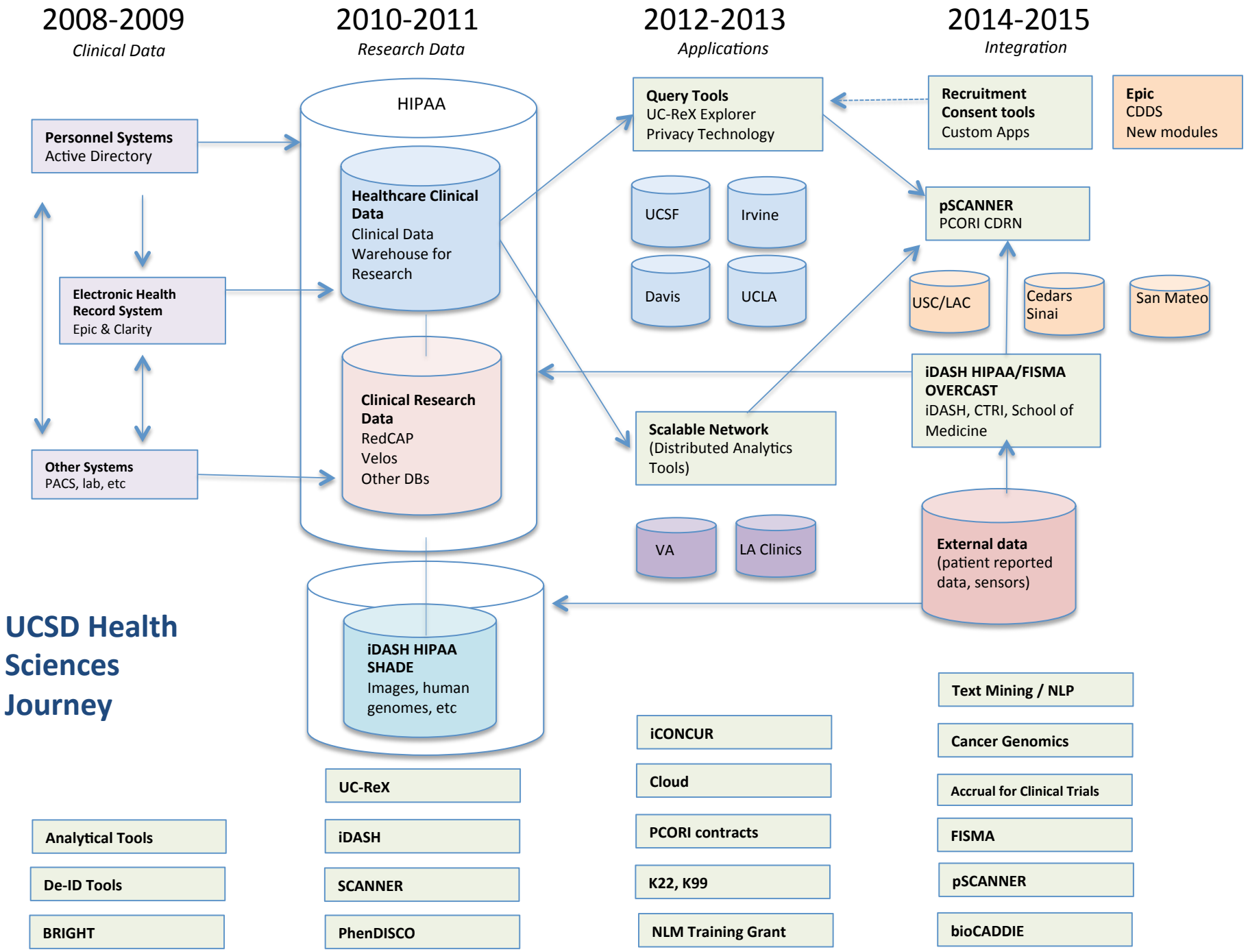
iDASH HIPAA-Cloud

Sharing and Protecting Human Subjects Data

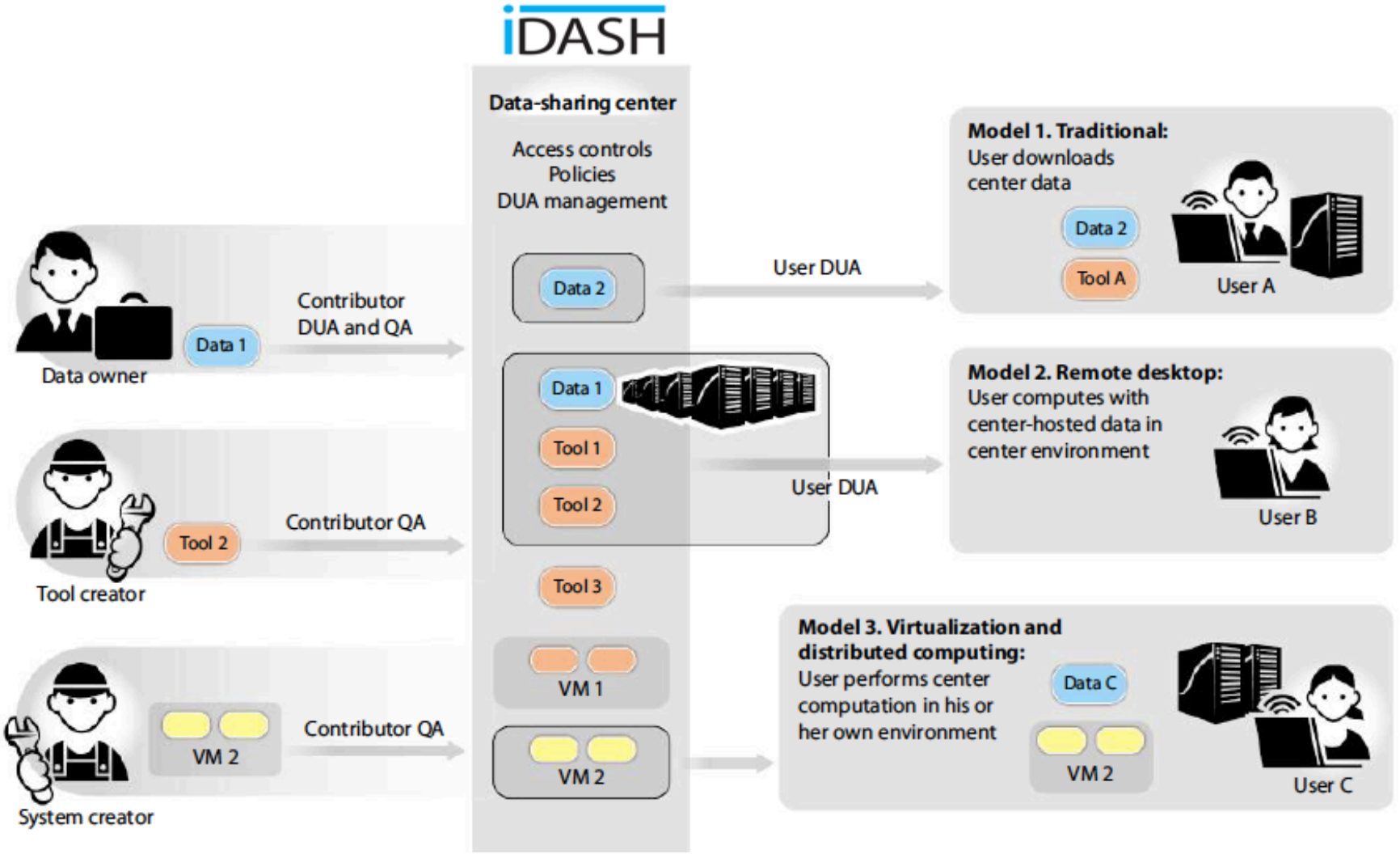
May 2015
IDI presentation

NIH U54 HL108460

Lucila Ohno-Machado, MD, MBA, PhD
Biomedical Informatics, University of California San Diego


2008-2009
Clinical Data
2010-2011
Research Data
2012-2013
Applications
2014-2015
Integration
HIPAA
Personnel Systems
Active Directory
Electronic Health Record System
Epic & Clarity
Other Systems
PACS, lab, etc
Healthcare Clinical Data
Clinical Data Warehouse for Research
Clinical Research Data
RedCAP
Velos
Other DBs
iDASH HIPAA SHADE
Images, human genomes, etc
Query Tools
UC-ReX Explorer
Privacy Technology
UCSF
Irvine
Davis
UCLA
Scalable Network
(Distributed Analytics Tools)
VA
LA Clinics
Recruitment Consent tools
Custom Apps
Epic
CDDS
New modules
pSCANNER
PCORI CDRN
USC/LAC
Cedars Sinai
San Mateo
iDASH HIPAA/FISMA OVERCAST
iDASH, CTRI, School of Medicine
External data
(patient reported data, sensors)
UCSD Health Sciences Journey
Analytical Tools
De-ID Tools
BRIGHT
UC-ReX
iDASH
SCANNER
PhenDISCO
iCONCUR
Cloud
PCORI contracts
K22, K99
NLM Training Grant
Text Mining / NLP
Cancer Genomics
Accrual for Clinical Trials
FISMA
pSCANNER
bioCADDIE

Models for 'Sharing' Data



Ohno-Machado L. To Share or Not To Share: That Is Not the Question. Science Translational Medicine, 2012 4(165)

Supported by the NIH Grant U54 HL108460 to the University of California, San Diego

UC ReX
University of California Research eXchange

Clinical Data Network – UC-ReX

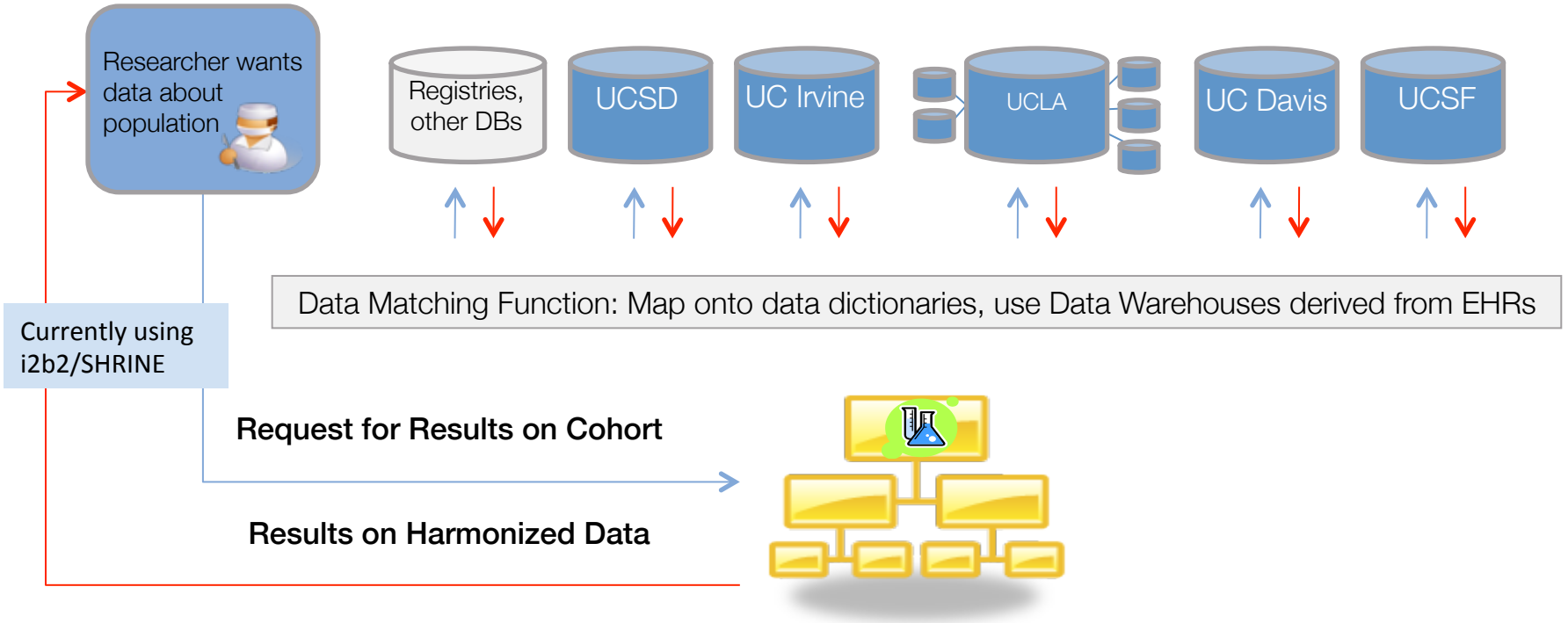
- Clinical Data Warehouses from 5 Medical Centers and affiliated institutions exchange (>13 million patients)
  - » Translational research
  - » Patient safety surveillance
  - » Quality improvement

UC-ReX was formed in 2010

Funded by the UC Office of the President

UC ReX
University of California Research eXchange

UC-Research eXchange

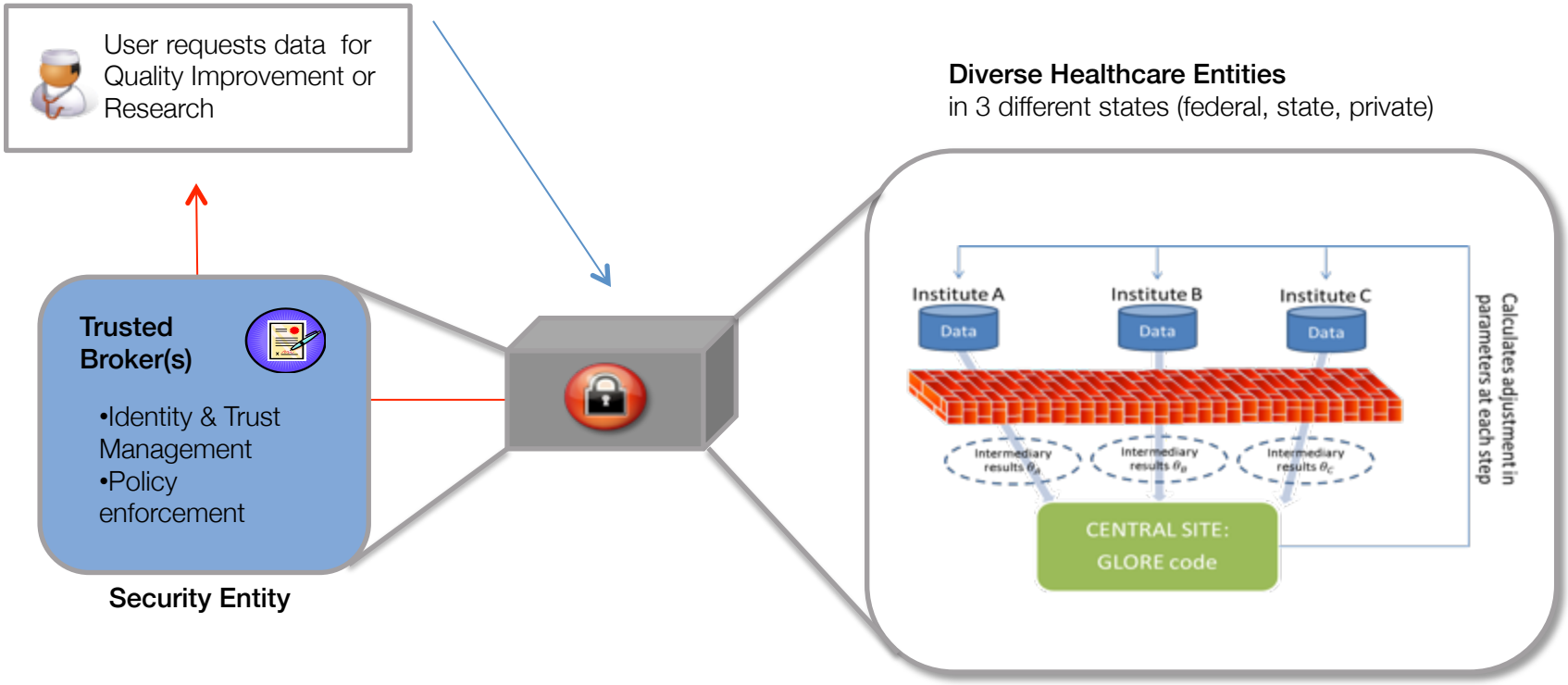


- Currently
  - Count queries (i.e., how many >65yo with afib taking dabigatran had a stroke?)
  - Demographics, meds, labs, diagnoses.Data quality checks
- Future steps
  - Individual-level patient data delivery (recruitment)
  - Cross-institutional analytics

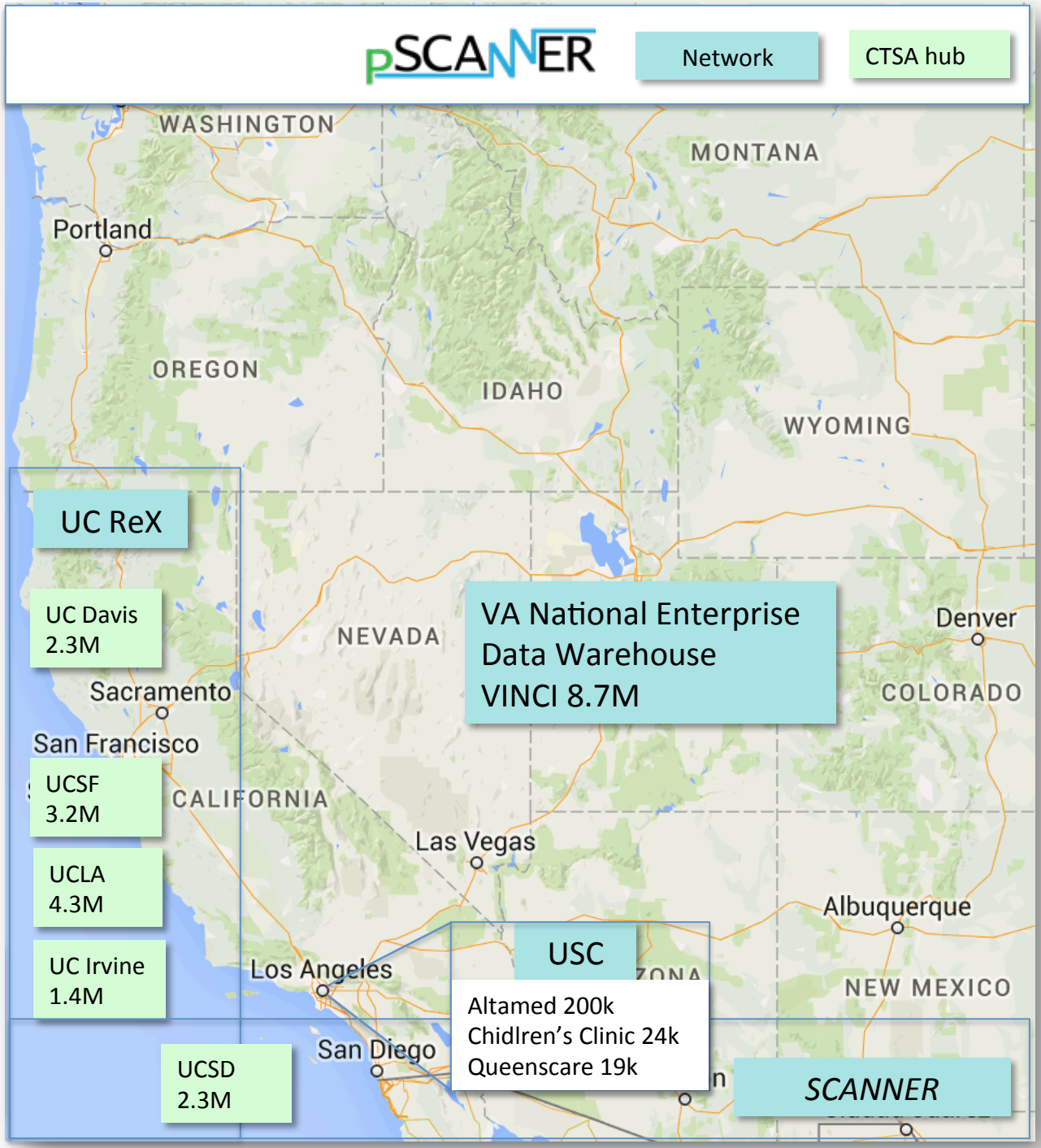
Funded by the UC Office of the President

Distributed computing

- Predictive modeling and adjustment for cofounders require lots of data
- Some institutions cannot move data outside their firewalls, we can bring computation to the data



Wu Y et al. Grid Binary LOgistic REgression (GLORE): Building Shared Models Without Sharing Data. *JAMIA*, 2012
Wang S et al. EXpectation Propagation LOgistic REgRession (EXPLORER): Distributed Privacy-Preserving Online Model Learning. *J Biomed Inf* 2013
Jiang W et al.. WebGLORE: A Webservice for Grid Logistic Regression. *Bioinformatics* 2014
Wu Y et al. Grid Multi-Category Response Logistic Models. *BMC Med Inform Dec Making* 2015



Phase 1

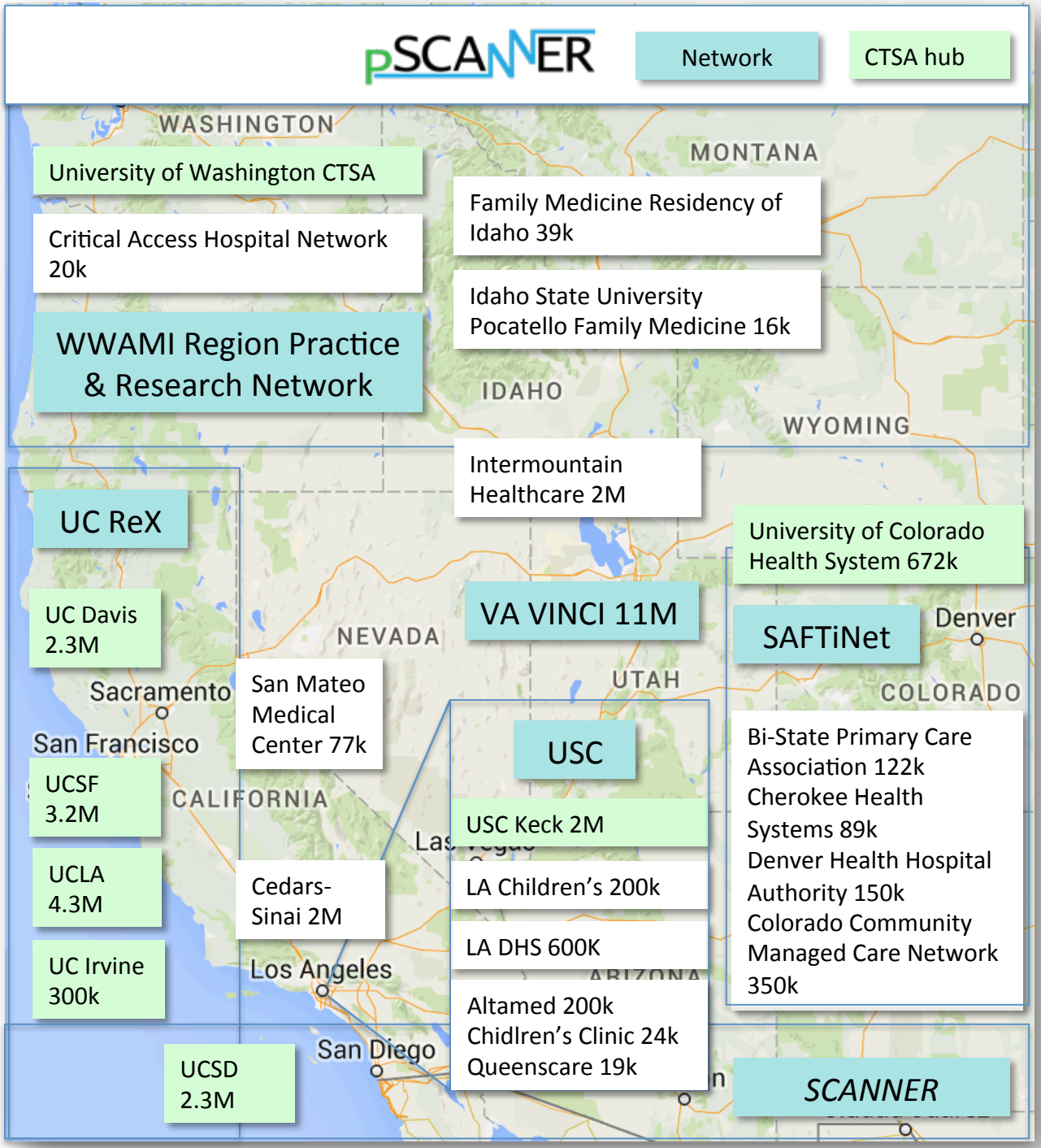
funded by pcori

21 Million people

Standardized data

Data governance

9 health systems



Phase 2

pcori

31 Million people

Same standards

Data governance

23 health systems

University of Texas Houston (NLP)

Health Information Privacy

HIPAA
- Understanding HIPAA Privacy
- HIPAA Administrative Simplification Statute and Rules
  - Omnibus HIPAA Rulemaking
  - Statute
  - Privacy Rule
  - Security Rule
  - Breach Notification Rule
  - Other Administrative Simplification Rules
  - Enforcement Rule
  - Combined Text of All Rules
- Enforcement Activities & Results
- How to File a Complaint
- News Archive
- Frequently Asked

Breaches Affecting 500 or More Individuals

As required by section 13402(e)(4) of the HITECH Act, the Secretary must post a list of breaches of unsecured protected health information affecting 500 or more individuals. These breaches are now posted in a new, more accessible format that allows users to search and sort the posted breaches. Additionally, this new format includes brief summaries of the breach cases that OCR has investigated and closed, as well as the names of private practice providers who have reported breaches of unsecured protected health information to the Secretary. The following breaches have been reported to the Secretary:

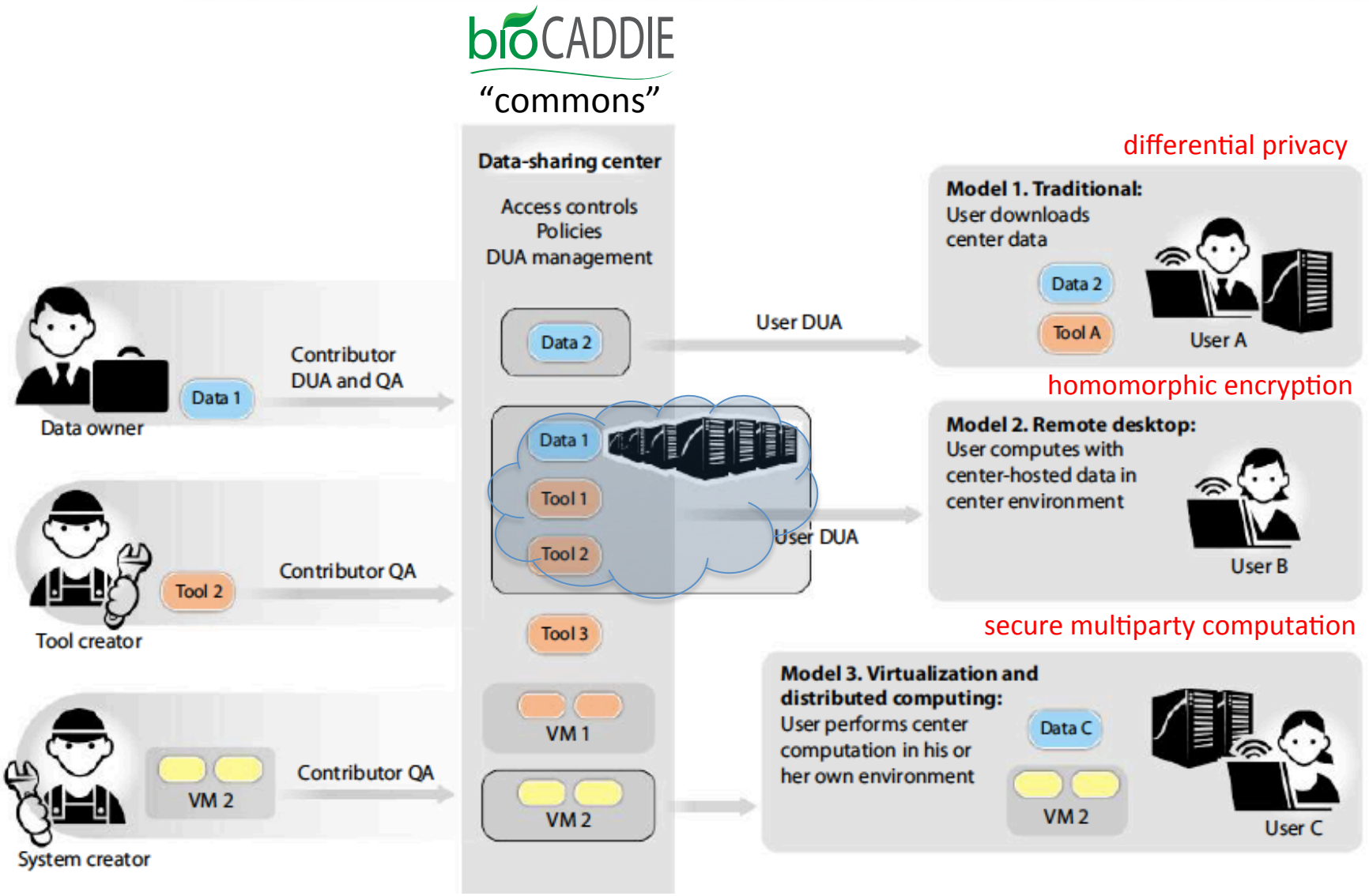
Full DataSet CSV format (18 KB) XML format (57 KB)

Select a column head to sort by that column. Select again to reverse the sort order. Select an individual record to display it in full below the table.

Filter: 556 records showing

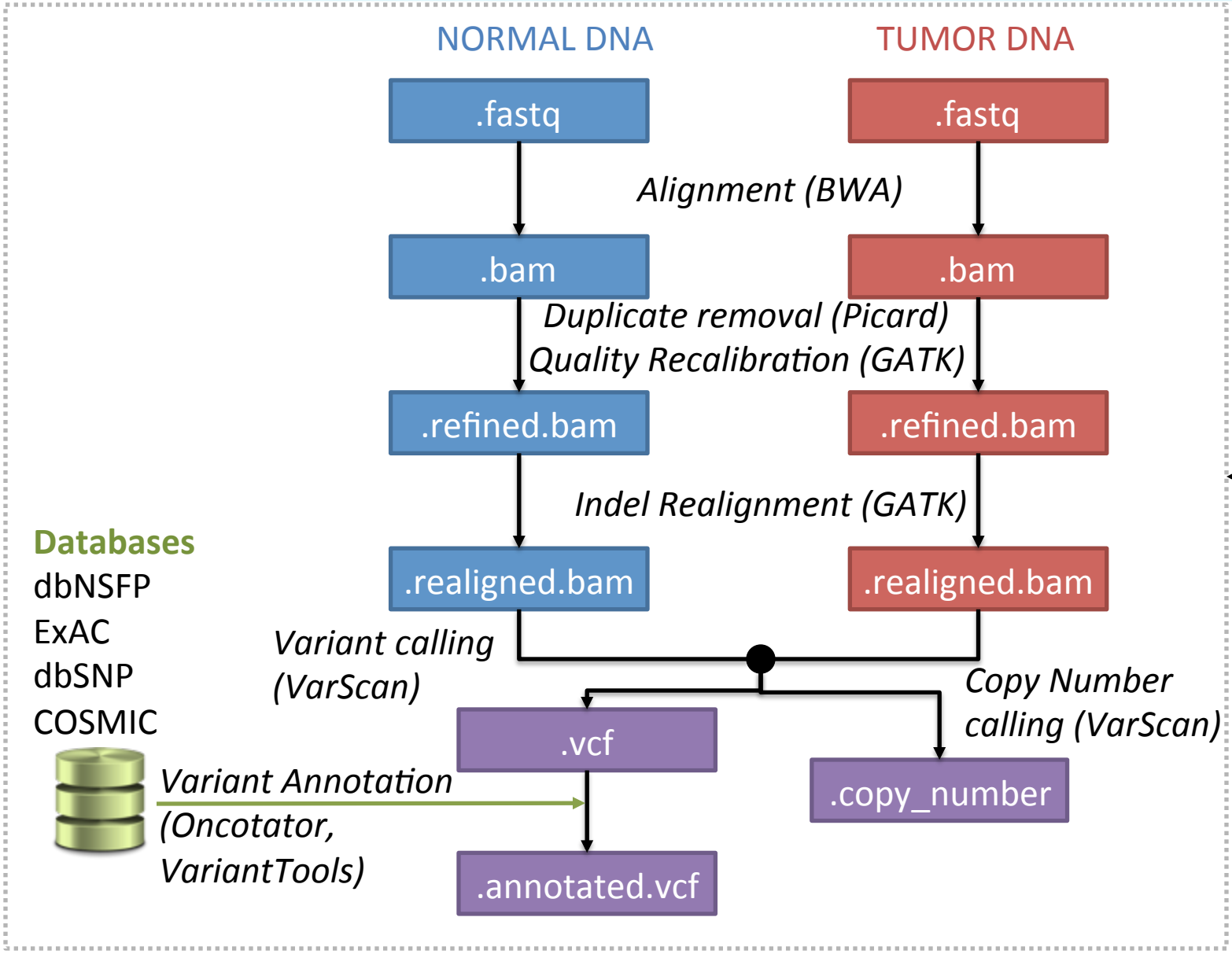
Name of Covered Entity	State	Individuals Affected	Date of Breach	Type of Breach	Location of Breached Info
1st response Medical Transpot Corp.	MD	552	06/15/2012-10/01/2012	Unauthorized Access/Disclosure	Desktop Computer
ABQ HealthPartners	NM	778	2012-12-20	Theft	Laptop
Accendo	AZ	175,350	2011-01-01	Unauthorized Access/Disclosure	Paper
Access Medical Group	PR	7,606	2012-01-11	Theft	Laptop

NIH Big Data to Knowledge



Ohno-Machado L. To Share or Not To Share: That Is Not the Question. Science Translational Medicine, 2012 4(165)

Whole Exome Pipeline Workflow

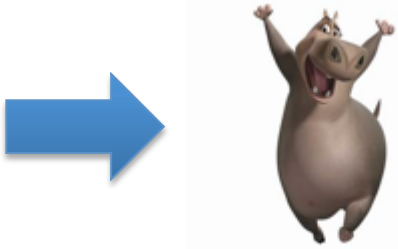


Supported by the NIH Grant U54 HL108460 to the University of California, San Diego

Genomes are Biometrics

Biometrics are
Protected Health Information (PHI)

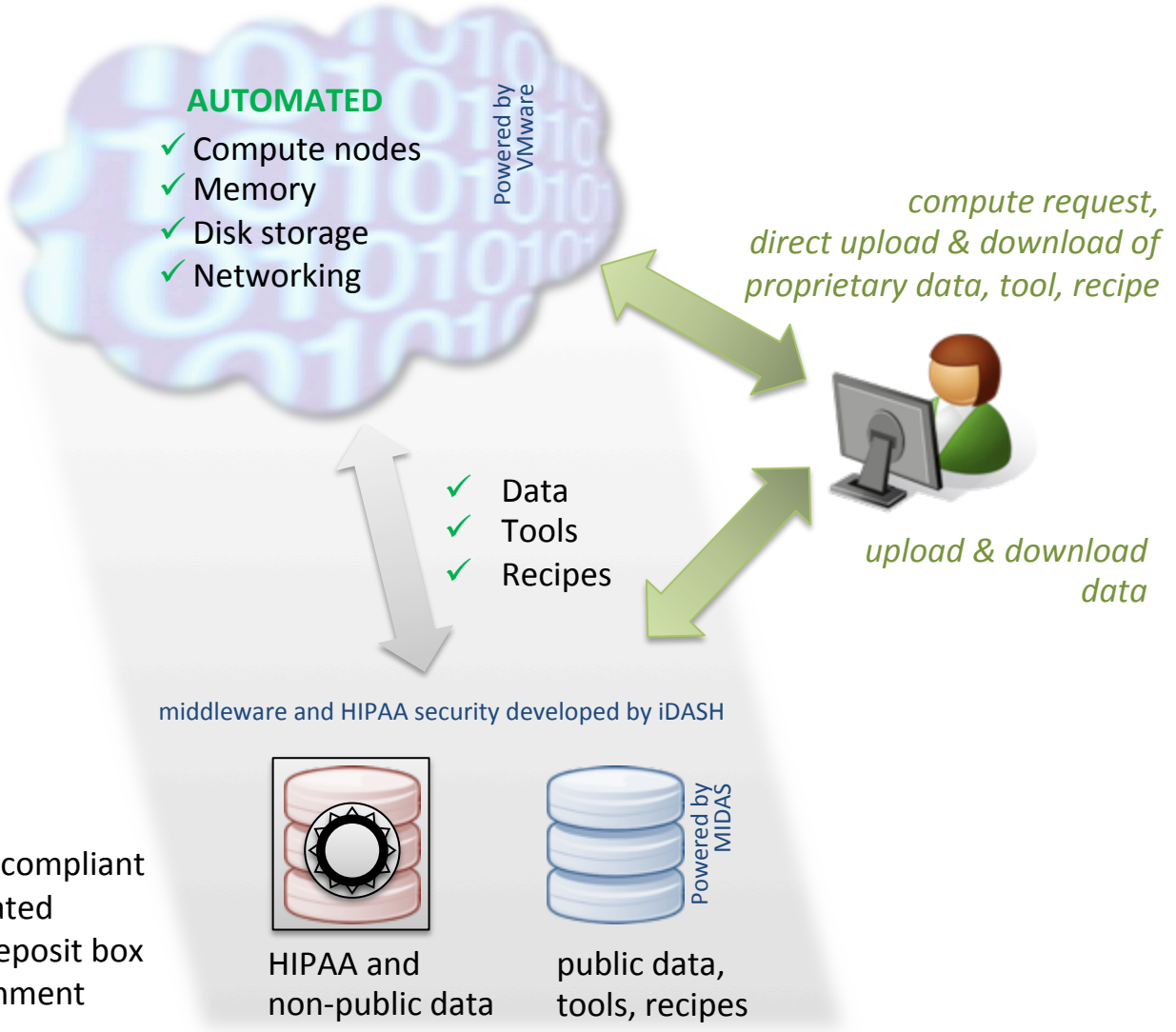
PHI requires HIPAA



- Biometrics require HIPAA

iDASH On-Demand Resources

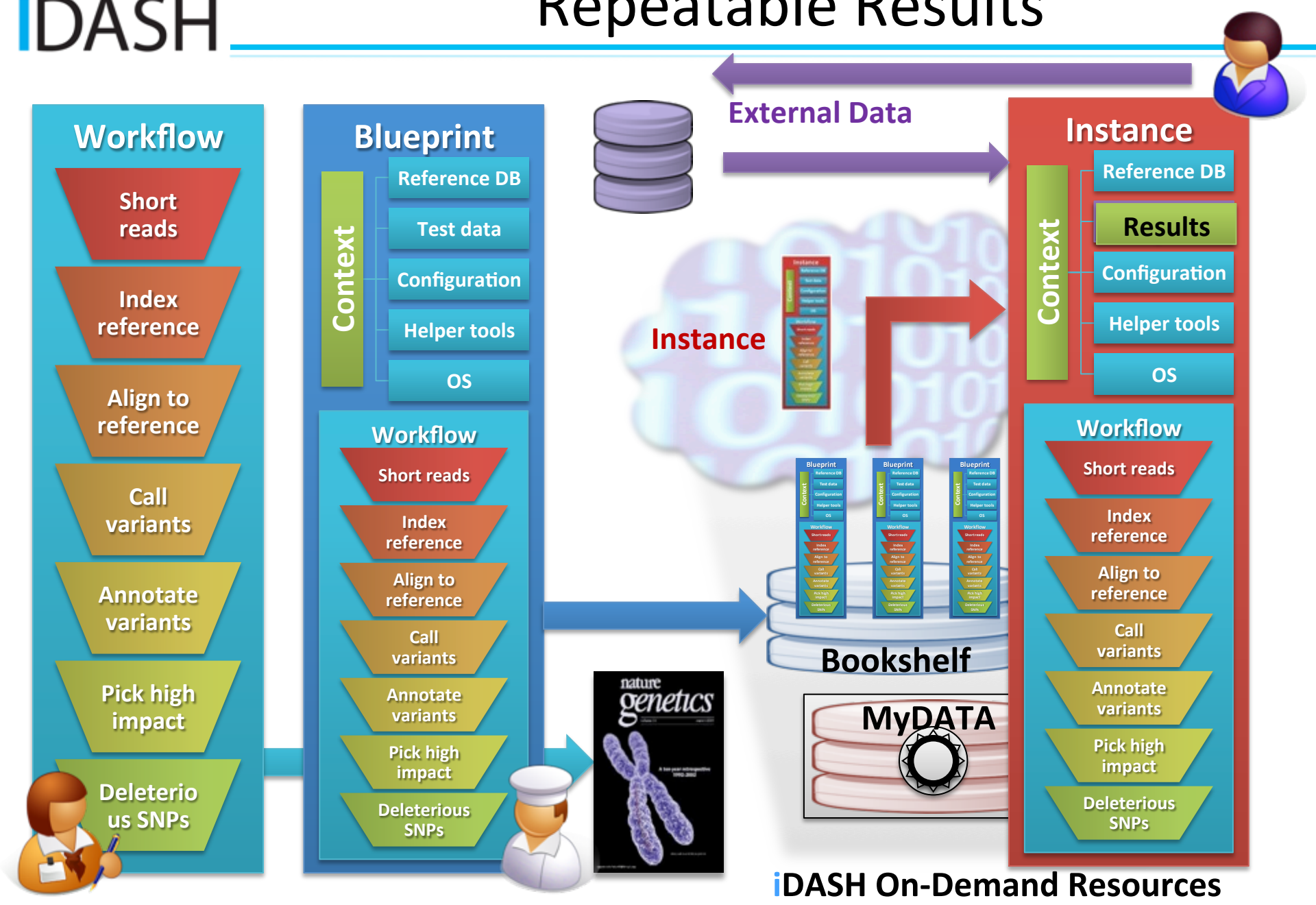
On-demand
Virtualized
Elastic
Resilient
Compute
And
Storage
Technology



Safe
HIPAA-compliant
Annotated
Data deposit box
Environment

Supported by the NIH Grant U54 HL108460 to the University of California, San Diego

Repeatable Results



iDASH On-Demand Resources

Supported by the NIH Grant U54 HL108460 to the University of California, San Diego

Plans

- Scale up the infrastructure (cores, disk, network)
  - » the plan was shared with IDI and there are many ways to collaborate and mutually benefit from a common vision
- Serve HIPAA-storage and compute needs of a larger UCSD community
  - » Expand library of workflows/pipelines
  - » FISMA ATO
- Hub infrastructure for large HIPAA-data networks
  - » Continue data extraction services and privacy technology research