

# Processing and Securing Research Data



# Presenters

Erik Deumens, PhD

Director

UF Research Computing

[deumens@ufl.edu](mailto:deumens@ufl.edu)

Marsha Pesch

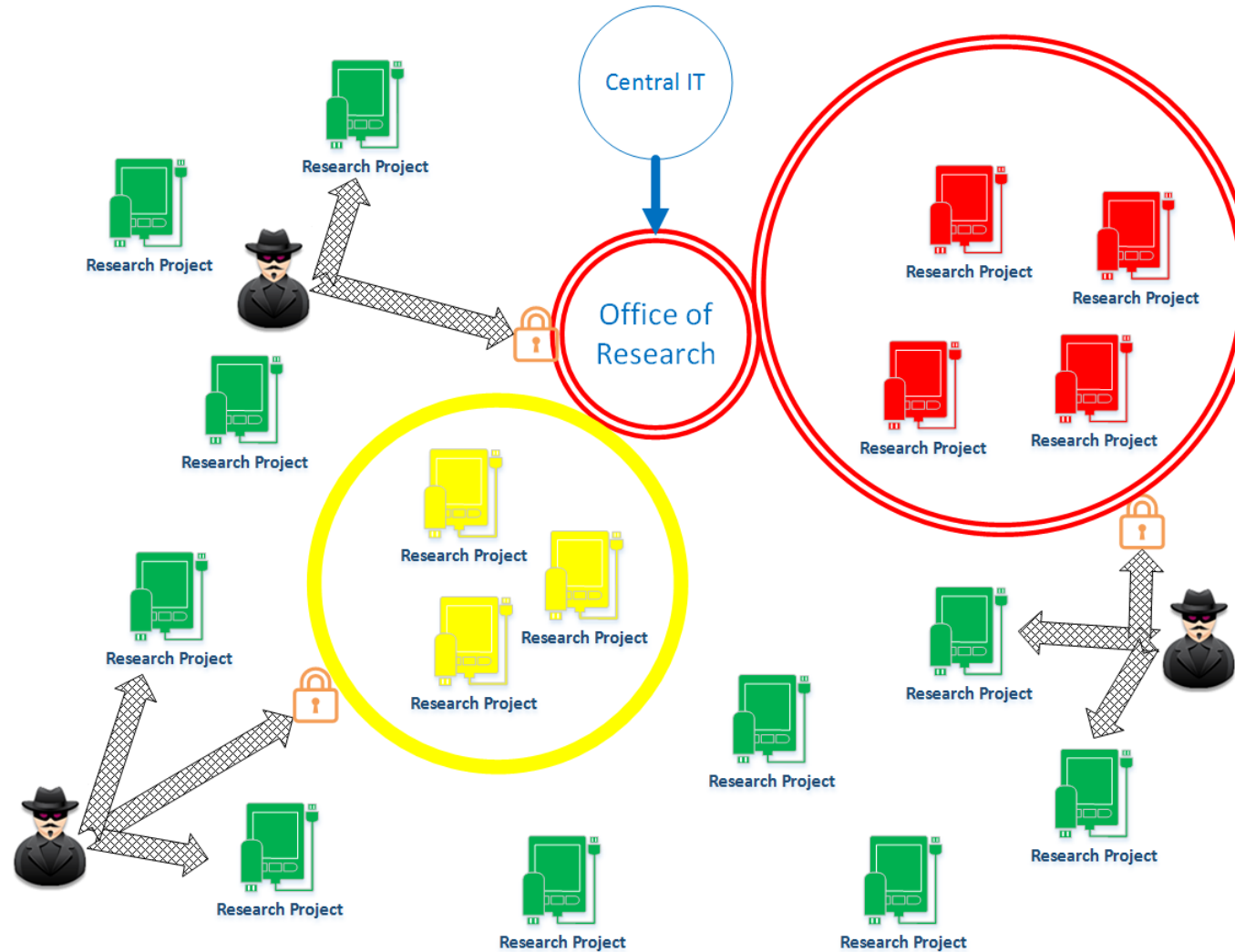
Assistant Director

Division of Research Compliance  
and Global Support

[mpesch@ufl.edu](mailto:mpesch@ufl.edu)



# Framing the Problem



# Role of College Research Administrators

- Office of Research deals with contracts
- Research Computing operates compliant infrastructures
  - But some workflows have parts that must be outside the compliant infrastructure
- College Research Administrators facilitate processes
  - Be aware of potential restrictions in RFP/BAA's
  - Be aware of faculty in your departments who process secure research
  - Prepare proposal budgets to handle IT security costs
  - Be aware of contract restrictions
  - Protect data received and marked as controlled
  - Understand controls in approved projects



# Regulations

Acronym	Regulation Title	Covers
ITAR	International Traffic in Arms (22 CFR 120-130)	Controls export and import of defense-related articles/services on the United States Munitions List (USML)
EAR	Export Administration Regulations (15 CFR 300-799)	Controls export and release of dual-use items listed on the Commerce Control List (CCL)
FISMA 2002 & FISMA 2014	Federal Information Security Management (Modernization) Act	Requires consistent standards to protect federal information and federal information systems
DFAR 252.204-7012	Safeguarding of Unclassified Controlled Technical Information	Required safeguards for Covered Defense Information (CDI) and Cyber Incident Reporting; NIST 800-171 applies; Safeguards required by Dec 2017
FAR 52.204-21	Basic Safeguarding of Contractor Information Systems	Required safeguards for information systems owned or operated by a contractor that processes, stores, or transmits federal information
32 CFR 2002	Controlled Unclassified Information	Required safeguards for federal information & systems: describes, defines, and provides guidance on the minimum protections for CUI

# Export Controlled Projects

- Fundamental Research – *Research in science, engineering, or mathematics, the results of which are intended to be published and shared broadly with the academic community (15 CR 734.8)*
- Fundamental Research can still involve export controlled technology or produce export controlled physical items.
- Controlled research is research which includes publication/dissemination restrictions or nationality restrictions.
- Controlled Research or items result in a TCP.

# Determining Controls

- Statements of control in contracts
  - Publication Restrictions
  - Nationality Restrictions
  - Determinations of Jurisdictions
  - Distribution Statements
- Export Control Reform
- ITAR – US persons only without license
- EAR – country specific for release based on CCL
- Additional contractual restrictions

# Technology Control Plans (TCP)

- Institutional Commitment
- Scope of Work
- Contractual Restrictions
- Jurisdiction and Classification Determination
- Physical Security
- Information Technology Security
- Project Personnel
- Certifications



# Initial Data Considerations

**Q1:** Is the PI collecting, storing, processing, analyzing, maintaining, or transporting any of the following data types?

**Q2:** What is the approximate total number of records in the dataset(s) that contain sensitive, restricted or regulated information?

**Q3:** There are many activities related to data that could introduce potential threats or risk to your research project and the institution. Several examples of data activities include:

- Primary data collection
- Secondary data analysis (e.g. merging data from existing datasets)
- Data transport to or from UF
- Data storage

Considering all activities related to the data, what are all the methods by which PI will collect, analyze, transmit or store data?

# Data Categorization

- DRCGS will determine if EAR or ITAR in conjunction with PI and contract clauses.
- Is Controlled Unclassified Information involved, if so NIST 800-171.
- Based on rules, what technical solutions are available?
  - ResVault
  - ResShield
- What does the PI and research team have to do?
  - Complete training & Rules of Behavior
  - Analyze data in the approved technical solutions
  - Follow data custodian chain of command

# Controlled Computing Environments

Restricted data is defined as data that is governed by laws, regulations, and/or contractual agreements (see the **UF Data Classification Policy**). To support research that involves restricted data, UF operates a Computing Environment for Restricted Data (CERD) to meet certain obligations to laws and contractual agreements. The details of UF's FISMA Program are described in the **CERD Handbook** and **specification of compliance and disclaimer**.



## ResearchVault

Computing Environment for Restricted Data (CERD) that is NIST 800-53 moderate and 800-171 compliant as of December 2017.



## ResearchShield

Computing Environment for Restricted Data (CERD) that is NIST 800-53 moderate as of November 2015. It will be CMS ARS 3.1 compliant by June 2018.

The **rates** for using the compliant environments for research with restricted data is calculated by counting computing and storage capacity.

<https://www.rc.ufl.edu/services/restricted-data/>



# Other Controlled Data

- Other restricted data
  - FERPA student data
  - HIPAA HITECH patient data – PHI – patient health information
  - PII – personal identifiable information
  - IP – intellectual property – is sensitive and is allowed on HiPerGator, but if you seek more controls then ResVault will provide that extra assurance e.g. to a pharmaceutical company funding the project



**HIPERGATOR**

**HiPerGator #3**  
MOST POWERFUL COMPUTER AT A U.S. PUBLIC UNIVERSITY

**HIPERGATOR HARDWARE SPECIFICATION SHEET**  
View the most recent version of the HiPerGator hardware spec sheet

HIPERGATOR 1.0:	HIPERGATOR 2.0:
<ul style="list-style-type: none"><li>◦ 21,000 cores</li><li>◦ 64 terabytes of RAM</li><li>◦ 2 petabytes of disk storage</li><li>◦ Maximum speed of 157 teraflops = 157 trillion floating point operations</li></ul>	<ul style="list-style-type: none"><li>◦ 30,000 cores</li><li>◦ 120 terabytes of RAM</li><li>◦ 1 petabyte of disk storage</li><li>◦ Maximum speed of 1,100 teraflops – 1,100 trillion floating point operations</li></ul>

<https://www.rc.ufl.edu/services/hipergator/>

# Budgeting for Data Services

## HiPer Gator Pricing

- Orange storage \$25/TB/year
- Blue storage \$140/TB/yr
- CPU \$44/NCU/yr
- GPU \$260/NGU/yr
- Tape backup \$78/TB/yr

## ResVault Pricing

- Replicated & backed-up storage \$220/TB/yr
- CPU for secure VM \$90/RNCU/yr
- GPU for secure VM \$260/RNGU/yr

Tips: Buying 5-year-life hardware can be cheaper

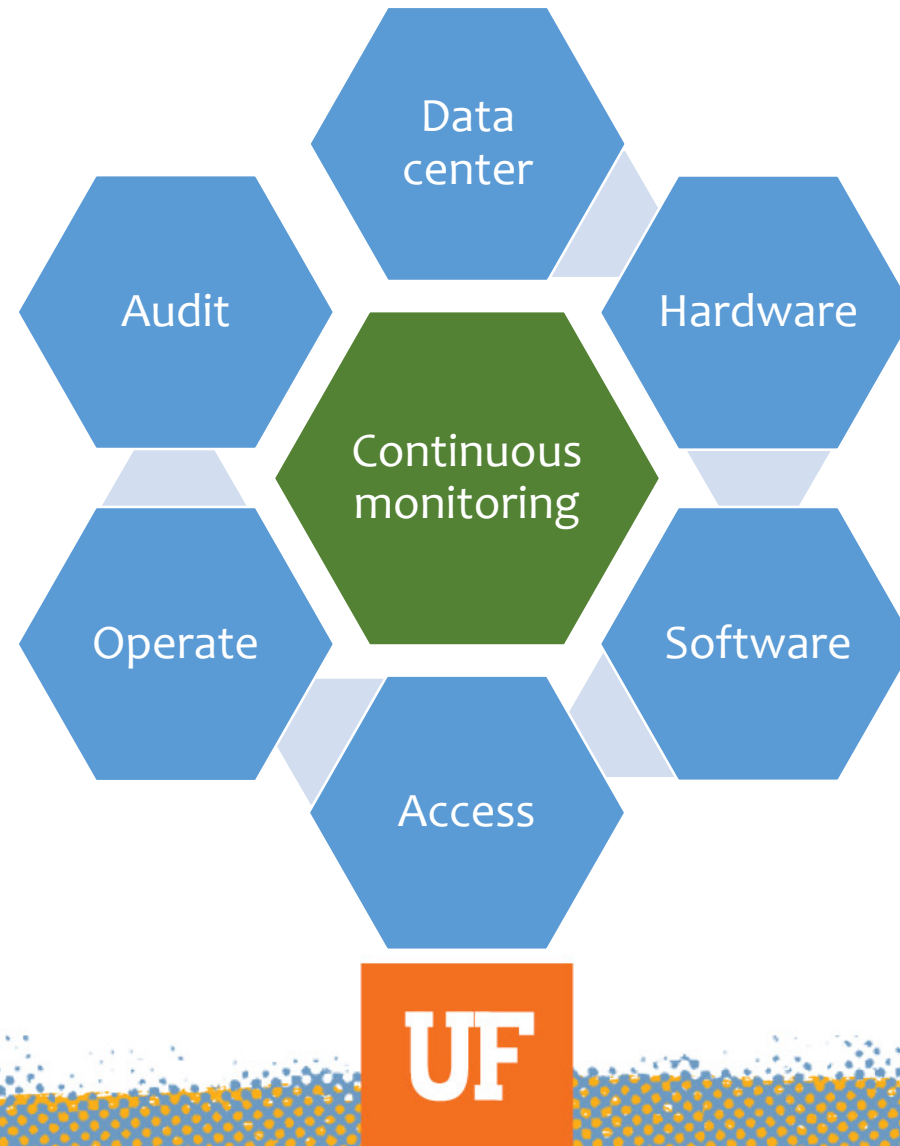
Find details at the Price Sheets under Services on the RC website <https://www.rc.ufl.edu/services/rates/>



# Individualized IT solutions

- General approach: leverage ResVault and add minimal extension
  - Example: a compute in a lab to test a circuit board and collect data, with the design and the data kept and processed in ResVault
- Exceptional needs for a complete system outside of ResVault because ResVault does not meet requirements
  - This takes (a lot) of time (3 to 6 months) and effort (several FTE)
  - Full Archer security review, all 260+ controls
  - Build system and implement all business processes to meet all 260+ controls
  - Pricing can be extensive and difficult to plan in advance

# Security Controls and Continuous Monitoring



# Additional Resources

- Research Computing Website: <https://www.rc.ufl.edu/>
- Division of Research Compliance and Global Support Website: <https://research.ufl.edu/compliance.html>

**Reach out to us, we're here to help:**

Division of Research Compliance &  
Global Support  
[exportcontrol@research.ufl.edu](mailto:exportcontrol@research.ufl.edu)

Phone: 352-392-9174

Research Computing Services  
Request Support Link  
<https://www.rc.ufl.edu/help/support-requests/>