

ROUND 1 - PROTECTING

Depends on where it's out

RD is wide bracket

Open Science vs Medical Data

Classify research data

PII, Research, Regulated

"Protect" preserve integrity
Preservation

Public availability?

Classification Question:

Not wall-based on particulars

Rules about where data can live

Classification are admin-centric

Regulated "Easier" due guidance

No guidance for microscope output

Data vs IP from publication

Protecting "Force fit" of device
to way normal enterprise ser

Cybersec + lifecycle - might
change depend on where you
are - e.g. Ligo + embargoed data

• Anonymized - net, host, logs
doing AI on those
Probs: student ids, et.

- IRB - consent of whole enterprise?
- UVA has policy but still needs IRB
GDPR makes this more fun.

- Research vs IT Net/Enterprise common
- Protect: Keep out vs Make available

Round 2 - Protecting

Enterprise → contain

Research → share

Doing research on enterprise
Data

How do we allow

Regulated vs Protected D.

Enterprise "Known"

RD - Hard to know what to expect

ED centralized; RD - spread out
hard to find

ED trends driving towards RD

e.g. Boss Intel, increasingly dispersed

Lines blurring

Clinical Trial

Protecting RZ

→ Subject vs generated data #2

Is this ED or RD?

Thinking @ ball of GDPR

→ QA data on patients

How diff from clinical trial

Ownership - e.g. student data

Student Data - repeated

→ RD - changing, evolving

Or is it?

Data access & storage methods.

Integrity of RD - conseq. of RD vs ED

Merging? Analytics on ED.

E.V. data lifecycle - always protected.

Using ED for Research

Enterprise App has business
rule (ACL) baked in
- Flat files don't

Availability

Protection data should inc
otherwise useless

- Mech to improve availability

- Value of data → where you
keep it open

- Faculty modified own data
"insider threat"

Protecting while sharing in
other institutions (Risk/Rew)

Data Governance big!

Right Granularity

Round 3 - Protecting

More than half don't think

→ RD is important to protect.

Just want to do their Research

Lot of talkover.

We want researchers to duplicate what we have.

"It's all open" MD5 sum

Students changing bit rat or
Changing data

Classes - could each other's
Don't understand change RD
= access to RD access to each;

TH: ZFA for HPC, R storage.

Failure to fund sec, which

Threatens to take over every

Budget

Community acceptance
of ZFL

- Govt don't often have sec
cost in budgets - e.g.

DMP

- Is sec heating + cooling?

Provided by F+A - need VPRs

to understand this

Ⓚ sec + Data storage in F+A

Core services

Need data break in

"Security slows my stuff down"