Cloud Computing for Large-Scale Sequencing

Nancy J. Cox, Ph.D.

The University of Chicago

http://genemed.uchospitals.edu

G C A C G G T T G T T

T C C C A G C T A G C

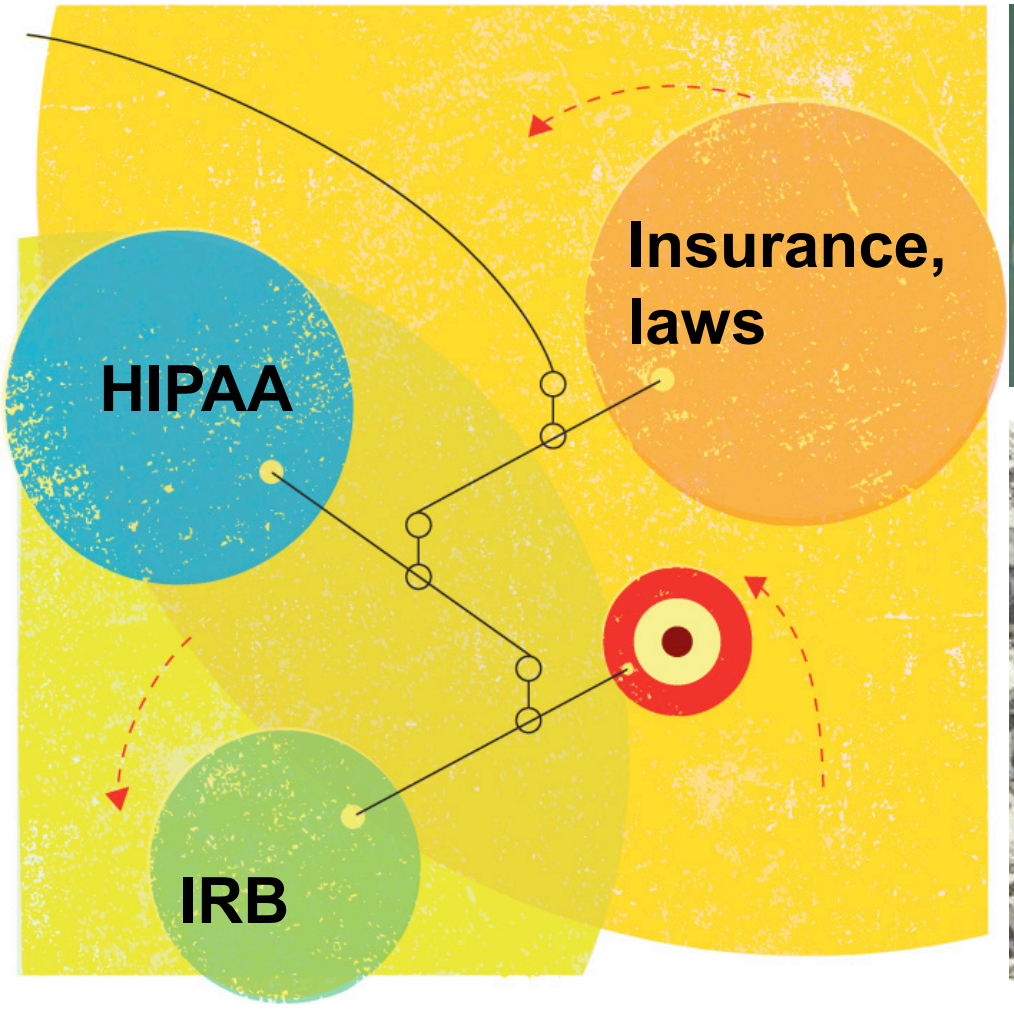
G G T T C G T A T C T

G C A C G G T T G T T

T T A A A C T T C C C

C C C T G G G C G T T

Challenges



$1000 Genome

$100,000 Genome Analysis

The sky won't fall; we used to worry about how we would deal with the deluge of data from GWAS

What is Cloud Computing?

- Anything and everything that is “out there” – not on your desktop or laptop
- Outsourcing
  - IT support
  - Hardware
  - Software
- A new way of doing business
  - Scale to what you need when you need it
  - Don’t own what you cannot use completely efficiently

Keys to Cloud Computing

- A revolution in the “virtualization” of computing architecture
  - Security and connectivity can be defined at the level of the processor
  - Expanded and contracted as needed
- Redundant and abundant (cheap) data storage
- Computational resources to process data are local to the data

Public Vs. Private Clouds

- Private clouds can be developed to meet any level of required security
  - HIPAA compliant clouds with EMR data
  - dbGaP-level security for –omics research data
- Amazon, Google, DropBox, etc
  - Extremely cost effective and very reliable for large-scale but short-term usage
  - Amazon hosting 1000 Genomes (and other large-scale –omics) data for computations

Get Your Amazon Grant Application in Soon!

Hello Researchers –

A quick note to update our progress – we received an enormous amount of requests for grants this quarter and it is taking us longer than expected to review. We will be contacting our recipients within the next week with approvals.

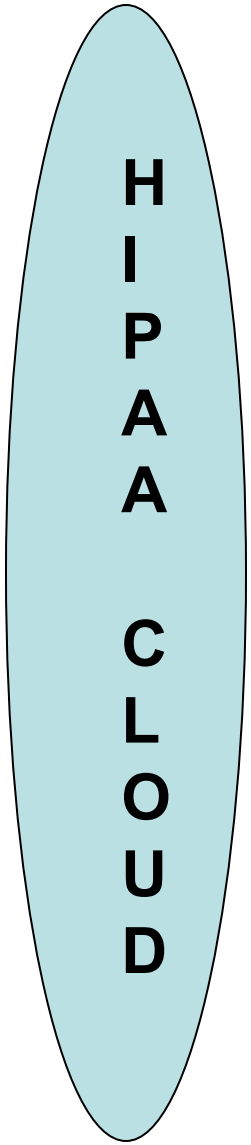
Thank you for your patience.

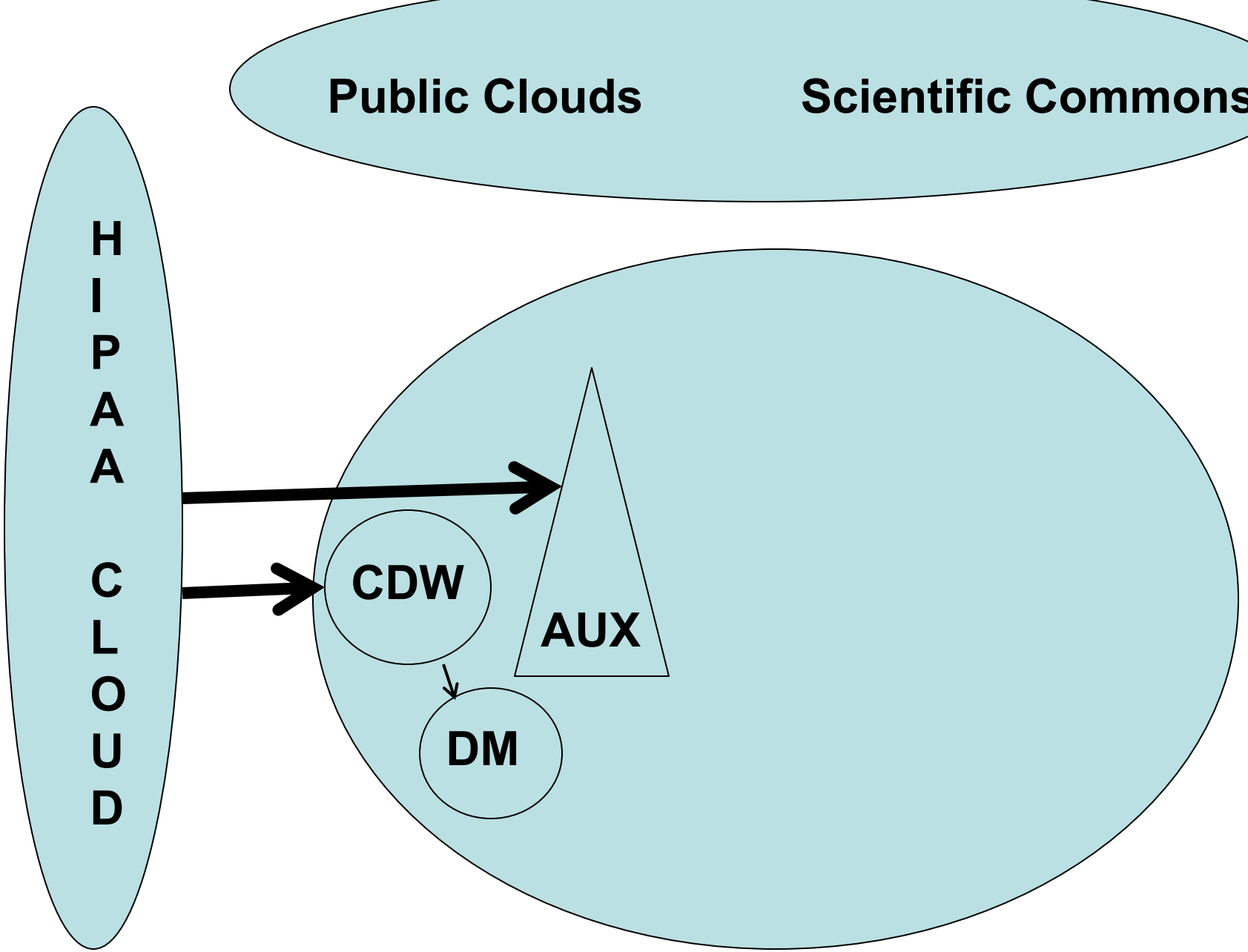
Best,

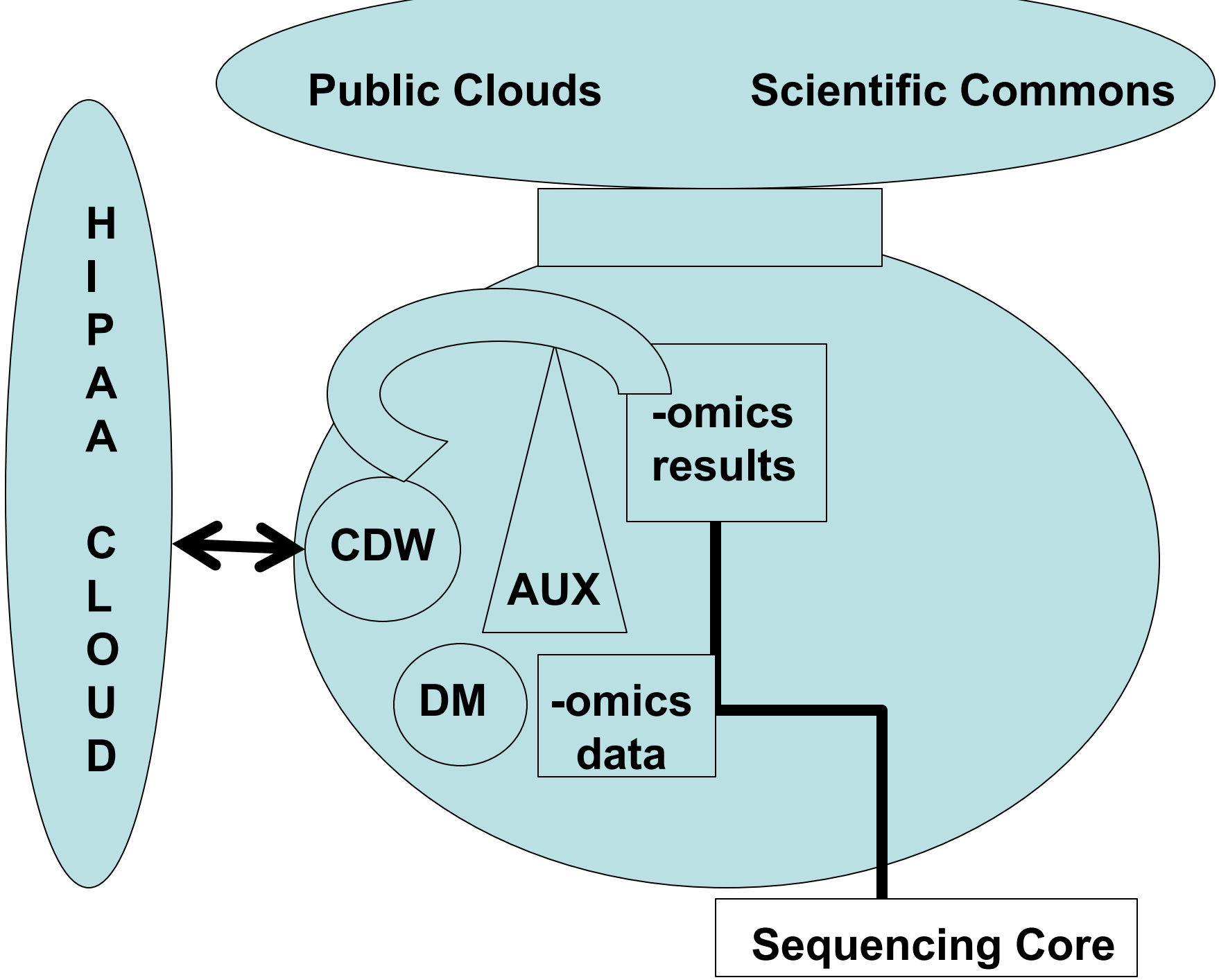
The AWS in Education Grants Team

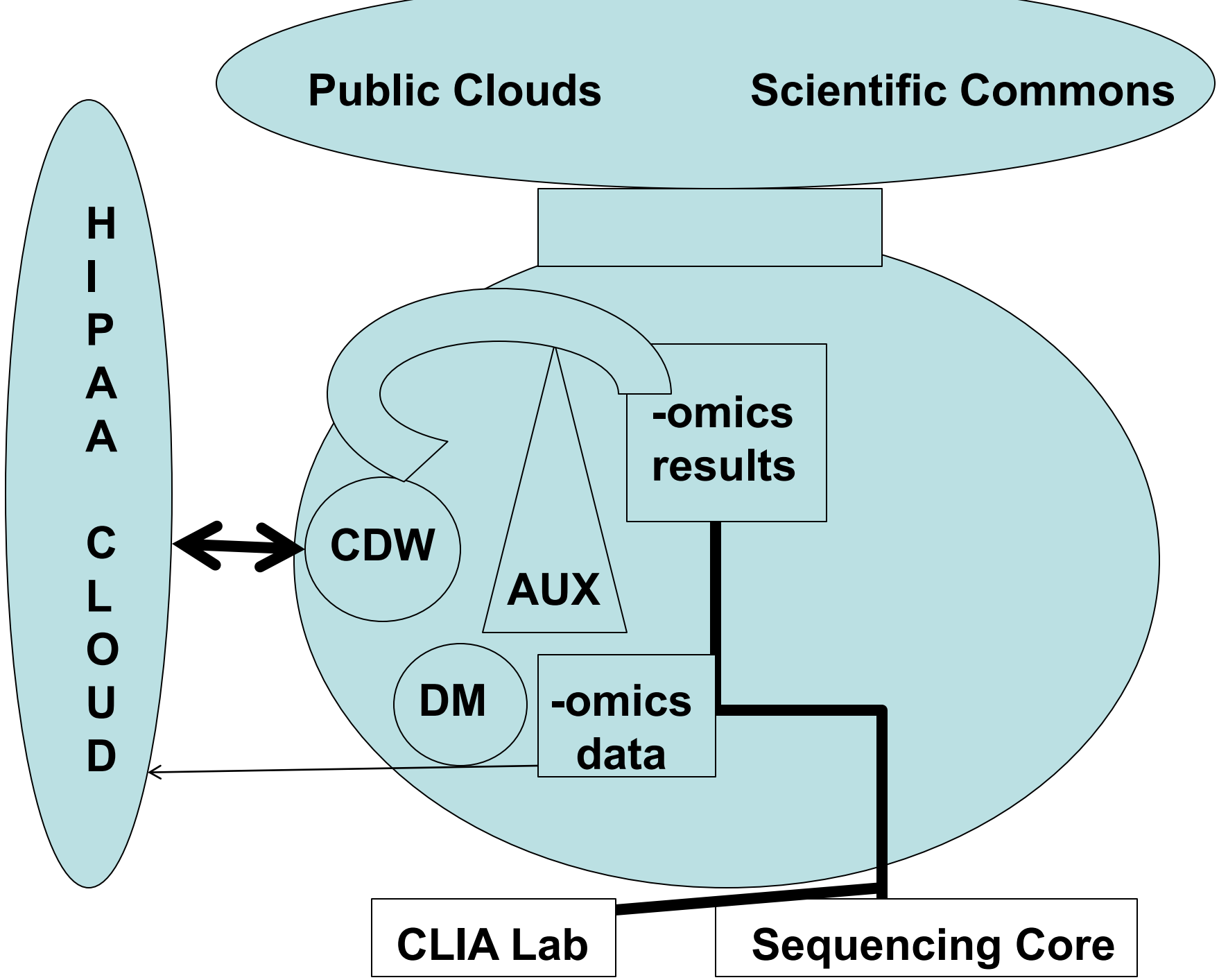

Public Clouds
Scientific Commons

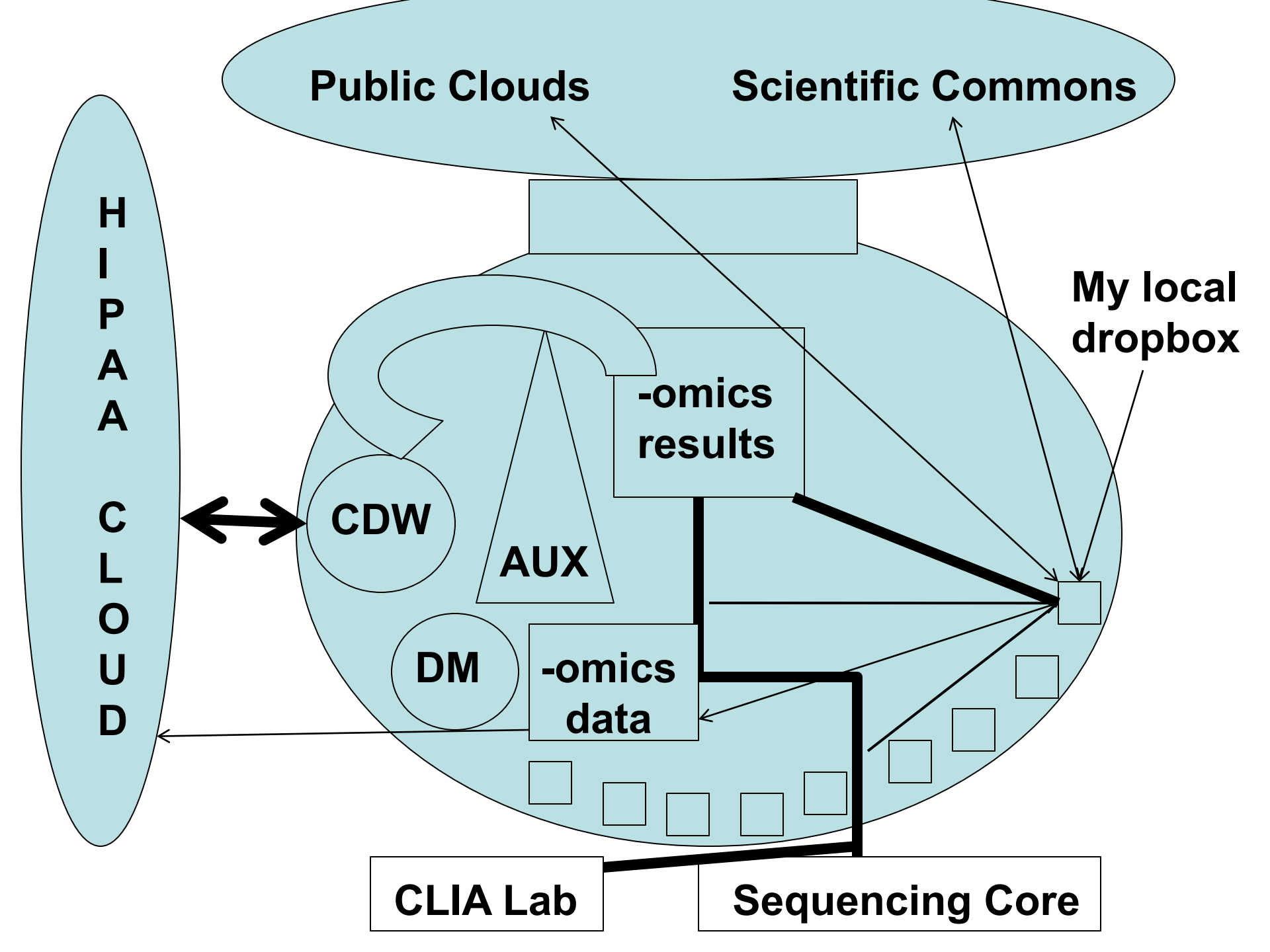

Public Clouds
Scientific Commons

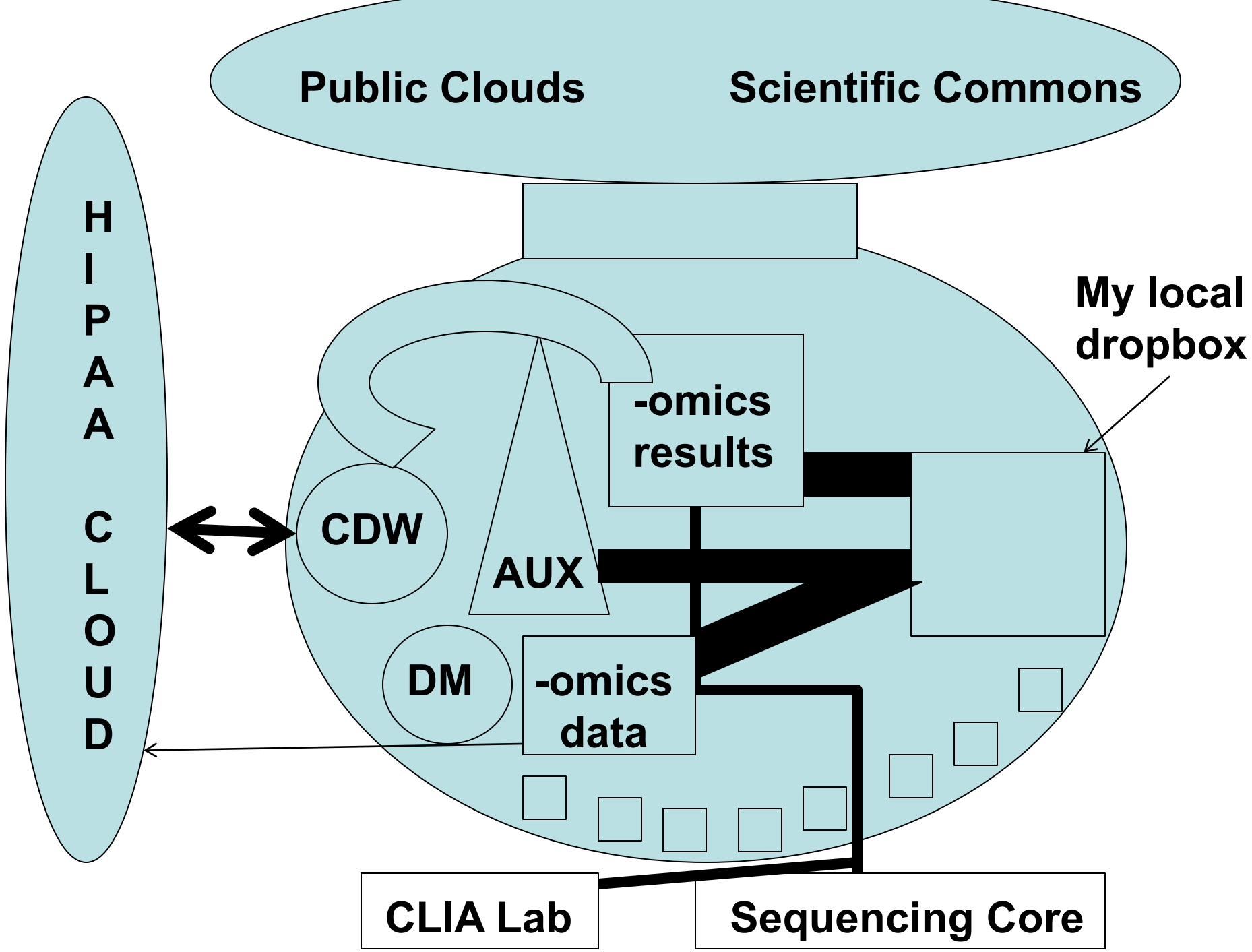

HIPAA CLOUD

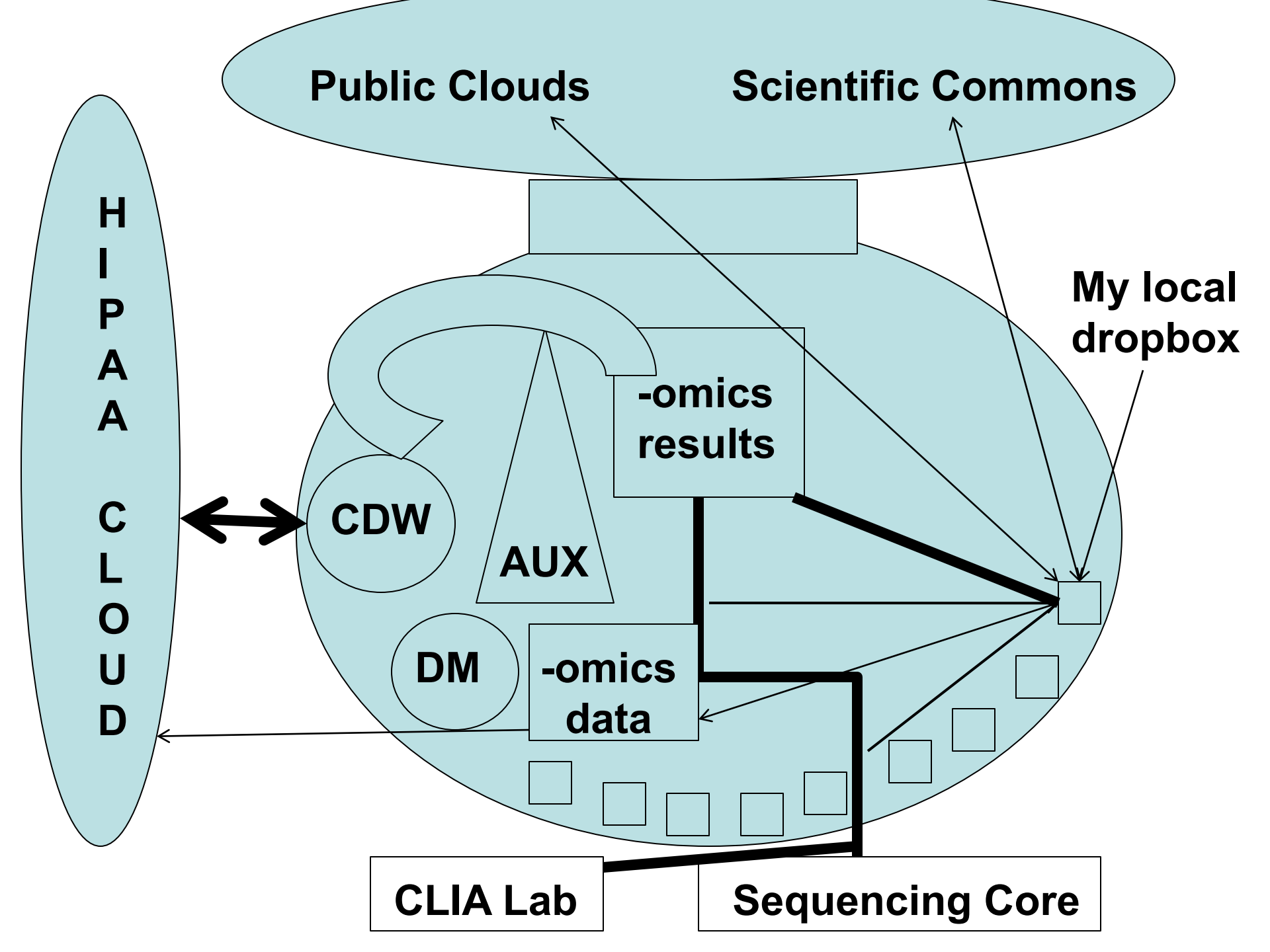

Public Clouds
Scientific Commons
HIPAA CLOUD
CDW
AUX
DM

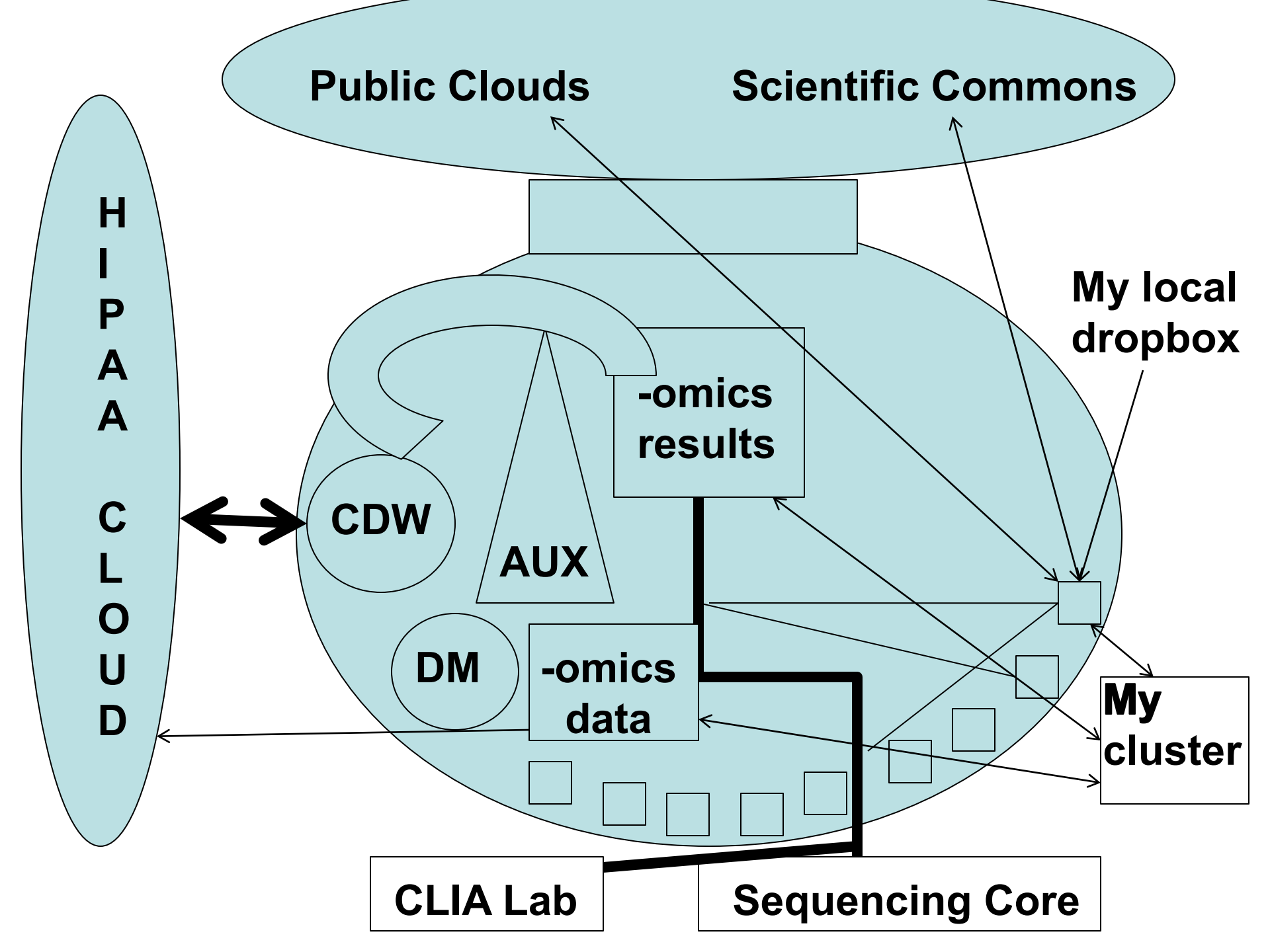

Public Clouds
Scientific Commons
HIPAA CLOUD
CDW
AUX
-omics results
DM
-omics data
Sequencing Core

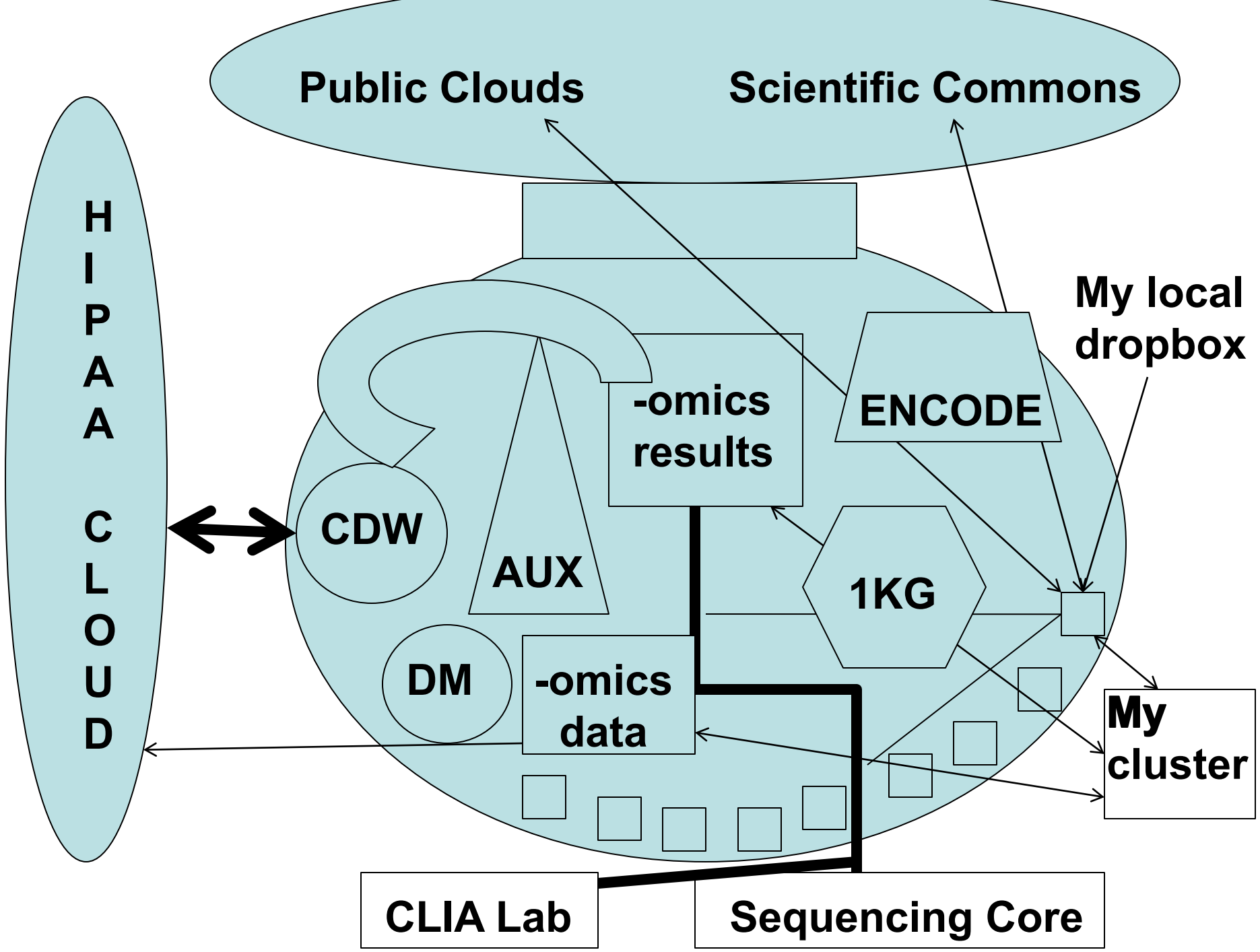

Public Clouds
Scientific Commons
HIPAA CLOUD
-omics results
CDW
AUX
DM
-omics data
CLIA Lab
Sequencing Core


Public Clouds
Scientific Commons
HIPAA CLOUD
My local dropbox
-omics results
CDW
AUX
DM
-omics data
CLIA Lab
Sequencing Core


Public Clouds
Scientific Commons
HIPAA CLOUD
My local dropbox
-omics results
CDW
AUX
DM
-omics data
CLIA Lab
Sequencing Core


Public Clouds
Scientific Commons
HIPAA CLOUD
My local dropbox
-omics results
CDW
AUX
DM
-omics data
CLIA Lab
Sequencing Core


Public Clouds
Scientific Commons
HIPAA CLOUD
My local dropbox
-omics results
CDW
AUX
DM
-omics data
My cluster
CLIA Lab
Sequencing Core

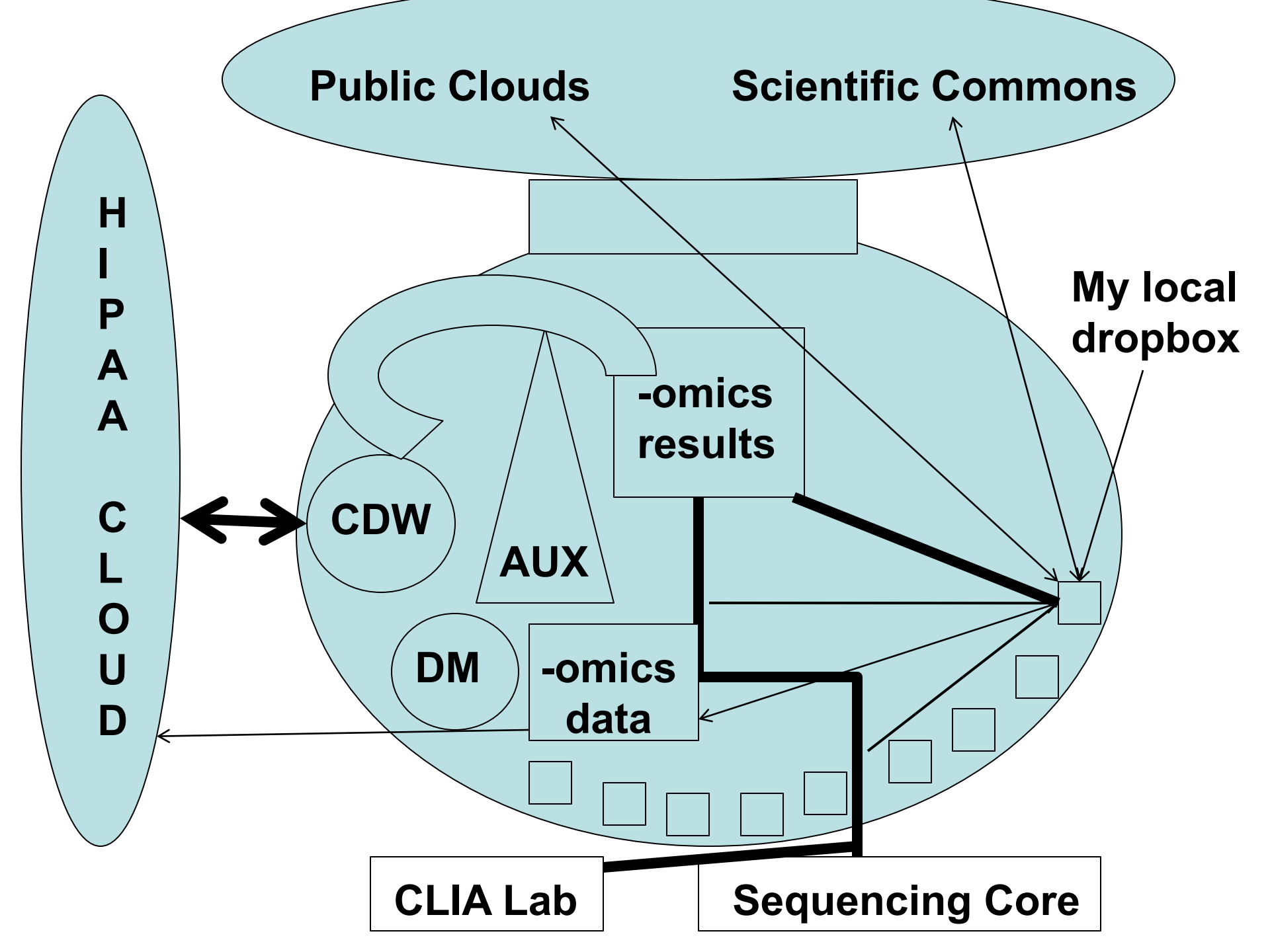

Public Clouds
Scientific Commons
HIPAA CLOUD
My local dropbox
-omics results
ENCODE
CDW
AUX
1KG
DM
-omics data
My cluster
CLIA Lab
Sequencing Core

Colleagues & Collaborators

Bob Grossman

Ian Foster

Kevin White


Public Clouds
Scientific Commons
HIPAA CLOUD
My local dropbox
-omics results
CDW
AUX
DM
-omics data
CLIA Lab
Sequencing Core