

Balancing Act: Big Data & Data Privacy

Ishwar Ramsingh, MBA, CISSP, CISM, CISA, CIPP/E, CIPM, CIPT, CHC, CHRC, CDPSE, AIGP

Executive Director, Research Privacy - Data Broker

Joey Casanova, BBA, CIPT, CHRC, CIP

Data Broker Manager, Research Privacy - Data Broker

Timothy Smith, PhD

Director, IT-Research Informatics Governance & Security-UHCORP

Tiffany-Ashley Disney, JD, MBA

Licensing Associate, Office of Technology Transfer

Disclosure of Financial Relationships

- **Ishwar Ramsingh**, MBA, CISSP, CISM, CISA, CIPP/E, CIPM, CIPT, CHC, CHRC, CDPSE, **Joey Casanova**, BBA, CIPT, CHRC, CIP, **Timothy Smith**, PhD, **Tiffany-Ashley Disney**, JD, MBA have no relevant financial relationships with ineligible companies to disclose.

The University of Miami Division of CME has mitigated all relevant financial relationships.

AI Use Notice

Microsoft Copilot was used to enhance formatting and improve the overall visual presentation.

All AI-assisted content was reviewed, edited, verified for accuracy, checked for bias/commercial influence, and approved by the presenter/faculty member before submission.

No PHI, PII, confidential learner data, proprietary faculty content, or non-approved sensitive information was entered into public AI tools.

IRB Grand Rounds 2026 - Virtual Series

Time	Topic	SPEAKER
September 17, 2026, 10-11 am	Set up for Success: Application, Reporting & System requirements	RIC, ORA, COI/ RSP,HSRO
October 15, 2026, 10-11 am	Balancing Act: Big Data & Data Privacy	Data Brokers, AI, OTT,HSRO
November 19, 2026, 10-11 am	Ready, Set, Comply!	IITSU, RQA,CTSI,HSRO
December 17, 2026, 10-11:30 am	Better Together: The Power of the Ancillary Network	EHS, Pathology, PRMC, HRSC, IBC,MCA, JHS,HSRO

Objectives

- Recognize data sensitivity and its regulatory implications
- Examine data in action—its power and impact
- Safeguard data in motion to ensure security and compliance

Data Protection in Practice: Recognizing Sensitivity, Examining Data in Action, and Safeguarding Data in Motion

Ishwar Ramsingh, MBA, CISSP, CISM, CISA, CIPP/E, CIPM, CIPT, CHC, CHRC, CDPSE

Executive Director, Research Privacy - Data Broker

Joey Casanova, BBA, CIPT, CHRC, CIP

Data Broker Manager, Research Privacy - Data Broker

Timothy Smith, PhD

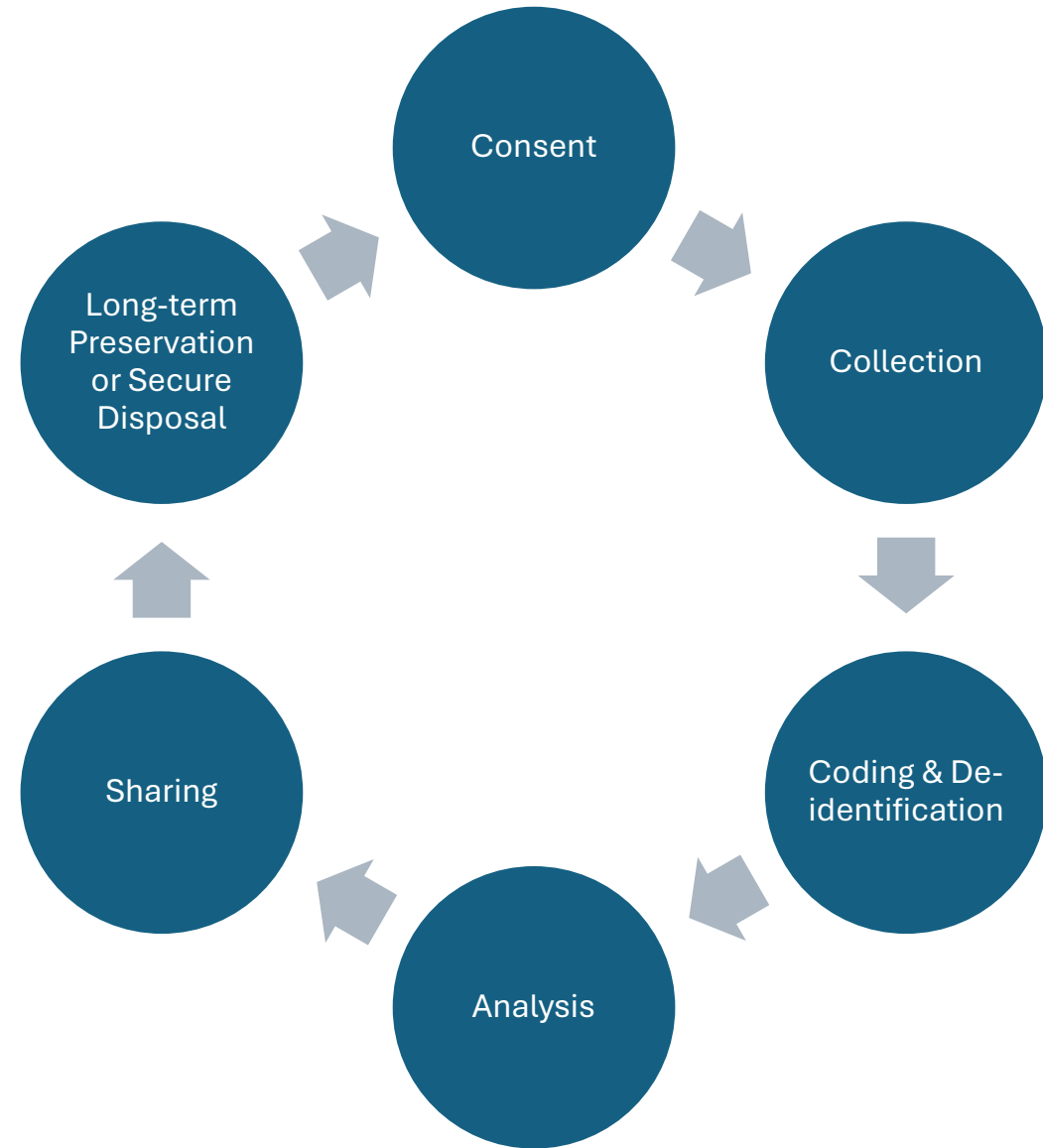
Director, IT-Research Informatics Governance & Security-UHCORP



Recognize Data Sensitivity

Start with what the data can reveal, not just what the file is called

One Specimen, Many Data Lives



The Balancing Act

Scientific value grows when data can be connected. Privacy risk can grow at the same time.

- Biospecimens create durable research value
- Linked datasets reveal patterns no single source can show
- Reuse, linkage, and sharing can outpace original expectations
- Trust is a research dependency, not a communications accessory

Sensitivity is Contextual

Risk depends on content, combinations, population, and intended use.

Direct Identifiers

Name, address, MRN, contact info

Indirect Identifiers

Dates, geography, rare traits, event patterns

Sensitive Attributes

Health, genetics, biometrics, socioeconomic factors

Linkage Potential

Multiple “safe” fields can become identifying together

U Data Classification Levels



UMIT Data Classification Policy. PolicyStat # 19626988

Public	Sensitive	Private	Confidential (Restricted)
Available to anyone without any legal restrictions on access or use.	Not approved for general or public distribution.	Considered proprietary and critical to the ongoing business continuity and operations of the University	University is under legal or contractual obligation to protect from disclosure, alteration or destruction.
- Tuition and fees - Annual reports - Press statements - External facing website and social media, blogs, etc. - Employee names, titles, work phone numbers, work address, email addresses	- Accounting and financial information not otherwise classified as Confidential data (internal use only) - Prospective student/applicant information - Prospective employee/applicant information	- Salaries - Financial transactions which do not include confidential data - Educational records required for business and educational purposes - Information that is related to a student, faculty, employee	- Medical research technology - Controversial research topics - Financial information - Donor names and account numbers - PHI, patient data, health and medical records. - Intellectual property. - Information covered by non-disclosure agreements.

All Research Data falls in one of these two categories

Does your research require Cybersecurity Maturity Model Certification (CMMC)?

- Determined by contract language (typically DARPA, CRADA or USACE) or Data Use Agreement
- CMMC Level 2 status means assessment against NIST SP800-171 requirements
- Certification of NIST SP 800-171 compliance at the proposal stage may be required. Examples:
 - *“Performer must achieve and post a valid CMMC Level 2 (Self-Assessment) status and current annual affirmation in the Supplier Performance Risk System (SPRS)”*
 - *“...evidence of compliance with NIST SP-800-171 Assessment”*
- Can be provided by
 - Self Assessment
 - Third-party certification

What is NIST SP 800-171, and why is compliance required for my project?

- National Institute of Standards and Technology (NIST) developed Special Publication (SP) 800-171 to establish guidelines for federal agencies when contracting with nonfederal organizations that store and share **sensitive data** not otherwise classified or protected by federal regulations
- Any research project requiring NIST SP 800-171 compliance must fulfill the steps for compliance ***before funding is awarded***.
- For each project, the Principal Investigator is responsible for establishing and maintaining NIST SP 800-171 compliance throughout the duration of the project.
- Need clear definition of the IT Resources to be used and for a System Security Plan (SSP)
 - Formal document that details the security requirements and describes the security controls in place or planned for meeting those requirements for the in-scope IT resources

How did this come about ?

- **Sept 24, 2025. Notice Number:** *NOT-OD-25-159 Security and Operational Standards for NIH Controlled-Access Data Repositories*
 - Modernize security standards provided in the [“NIH Security Best Practices for Controlled-Access Data Subject to the NIH Genomic Data Sharing \(GDS\) Policy”](#) and
 - Established minimum expectations for access to controlled-access data by developers.

<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-25-159.html>

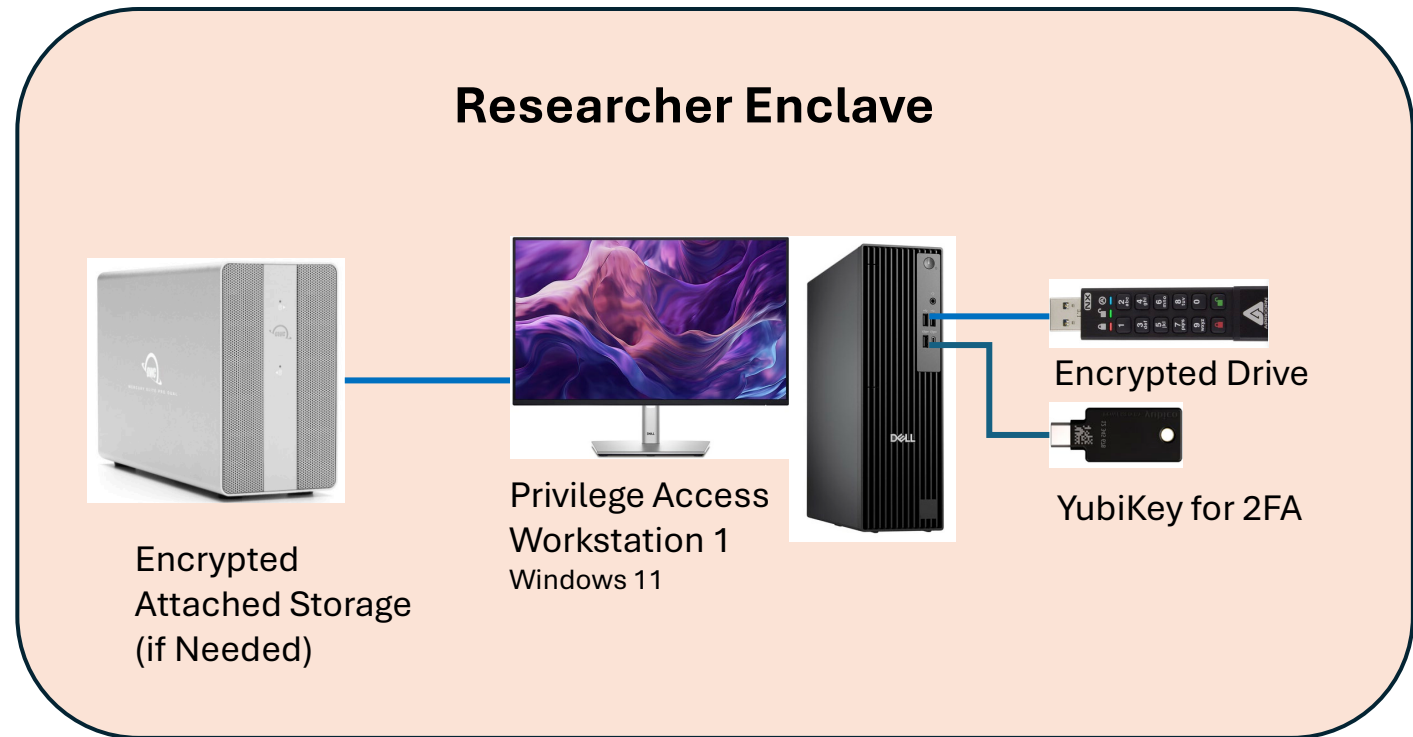
NIH Repositories

- Data access requests, i.e., NIH Controlled-Access Data Repositories (NIH CADRs).
 - dbGAP, SEER, NINDS, NIMH, NIDDK, etc.
- NIH has established **more stringent** and detailed security and operational standards
- Bottom Line: IT resources used to process or store this data must conform to NIST SP800-171 requirements

<https://grants.nih.gov/policy-and-compliance/policy-topics/sharing-policies/accessing-data/scientific>

Compliance with NIST SP800-171 requires special security posture

- Compliance achieved with
 - **separate enclave**
 - not connected to UM network
- Dedicated hardware
 - Privileged Access
 - Workstation (PAW)
 - encrypted external
 - hard drive
 - YubiKey for 2FA
- PAW configured to special STIG standard



Biospecimens – More Than Physical Objects

- A vial may be linked to clinical, demographic, and consent data
- Derived data can persist after the original specimen is depleted
- Genomic characteristics can implicate biological relatives
- Future analytical methods may reveal information not anticipated today

Coded Is Not Anonymous

Coded/Pseudonymized

- Identifiers are replaced with a code
- A key or other means can reconnect records to an individual
- Treat as sensitive and control the key separately

Anonymized/De-Identified

- The governing standard matters
- Residual linkage and inference risk should be assessed, particularly for rich or rare datasets

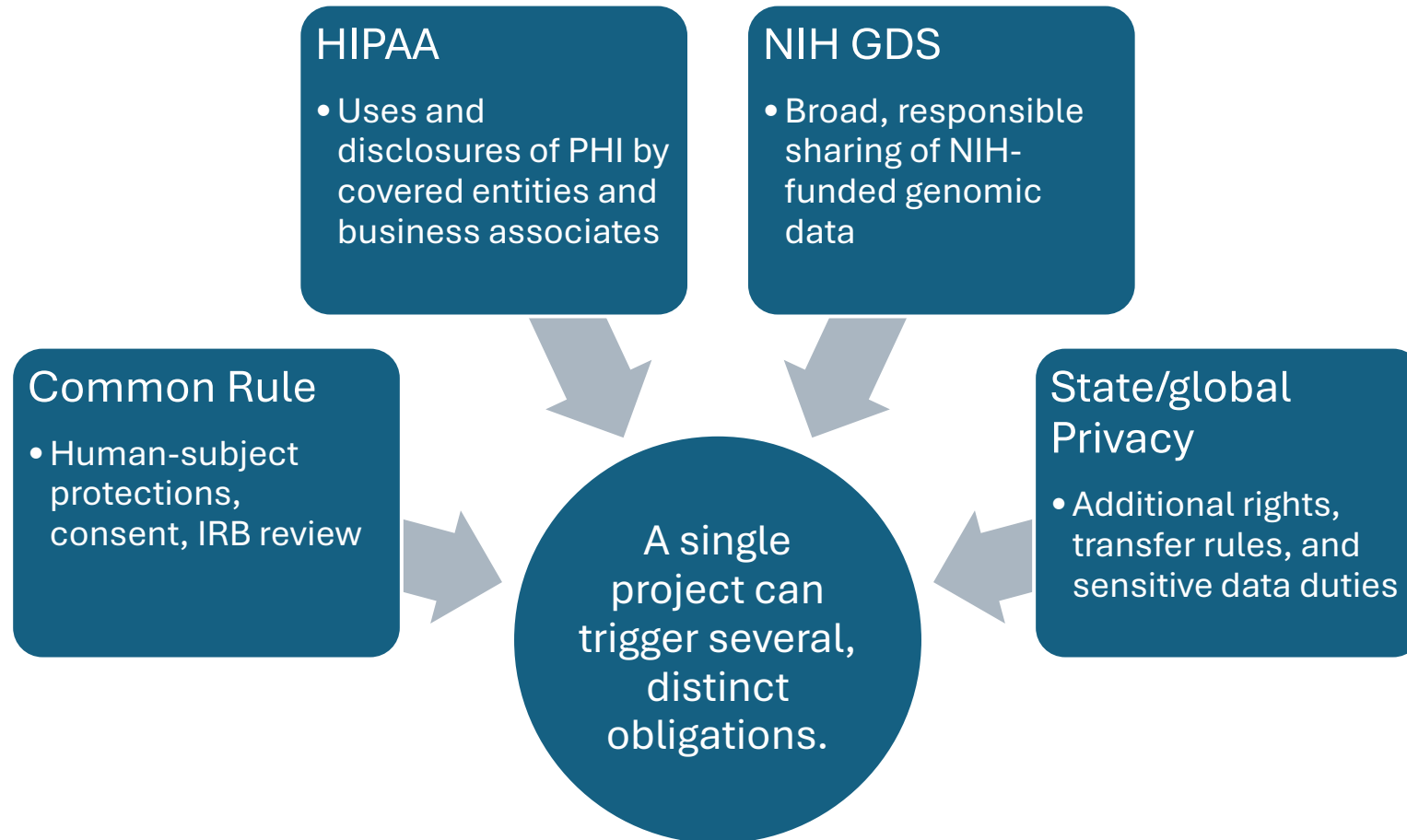
Re-identification Is a Systems Problem

- Rare combinations can single out people
- External datasets can change yesterday's risk assessment
- Genomic and longitudinal data resist simple anonymization
- Controls should address access, purpose, environment, and outputs

Consent Is a Data-Use Boundary

- What future uses were explained?
- Was broad sharing described?
- Are commercial uses compatible with the consent?
- Does withdrawal stop new use, future sharing, or both?
- Are return-of-results expectations clear?

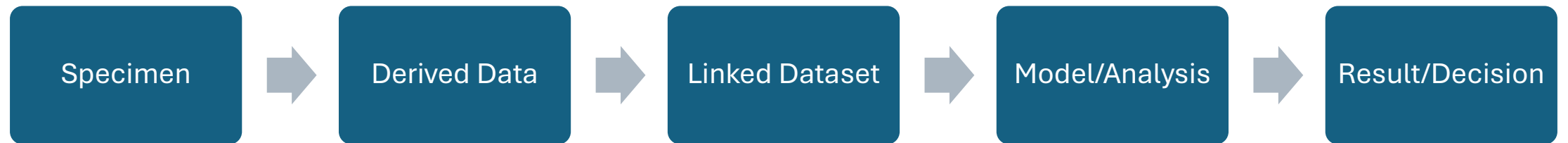
Regulatory Overlay



Examine Data in Action

Value emerges across linkage, analytics, AI and secondary use

From Specimen to Insight



Every analytical step creates new data, new decisions, and new risk



Why Big Data Changes the Research Equation

Scale

More records and modalities

Linkage

Richer context across sources

Velocity

Continuous or near-real-time flows

Reuse

New questions beyond original study

Big data -3Vs

Volume (amount) – 33 ZB to 175 ZB (2018-2025)

Velocity (speed of generation and processing)

Variety (types and formats).

1 ZB = 1,000,000,000 TB

1 TB = 1,000 GB

4 billion laptops (250 GB) - 1 ZB

1000,000,000,000 GB

Data Linkage: Both an Opportunity and a Risk



Scientific Power

Combine clinical outcomes, genomics, imaging, environmental exposure, behavior and social context to reveal interactions and support validation

Privacy Impact

Linkage can increase identifiability, expose sensitive relationships, create group harms, and enable uses participants may not reasonably expect

Secondary Use is Not a Single Decision

Consent
Compatibility

IRB/Regulatory
Status

Data-use
Limitations

Repository
Terms

Minimum
Necessary
Scope

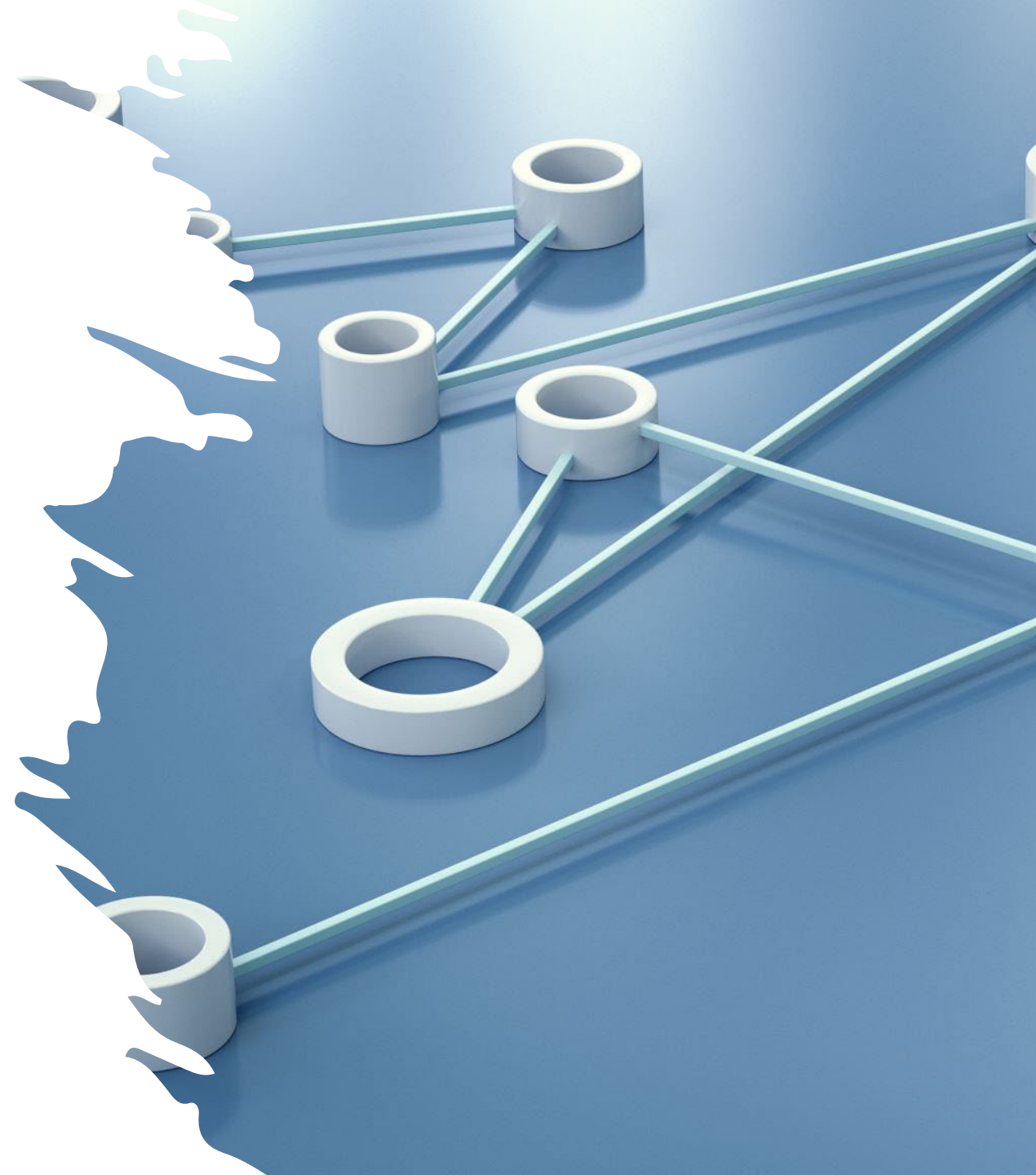
Return or
Destruction
Obligations

AI Expands the Data Boundary

- Training data may be copied, transformed, cached, or logged
- Model outputs can disclose or infer sensitive attributes
- Bias can be introduced by missing data and representation gaps
- Vendor terms and technical architecture affect the disclosure analysis
- Researchers remain accountable for validity and appropriate use

Federated Learning: Data Local, Risk Still Shared

- Participant-level data can remain within local environments
- Model updates and orchestration still require security analysis
- Local access, logging and oversight remain essential
- Aggregation does not eliminate inference, integrity, or contractual risk



Data Quality - Privacy Component

Accuracy

Wrong data can produce wrong decisions about a person

Representativeness

Underrepresentation can shift burdens and benefits

Provenance

Unknown lineage weakens permissible-use and validity claims

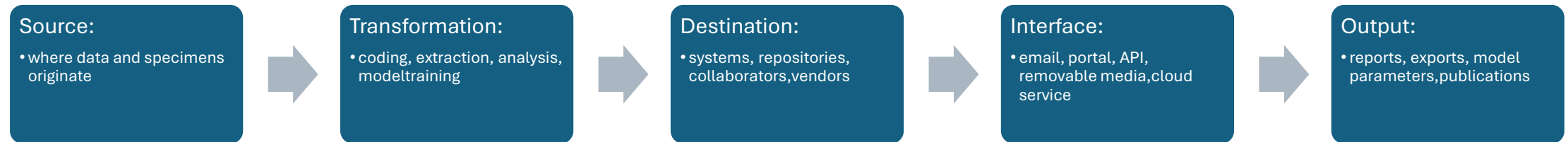
Drift

Data and populations change after deployment

Safeguard Data in Motion

Build controls around the real flow, including people, platforms, vendors, and outputs

Map the Actual Data Flow



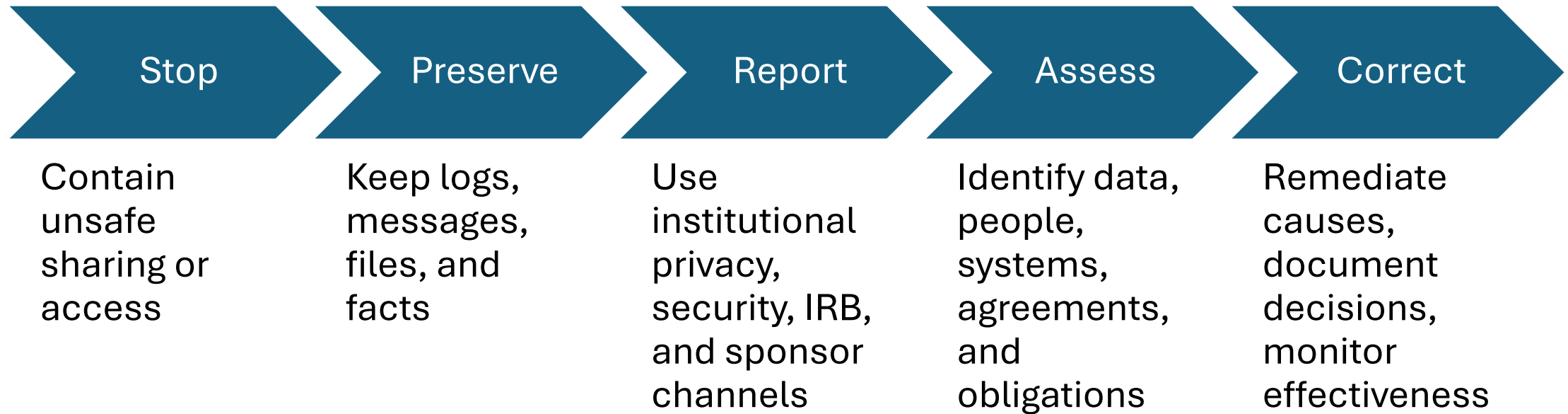
Vendor and Platform Due Diligence

- Where is the data stored and processed?
- Is submitted data retained, logged, or used to improve a service?
- Who are the subprocessors?
- What security, privacy, and audit evidence is available?
- How are incidents, deletion, export, and termination handled?
- Does the contract match the protocol, consent, and approved workflow?

Operational Safeguards Checklist

- Approved storage and transfer tools
- Encryption in transit and at rest
- Data minimization and field-level review
- Output checking before release
- Role-based access and timely deprovisioning
- Separate protection for code keys
- Audit trails and periodic access review
- Documented retention and secure disposal

Incident Readiness



Data Sharing and Compliance: Navigating DUAs, AI, and IRB Oversight

Tiffany-Ashley Disney, JD, MBA

Licensing Associate, Office of Technology Transfer

Overview of Office of Technology Transfer

The Office of Technology Transfer (OTT) is dedicated to assessing, protecting and commercializing UM owned Innovations through licensing to startups and established companies. Together with the rest of the U Innovation team, the OTT supports knowledge translation and commercialization of novel ideas, research discoveries, and innovations that improve the lives of our local, national and international communities.

Overview of Office of Technology Transfer

OTT is responsible for:

- Evaluating Innovation Disclosure Forms and working with applicable personnel to determine commercial potential and the most appropriate mechanism, if any, for protecting each Innovation.
- Working with Applicable Personnel and internal stakeholders to identify and engage potential commercial partners for their Innovations.
- The negotiation and execution of agreements pertaining to the development and commercialization of Innovations, including:
 - Exclusive and/or Non-exclusive licenses
 - Assignments
 - Outbound Material Transfer Agreements (MTA)
 - Confidential Disclosure Agreements (CDA)
 - Outbound Data Use Agreements (DUA)

How to Submit an Outbound DUA Request to the Office of Technology Transfer

To access the agreement portal:

- Go to <https://miami.inteum.com/miami/Minuet/InventorPortal/>
- Log in with your CaneID and password.
 - If this is your first time logging in, you will be prompted to create an account.
 - After the account is created, you will be redirected to the dashboard.
- Select “Request an Agreement” from the menu on the left.
- Select Outbound Data Use and follow the prompts for the requested information.
- If you have general inquiries or questions, please email techtransfer@med.miami.edu.

The IRB Controls What Data Can Be Sent

- **Must have an approved IRB prior to initiating the data transfer and signing the data use agreement**
- Additionally, the IRB must state that the transfer is allowed to the type of party
 - Only states academic collaborators, data cannot be transferred to industry partners
 - Only states a specific academic institution, data cannot be transferred to additional academic institutions
- For Jackson Data, the IRB must acknowledge the Master DUA between Jackson and UM

Special Considerations for AI

- IRB must ***specifically*** allow for the data to be used for machine learning, training, or model development
 - Even if the data is de-identified, the IRB must specify this use
- The more specific the use and scope of work is outlined the better
 - If the IRB states that the data will be used for machine validation, that is not an adequate description if the data will be used for model development

Data Use Agreement Versus License

- If data is being sent to a commercial entity, the IRB must allow for commercial use
 - Standard informed consent is likely inadequate
- Data to commercial entities should be under a non-exclusive license agreement versus a data use agreement

Sample of Language for AI Use

Your data may be used to help train and further develop the AI. Once shared, information cannot be withdrawn from the AI learning software and shared information may be available for future research and commercial interests. Your data will be kept confidential and any information used for development will not contain any identifiable information.

The use of your data and biospecimens may lead to new tests, drugs, devices, or other products or services with commercial value. These products or services could be patented and licensed. There are no plans to provide any payment to you should this occur.

Continuing Medical Education (CME)

Below please find the claim credit link for session scheduled for 10.15.2026: <https://miami.cloud-cme.com/cme/ClaimCredit?P=850&eventid=18756>



2026 IRB Grand Rounds - Continuing Nursing Education (CNE) Credit Registration Form



The University of Miami School of Nursing and Health Studies is accredited as a provider of nursing continuing professional development by the American Nurses Credentialing Center's Commission on Accreditation.



Upcoming IRB Grand Round

November 19, 2026 (10-11 AM)

Ready, Set, Comply!

1. Describe available resources to support protocol development
2. Maintain protocol compliance throughout the study lifecycle
3. Share practical tips and best practices for staying compliant

Speakers: IITSU, RQA, CTSI

*Registration information will be shared via **the flyer and the HSRO website.***

Thank you!

For additional information about today's presentation,

Please contact curatingconnection@miami.edu and reference

“IRB Grand Round 2026”.