



THE UCLA CLINICAL AND TRANSLATIONAL SCIENCE INSTITUTE

HONEST BROKER SERVICES AND DATA ACCESS

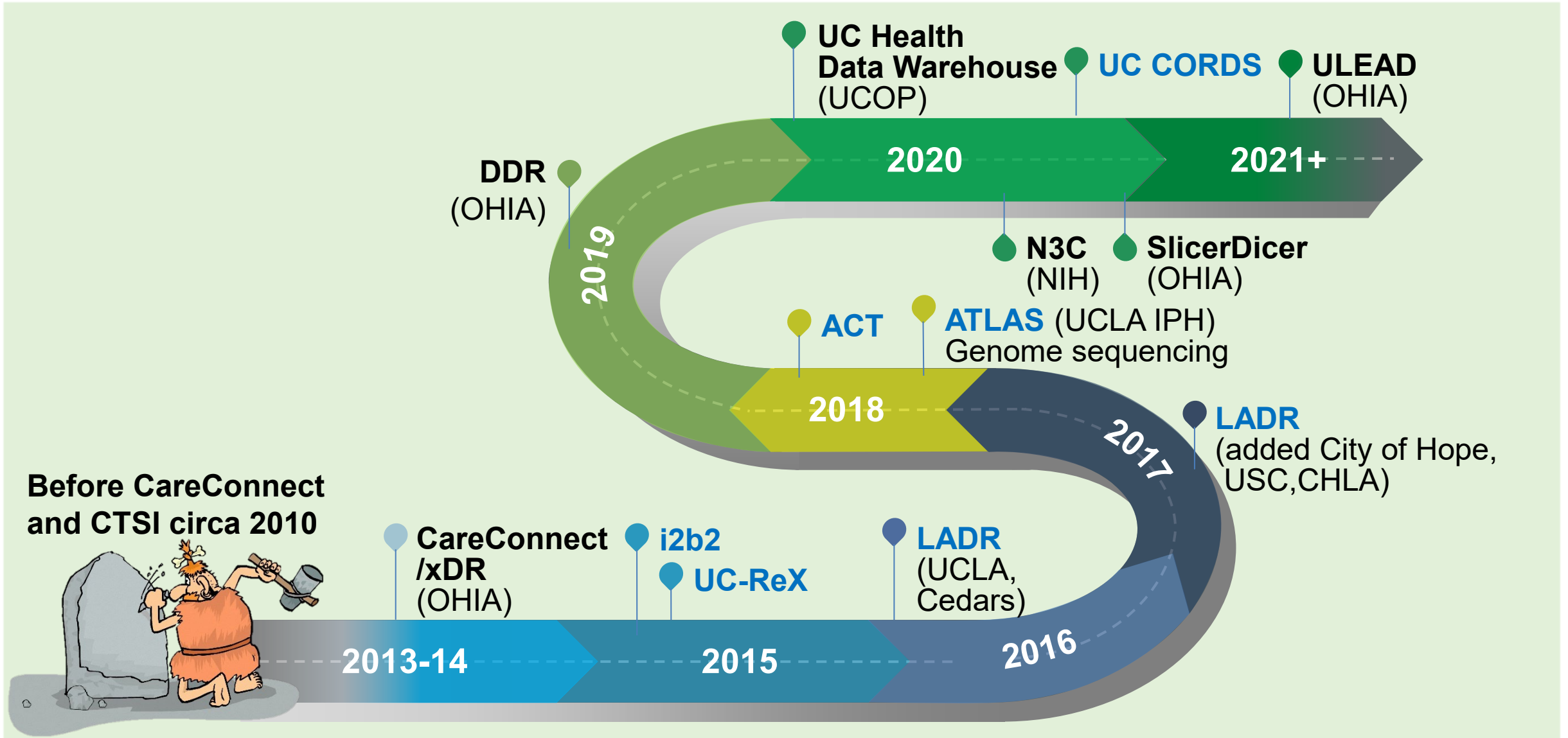
May 24, 2021

Arash Naeim, MD PhD
Senior Associate Director, UCLA CTSI
Chief Medical Officer for Clinical Research, UCLA Health

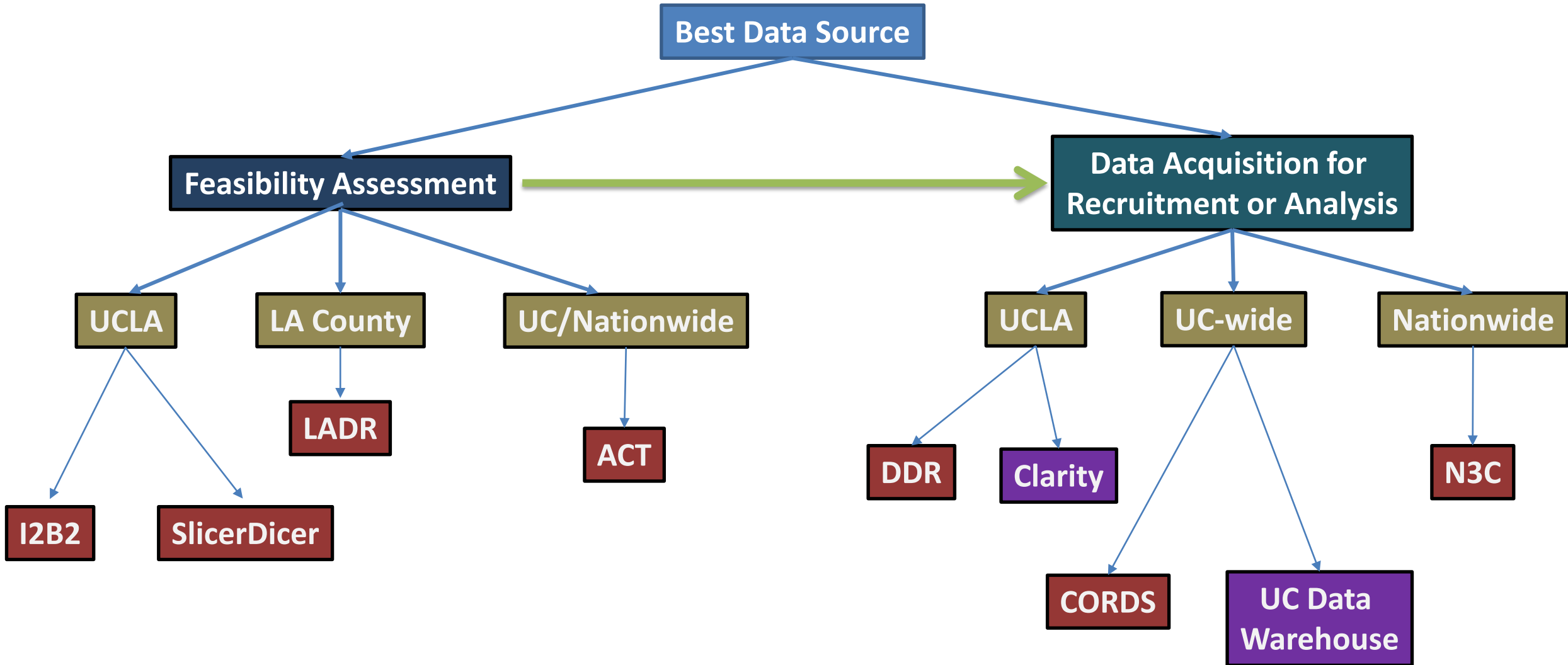
www.ctsi.ucla.edu ■ bringing CTSI innovations to the greatest health needs in Los Angeles



Data Access History: Clinical Data Request Tools



Navigation to the Best Data Source



CORDS=COVID Research Data Set
N3C=National COVID Cohort Collaborative

Managing UCLA Health Data

Best Data Management Solution

```
graph TD; A[Best Data Management Solution] --> B[ULEAD]; A --> C[REDCap]; A --> D[HIPAA-Safe Server/Box]
```

ULEAD

REDCap

**HIPAA-Safe
Server/Box**



Data Workflow Re-engineering: Compliance Fast Track Approval

www.ctsi.ucla.edu ■ bringing CTSI innovations to the greatest health needs in Los Angeles



ULEAD=FASTEST

- **Compliance does not need to review** since not considered a data release if access is just by UCLA investigators [Data can not leave ULEAD]
- **IRB does not need to review if De-identified or Limited Dataset** since it would not be considered human subjects unless it is a rare disease.

Phase 1 will allow the following activities:

CTSI

1. Ability to provide data to researchers in ULEAD (flat files)

Researchers

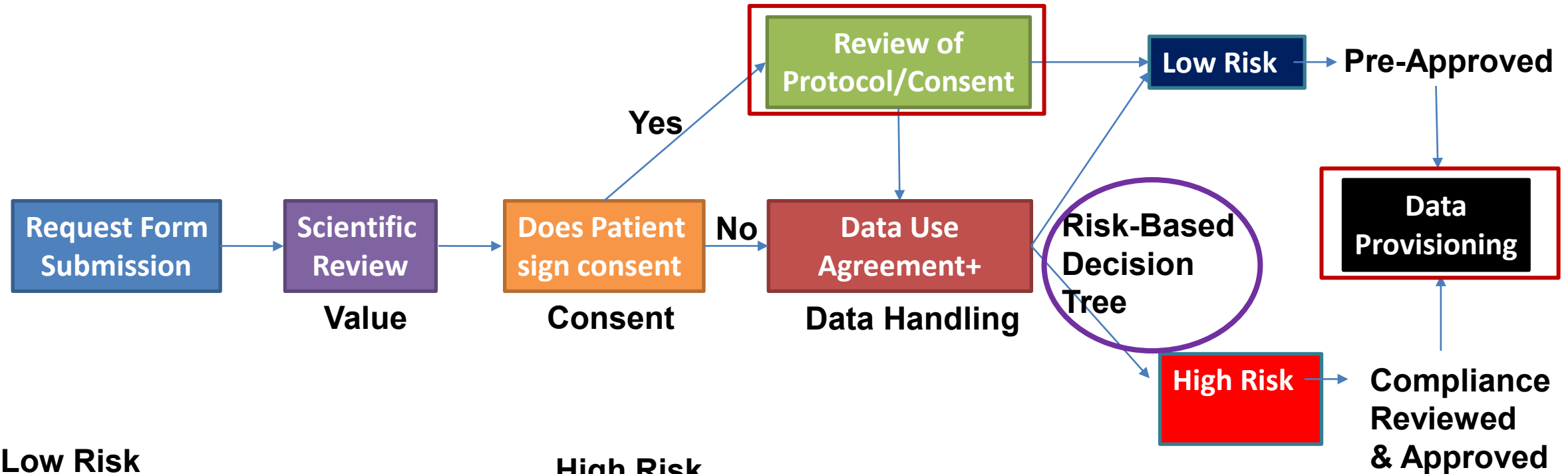
1. Ability to copy data provided by CTSI into a User Folder within ULEAD
2. Ability to bring other data into ULEAD
3. Ability to connect to approved data sources
4. Ability to blend data from CTSI with other data sources. E.g. Cancer Registry, images, etc.
5. Ability to load data into a database (Microsoft Access)
6. Ability to analyze and visualize data
7. Limited ability to develop, train and test ML models
8. Ability to summarize results and prepare reports/graphs
9. Ability to share summarized data with the UCLA Health research community, within ULEAD

ULEAD Specifications

Phase one of the ULEAD platform will offer each user the features listed below, and will support 250 concurrent users.

- Virtual Desktop: Compute: 4 CPUs (10~ GHz), RAM: 32 GB. Can be increased on a case-to-case basis
- Database: MS Access
- Software: Python, R, SQL Server, SAS
- Applications: Anaconda, Jupyter Notebook, R Studio, Visual Studio, SSMS, MS Office (Word, Excel, Powerpoint, Access), Pycharm, PowerBI, Adobe Acrobat, SAS, MATLAB.
 - ***For licensed software such as SAS and MATLAB, users must bring their own license***
- Personal Storage: User Folder within ULEAD. 50GB+ storage
- Collaboration: Team Folders within ULEAD
- Data Transfer In & Out: Data Transfer Folders within ULEAD
- Internet access: Websites approved by OHIA
- Interfaces: Allowed with approved databases and applications

Compliance Fast Track Pilot



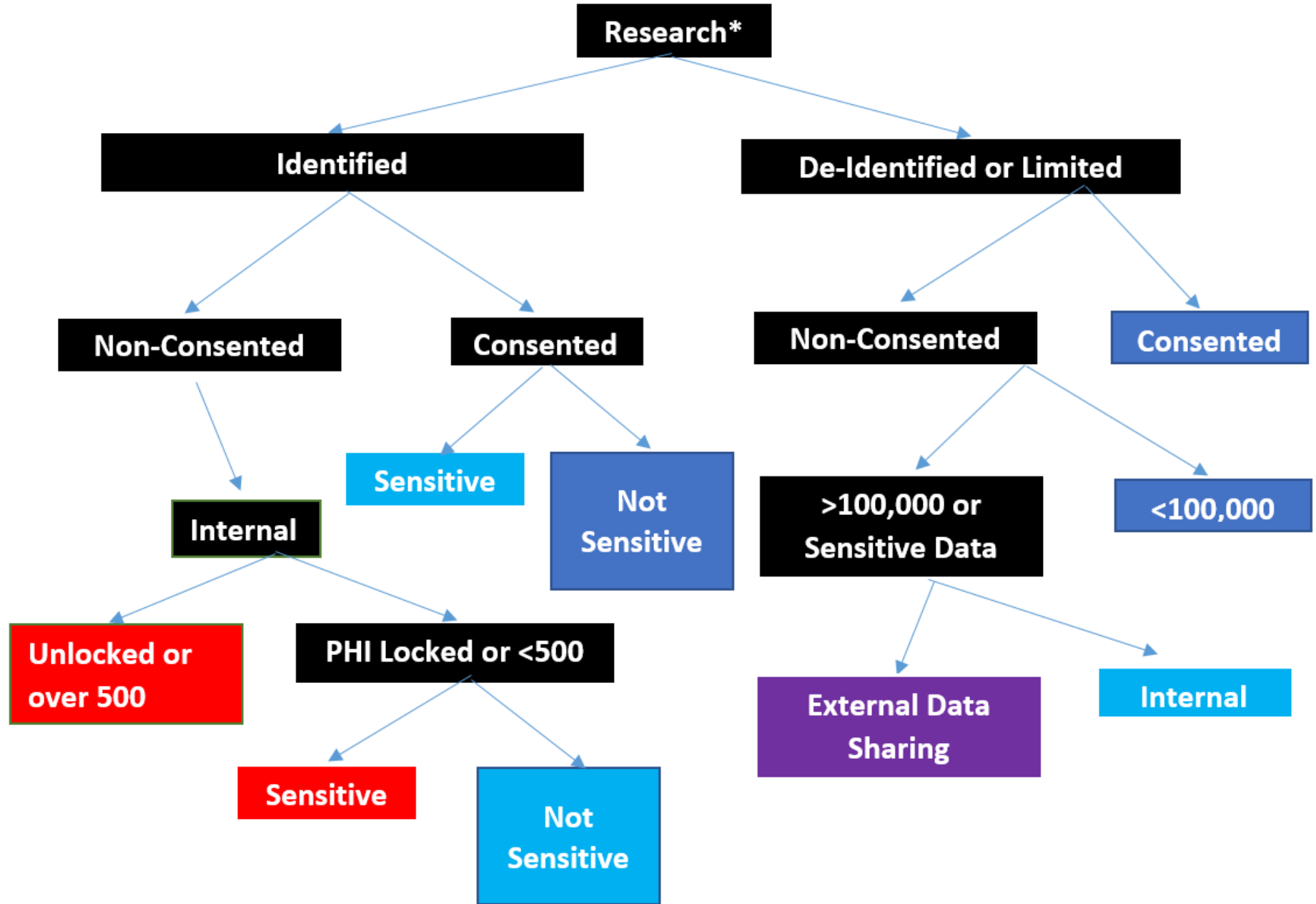
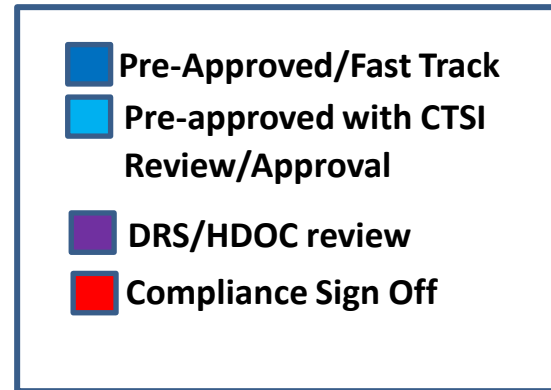
Low Risk

1. Consent Given (any type of data)
2. De-identified data <100,000
3. Limited Data <100,000
4. Identified (Locked Down)
5. Identified <500

High Risk

1. High Volume >100,000
2. PHI >500 (not locked down)
3. Sensitive Data

Decision Tree



Sensitive Data Elements (Identified Data Sets)

- 1. Mental Health Disorder**
- 2. Eating Disorder**
- 3. Substance Abuse Disorder**
- 4. Sexually Transmitted Disease (including HIV)**
- 5. Child Abuse**
- 6. Domestic Abuse**
- 7. Sexual Abuse**
- 8. Abortion**
- 9. Transplant Donor-Recipient**
- 10. Adoption**
- 11. Financial Billing Data**

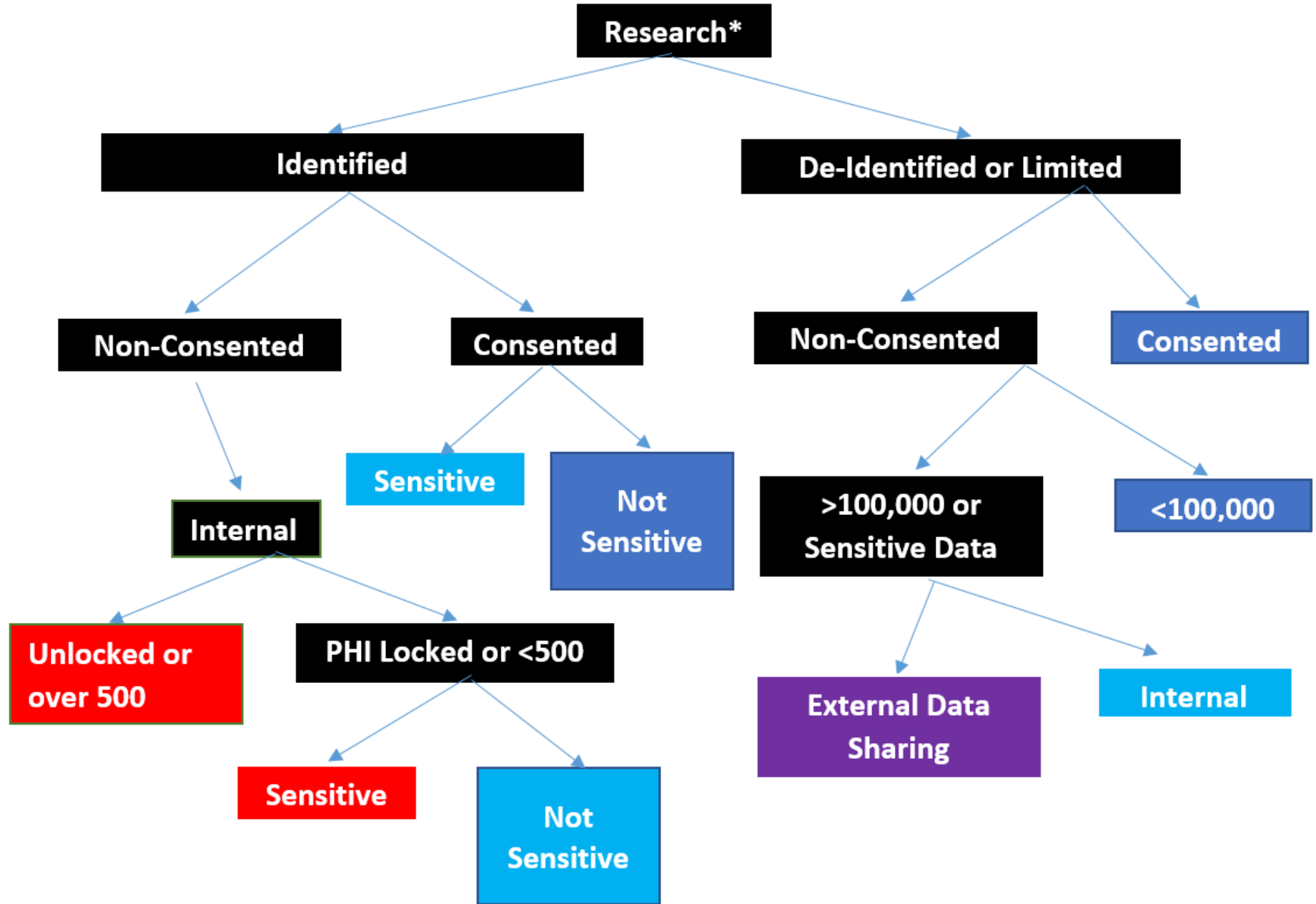
Less Sensitive:

- De-identified/Limited Dataset**
- Recreational use in Social History Fields**
- Global Problem List, Past Medical History, or Diagnosis Fields**

Decision Tree

Legend:

- Pre-Approved/Fast Track
- Pre-approved with CTSI Review/Approval
- DRS/HDOC review
- Compliance Sign Off



Detailed Use Cases to Risk Stratify

Case #	Decision Tree Drivers						
	Data Origin	Type of Data	Consent	Volume	Sensitive Data	PHI Locked Down or less than 50	Sharing
9	Health Data	Deidentified	Yes	N/A	No	N/A	No
10	Health Data	Deidentified	Yes	N/A	No	N/A	Yes
13	Health Data	Deidentified	No	<100,000	No	N/A	No
14	Health Data	Deidentified	No	<100,000	No	N/A	Yes
17	Health Data	Deidentified	No	>100,000	No	N/A	No
18	Health Data	Deidentified	No	>100,000	No	N/A	Yes
21	Health Data	Limited	Yes*	N/A	No	N/A	No
22	Health Data	Limited	Yes*	N/A	No	N/A	Yes
25	Health Data	Limited	No	<100,000	No	N/A	No
26	Health Data	Limited	No	<100,000	No	N/A	Yes
29	Health Data	Limited	No	>100,000	No	N/A	No
30	Health Data	Limited	No	>100,000	No	N/A	Yes
33	Health Data	Identified	Yes*	N/A	No	No	No
34	Health Data	Identified	Yes*	N/A	No	No	Yes
37	Health Data	Identified	No	Locked or <500	No	Yes	No
39	Health Data	Identified	No	>500	No	No	No
41	Health Data	Identified	No	>0	No	No	Yes

Risk Assessment

Yes No

1. Identified
 - Locked or <500
2. Unconsented
3. Sensitive
4. Large #
5. External sharing

Risk Assessment

Data Use Agreements

- Summary of the Data Request and/or Release Page 1
- Summary of the Risk Assessment Page 2
- Data Use Agreement Page 3-6
- Chair Attestation Page 6

Chair Attestation:

The above Principal Requestor:

- is in good standing to perform the proposed project associated with this data request; and
- has the resources and capability to manage the data being requested.

By signing this attestation, the Department/Division will be responsible to ensure the above Principal Requestor adheres to this Data Use Agreement and will share in the financial responsibility associated with a data breach that stems from a violation of this agreement.

Chair/Division Chief Name

Chair/Division Chief Signature

Date

Data Elements Template and Data Dictionary

- **Fastest – preset data (top 100-150 data elements)**
- **Faster – preset data + additional pre-programmed data element**
- **Fast – All pre-programmed data elements**

- **Slow – Custom request of well defined data**
- **Slower – Custom request of ill defined data**
- **Slowest – Custom request of ever changing and non-defined data**