v.Hospital_Admission_Date >= to_date('2012-12-01', 'yyyy-mm-dd')
```

New code

- 0 Joins
- 0 Calculations
- 14 Lines of Code

Testimonials

"With CHOP blocks, many of the key information were easy to find, simple to code, and quick to run!"

"I've been **using blocks like crazy** and there are three major reasons – speed, consistency, and brevity.

"Blocks … made it super **easy and efficient** to find ED visits and IP admissions for my patients, and had a ton of other relevant info..."

"It's a relief to know that this data has been **validated** and that I won't be sending my customers bad data on account of a silly mistake."

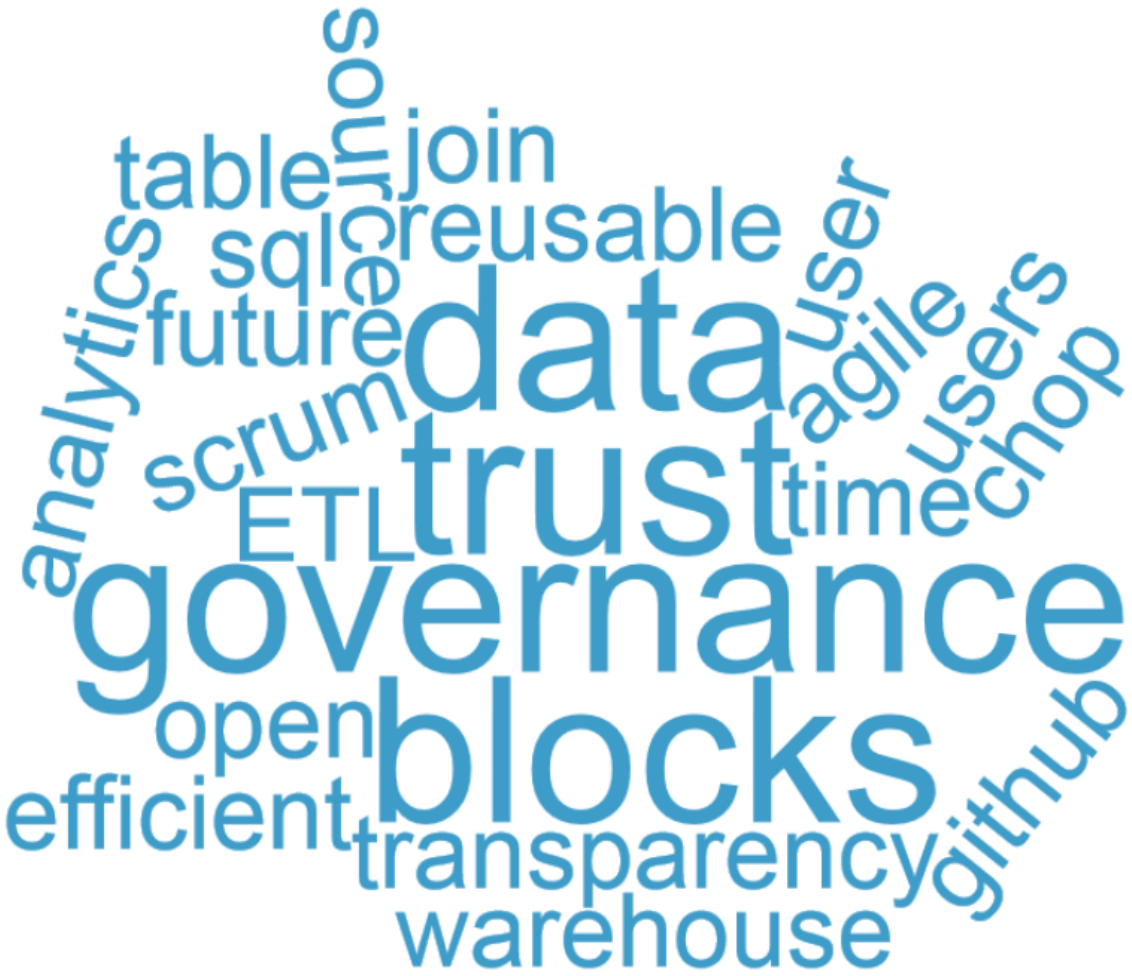
Children's Hospital of Philadelphia℠

Future of data at CHOP

- Blocks will serve as foundation to Self Service BI tools
- Data consumers of all technical levels will have ability to access block data
- Access to the underlying data warehouse will be mostly replaced by blocks

Children's Hospital of Philadelphia℠

Thank you!



Children's Hospital of Philadelphia℠