

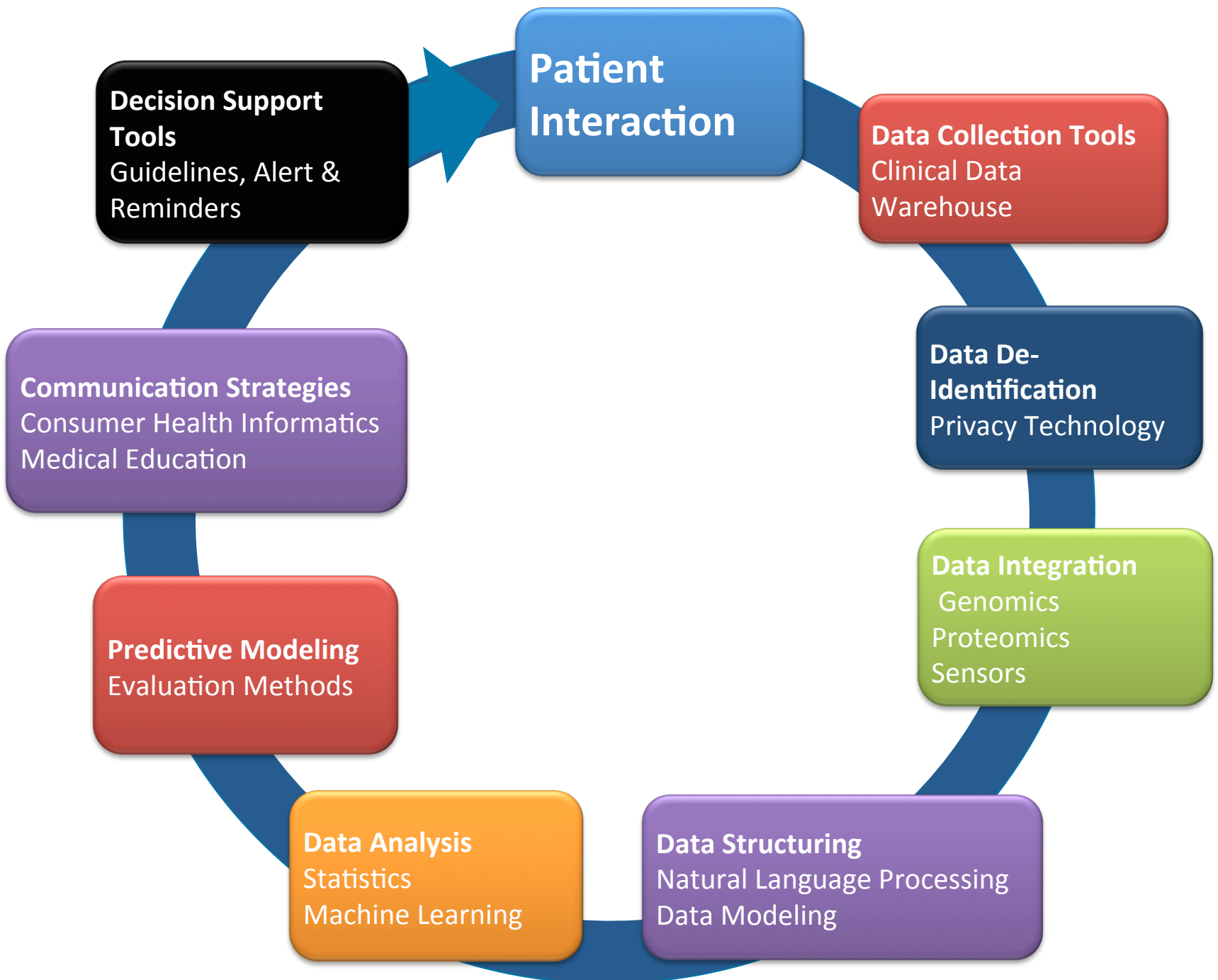
Patient Privacy and Research on Genomes

March 16, 2015 - La Jolla, CA

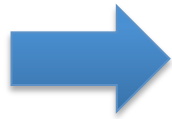


Lucila Ohno-Machado, MD, PhD

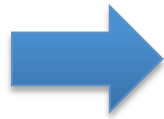
funded by NIH U54 HL108460, R01HG007078, K99HG008175



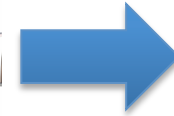
-
- Personalized medicine
- Personal genomics
- Family history data
- Pharmacogenomics
- EHRs



Biometrics are
Protected Health Information (PHI)



PHI requires HIPAA
protection



- Biometrics require HIPAA protection

- Security
 - » Risk of disclosure
 - » Liability
- Privacy
 - » Risk of re-identification
 - » Informed consent
- Presidential Commission for the Study of Bioethical Issues. Privacy and progress in whole genome sequencing. <http://www.bioethics.gov> covers genome sequencing, exome sequencing, genome-wide SNV analysis, and data from large scale genomic studies

HIPAA 'De-identified' data

- removal of 18 identifiers, such as dates, biometrics, names, etc.
- expert certification of low risk of re-identification
- 'Limited' data sets have dates

Health Information Privacy

[Office for Civil Rights](#)[Civil Rights](#)[Health Information Privacy](#)[OCR Home](#) > [Health Information Privacy](#) > [HIPAA Administrative Simplification Statute and Rules](#) > [Breach Notification Rule](#)

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



1M people longitudinal cohort
will want privacy protection

Lots of data to protect

EHR Group recommendations

- Patient participation: Blue button rights (Synch for Science)
- Patient portals: patient-reported data
- Security/privacy of EHR systems

- Genomics, imaging, sensors
- EHR data already harmonized to national standards
- Point of care contact via EHR, mHealth
- Health gaming, cloud, privacy
- Biorepositories

- Imagine a public data set of genomes about a certain disease (e.g., Alzheimer's)
- Can you find out if your neighbor has Alzheimer's?
- Same problem happens with clinical data, which can be re-identified to a target patient
 - » Dates of visits
 - » Combination of patient characteristics

- High dimensional data (e.g., whole genomic sequencing)
- Privacy rules (PHI, including genomic data need to be protected under HIPAA)
- Institutional processes (internal review boards)
- Patient information consent
- Clinical-genomic data research networks
 - » Secure multiparty computation
 - » Homomorphic encryption

Jiang X, Sarwate A, Ohno-Machado L, Privacy technology to support data sharing for comparative effectiveness research: A systematic review, *Medical Care*, 51(8 Suppl 3):S58-65, 2013. PMID: 23774511

adapted from Dr. Xiaoqian Jiang

- Share access to data and computation
- Train the new generation of data scientists
- Provide innovative software, platform, and infrastructure
- Protect privacy

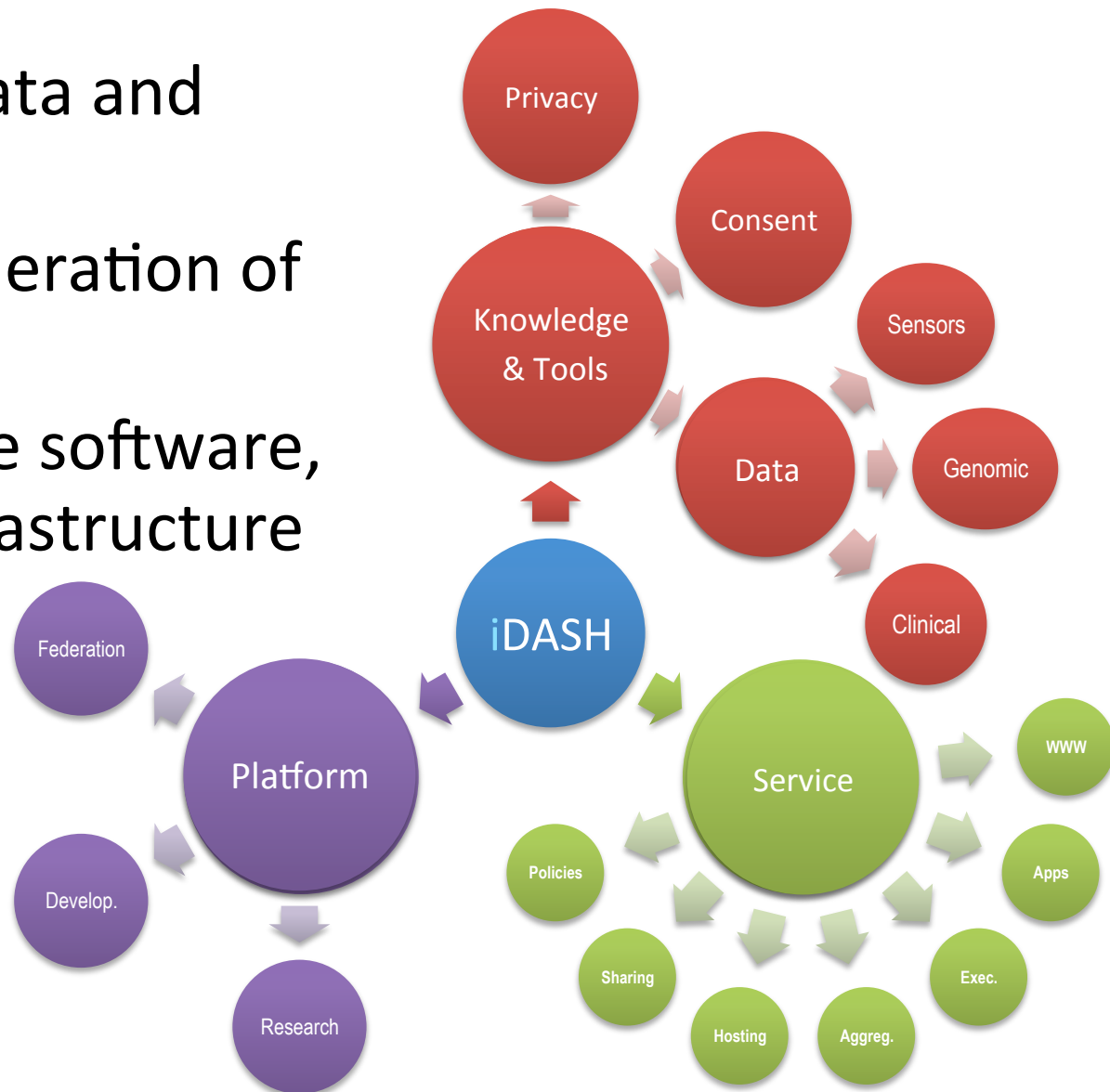
Develop

» Algorithms

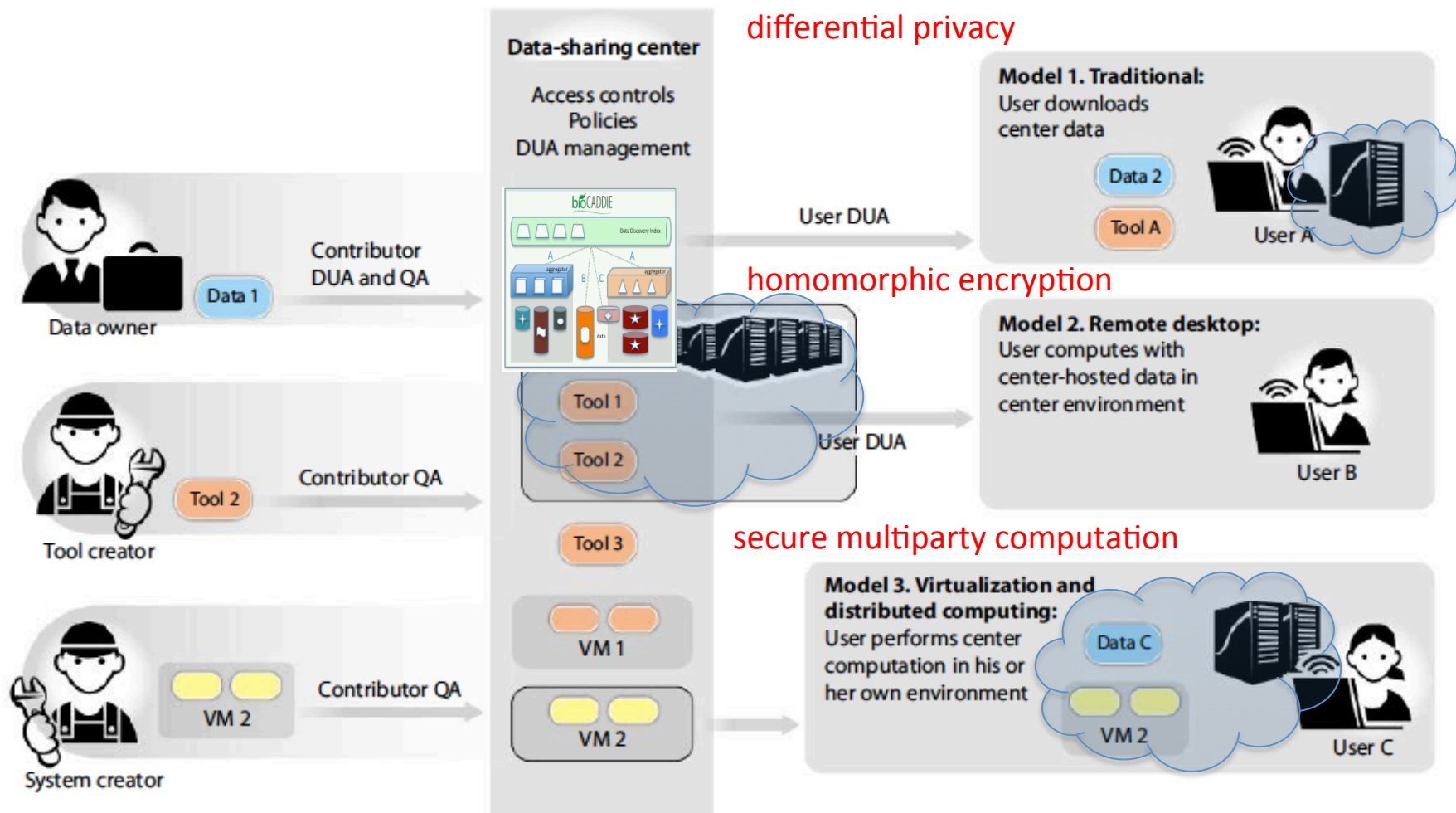
» Tools

» Infrastructure

» Policies

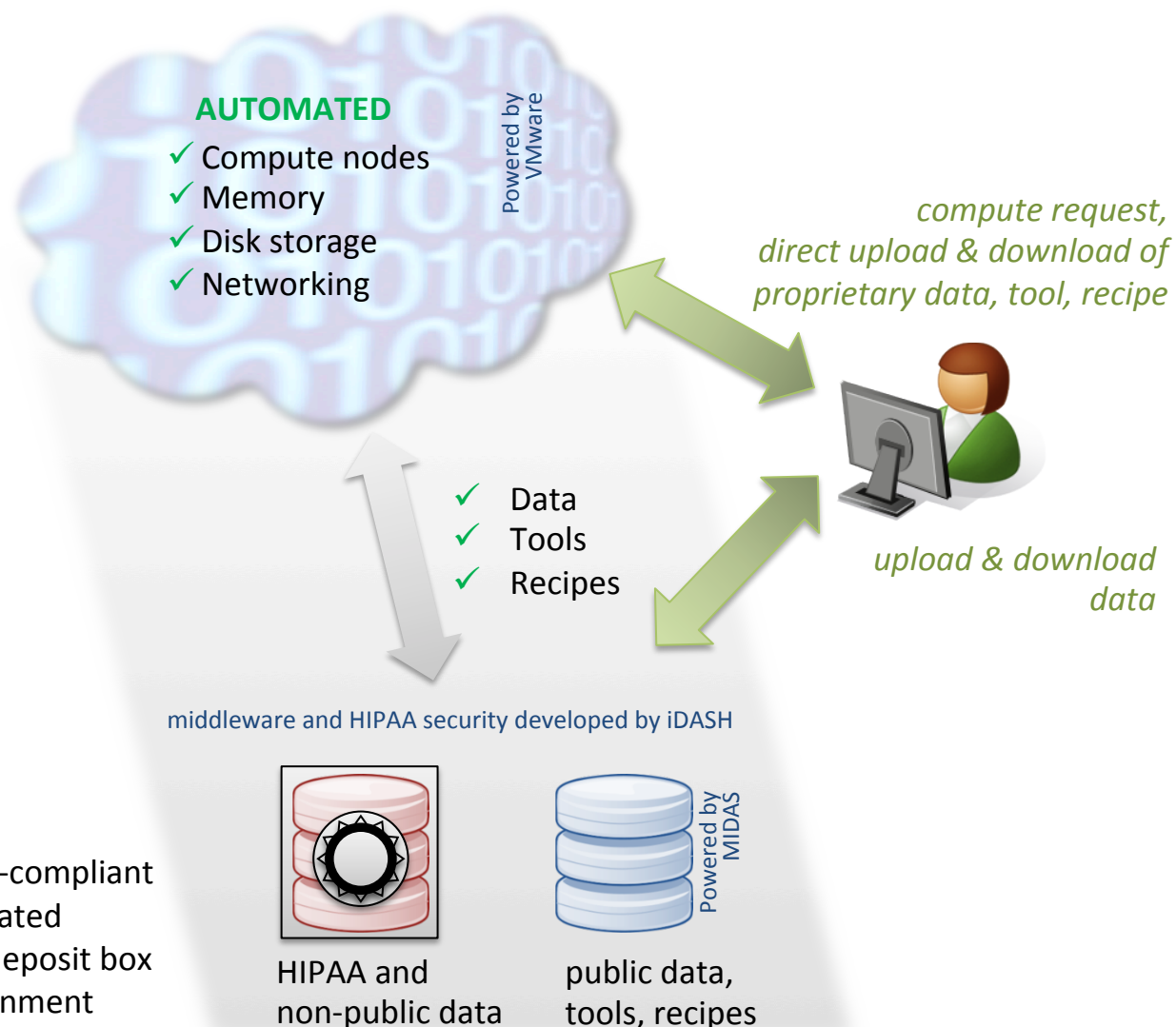


iDASH "commons"



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

On-demand
Virtualized
Elastic
Resilient
Compute
And
Storage
Technology

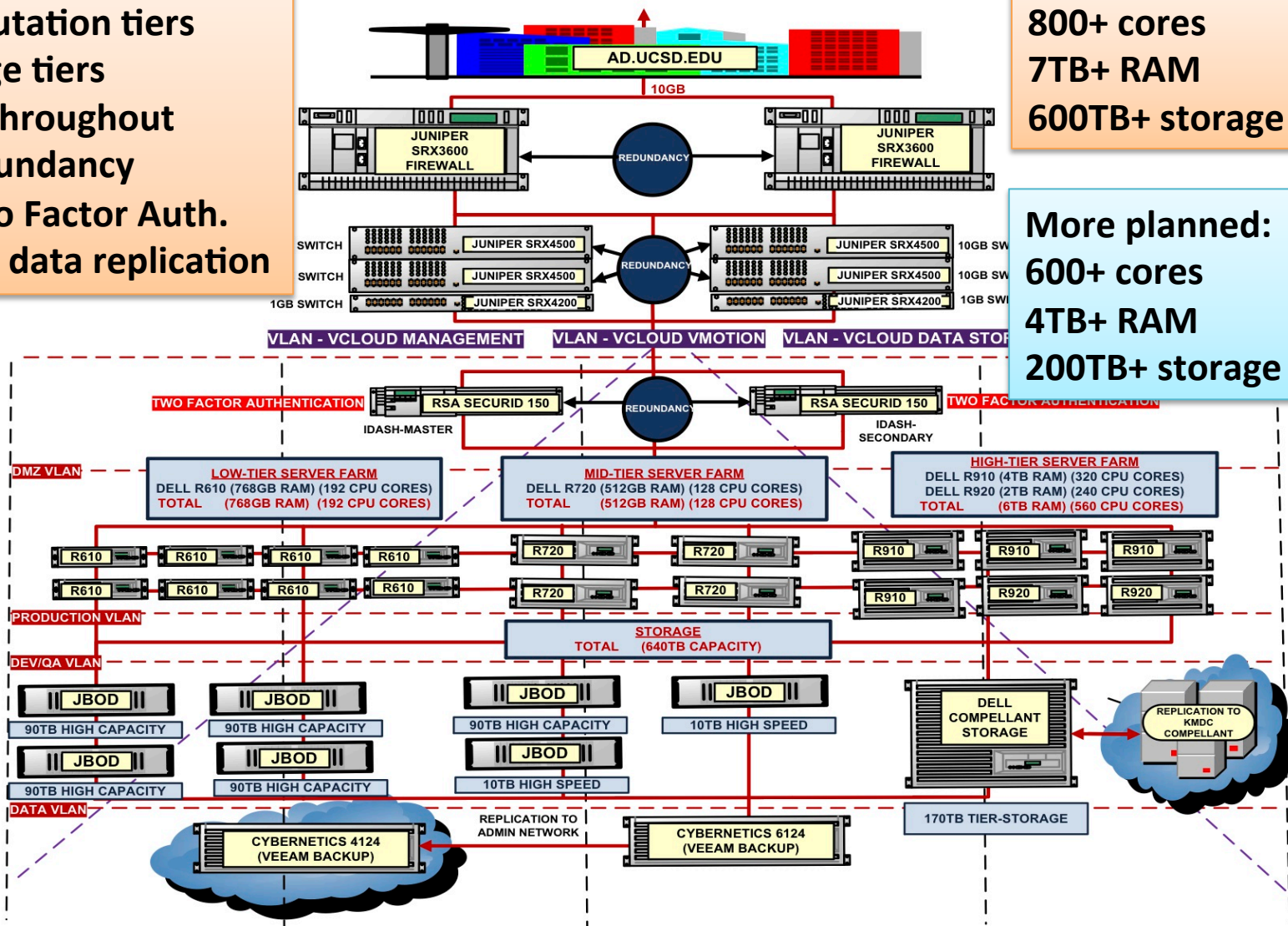


Safe
HIPAA-compliant
Annnotated
Data deposit box
Environment

3 computation tiers
3 storage tiers
10GbE throughout
Full redundancy
RSA Two Factor Auth.
Remote data replication

800+ cores
7TB+ RAM
600TB+ storage

More planned:
600+ cores
4TB+ RAM
200TB+ storage



2008-2009

Clinical Data

2010-2011

Research Data

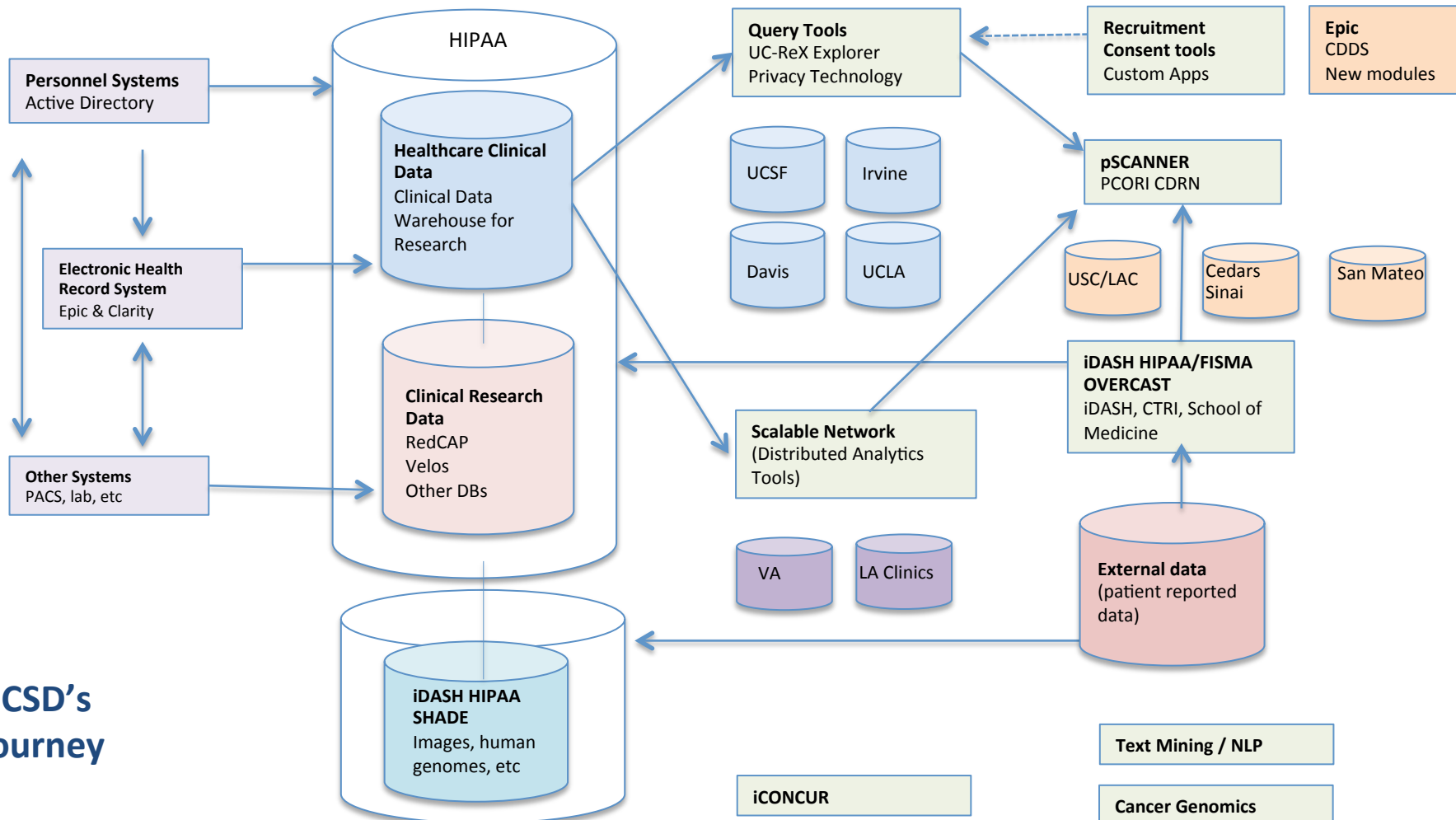
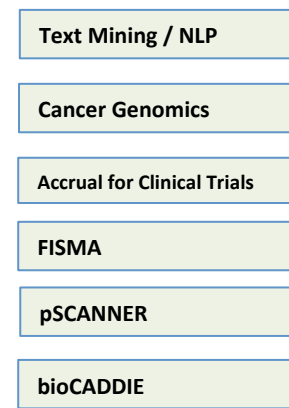
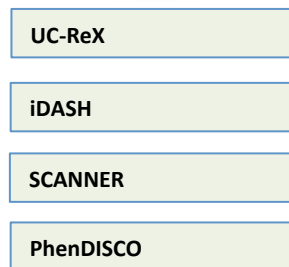
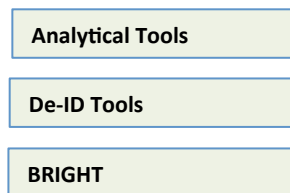
2012-2013

Applications

2014-2015

Integration

UCSD's Journey

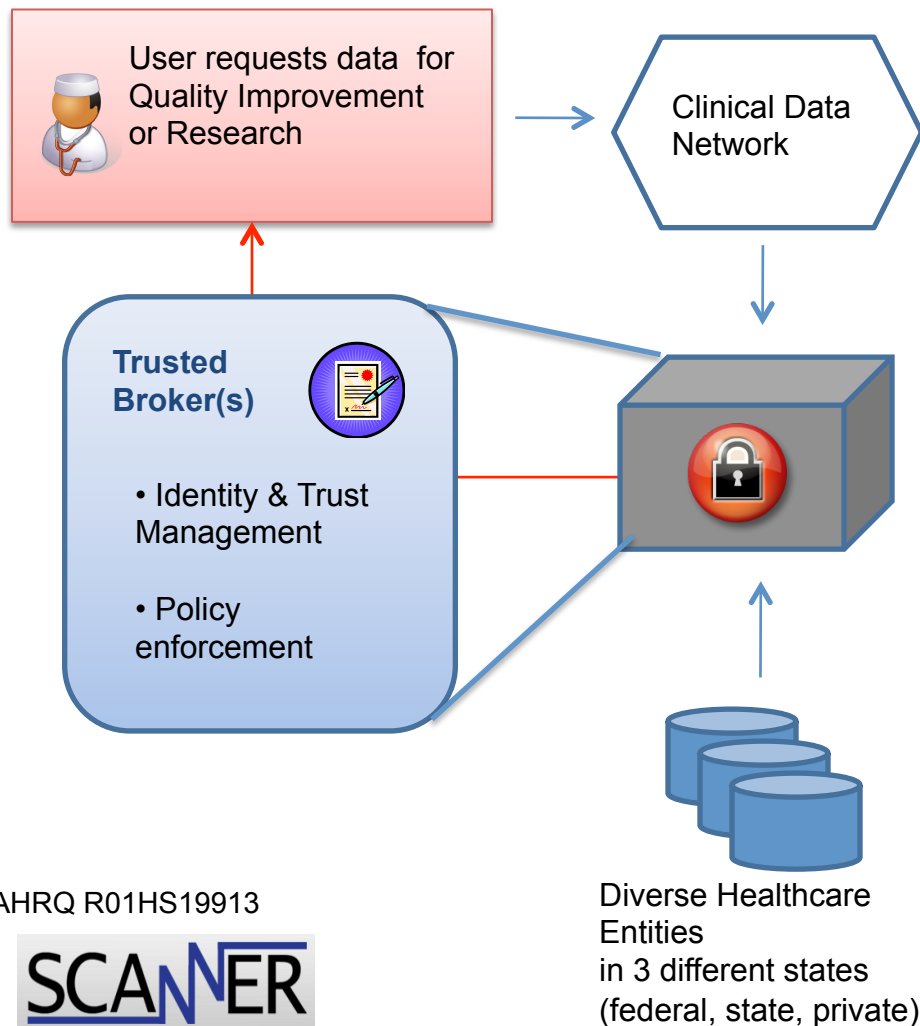


iDASH Clinical Data Network – UC-ReX



- Clinical Data Warehouses from 5 Medical Centers and affiliated institutions exchange (>12 million patients)
- Improve patient safety surveillance, quality improvement, translational research

Big Healthcare Data, Big Clinical Sequencing Data



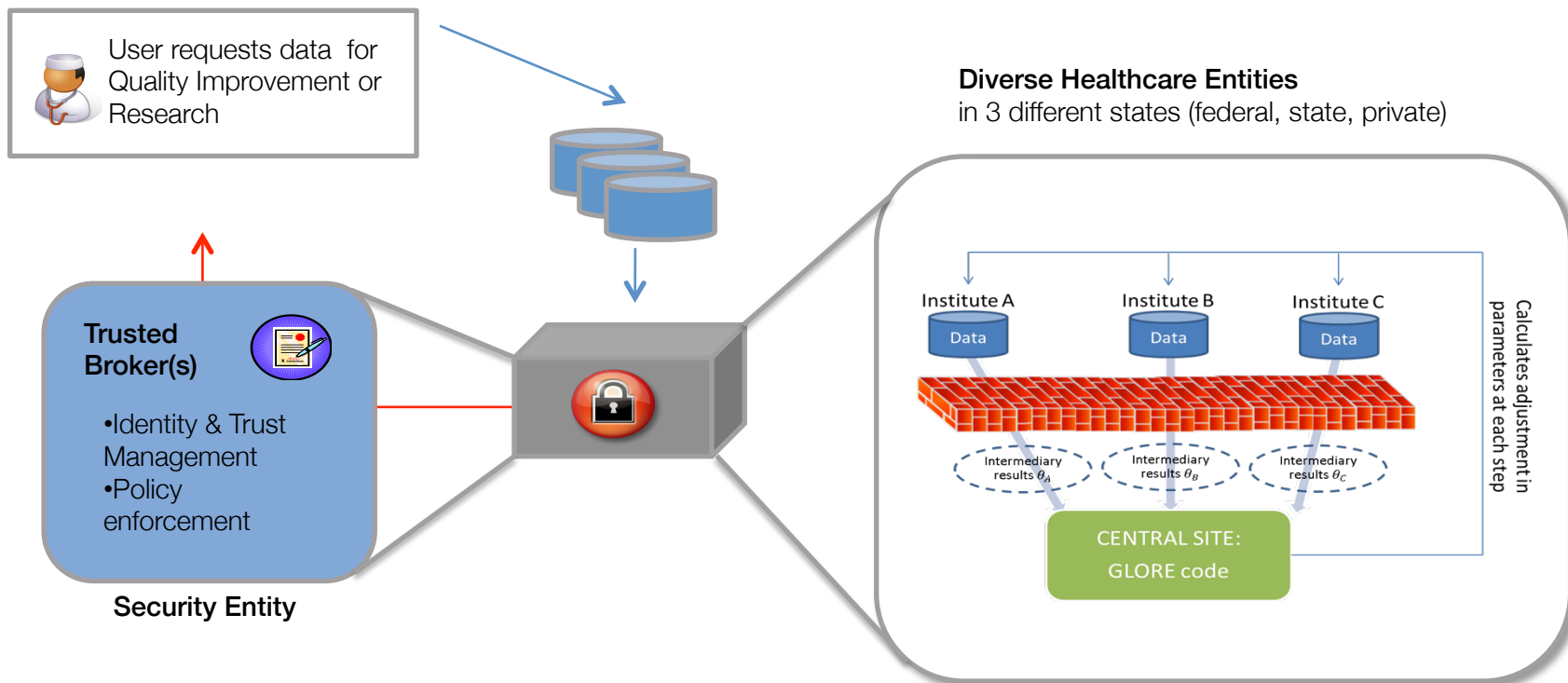
How many patients over 65 are on *Warfarin* or *Dabigatran*?

What are the *major and minor bleeding rates* for patients on these drugs, adjusted for co-morbidities?

Which CYP2C9 variants are most influential?

- Kim K, et al. Development of a Privacy and Security Policy Framework for a Multi-State Comparative Effectiveness Research Network. *Medical Care* 2013
- Jiang X et al. Privacy Technology to Support Data Sharing for Comparative Effectiveness Research: A Systematic Review. *Medical Care* 2013
- Kim KK et al. Data Governance Requirements for Distributed Clinical Research Networks: Triangulating Perspectives of Diverse Stakeholders. *JAMIA* 2014

- 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



AHRQ R01HS19913 / EDM forum

SCANNER

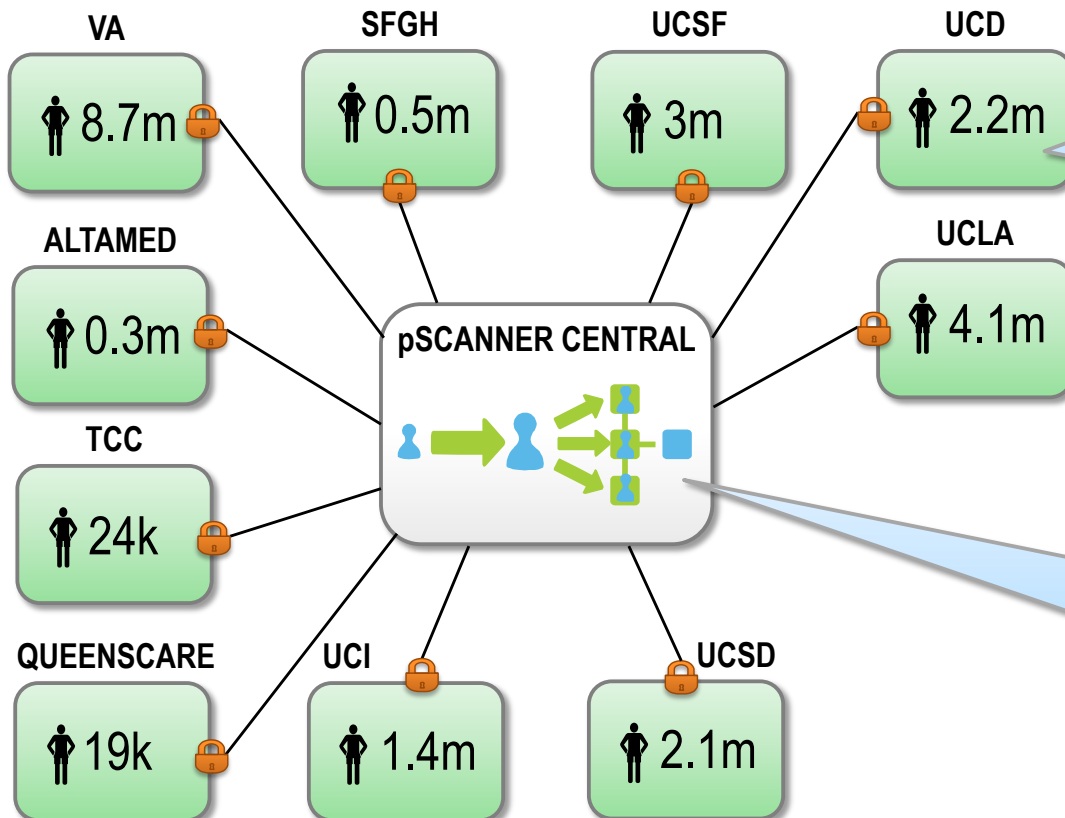
Scalable National Network for
Comparative Effectiveness Research

Distributed regression models

Wu Y et al. Grid Binary LOGistic REGression (GLORE):
Building Shared Models Without Sharing Data. JAMIA 2012

patient-centered SCAlable National
Network for Effectiveness Research

21 M people: 5 UCs, National VA (VINCI resource), 3 Federally Qualified Health Systems in LA (USC), RAND
3 Cohorts: Kawasaki Disease, Congestive Heart Failure, Obesity



New Partners
(~9 M people)

- Cedars-Sinai
- USC Keck / LA Children's
- San Mateo Med Ctr
- Intermountain Healthcare
- U Colorado
- U Washington-led WWAMI practice-based network
- SAFTiNet
 - FQHCs

Legend:



Site



Patient count

TCC: The Children's Clinic
SFGH: San Francisco General Hospital

VM: Virtual Machine
OMOP: Observational Medical Outcomes Partnership

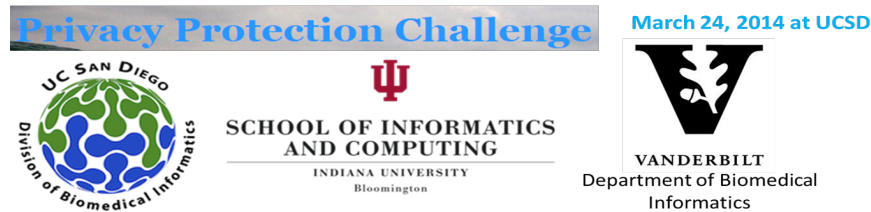


iDASH'14

First Privacy Protection Challenge

Evaluate solutions of **guaranteed privacy protection** for protecting the **output** of genomic data analysis

- Task 1: Privacy-preserving SNP Data Sharing
- Task 2: Privacy-preserving release of top K most significant SNPs



Volume 14 Supplement 1

Critical Assessment of Data Privacy and Protection (CADPP)

Research

3rd iDASH Privacy Workshop

San Diego, CA, USA

24 March 2014

Edited by Xiaoqian Jiang, Lucila Ohno-Machado, Bradley Malin, Shuang Wang, Haixu Tang, Li Xiong and Xiaofeng Wang

Publication of this supplement has not been supported by sponsorship. Information about the source of funding for publication charges can be found in the individual articles. Articles have undergone the journal's standard peer review process for supplements. The Supplement Editors declare that they have no competing interests.

[Conference website](#)

Research article [Open Access](#)

A community assessment of privacy preserving techniques for human genomes

Xiaoqian Jiang, Yongan Zhao, Xiaofeng Wang, Bradley Malin, Shuang Wang, Lucila Ohno-Machado, Haixu Tang
BMC Medical Informatics and Decision Making 2014, **14**(Suppl 1):S1 (8 December 2014)

[Abstract](#) | [Full text](#) | [PDF](#) | [PubMed](#)

Research article [Open Access](#)

Differentially private genome data dissemination through top-down specialization

Shuang Wang, Noman Mohammed, Rui Chen
BMC Medical Informatics and Decision Making 2014, **14**(Suppl 1):S2 (8 December 2014)

[Abstract](#) | [Full text](#) | [PDF](#) | [PubMed](#)

Research article [Open Access](#)

Scalable privacy-preserving data sharing methodology for genome-wide association studies: an application to iDASH healthcare privacy protection challenge

Fei Yu, Zhanglong Ji
BMC Medical Informatics and Decision Making 2014, **14**(Suppl 1):S3 (8 December 2014)

[Abstract](#) | [Full text](#) | [PDF](#) | [PubMed](#)

iDASH'15 Privacy Protection Challenge

- Focus on secure outsourcing and secure data analysis in a distributed setting (humangenomeprivacy.org)



IDASH PRIVACY & SECURITY WORKSHOP 2015

SECURE GENOME ANALYSIS COMPETITION

MARCH 16, 2015
8:30am - 3:00pm
UC SAN DIEGO

Biomedical Research Facility II 5A03

[ENTER THE COMPETITION](#)

• See Media and news coverage in [GenomeWeb](#)
• CADPP15 is supported by [iDASH](#) U54HL108460, iDASH linked R01HG007078 and NHGRI K99HG008175

The poster features a large, stylized clock face on the right side, with the hands and numbers rendered in a modern, geometric style. The text is arranged in a clean, professional layout with a mix of bold and regular fonts.

- Task 1: Homomorphic encryption (HME) based secure genomic data analysis
- Task 2: Secure comparison between genomic data in a distributed setting



At UCSD since 2009, funded by
NIH U54, UL1, U24, UH3, R21, U01, T15, R00, K22, K99, D43, UCBRAID/OP, PCORI, NVIDIA

Research Education
Service *Change*

THANK YOU

Skaggs School of Pharmacy
& Pharmaceutical Sciences