

Understanding Epic EHR Data from MSDW and Caboodle

Presentation to The Institute for Genomic Health

June 16, 2022

Timothy Quinn, PhD

Scientific Computing and Data

Icahn School of Medicine at Mount Sinai

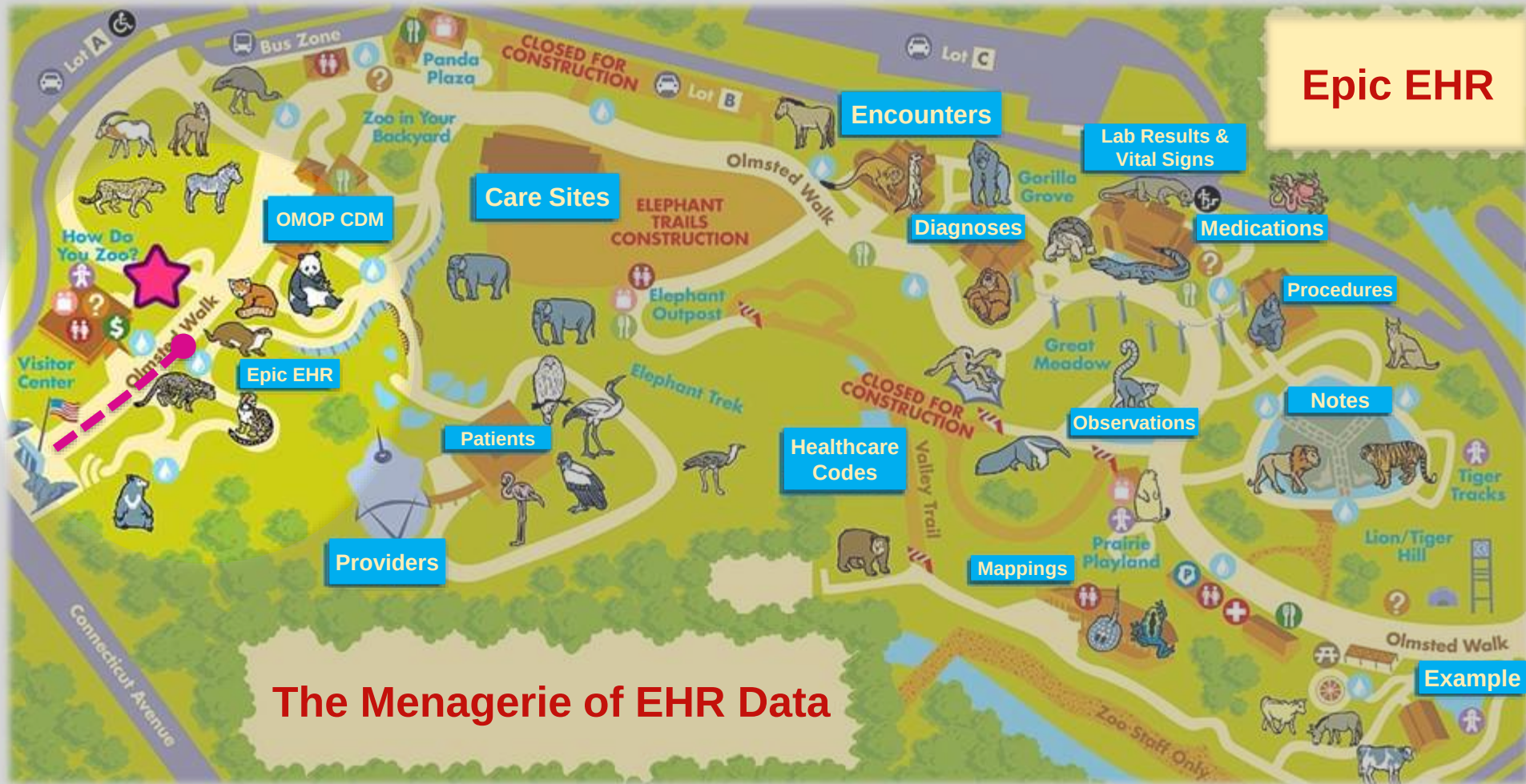


Icahn
School of
Medicine at
**Mount
Sinai**

Agenda

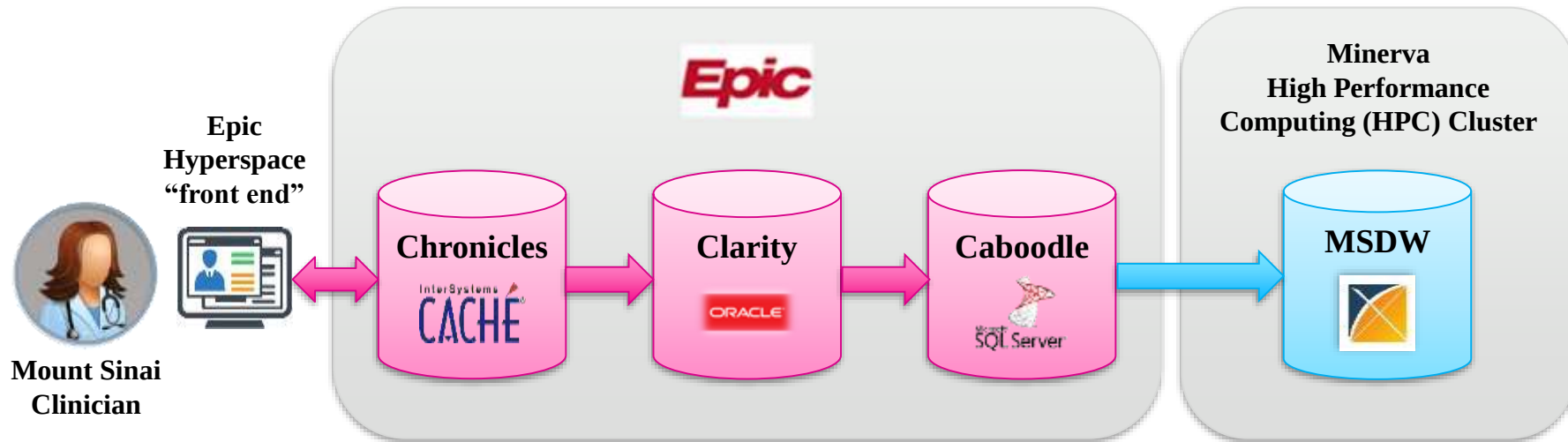
- ▶ Data Pipeline from Epic EHR to MSDW
 - What is a Data Warehouse?
- ▶ OMOP Common Data Model
 - What is a Data Model?
- ▶ EHR Data and Where to Find Them
 - Content Standardization via Mapping to Healthcare Codes
- ▶ What's Coming Next?

Epic EHR

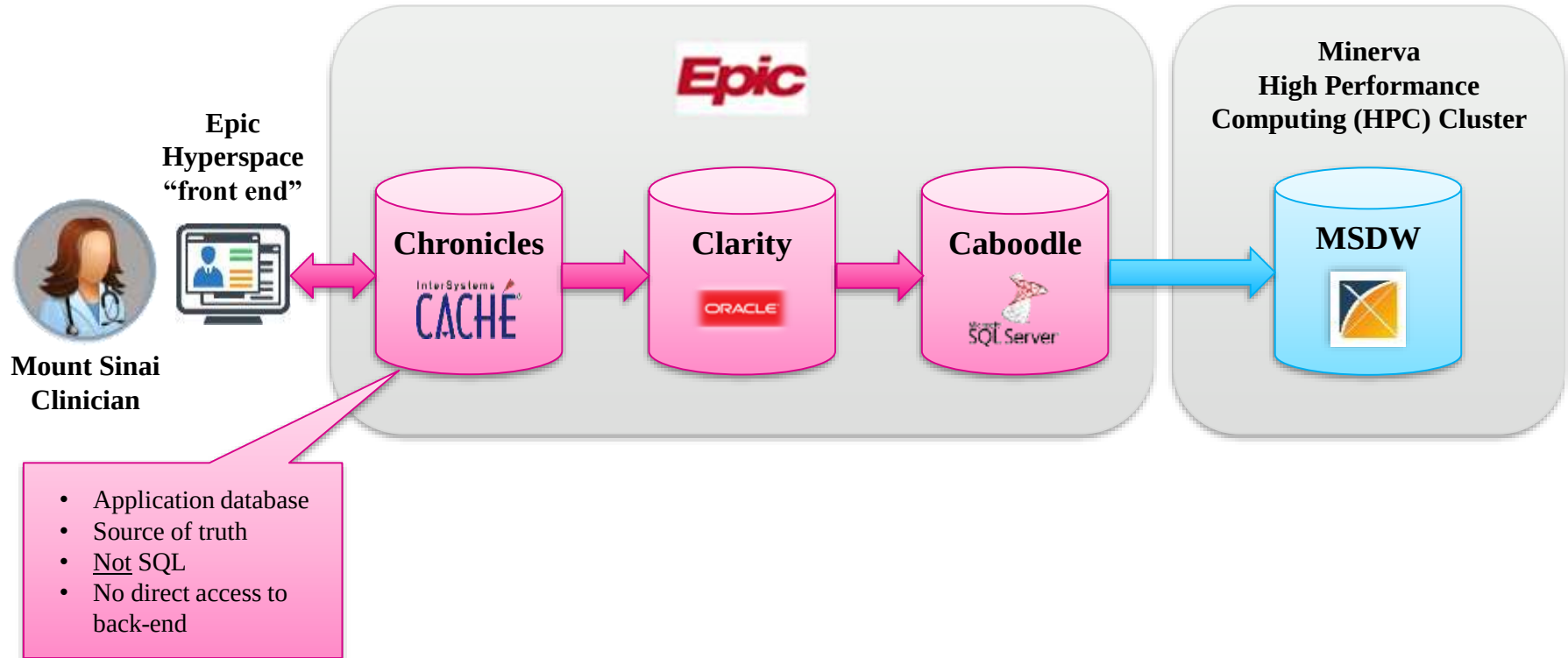


The Menagerie of EHR Data

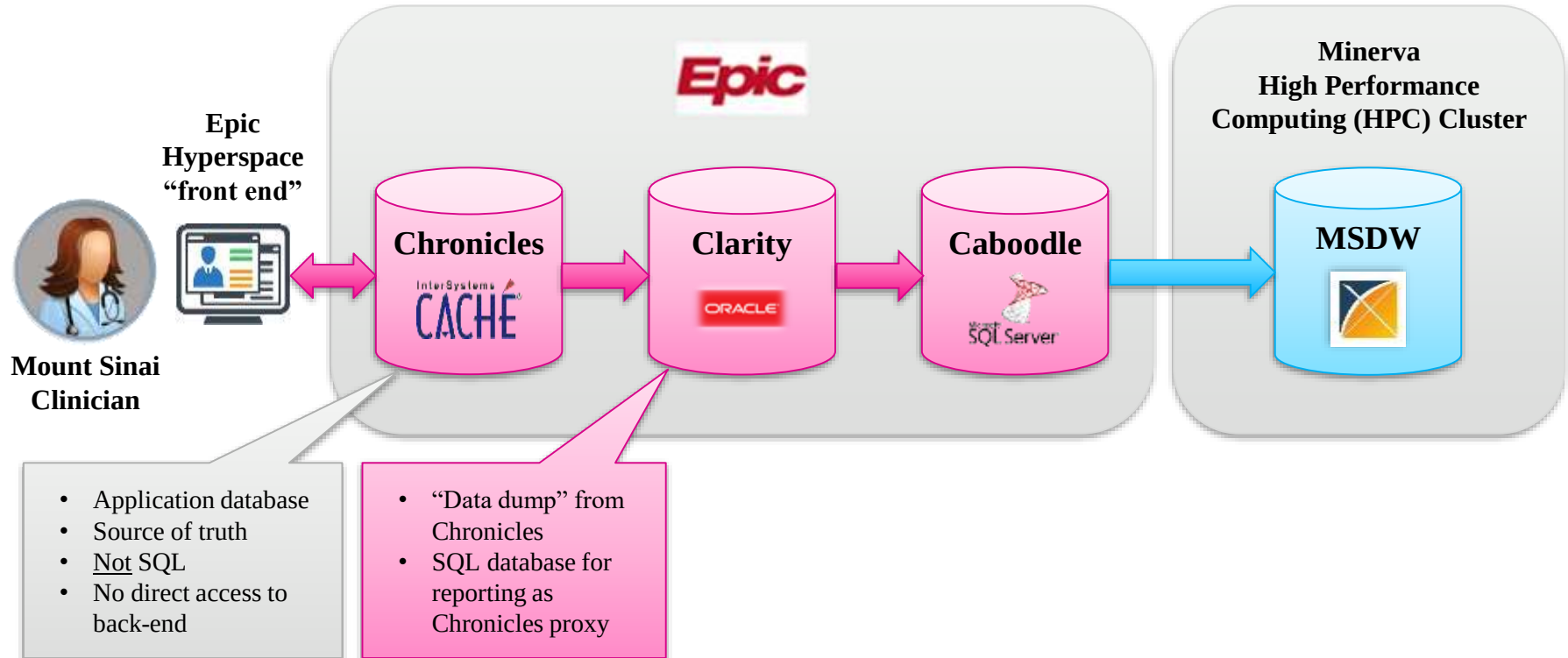
Data Pipeline from Epic EHR to MSDW



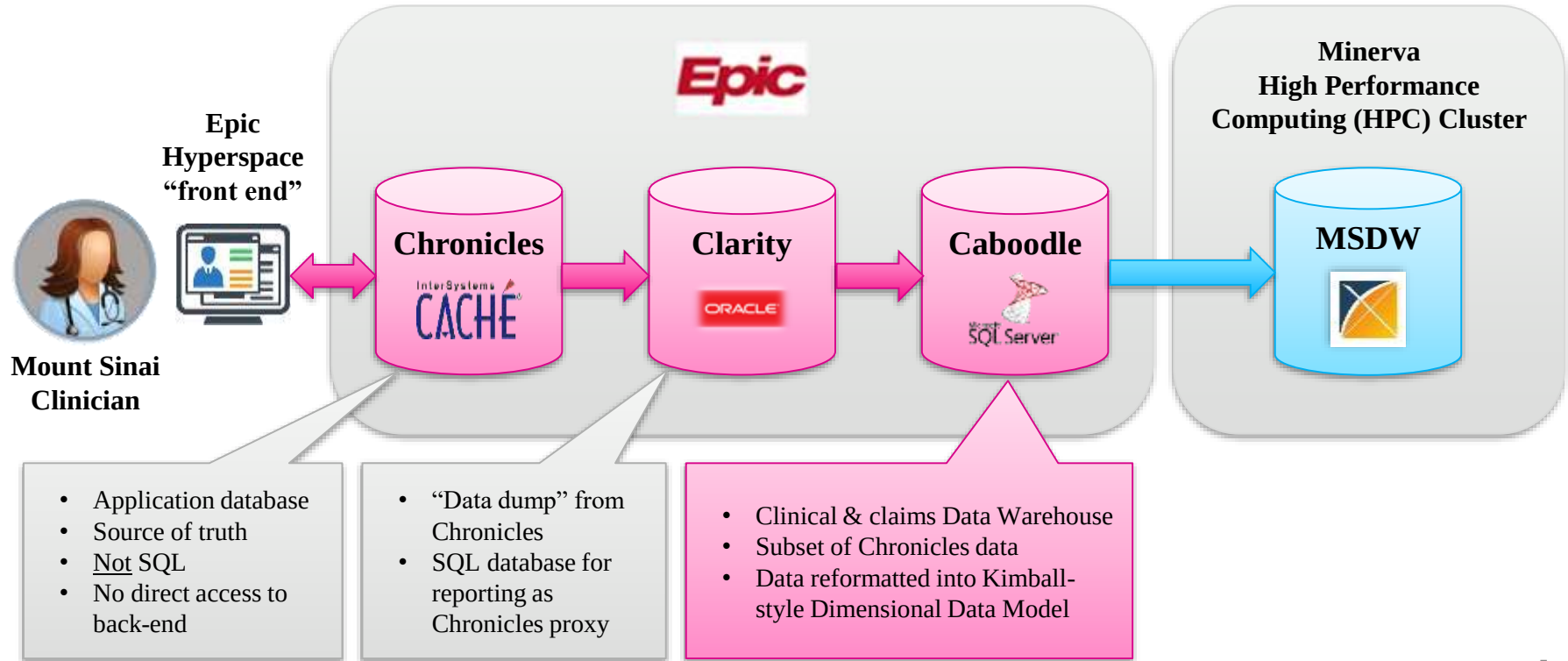
Data Pipeline from Epic EHR to MSDW



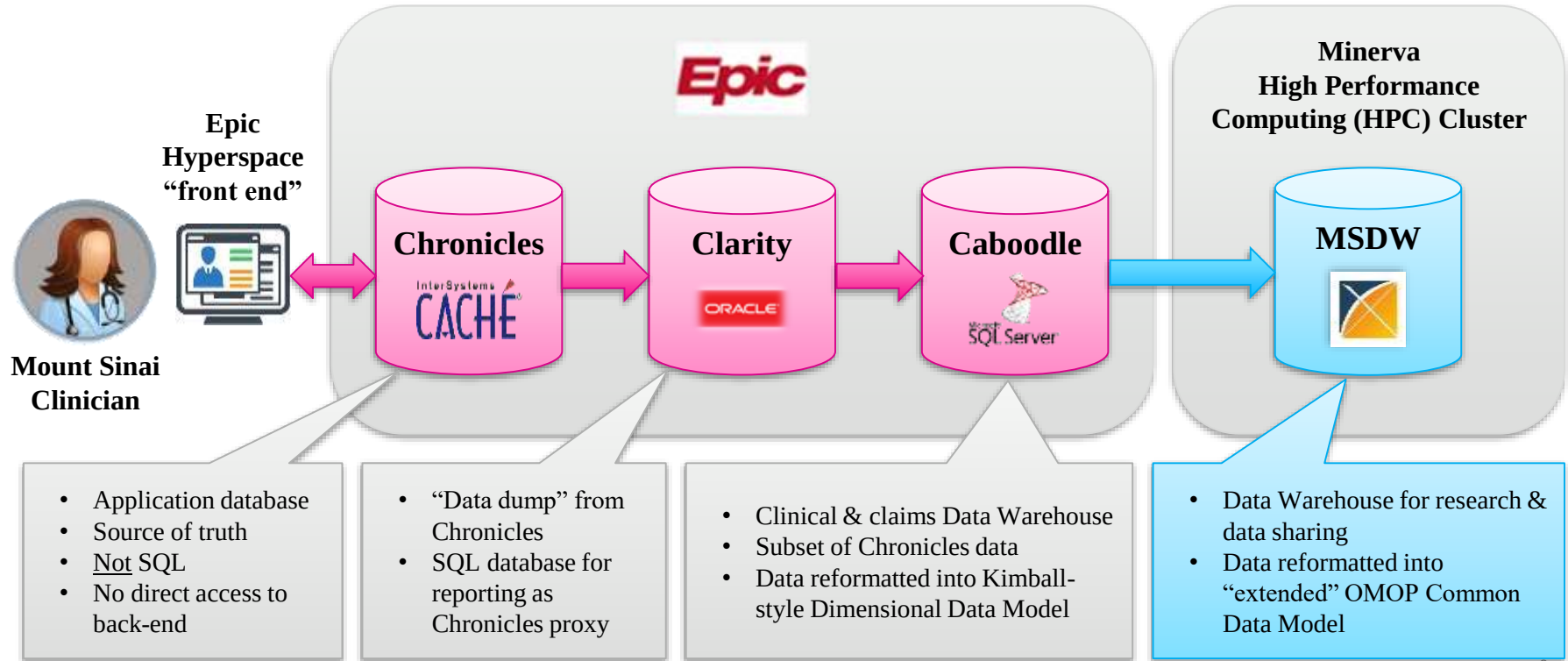
Data Pipeline from Epic EHR to MSDW



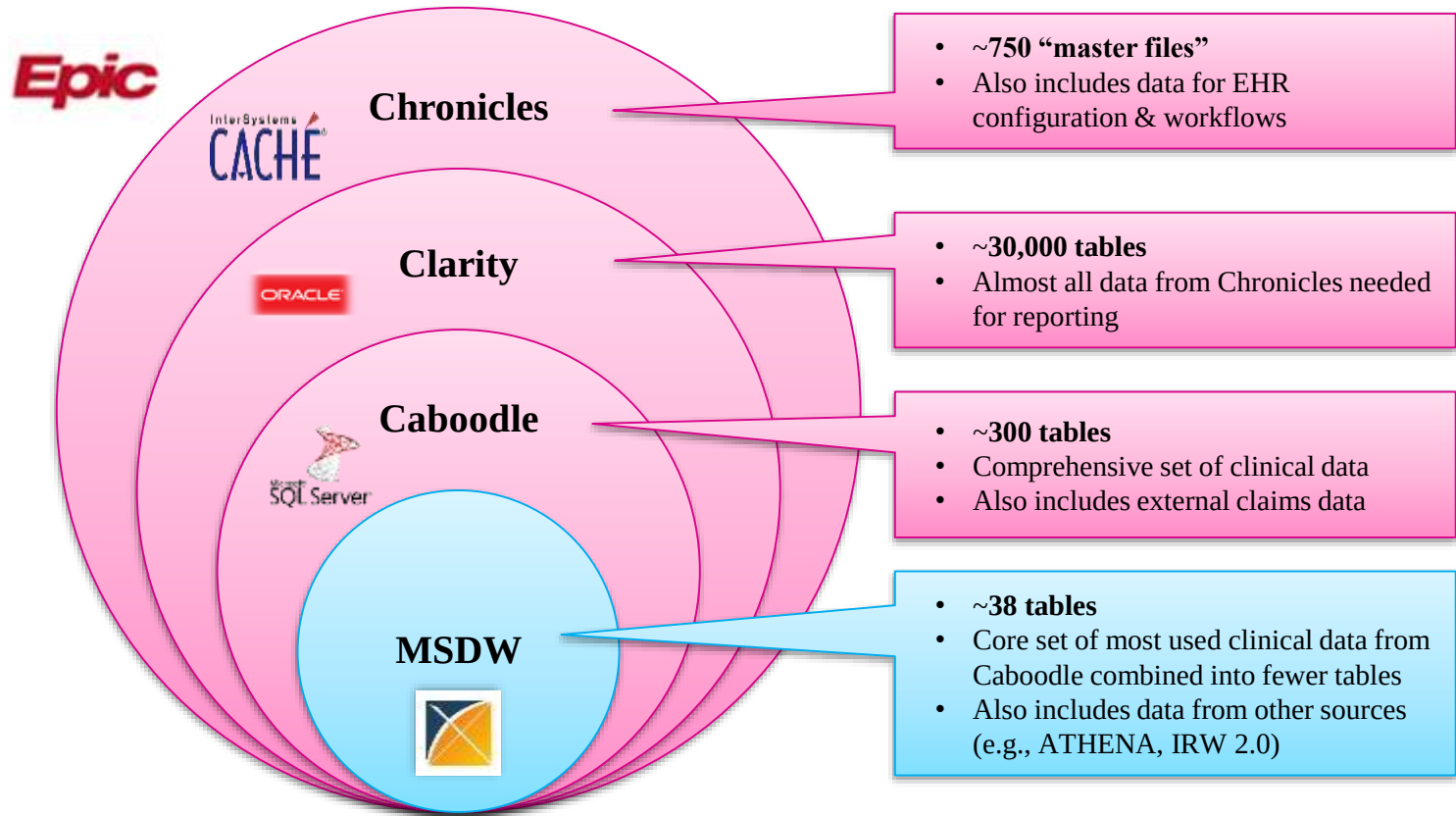
Data Pipeline from Epic EHR to MSDW



Data Pipeline from Epic EHR to MSDW



Venn Diagram of Data Contents



What is a Data Warehouse?

What is a Data Warehouse?



Data Set



Database



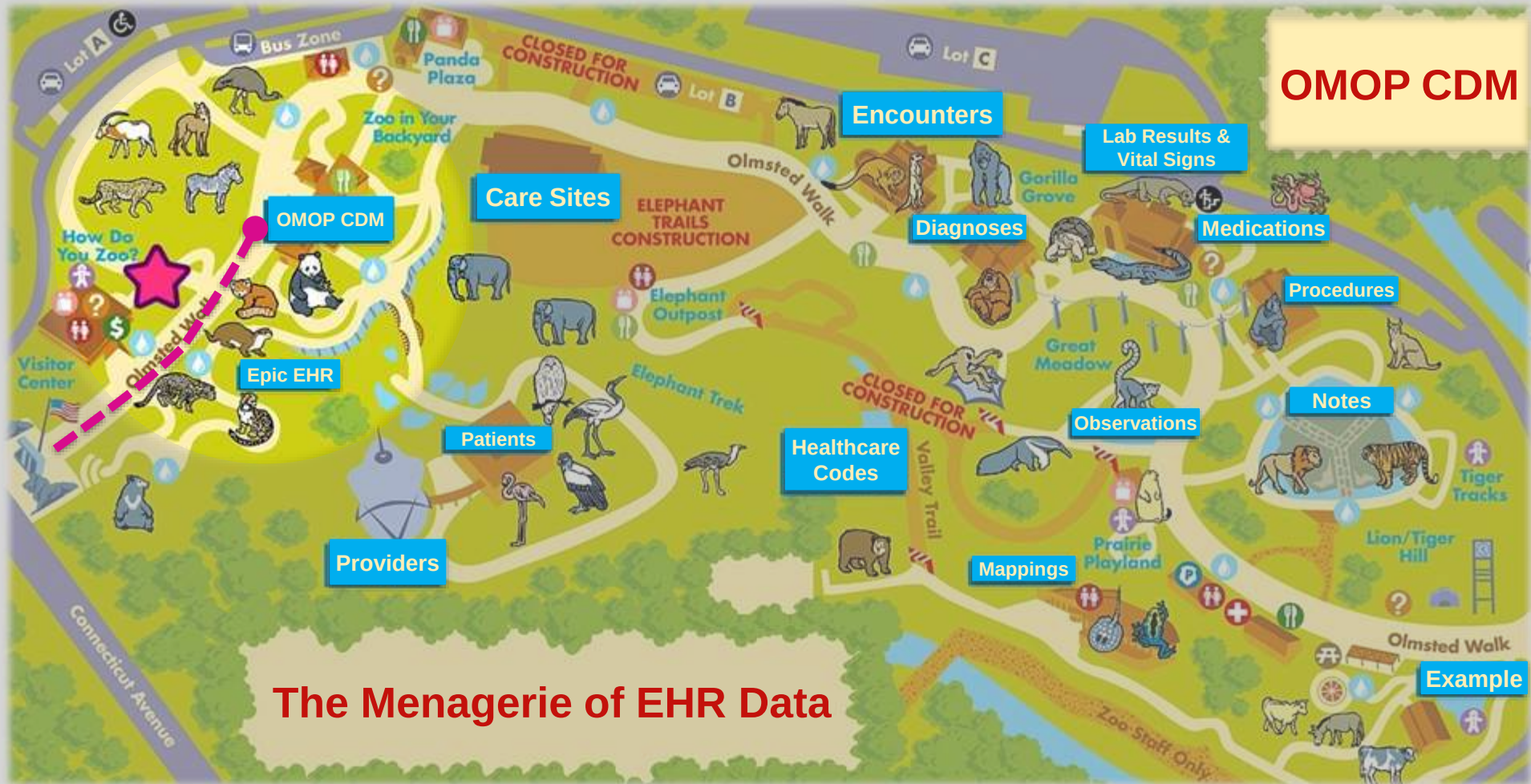
Data
Warehouse



What is a Data Warehouse?

	The data in a Data Warehouse are...	Technical Jargon
1.	Combined from more than one source / system into a centralized repository	<i>"Data integration" or "ETL"</i>
2.	Reformatted / standardized into a data model	<i>"Data transformation"</i>
3.	Loaded into a data model designed to provide value rather than mirror the source systems' formats	E.F. Codd's <i>"Database normalization"</i> Kimball's <i>"Dimensional data modeling"</i>
4.	Loaded in such a way to capture & store changes made over time...	<i>"Change data capture"</i>
5.	...to enable querying the data set as it existed at any time in the past	<i>"Bi-temporal versioning"</i>
6.	Updated incrementally , rather than rebuilt / reloaded from scratch	<i>"Incremental" or "delta" loads</i> rather than <i>"truncate and reload"</i>
7.	Loaded with an auditing framework to monitor & record what data processing happened and when	<i>"Audit balance controls"</i>

OMOP CDM



The Menagerie of EHR Data

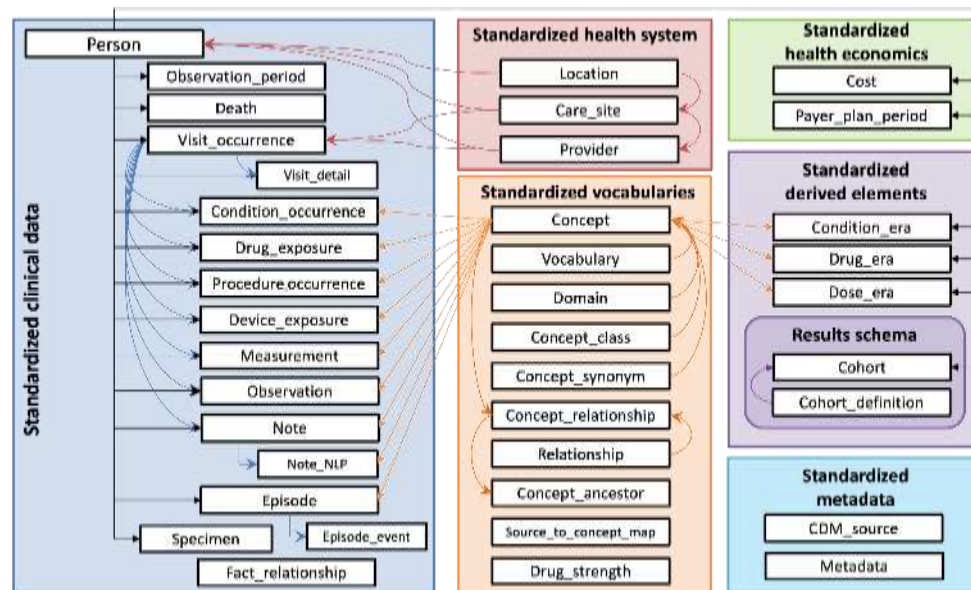
OMOP Common Data Model

► Observational Medical Outcomes Partnership (OMOP)

- Public-private partnership
- 5-year program for surveillance of medical products
- Claims & EHR data
- Chartered by FDA during 2009-2013
- Transitioned to OHDSI consortium in 2013

► Common Data Model (CDM)

- Standardizes both structure & content for interoperability
- Now used by many regional, national, and international research programs & networks



<https://ohdsi.github.io/CommonDataModel/>

What is a Data Model?

An Analogy for Data?

Money's a matter of functions four,
A Medium, a Measure, a Standard, a Store.

based on the work of economist William Stanley Jevons
(1835 - 1882)

Sources:

Milnes, Alfred (1919). *The economic foundations of reconstruction*. Macdonald and Evans. p. 55.
<https://archive.org/details/in.ernet.dli.2015.22790/page/n67>

Accessed: June 11, 2022

Jevons, W. S. (1875). *Money and the Mechanism of Exchange*. D. Appleton and Co.

What is a Data Model?

The container for data is a **table** or matrix

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

What is a Data Model?

Columns are data elements, fields, or variables

The container for data is a **table** or matrix

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

What is a Data Model?

Columns are data elements, fields, or variables

Columns have **data types**: text, date, code, number

The container for data is a **table** or matrix

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

What is a Data Model?

Columns are data elements, fields, or variables

Columns have **data types**: text, date, code, number

The container for data is a **table** or matrix

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	45
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Rows are transactions, records, or observations

“Grain” = What each row represents IRL

What is a Data Model?

Columns are data elements, fields, or variables

Columns have **data types**: text, date, code, number

The container for data is a **table** or matrix

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Rows are transactions, records, or observations

“Grain” = What each row represents IRL

A data element can be **required** or **optional**

What is a Data Model?

Columns are data elements, fields, or variables

Columns have **data types**: text, date, code, number

The container for data is a **table** or matrix

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Rows are transactions, records, or observations

“Grain” = What each row represents IRL

A data element can be **required** or **optional**

Relationships among data elements and records

What is a Data Model?

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	<i>null</i>	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Name	DOB	DOD	Address	Race Code	Fill Date	Prescription Num
Jane Doe	7/4/1969	<i>null</i>	456 Winding Ln	2054-5	3/23/2010	100200300
Jane Doe	7/4/1969	<i>null</i>	456 Winding Ln	2054-5	6/1/2013	400500600
Jane Doe	7/4/1969	<i>null</i>	456 Winding Ln	2054-5	10/19/2015	700800900
Jane Doe	7/4/1969	<i>null</i>	456 Winding Ln	2054-5	1/17/2016	999555222
Jane Doe	7/4/1969	<i>null</i>	456 Winding Ln	2054-5	4/16/2016	777444111

Many data elements
duplicated

What is a Data Model?

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	40
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Name	DOB	DOD	Address	Race Code	Fill Date	Prescription Num
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	3/23/2010	100200300
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	6/1/2013	400500600
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	10/19/2015	700800900
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	1/17/2016	999555222
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	4/16/2016	777444111

Many data elements
duplicated

Patient's address
has changed

What is a Data Model?

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	45
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Patient has many encounters

Grain = 1 encounter per row

Name	DOB	DOD	Address	Race Code	Fill Date	Prescription Num
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	3/23/2010	100200300
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	6/1/2013	400500600
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	10/19/2015	700800900
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	1/17/2016	999555222
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	4/16/2016	777444111

Many data elements duplicated

Patient's address has changed

What is a Data Model?

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	45
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Patient has many encounters

Grain = 1 encounter per row

Name	DOB	DOD	Address	Race Code	Fill Date	Prescription Num
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	3/23/2010	100200300
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	6/1/2013	400500600
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	10/19/2015	700800900
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	1/17/2016	999555222
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	4/16/2016	777444111

Many data elements duplicated

Patient's address has changed

Patient has many prescription fills

Grain = 1 prescription fill per row

What is a Data Model?

Name	DOB	DOD	Address	Race Code	Enc Date	Encounter	Age
Jane Doe	7/4/1969	null	123 Main St	2054-5	3/12/2010	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	9/25/2012	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	4/30/2013	ED Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	5/23/2013	Clinic Office Visit	43
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/11/2015	ED Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	10/12/2015	Hospitalization	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	11/1/2015	Clinic Office Visit	46
Jane Doe	7/4/1969	null	123 Main St	2054-5	1/13/2016	Clinic Office Visit	46

Patient has many encounters

Grain = 1 encounter per row

Name	DOB	DOD	Address	Race Code	Fill Date	Prescription Num
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	3/23/2010	100200300
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	6/1/2013	400500600
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	10/19/2015	700800900
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	1/17/2016	999555222
Jane Doe	7/4/1969	null	456 Winding Ln	2054-5	4/16/2016	777444111

Many data elements duplicated

Patient's address has changed

Patient has many prescription fills

Grain = 1 prescription fill per row

Different grains!

Cannot combine into one table

The Data Model Solution!

Pt ID	Name	DOB	DOD	Race Code
A9999	Jane Doe	7/4/1969	<i>null</i>	2054-5

Pt ID	Address	Start Date	End Date
A9999	123 Main St	1/1/1800	7/20/2011
A9999	456 Winding Ln	7/21/2011	12/31/9999

Create a **separate table** for each **distinct grain** in the data

Pt ID	Enc Date	Encounter	Age
A9999	3/12/2010	Clinic Office Visit	40
A9999	9/25/2012	Clinic Office Visit	43
A9999	4/30/2013	ED Visit	43
A9999	5/23/2013	Clinic Office Visit	43
A9999	10/11/2015	ED Visit	46
A9999	10/12/2015	Hospitalization	46
A9999	11/1/2015	Clinic Office Visit	46
A9999	1/13/2016	Clinic Office Visit	46

Pt ID	Fill Date	Prescription Num
A9999	3/23/2010	100200300
A9999	6/1/2013	400500600
A9999	10/19/2015	700800900
A9999	1/17/2016	999555222
A9999	4/16/2016	777444111

The Data Model Solution!

Eliminate duplicated
data elements

Pt ID	Name	DOB	DOD	Race Code
A9999	Jane Doe	7/4/1969	<i>null</i>	2054-5

Pt ID	Address	Start Date	End Date
A9999	123 Main St	1/1/1800	7/20/2011
A9999	456 Winding Ln	7/21/2011	12/31/9999

Create a **separate table** for
each **distinct grain** in the data

Pt ID	Enc Date	Encounter	Age
A9999	3/12/2010	Clinic Office Visit	40
A9999	9/25/2012	Clinic Office Visit	43
A9999	4/30/2013	ED Visit	43
A9999	5/23/2013	Clinic Office Visit	43
A9999	10/11/2015	ED Visit	46
A9999	10/12/2015	Hospitalization	46
A9999	11/1/2015	Clinic Office Visit	46
A9999	1/13/2016	Clinic Office Visit	46

Pt ID	Fill Date	Prescription Num
A9999	3/23/2010	100200300
A9999	6/1/2013	400500600
A9999	10/19/2015	700800900
A9999	1/17/2016	999555222
A9999	4/16/2016	777444111

The Data Model Solution!

Eliminate duplicated
data elements

Pt ID	Name	DOB	DOD	Race Code
A9999	Jane Doe	7/4/1969	<i>null</i>	2054-5

Pt ID	Address	Start Date	End Date
A9999	123 Main St	1/1/1800	7/20/2011
A9999	456 Winding Ln	7/21/2011	12/31/9999

Track address
changes over time

Create a **separate table** for
each **distinct grain** in the data

Pt ID	Enc Date	Encounter	Age
A9999	3/12/2010	Clinic Office Visit	40
A9999	9/25/2012	Clinic Office Visit	43
A9999	4/30/2013	ED Visit	43
A9999	5/23/2013	Clinic Office Visit	43
A9999	10/11/2015	ED Visit	46
A9999	10/12/2015	Hospitalization	46
A9999	11/1/2015	Clinic Office Visit	46
A9999	1/13/2016	Clinic Office Visit	46

Pt ID	Fill Date	Prescription Num
A9999	3/23/2010	100200300
A9999	6/1/2013	400500600
A9999	10/19/2015	700800900
A9999	1/17/2016	999555222
A9999	4/16/2016	777444111

The Data Model Solution!

Eliminate duplicated data elements

Pt ID	Name	DOB	DOD	Race Code
A9999	Jane Doe	7/4/1969	null	2054-5

Pt ID	Address	Start Date	End Date
A9999	123 Main St	1/1/1800	7/20/2011
A9999	456 Winding Ln	7/21/2011	12/31/9999

All linked together with a **Patient ID**

Track address **changes** over time

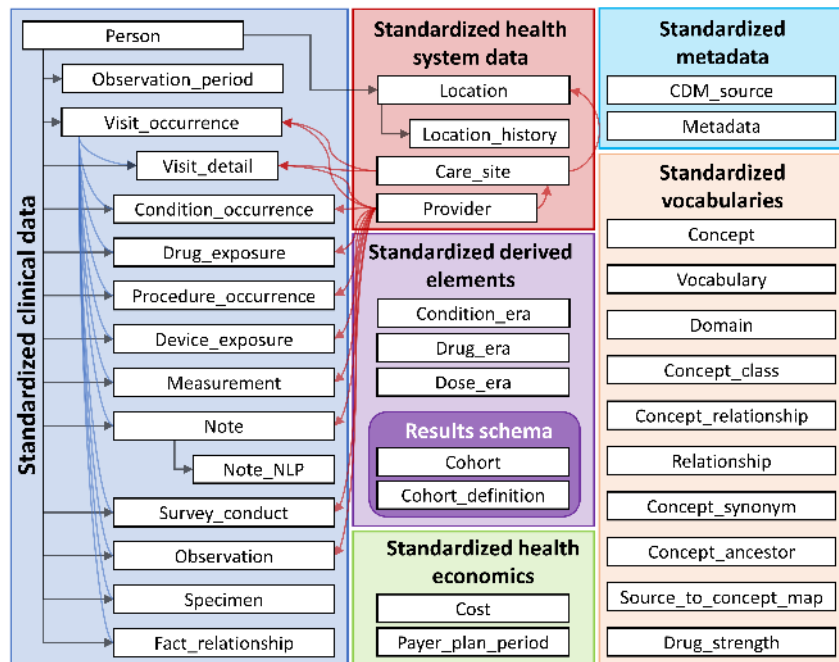
Create a **separate table** for each **distinct grain** in the data

Pt ID	Enc Date	Encounter	Age
A9999	3/12/2010	Clinic Office Visit	40
A9999	9/25/2012	Clinic Office Visit	43
A9999	4/30/2013	ED Visit	43
A9999	5/23/2013	Clinic Office Visit	43
A9999	10/11/2015	ED Visit	46
A9999	10/12/2015	Hospitalization	46
A9999	11/1/2015	Clinic Office Visit	46
A9999	1/13/2016	Clinic Office Visit	46

Pt ID	Fill Date	Prescription Num
A9999	3/23/2010	100200300
A9999	6/1/2013	400500600
A9999	10/19/2015	700800900
A9999	1/17/2016	999555222
A9999	4/16/2016	777444111

OMOP Common Data Model Requirements

1. Standardize **data structure** via common format



2. Standardize **data content** via mapping EHR codes to standard healthcare vocabularies

OMOP Domain	Standard Vocabularies	Non-standard Vocabularies
Condition	SNOMED-CT	ICD-10-CM, ICD-9-CM
Drug	RxNorm, CVX	ATC, NDC, Multum
Measurement	LOINC	SNOMED-CT, Nebraska Lexicon
Procedure	CPT4, HCPCS, ICD-10-PCS	ICD-9-Proc
Observation	SNOMED-CT, LOINC	ICD-10-CM, ICD-9-CM
Race, Ethnicity	OMOP Race, OMOP Ethnicity	SNOMED-CT, Nebraska Lexicon
Provider (Specialty)	NUCC, Medicare Specialty	SNOMED-CT, Nebraska Lexicon
Route	SNOMED-CT	Nebraska Lexicon
Unit	UCUM	SNOMED-CT, Nebraska Lexicon

MSDW has an “Extended” OMOP CDM

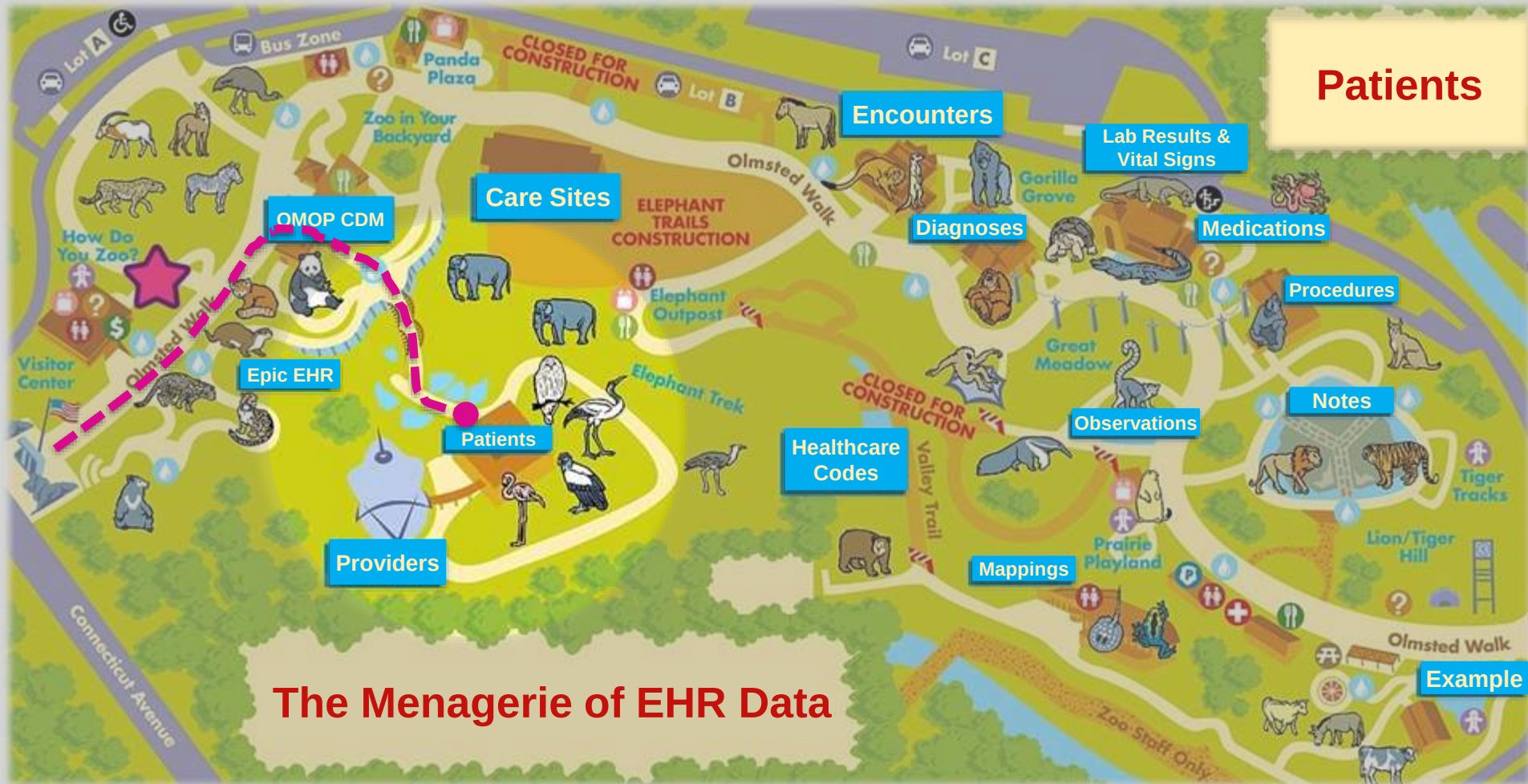
Column Type	OMOP CDM	MSDW Extension Tables
Standard OMOP columns	434	<i>n/a</i>
De-identification columns that mask standard OMOP PHI columns	170	0
Extension columns from Caboodle	291	31
Data lineage & ETL audit columns	717	37
TOTAL:	1,612	68

36 of 38 OMOP tables

cdc_race_ethnicity_xtn
provider_attribute_xtn

EHR Data and Where to Find Them

Patients

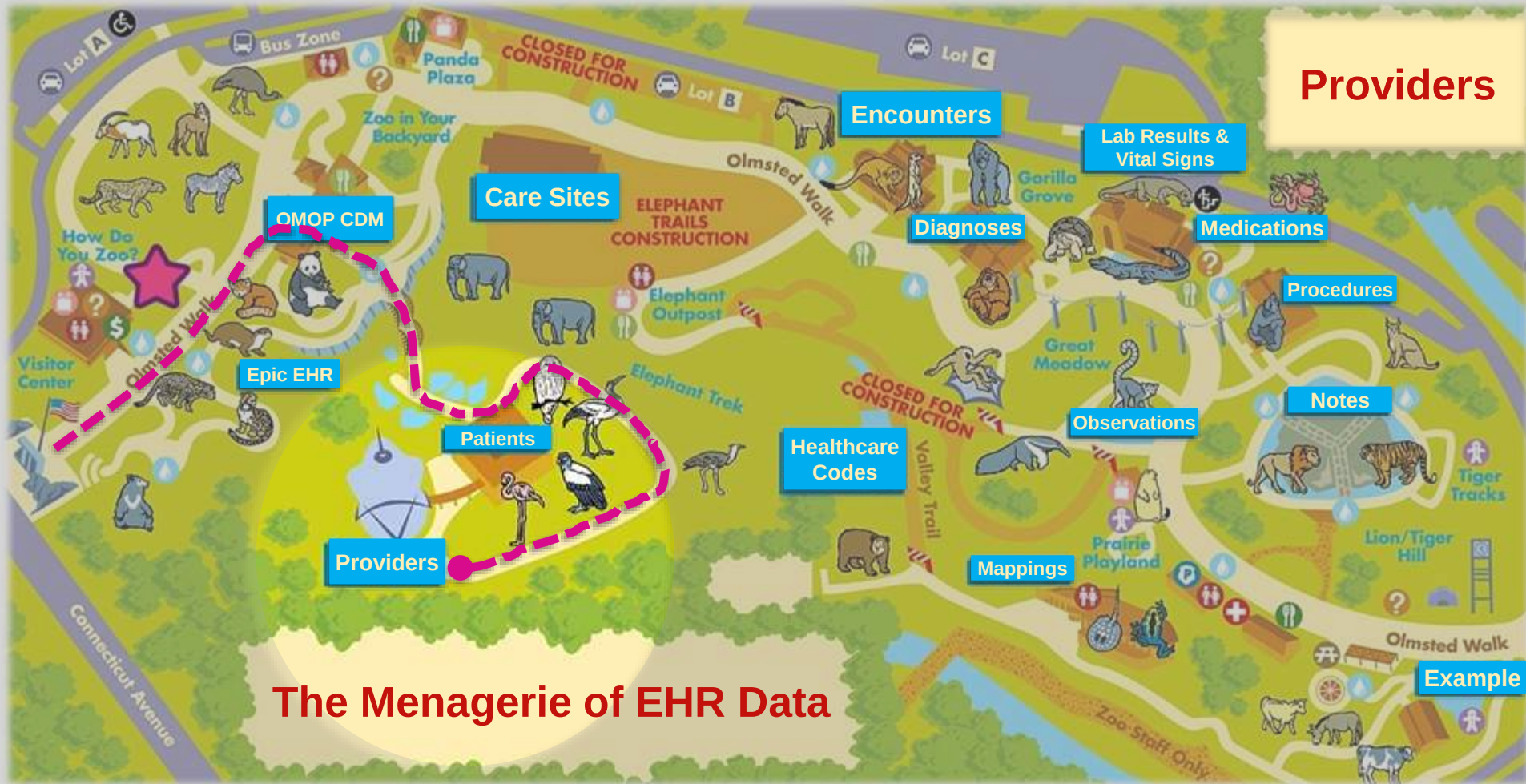


The Menagerie of EHR Data

Patients

observation		person	location		death	
Patient A	Race	Patient A	Patient A	Home Address A	Patient A	Death date
	Ethnicity					
	Language Preference					
	Sexual Orientation					
Patient B	Race	Patient B	Patient B	Home Address B		
	Marital Status					
	Gender Identity					
Patient C	Ethnicity	Patient C				
	Religious Affiliation					

Providers

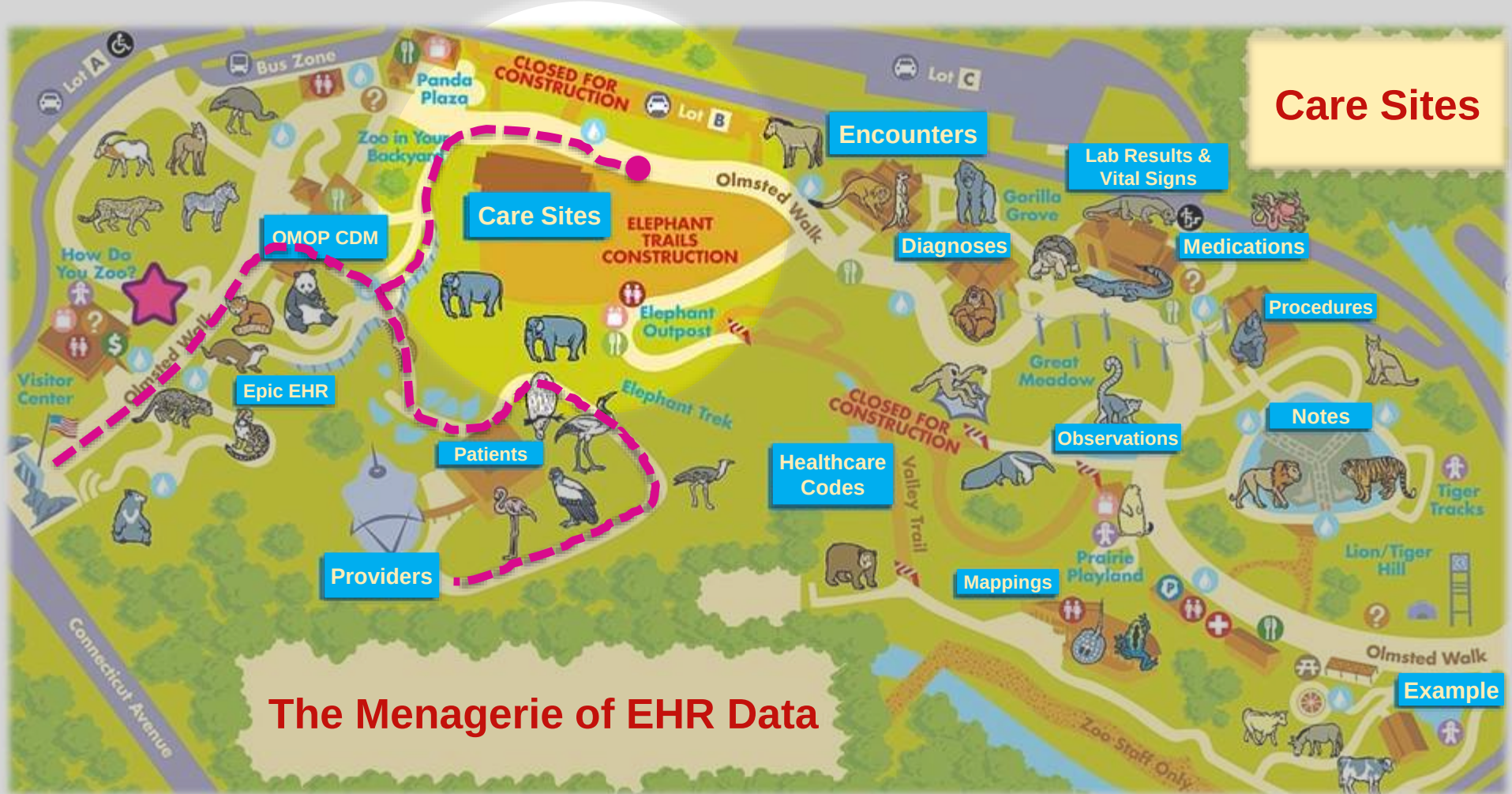


The Menagerie of EHR Data

Providers

provider_attribute_xtn		provider	care_site		location	
Physician A	Specialty 1	Physician A	Physician A	Office A	Office A	Address A
	Specialty 2					
Physician B	Specialty 3	Physician B	Physician B	Office B	Office B	Address B
	Specialty 4					
Physician C	Specialty 2	Physician C				
Clinician D	Specialty 5	Clinician D				
	Specialty 6					
Clinician E	Specialty 7	Clinician E				

Care Sites

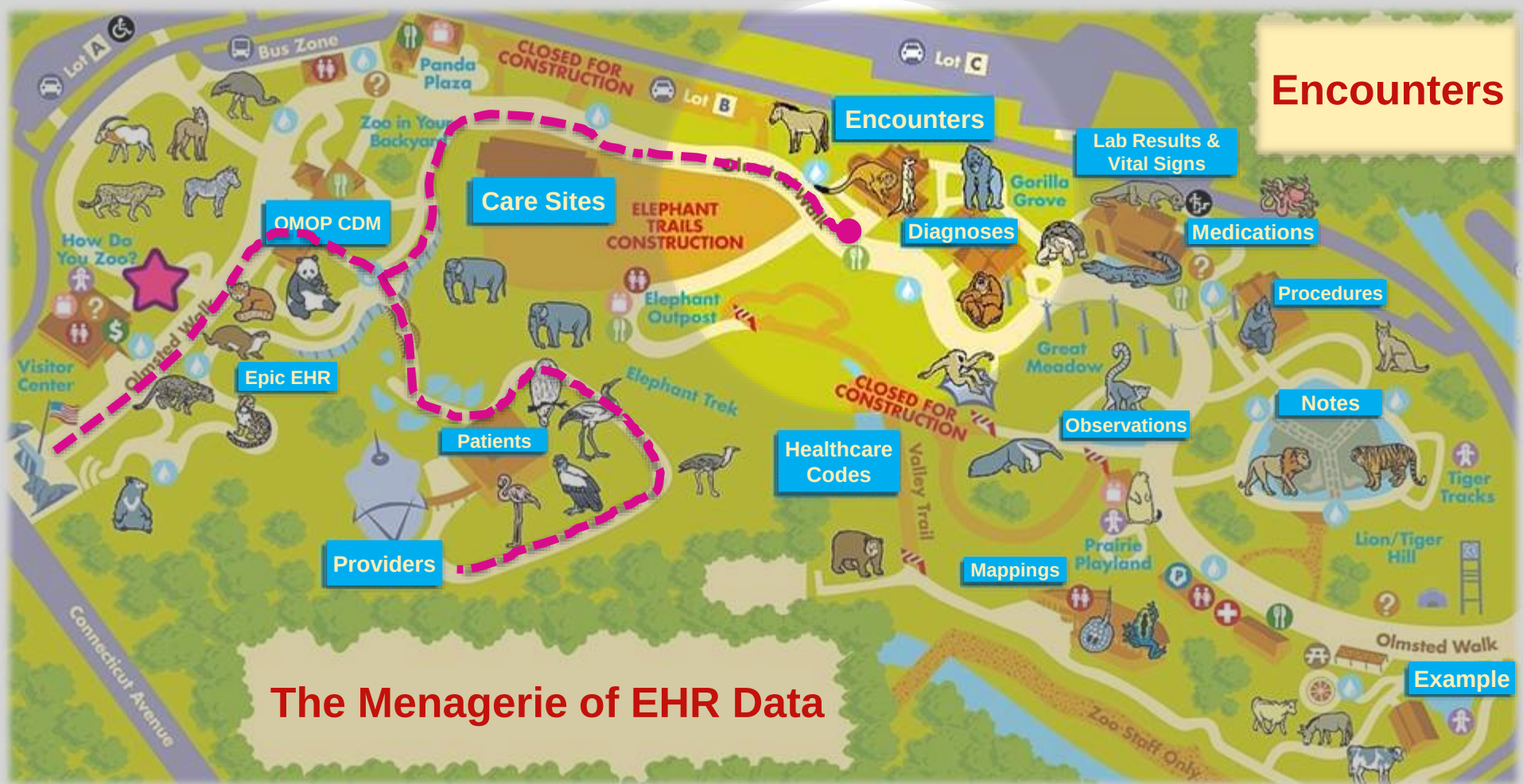


The Menagerie of EHR Data

Care Sites

location		care_site	Definition
Service Area A	Address A	Service Area A	Legal entity
Parent Location B	Address B	Parent Location B	Collection of “locations” grouped together for financial accounting & billing purposes
Location C	Address C	Location C	Actual organizational unit, which often (but not always) represents a physical building
Department D	Address D	Department D	Org unit where patient care is actually delivered, according to Epic’s workflows
		Care Area E	Subdivision of a department, usually a grouping of patient rooms
		Room F	Where an inpatient is assigned, used for inpatient census and ADT transactions
		Bed G	Used when a room has more than one bed (i.e., is shared)
Place of Service H	Address H	Place of Service H	Legal entity or subdivision used for billing purposes

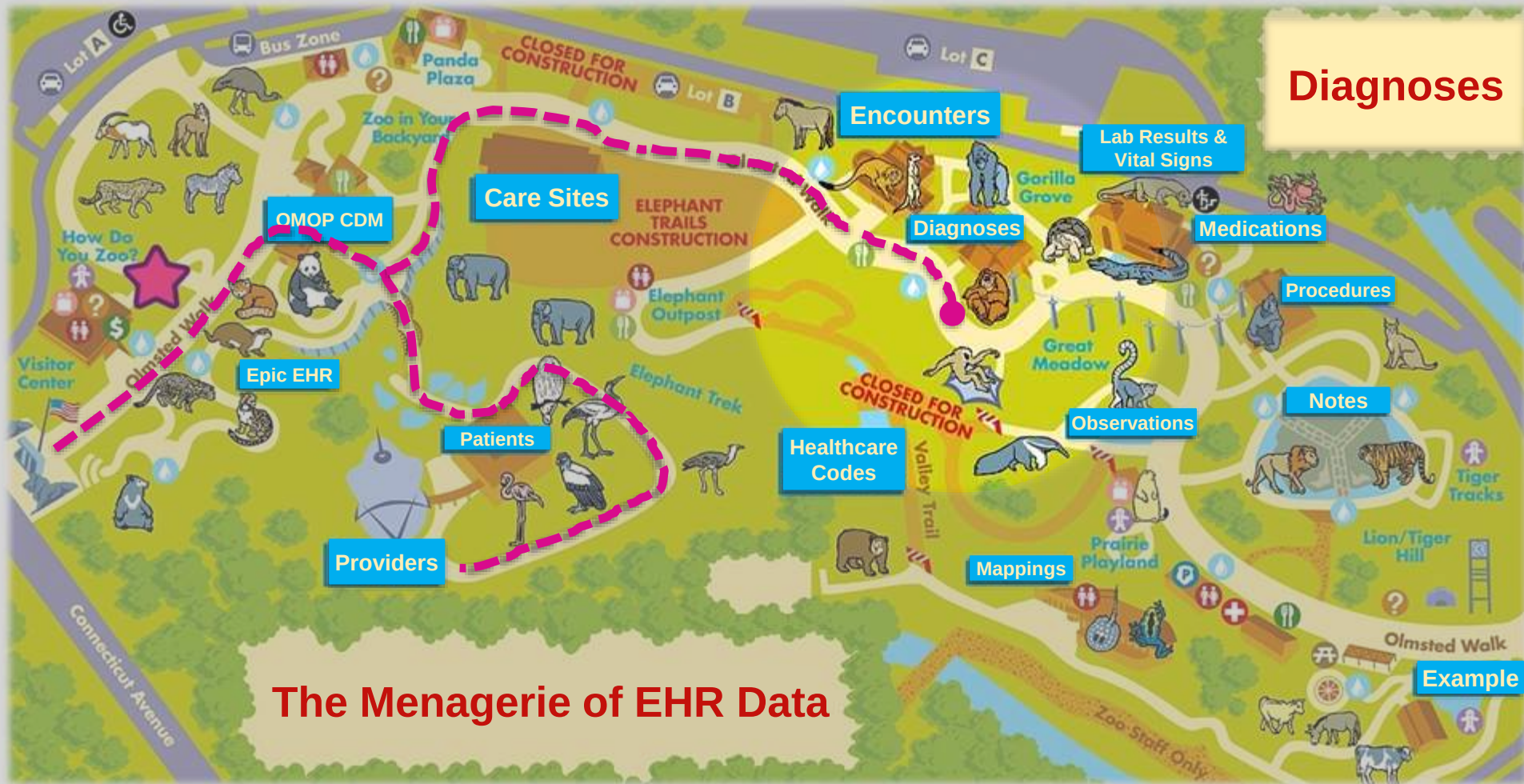
Encounters



Encounters

ED Visit	Hospital Admission	Encounter	visit_occurrence	visit_detail*
	Patient A Inpt Hospitalization 1	Patient A Inpt Hospitalization 1	Patient A Inpt Hospitalization 1	Patient A Care Site Z Care Site Y
Patient B ED Visit 2	Patient B Inpt Hospitalization 2	Patient B Inpt Hospitalization from ED Visit 2	Patient B Inpt Hospitalization from ED Visit 2	Patient B Care Site X Care Site W Care Site V
Patient C ED Visit 3		Patient C ED Visit 3	Patient C ED Visit 3	*Under construction
Patient A ED Visit 4		Patient A ED Visit 4	Patient A ED Visit 4	
		Patient D Hosp. Outpt Visit 5	Patient D Hosp. Outpt Visit 5	
		Patient E Outpatient Visit 6	Patient E Outpatient Visit 6	
		Patient B Telehealth Visit 7	Patient B Telehealth Visit 7	
		Patient F Mobile Unit Visit 8	Patient F Mobile Unit Visit 8	
		Patient D Patient Contact 9		
		Patient G Documentation 10		

Diagnoses

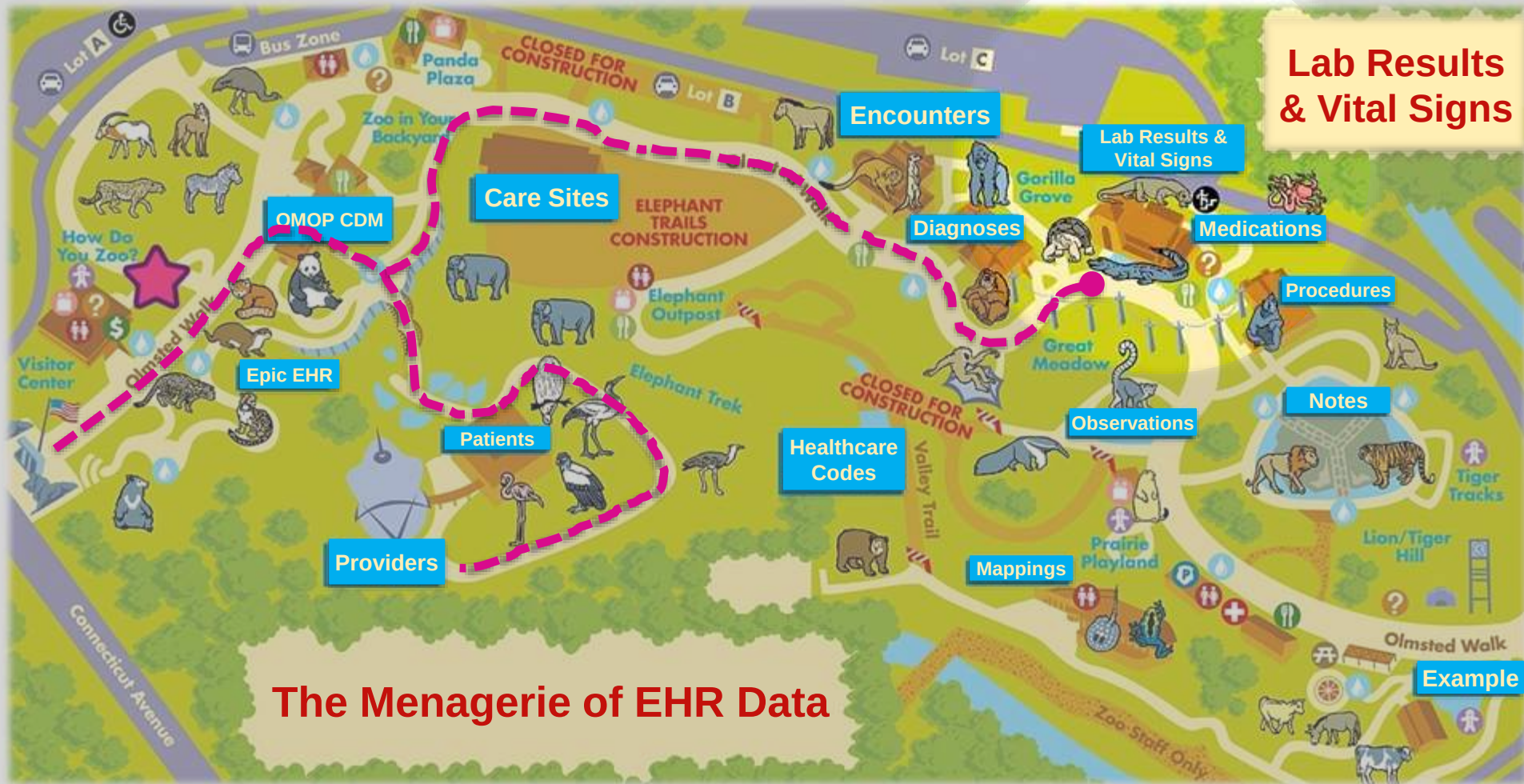


The Menagerie of EHR Data

Diagnoses

visit_occurrence		condition_occurrence		observation
Patient A	Outpatient Visit 1			Patient A Past Medical History F6 Past Medical History K11
Patient B	Outpatient Visit 2	Patient B Outpatient Visit 2	Encounter Diagnosis D4 Encounter Diagnosis H8	Patient B Past Medical History D4 Past Medical History L12
Patient C	Inpt Hospitalization 3	Patient C Inpt Hospitalization 3	Encounter Diagnosis A1 Encounter Diagnosis B2 Hospital Problem C3 Hospital Problem D4 Problem List E5	
Patient D	Inpt Hospitalization 4	Patient D Inpt Hospitalization 4	Encounter Diagnosis F6 Billing Diagnosis G7	
		Patient E	Problem List I9 Problem List J10	

Lab Results & Vital Signs

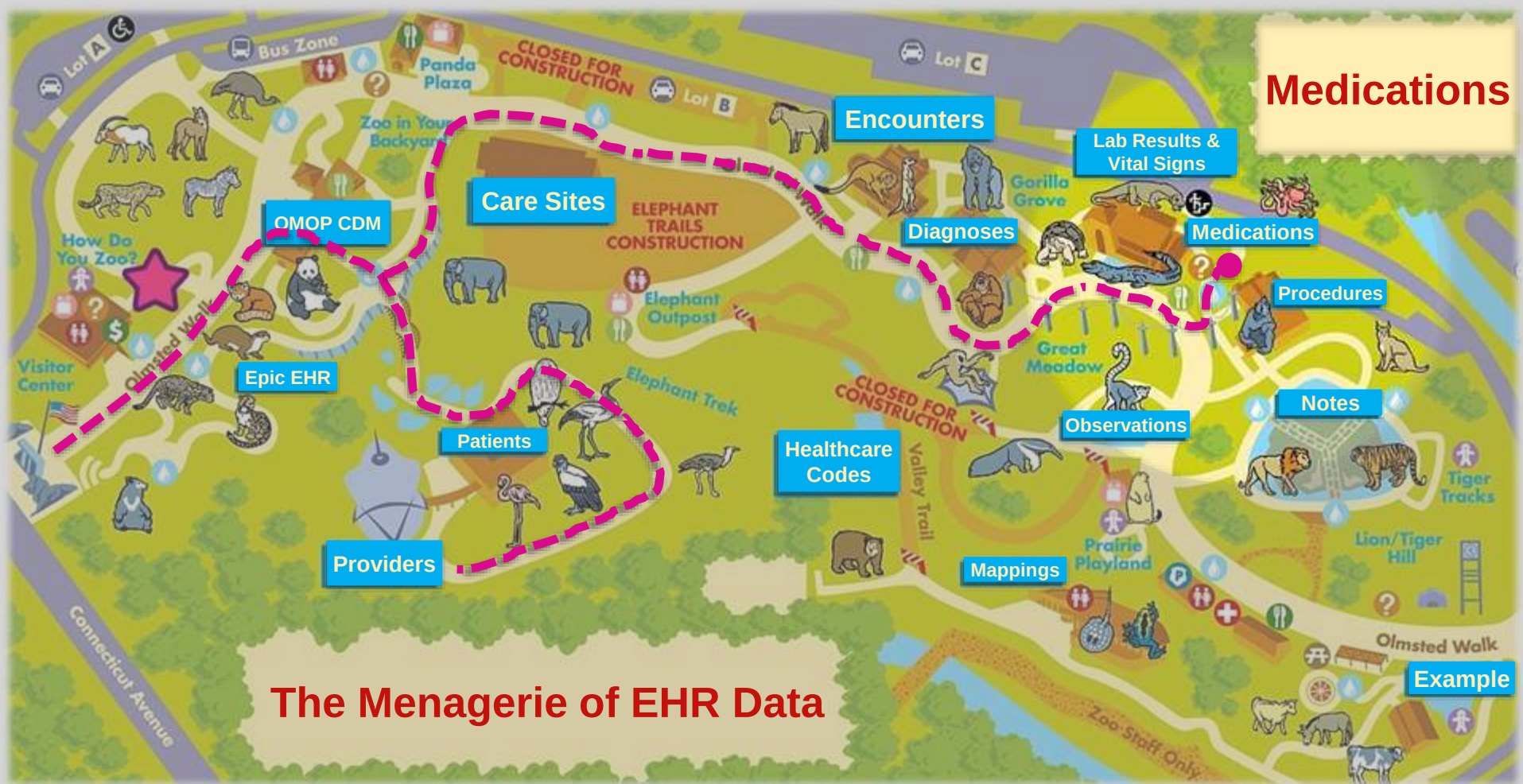


The Menagerie of EHR Data

Lab Results & Vital Signs (measurements)

Lab Order		Lab Component Result		measurement		Flowsheet	
Patient A	Lab Order 1	Lab Order 1	WBC Count	Patient A	WBC Count		
			RBC Count		RBC Count		
			Hct Result		Hct Result		
			Hgb Result		Hgb Result		
Patient B	Lab Order 2	Lab Order 2	Glucose Result	Patient B	Glucose Result		
			Calcium Result		Calcium Result		
			Sodium Result		Sodium Result		
			BUN Result		BUN Result		
				Patient C	Height	Patient C	Height
					Weight		
					Temp		
				Patient D	Metric 3	Patient D	Metric 3
				Patient E	Metric 4	Patient E	Metric 4

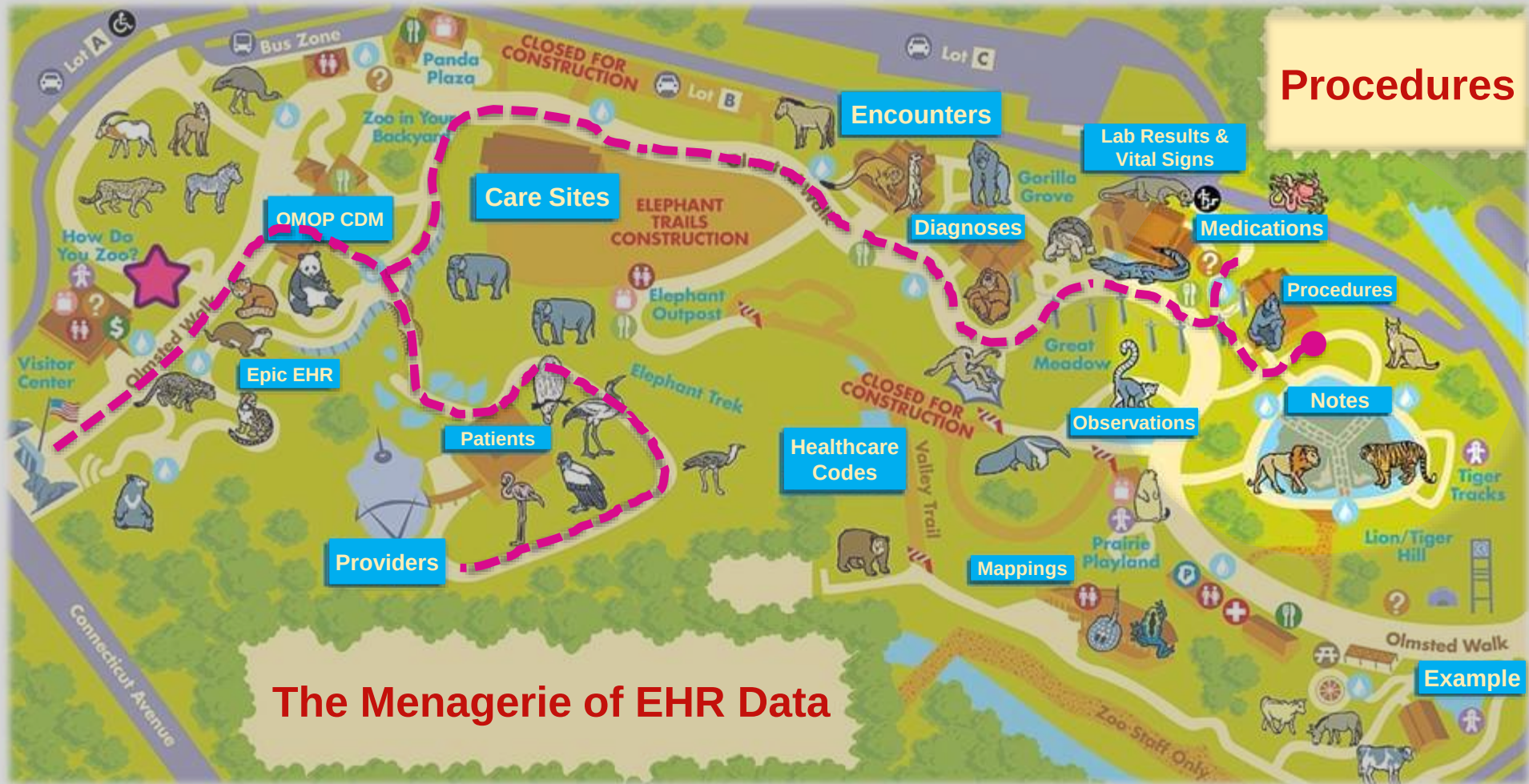
Medications



Medications

Medication Order		Medication Dispense		Medication Administration		Immunization Event		drug_exposure		
						Patient A	Vaccine Z	Patient A	Immuniz Z1	Vaccine Z
Patient B	Inpt Med Order 2					Inpt Med Order 2	Vaccine Y	Patient B	Inpt Med Order 2	Vaccine Y
Patient C	Inpt Med Order 3			Inpt Med Order 3	Med Admin 3-1			Patient C	Med Admin 3-1	Drug X
									Med Admin 3-2	Drug X
Patient D	Inpt Med Order 4	Inpt Med Order 4	Dispense 4-1	Inpt Med Order 4	Med Admin 4-1			Patient D	Med Admin 4-1	Drug W
			Dispense 4-2		Med Admin 4-2				Med Admin 4-2	Drug W
Patient E	Inpt Med Order 5	Inpt Med Order 5	Dispense 5-1					Patient E	Dispense 5-1	Drug V
Patient F	Outpt Med Order 6	Outpt Med Order 6	Dispense 6-1					Patient F	Dispense 6-1	Drug U
Patient G	Outpt Med Order 7	Outpt Med Order 7	Dispense 7-1					Patient G	Dispense 7-1	Drug T
			Dispense 7-2						Dispense 7-1	Drug T
Patient H	Outpt Med Order 8							Patient H	Outpt Med Order 8	Drug S

Procedures



The Menagerie of EHR Data

Procedures

Procedure

Procedure 1
Procedure 2
Procedure 3

Surgical Procedure

Surgical Procedure 4
Surgical Procedure 5

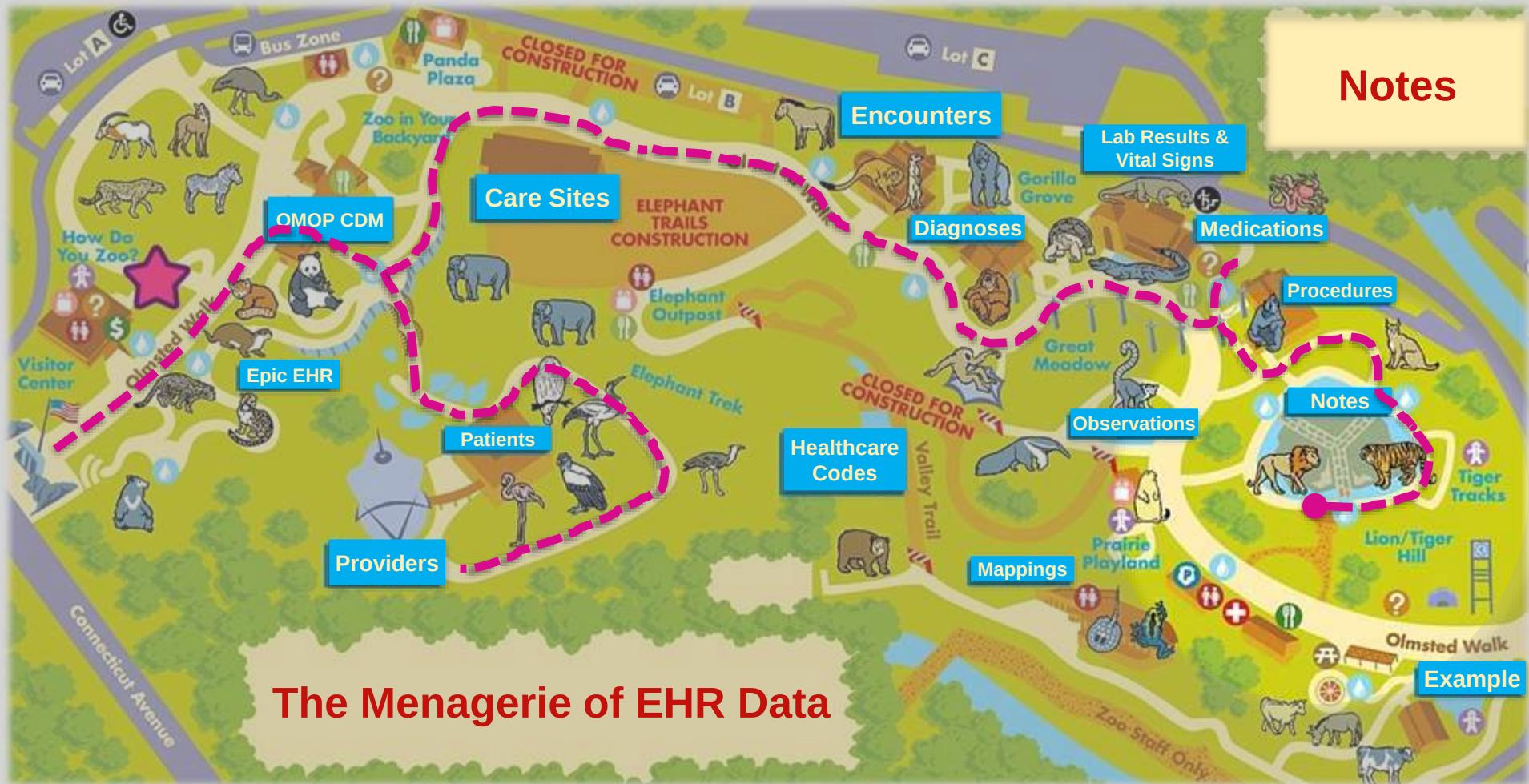
procedure_occurrence

Patient A	Procedure 1
Patient B	Procedure 2
	Procedure 3
Patient C	Surgical Procedure 4
Patient D	Surgical Procedure 5
Patient E	Inferred Procedure 6

Flowsheet

Patient E	Flowsheet Metric 6-1
	Flowsheet Metric 6-2
	Flowsheet Metric 6-3

Notes



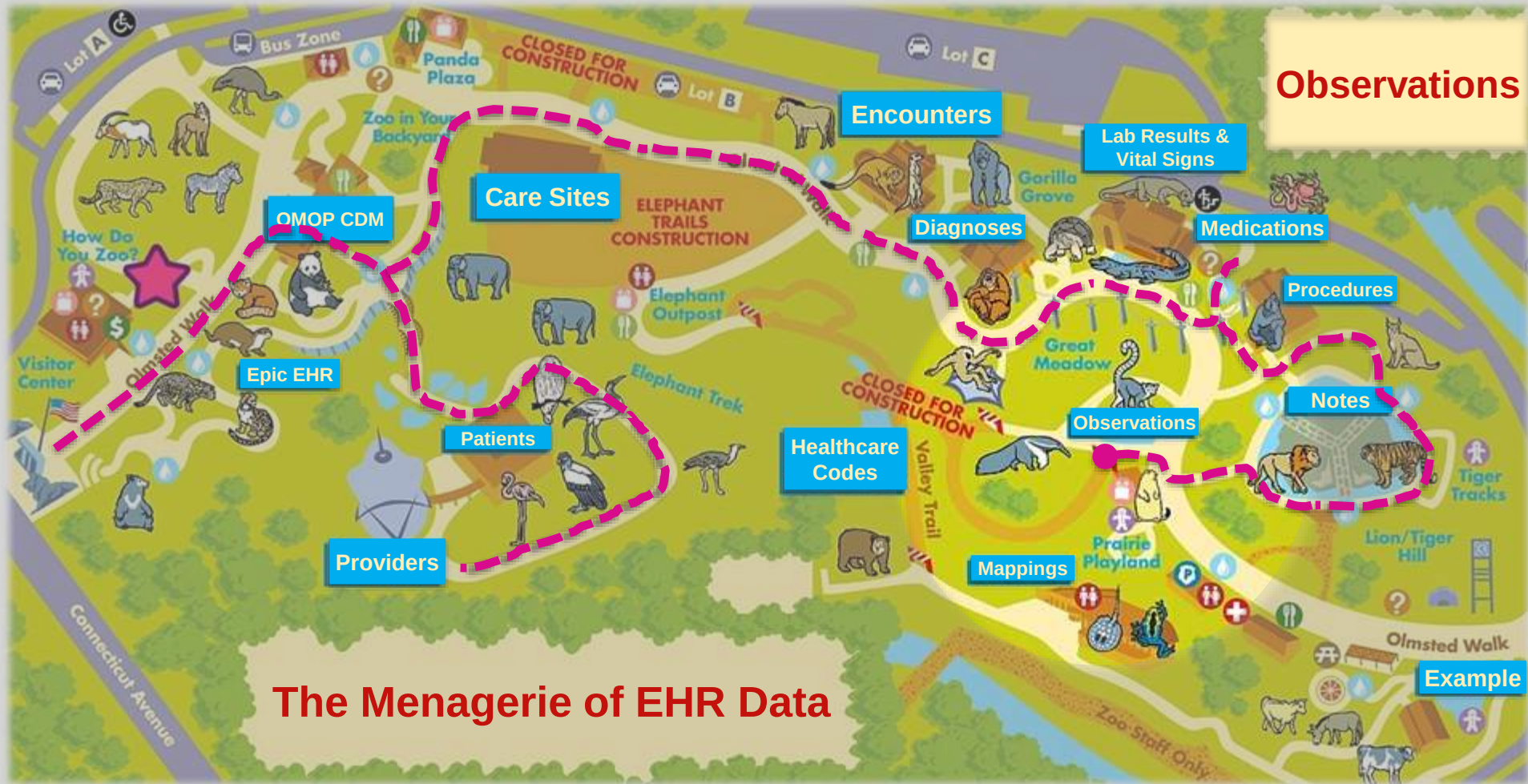
The Menagerie of EHR Data

Notes

Clinical Note	Lab Result Text	Imaging Study Text*	note
Patient A Clinic Note			Patient A Clinic Note
Patient B Progress Note			Patient B Progress Note
A&P Note			A&P Note
Procedure Report			Procedure Report
Periop Record			Periop Record
Consult Note			Consult Note
Discharge Summary			Discharge Summary
	Patient B Pathology Report		Pathology Report
	Lab Result Note		Lab Result Note
		Patient B Imaging Report Narrative	Imaging Report Narrative
		Imaging Report Impression	Imaging Report Impression
		Patient C Imaging Report Narrative	Patient C Imaging Report Narrative
		Imaging Report Impression	Imaging Report Impression
		Imaging Report Addenda	Imaging Report Addenda

****Under construction***

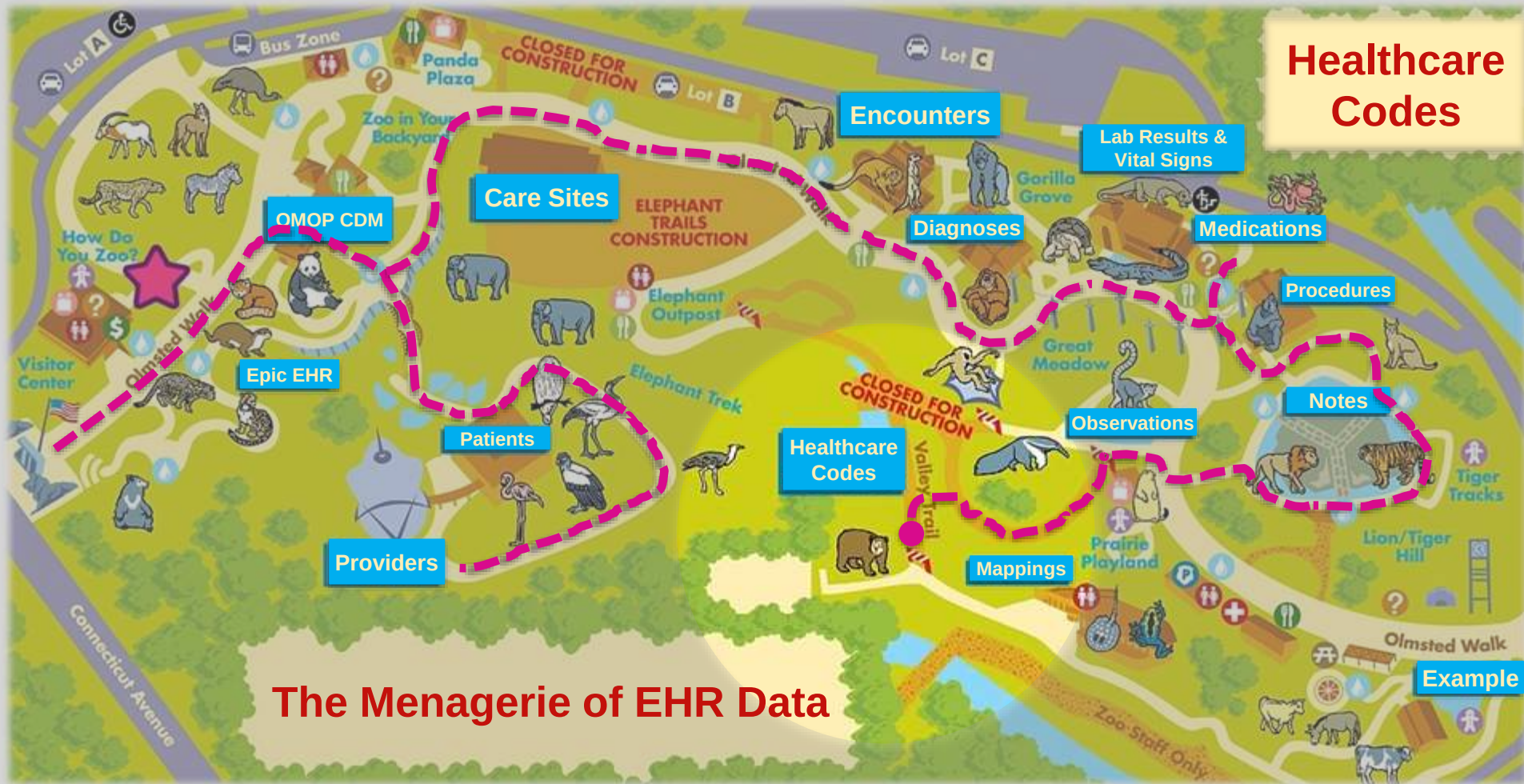
Observations



Observations

Patient Demographics		Surgical History		Social History		Family History		Patient Allergy		observation	
Patient A	Race									Patient A	Race
	Ethnicity										Ethnicity
	Language Preference										Language Preference
	Sexual Orientation										Sexual Orientation
Patient B	Race									Patient B	Race
	Marital Status										Marital Status
	Gender Identity										Gender Identity
Patient C	Ethnicity									Patient C	Ethnicity
	Religious Affiliation										Religious Affiliation
		Patient A	Procedure Z							Patient A	Procedure Z
			Procedure Y					Procedure Y			
		Patient B	Procedure X							Patient B	Procedure X
			Procedure W					Procedure W			
				Patient C	Soc Hx Item 1-1					Patient C	Soc Hx Item 1-1
			Soc Hx Item 1-2						Soc Hx Item 1-2		
			Soc Hx Item 1-3						Soc Hx Item 1-3		
						Patient D	Fam Hx 2-1			Patient D	Fam Hx 2-1
					Fam Hx 2-2				Fam Hx 2-2		
								Patient E	Allergy 3-1	Patient E	Allergy 3-1

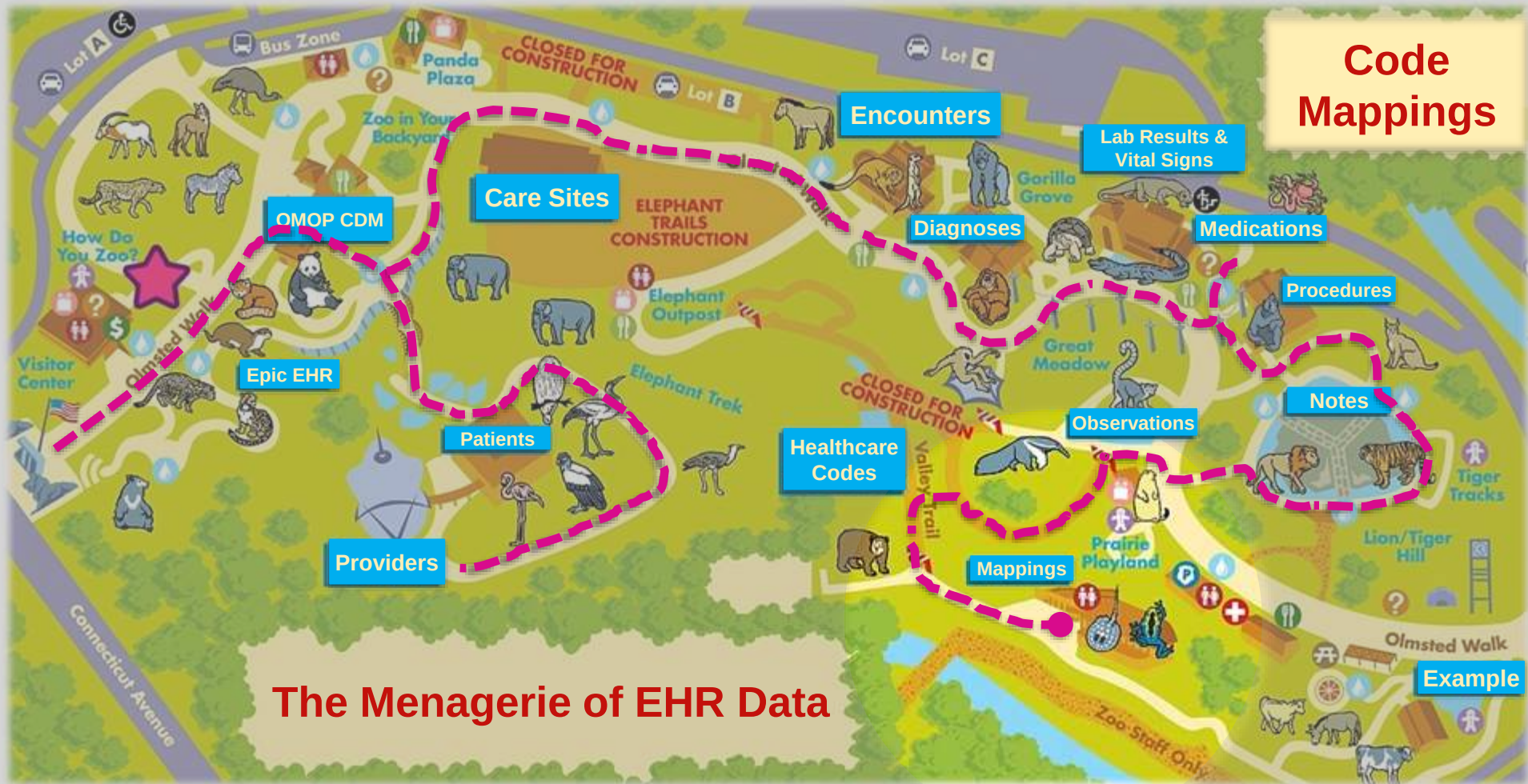
Healthcare Codes



Healthcare Vocabularies & Codes

vocabulary	concept_class	concept				domain
SNOMED	Std Class A-1	Concept 1	Std Class A-1	SNOMED	Diagnosis	Diagnosis
	Std Class A-2	Concept 2	Std Class A-2		Diagnosis	
	Std Class A-3	Concept 3	Std Class A-3		Observation	Observation
LOINC	Std Class B	Concept 4	Std Class B	LOINC	Measurement	Measurement
Epic Dx	Src Class C-1	Concept 2100100100	Src Class C-1	Epic Dx	Diagnosis	Diagnosis
	Src Class C-2	Concept 2200200200	Src Class C-2		Diagnosis	
MSDW Type	Src Class D	Concept 2300300300	Src Class D	MSDW Type	Type Concept	Type Concept

Code Mappings

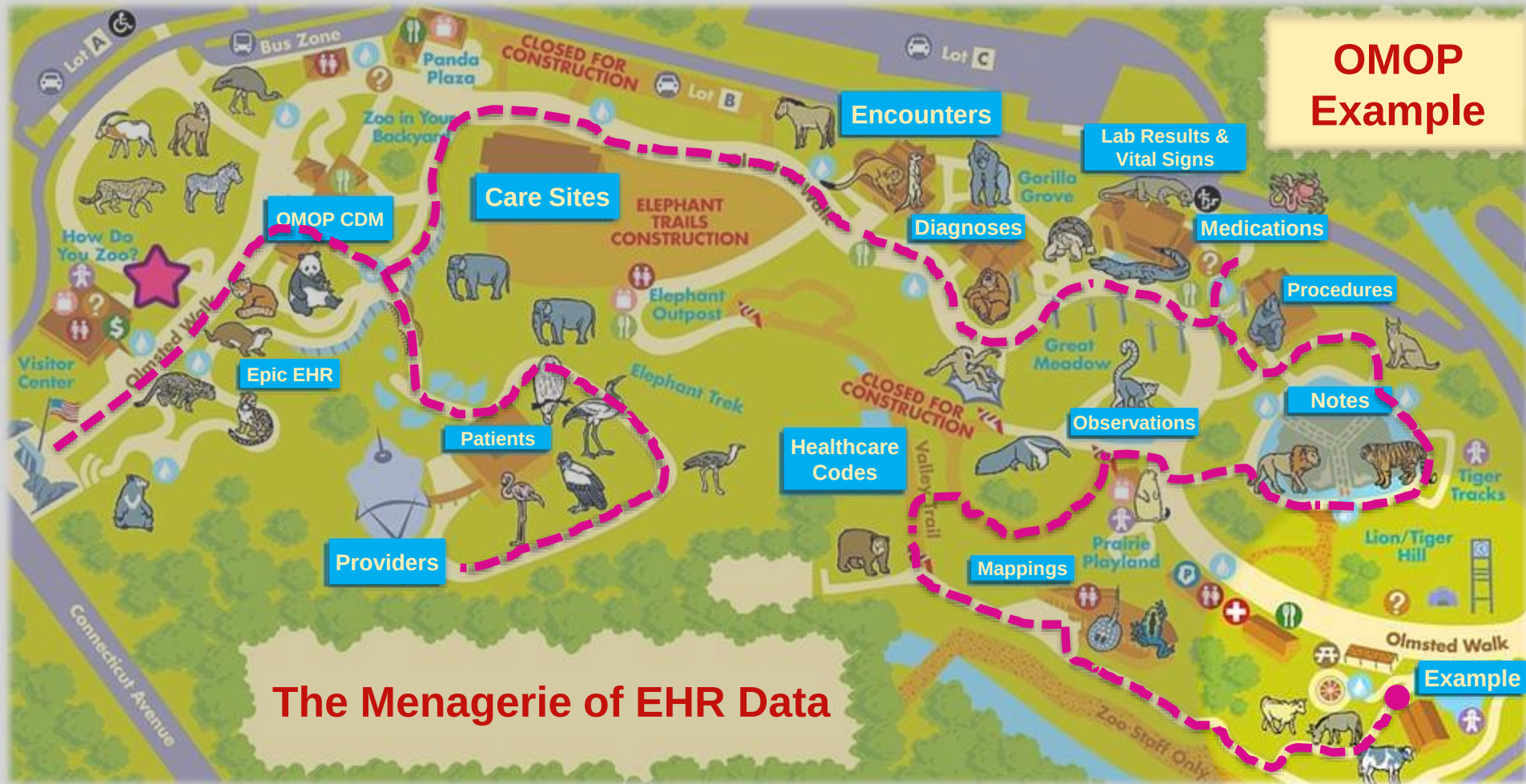


The Menagerie of EHR Data

Content Standardization via Concept Mapping

concept (1)		concept_relationship			concept (2)		relationship
Epic Dx	Concept 2100100100	Concept 2100100100	Maps to	Concept 1	Concept 1	SNOMED	Maps to
		Concept 2100100100	Maps to non-standard	Concept 300	Concept 300	ICD10CM	Maps to non-standard
	Concept 2200200200	Concept 2200200200	Maps to	Concept 2	Concept 2	SNOMED	Maps to
		Concept 2200200200	Maps to non-standard	Concept 400	Concept 400	ICD10CM	Maps to non-standard
SNOMED	Concept 1	Concept 1	Mapped from	Concept 2100100100	Concept 2100100100	Epic Dx	Mapped from
	Concept 2	Concept 2	Mapped from	Concept 2200200200	Concept 2200200200		
ICD10CM	Concept 300	Concept 300	Mapped from non-std	Concept 2100100100	Concept 2100100100		Mapped from non-std
	Concept 400	Concept 400	Mapped from non-std	Concept 2200200200	Concept 2200200200		

OMOP Example



The Menagerie of EHR Data

OMOP Data Example (*synthetic*)

condition_occurrence	Value	Definition
condition_type_concept_id	32827	OMOP's standard record type "EHR encounter record"
xtn_condition_type_source_concept_id	2000000129	MSDW's source record type "Encounter Diagnosis"
visit_occurrence_id	999888777	Unique identifier for the encounter
person_id	987654321	Unique identifier for the patient
provider_id	123456789	Unique identifier for the provider recording the diagnosis
condition_concept_id	4193704	OMOP's identifier for SNOMED code 313436004
condition_source_concept_id	2000602205	MSDW's identifier for Epic diagnosis code 521601
condition_start_date	1/1/2022	The date of condition onset or documentation per provider
condition_end_date	NULL	The date on which the condition resolved (if any)

concept_relationship	Value	Definition
concept_1	2000602205	MSDW's identifier for Epic diagnosis code 521601
relationship_id	Maps to non-standard	Text string denoting the type of mapping relationship between concept_id_1 & concept_id_2
concept_2	35206882	OMOP's identifier for ICD-10 code E11.9

OMOP Data Example (*synthetic*)

condition_occurrence	Value
condition_type_concept_id	32827
xtn_condition_type_source_concept_id	2000000129
visit_occurrence_id	999888777
person_id	987654321
provider_id	123456789
condition_concept_id	4193704
condition_source_concept_id	2000602205
condition_start_date	1/1/2022
condition_end_date	NULL

concept_relationship	Value
concept_1	2000602205
relationship_id	Maps to non-standard
concept_2	35206882

concept			
concept_id	vocabulary_id	concept_code	concept_name
32827	Type Concept	OMOP4976900	EHR encounter record
2000000129	MSDW Src Rec Type	Encounter Diagnosis	Encounter Diagnosis

4193704	SNOMED	313436004	Type 2 diabetes mellitus without complication
2000602205	EPIC EDG .1	521601	Type 2 diabetes mellitus without complications

2000602205	EPIC EDG .1	521601	Type 2 diabetes mellitus without complications
------------	-------------	--------	--

35206882	ICD10CM	E11.9	Type 2 diabetes mellitus without complications
----------	---------	-------	--

What's Coming Next?

MSDW Work in Progress

Under Construction (as of June 16, 2022)	Roadmap Ideas	OMOP Tables Not Used
• Admission, Discharge & Transfer (ADT) Transactions • Radiology Reports • Health Risk Assessments • PHQ-9 • Historical MRNs • Improved patient record linking • Integration with IRW 2.0 • Imaging study DICOM headers • Framework for Epic upgrades • Next Epic upgrade Nov 6, 2022 • Automated data quality checks • OMOP version 5.4 • Support for limited data set in data marts	• Social Security Death Index (SSDI) • Billing data from MSX • Pregnancy episodes • L&D data elements • Ultrasounds • Geocoding patient addresses • G-CDM genomics extension tables for variants • Digitized pathology slides	• specimen • device_exposure • note_nlp • episode • episode_event • cost • payer_plan_period • survey_conduct

Thank You!
For help and more information:
msdw.mountsinai.org

