



University of California
San Francisco

ZSFG

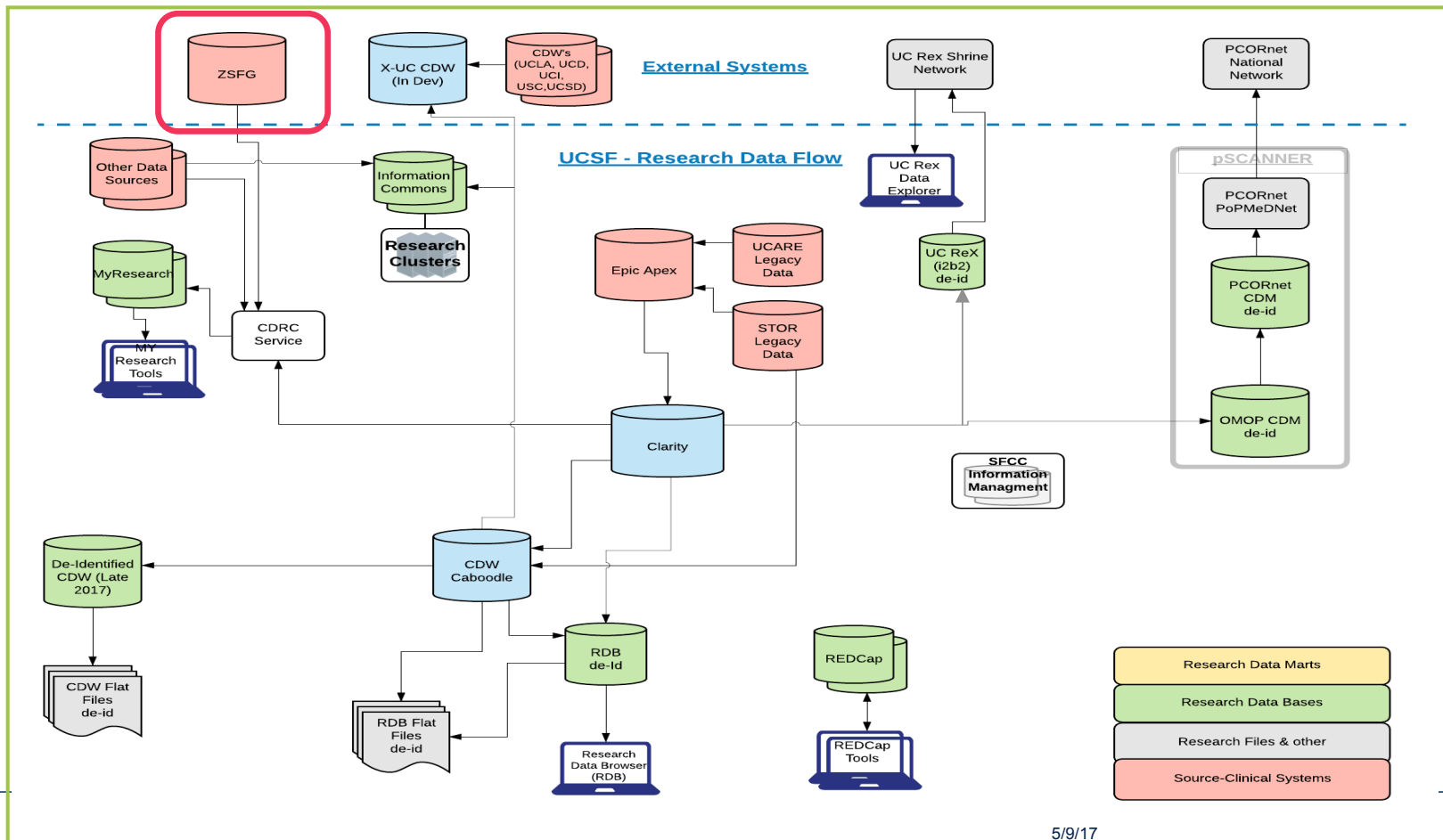
Rick Larsen
UCSF IT – Academic Research Systems (ARS)

Zuckerberg San Francisco General (ZSFG)

Scenario	Self-serve (Free)	Consult Required (May have recharge)	IRB needed	Requires MyResearch account or other secure environment	Includes clinical notes	UC Health data available in addition to UCSF data
Counts	Yes	No	No	No	No	Yes
De-ided data	Yes	No	No	No	No	Yes *
Limited data	No	Yes	Yes	Yes	No	Yes *
Id'ed data	No	Yes	Yes	Yes	Yes	Yes *
Recruit	No	Yes	Yes	Yes	No	Yes *

* Process to obtain data from UC Health sites can be very lengthy

ZSFG



ZSFG Data Sources Available for Research

- Invision (procedures, diagnoses, visits, demographics) 1990 ->
- LCR – Lifetime Clinical Record (labs, notes, meds) 1997 ->
- eCW – OP charting system (clinics – provides meds, notes, some labs) 2012 ->
- Jail Health- for SF inmate population
- Community Behavioral Health- includes the Avatar charting system
- Laguna Honda – long term stay facility



University of California
San Francisco

APeX and Clarity

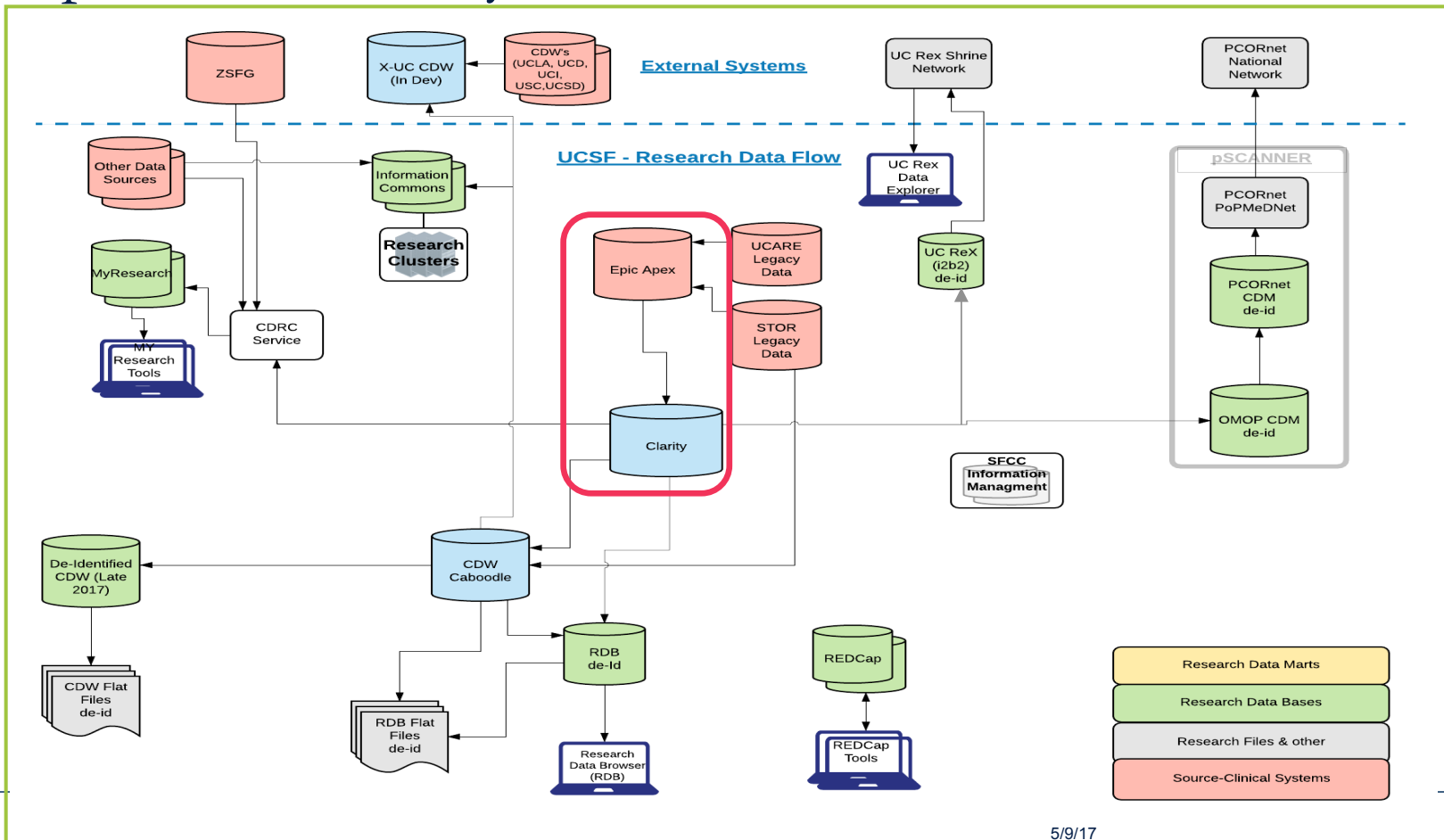
Dana Ludwig
Enterprise Information and Analytics

Epic APeX/Clarity

Scenario	Self-serve (Free)	Consult Required (May have recharge)	IRB needed	Requires MyResearch account or other secure environment	Includes clinical notes	UC Health data available in addition to UCSF data
Counts	Yes	No	No	No	No	Yes
De-ided data	Yes	No	No	No	No	Yes *
Limited data	No	Yes	Yes	Yes	No	Yes *
Id'ed data	No	Yes	Yes	Yes	Yes	Yes *
Recruit	No	Yes	Yes	Yes	No	Yes *

* Process to obtain data from UC Health sites can be very lengthy

Epic APeX/Clarity



APeX - A Suite of Medical Center Management Software Applications

- Licensed from Epic Systems
- Operational System for:
 - Patient Care - a comprehensive electronic medical record
 - Patient Billing
 - Lab, pharmacy, radiology, Operating Room, Emergency Department
 - ADT, Materials Management
- Went live around 6/1/2012, though staggered by departments over time
- Data Storage
 - Chronicles - a hierarchical time-oriented database created by Epic
 - Written in Intersystems Cache = MUMPS
 - As of 2010:
 - 772 Master Files - more like a tree than table
 - 109,982 Items - like fields or column

APeX - Scope of UCSF Data

- Includes
 - Since 6/1/2012 - EHR data - most everything that is known about the patient as text, both structured and unstructured.
 - Before 6/1/2012 - Most labs and clinical notes - migrated from legacy systems
 - Some EHR documents from outside UCSF are scanned PDF's and not stored as text.
 - Extensive user-defined data:
 - Flowsheets
 - Smartdata
- Excludes
 - Diagnoses, medications, procedures prior to 6/1/2012
 - Images - contains hyperlinks to images in PACS image servers
 - Biometrics from ICU devices - on separate systems

Can we get the data directly from APeX?

- Major challenges
 - Querying Chronicles could slow down patient care
 - Would require MUMPS programming
- Clarity - a better data source for most research applications

Clarity

- Relational database extract of Chronicles data
- Updated on a nightly basis
- 12,531 Tables
- 106,283 Columns
- Most APeX clinical data can be found in Clarity
 - Excludes a few things like mouse click logs
- Challenges of Clarity:
 - No one person, even at Epic, understands all the data
 - Many clinical facts are in unstructured clinical notes
 - Epic Clarity documentation does not document UCSF specific uses of the data
 - Forest and Trees problem - too many tables to find what you want.
 - HIPAA privacy rules for research make data governance complicated

Web-based Data Documentation - by Epic

The screenshot shows the Epic web-based data documentation interface. The top navigation bar is green and contains the following items: Home, Dictionary, Dictionary Editor, Executions, Work Queue (99+), Configuration, a search icon, and Star Development with CDW-production copy. The left sidebar is titled "Filter components..." and lists various components, with "FlowsheetValueFact" selected. The main content area displays the details for "FlowsheetValueFact".

FlowsheetValueFact

The flowsheet value fact contains information about measurements from flowsheets. Each row represents a flowsheet value. [+ Notes](#)

Properties

Granularity	A flowsheet value
Profile tables	FlowsheetValueProfileDim
Links to	DateDim , EmployeeDim , EncounterFact , FlowsheetRowDim , FlowsheetTemplateDim , FlowsheetValueProfileDim , PatientDim , TimeOfDayDim
Links from	none
Change type	Type 1

Sources

- ▶ **Non-Epic**
Unspecified custom extracts

Columns [Expand All](#)

- ▶ **FlowsheetValueKey** : 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 f...
- ▶ **Id** : string
This column should no longer be used and will be removed in a future version. Source ID for th