# Privacy Threat Modeling Genomic Workflows

Stuart Shapiro PhD, CIPP/US, CIPP/G

Principal

Nemo Jr Consulting

14 November 2025

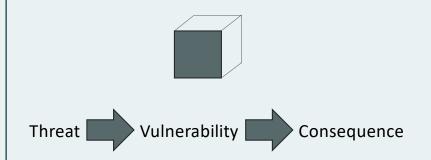
© 2025 Stuart S. Shapiro

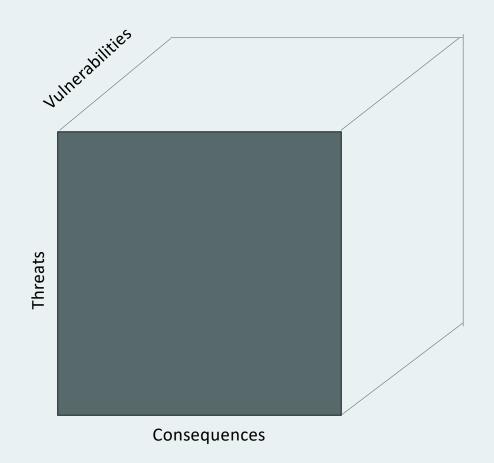
### Overview

- Privacy threat modeling
- NIST NCCoE Genomic Data Project
- Privacy threat modeling approach
- What are we doing?
- What could go wrong?
- What do we do about it?
- Conclusion

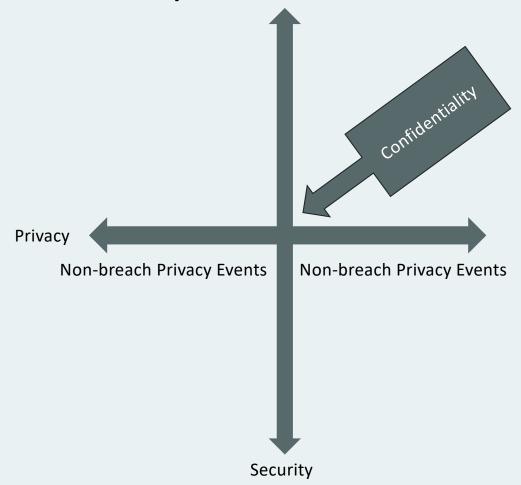
# Privacy Threat Modeling

### **Risk Composition**





### Privacy vs. Security Risks



### Privacy vs. Security Threats

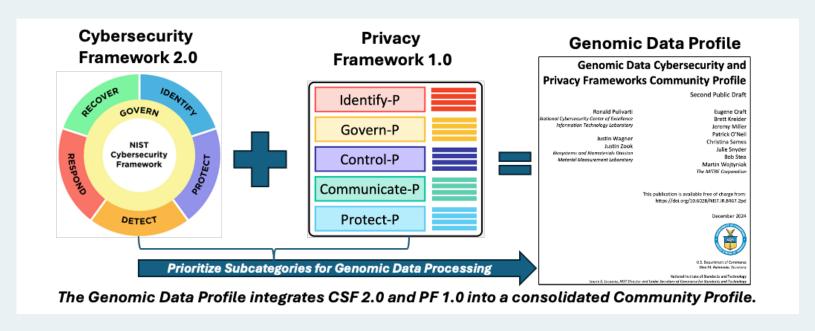


NIST National Cybersecurity Center of Excellence (NCCoE)Genomic Data Project

### **Project Summary**

- Risk-based cybersecurity and privacy guidelines for the genomic data community
- · Cross-stakeholder engagement
  - Government
  - Academia
  - Industry
- Genomic Data and Privacy Frameworks Community Profile
- Privacy Enhancing Technologies (PETs) Testbed
  - Initial focus: privacy-preserving federated learning
- NIST Special Publication (SP) 1800-43, Genomic Data Threat Modeling
  - Volume A: Executive Summary
  - Volume B: Cybersecurity Threat Modeling
  - Volume C: Privacy Threat Modeling

## Genomic Data Cybersecurity and Privacy Frameworks Community Profile

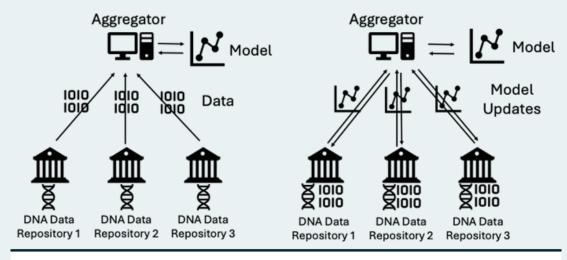


NIST NCCOE, "Cybersecurity and Privacy of Genomic Data Factsheet." Available at https://www.nccoe.nist.gov/sites/default/files/2024-12/genomics-fact-sheet.pdf

### **PETs Testbed**

### **Centralized Learning**

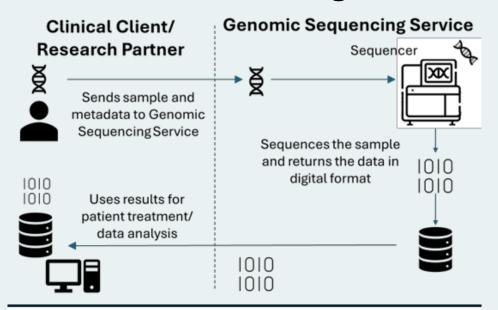
### **Federated Learning**



### **Centralized versus Federated Learning**

**Federated learning** trains machine learning models across multiple nodes. **PPFL** is a set of techniques to limit sharing private information across nodes.

### Genomic Data Threat Modeling

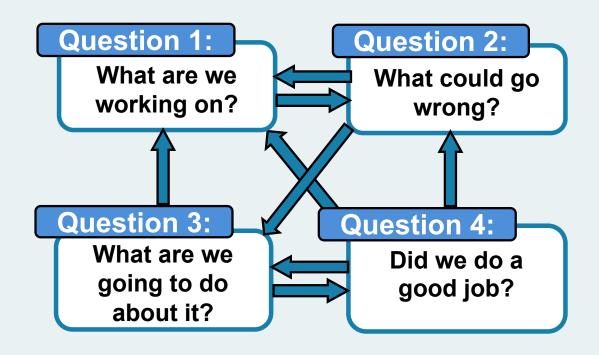


### The Genomic Data Sequencing Workflow

The **Clinical Client/Research Partner** sends a DNA sample to a **Genomic Sequencing Laboratory** that returns the digital results for patient treatment or further analysis.

### Privacy Threat Modeling Approach

### Four Questions Framework



 $https://pages.nist.gov/nccoe-genomic-data-threat-modeling/Vol\_C/Appendix/appendixC.html \\$ 

### Strategic View

### STRIDE

STRIDE Element	Description
Spoofing	Tricking a system into believing a false entity is a true entity
Tampering	Intentional modification of a system or data in an unauthorized manner
Repudiation	Disputing the authenticity of an action taken
Information Disclosure	Exposing information intended to have restricted access levels
Denial of Service (DoS)	Blocking legitimate access to the functionality of a system by malicious process(es)
Elevation of Privilege (EoP)	Gaining access to functions to which an attacker should not normally have access according to the intended security policy

The MITRE Corporation and Medical Device Innovation Consortium (MDIC) (2021), Playbook for Threat Modeling Medical Devices. Available at https://www.mitre.org/sites/default/files/2021-11/Playbook-for-Threat-Modeling-Medical-Devices.pdf

### Strategic View

### LINDDUN

LINDDUN Element	Description
Linking	Learning more about an individual (or a group) by associating related data items with one another
Identifying	Identity of an individual can be learned through leaks, deduced, or inferred when that is undesirable
Non-repudiation	An individual is unable to deny certain claims pertaining to them as a result of data collected, data shared, or a system action taken by the individual or others  Note that this directly conflicts with the repudiation threat type in STRIDE
Detecting	Becoming aware of an individual's involvement, membership, or participation via observation of relevant information
Data Disclosure	Avoidable transfer of an individual's data across a boundary, whether intended or unintended
Unawareness and unintervenability	Insufficiently informing, involving, or empowering the individual with respect to their role and relation to the system
Non-compliance	Lack of adherence to statutory or regulatory requirements or to standards or best practices

15

### **Tactical View**

### MITRE ATT&CK® ENTERPRISE FRAMEWORK

RECONNAISSANCE 10 techniques	RESOURCE DEVELOPMENT 8 techniques	INITIAL ACCESS 10 techniques	EXECUTION 14 techniques	PERSISTENCE 20 techniques	PRIVILEGE ESCALATION 14 techniques	DEFENSE EVASION 43 techniques	CREDENTIAL ACCESS 17 techniques	DISCOVERY 32 techniques	LATERAL MOVEMENT 9 techniques	COLLECTION 17 techniques	COMMAND AND CONTROL 18 techniques	EXFILTRATION 9 techniques	IMPACT 14 techniques
tive Scanning	Acquire Infrastructure	■ Valid Accounts ≡		Scheduled Task/Job		■ Modify Auther		System Service Discovery	Remote Services =	Data from Local System	Data Obfuscation =	Exfiltration Over Other	Data Destruction
ather Victim		Replication Through	Windows Management		Valid Accounts			k Sniffing	Software Deployment	Data from Removable	Fallback Channels	Network Medium	Data Encrypted for Impa
ost Information	Compromise Infrastructure		Instrumentation		Hijack Execution Flow			Application Window	Tools	Media	Application Layer Protocol ≡		Service Stop
lather Victim Identity		Trusted Relationship	Software Deployment			■ Direct Volume Access	input copture	Discovery	Replication Through Removable Media	-4	,	Data Transfer Size Limits	Inhibit System Recovery
formation		cabbal commonwhammer -	Tools		., ., .,	Rootkit		System Network Configuration Discovery	Internal Spearphishing		Communication Through Removable Media	Exfiltration Over C2 Channel	Defacement Firmware Corruption
ather Victim Network Iformation			Shared Modules User Execution			Obfuscated Files  or Information	Two-Factor Authentication Interception			Screen Capture Email Collection	Web Service		Resource Hijacking
	Acquire Access		Exploitation for Client				Exploitation for	System Owner/User Discovery	Use Alternate Authentication Material	Clipboard Data	Multi-Stage Channels	Exfiltration Over Physical Medium	Network Denial of Servi
Sather Victim Org Information	Acquire Access		Execution	External Remote Services	Process Injection		Exploitation for Credential Access	System Network	Lateral Tool Transfer	Automated Collection	Ingress Tool Transfer	,	Endpoint Denial of Servi
Phishing for Information		1 maning —			<ul> <li>Access Token Manipulation</li> </ul>		Steal Web Session Cookie	Connections Discovery	Taint Shared Content	Audio Capture		Exfiltration Over Web Service	System Shutdown/Rebo
Search Closed Sources			Command and Scripting		Abuse Elevation Control Me			Permission Groups	Exploitation of Remote	Video Capture			Account Access Remova
Search Open :		Content Injection	Interpreter =	Browser Extensions	Domain or Tenant Policy Mo	dification =	Credentials from	Discovery	Services	Browser Session Hijacking	Remote Access Software	Exfiltration Over	Disk Wipe
Fechnical Databases			Native API	Traffic Signaling	■ Escape to Host	Modify Registry	Password Stores	File and Directory	Remote Service Session _	Data from Information	■ Dynamic Resolution ■		Data Manipulation
Search Open =			Inter-Process	BITS Jobs	Exploitation for Privilege	Trusted Developer Utilities ≡	Steal or Forge Kerberos	Discovery	Hijacking =	Repositories	Non-Standard Port	Transfer Data to	Financial Theft
Vebsites/Domains			Communication =		■ Escalation	Proxy Execution	Tickets	Peripheral Device		Adversary-in-the-Middle	■ Protocol Tunneling	Cloud Account	
Search Victim-Owned			Container Administration	Component		Traffic Signaling =	Forced Authentication	Discovery		Archive Collected Data	■ Encrypted Channel ■		
Websites			Command	Pre-OS Boot	■	Signed Script Proxy	Steal Application	Network Share Discovery		Data from Network	Non-Application Layer		
Has sub-techniques			Deploy Container	Compromise Client		Execution	Access Token	Password Policy Discovery		Shared Drive	Protocol		
= rias sub-teciniques			Serverless Execution	Software Binary		Rogue Domain Controller	Adversary-in-the-Middle ≡			Data from Cloud Storage	Hide Infrastructure		
			Cloud Administration Command	Implant Internal Image	_	Indirect Command Execution		Discovery			Content Injection		
			Communo	Modify Authentication : Process	≡	BITS Jobs	Multi-Factor Authentication Request Generation	Virtualization/Sandbox Evasion		Repository			
					-	XSL Script Processing		Cloud Service Dashboard					
				Power Settings		Template Injection	Steal or Forge Authentication Certificates	Software Discovery =					
						File and Directory		Query Registry					
						Permissions Modification		Remote System Discovery					
						Virtualization/Sandbox ≡		Network Service Scanning					
						Evasion		Process Discovery					
						Unused/Unsupported		System Information					
						Cloud Regions		Discovery					
						Use Alternate  Authentication Material		Account Discovery					
								System Time Discovery					
						Impair Defenses  Hide Artifacts		Domain Trust Discovery					
						Masquerading =		Cloud Service Discovery					
						Deobfuscate/Decode Files		Container and Resource Discovery					
			A = = 0	01/0		or Information		Cloud Infrastructure					
	N	NTRE	Al I&C	JK .		Signed Binary Proxy		Discovery					
						Execution		System Location Discovery  Cloud Storage Object					
	Fr	nterprise	Framey	<i>y</i> ork		Exploitation for Defense Evasion		Discovery					
						Execution Guardrails		Group Policy Discovery Debugger Evasion					
	atta	ck.mitre.org				Modify Cloud Compute =		Device Driver Discovery					
						Pre-OS Boot		Log Enumeration					
						Subvert Trust Controls							
						Build Image on Host							
						Deploy Container							
						Modify System Imper							

© 2024 MITRE - Framework version v15, current as of April 2024

MITRE | SOLVING PROBLEMS FOR A SAFER WORLD

### **Tactical View**

All Surve	PANOP	TIC™ V2	2.0	Privacy Activities (1/2)			
Dut of   Sequence						Quality	
PA01.03	Out of	Out of	Application or	Insufficient	Implicit		
PA01.04				Insufficient	Re-	Unvetted data	
PAGE 105			Tracking & affording	Undermining or interfering with	Identifier	Unvetted data	
			Sniffing & affording	Detection		Bias of date not	
Mileseding  No opt-out/his   External properties   External properties   External paper opt-out-out-out-out-out-out-out-out-out-ou				Misconfigured	Pseudo-	Unvetted date	
No award opt- active ac	Misleading/		External			Unvetted	
No granular   Soliciting & Insufficient   Insuff		No overed opt-			Compulsory self-	Unvetted downstream practices	
Inherited 2" party Data Collects 1" Insufficiently de decisited 1 Insufficiently de decisited 2 Insufficiently aparty Data cut of scope 3" party Data cut of scope 2 Insufficient 2 Insuff		No granular	Soliciting & affording			Insufficient communication of downstream	
3" party società 2" scope scop			2 <sup>nd</sup> party solicits 1 <sup>st</sup>			Data insufficiently de-	
3" perty  Solicite fri of soape un out of soape un			3" party solicits 2"			Data out of	
Recording			3" party solicits 1"			Data action out	
Transaction						Data collection	
Biological Class sharing out of scope o						Data processing	
Extraction Insufficient agreed usage restrictions PAGS.13 Logal			Biological			Data sharing ou	
			Extraction			Insufficient agreed usage	

17

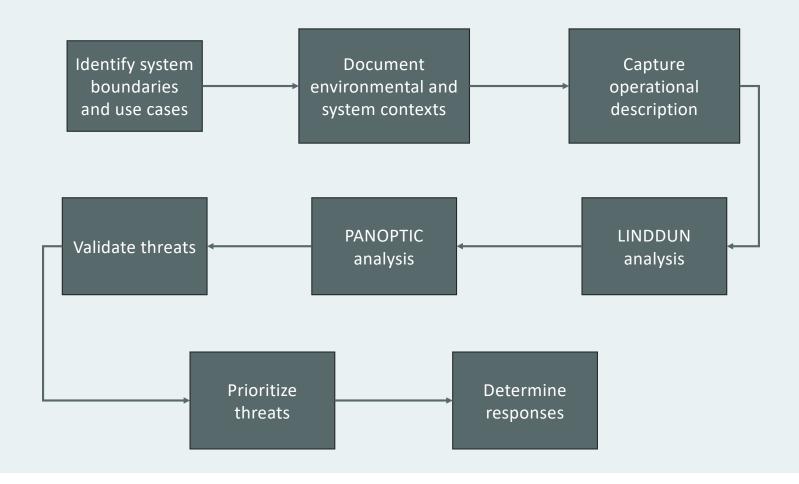
https://ptmworkshop.gitlab.io/#/panoptic

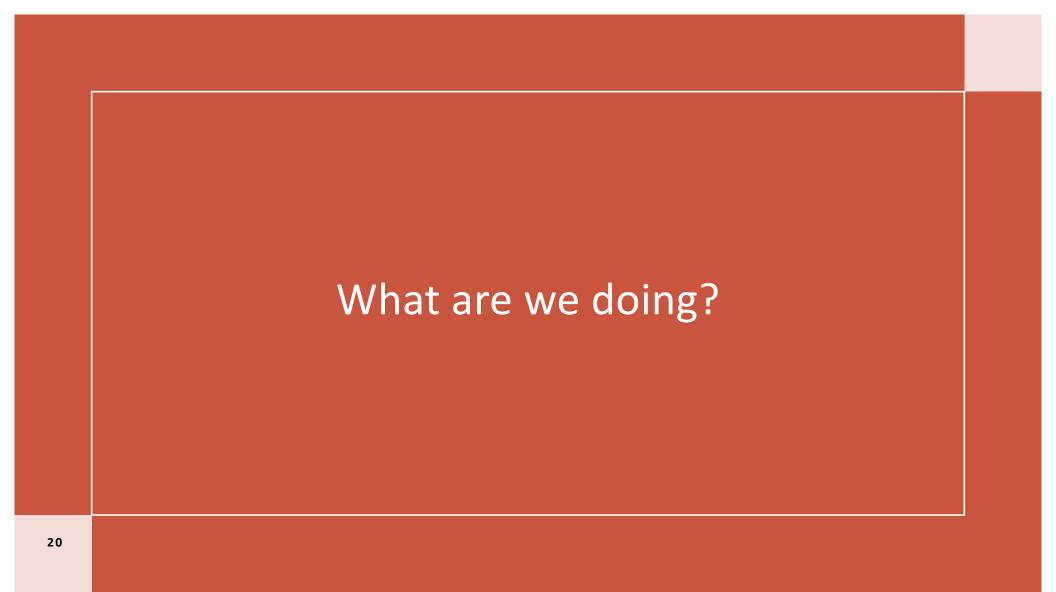
### **Supporting Players**

- NIST Privacy Risk Assessment Methodology (PRAM)
- NIST Privacy Framework v1.0
- NIST Genomic Data Profile
- NIST Special Publication 800-53r5
- NIST Privacy Framework SP 800-53r5 mapping

### Process Flow (Questions 1-3)

19





### **Environmental Context**

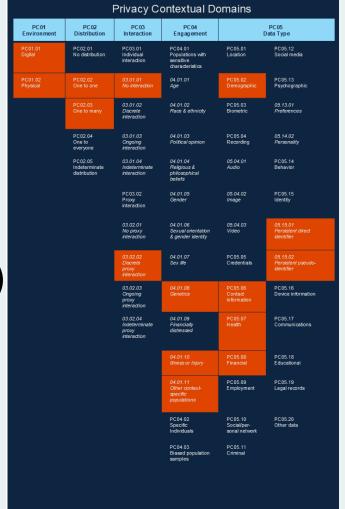
- PRAM Worksheet 1
  - Task 1: Frame organizational objectives
    - Mission needs, functional capabilities, privacy-related goals
  - Task 2: Frame organizational privacy governance
    - Legal environment, privacy-related commitments and policies, risk tolerance
- PRAM Worksheet 2
  - Contextual factors for organization
  - Contextual factors for individuals
- For each use case: clinical and research

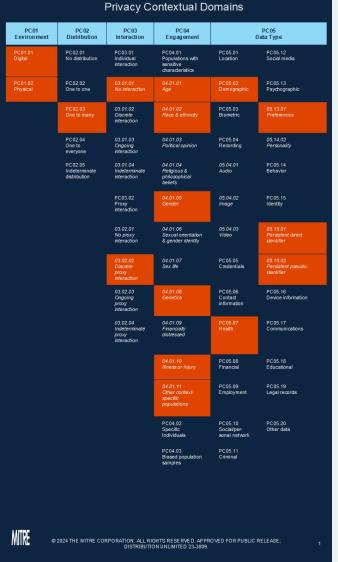
### System Context (1/2)

- PRAM Worksheet 2
  - System privacy capabilities for
    - Predictability
    - Manageability
    - Disassociability
  - System contextual factors
- For each use case: clinical and research
- Privacy threat: "[A]ny circumstance or event with the potential to compromise the predictability, manageability, and/or disassociability of systems involving data associated with individuals."\*

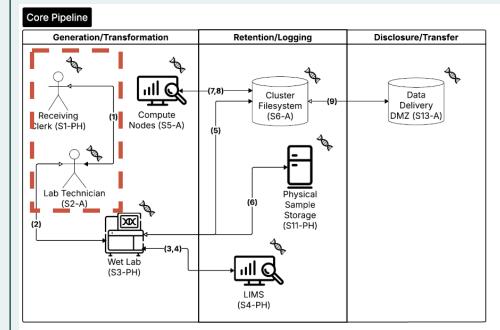
System Context (2/2)

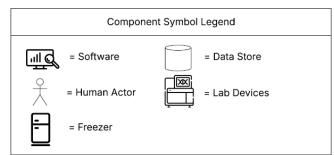
23





### Operational Description\*





### Data Action Key

- Receiving Clerk delivers the biological sample to Lab Technician
- Lab Technician runs the biological sample through the Wet Lab to be transformed into digital data
- The LIMS is used to track the sample before being de-identified so that the results can then be re-identified after the data has been generated
- The LIMS returns a de-identified ID to be used for the sample
- 5. Digital data are transferred to the Cluster Filesystem for use and retention
- Excess physical samples are moved to Physical Sample Storage for retention according to contextual physical sample retention needs
- 7. Compute Nodes operate on data to transform it into objective-specific data
- Objective-specific data is moved from the Compute Nodes and saved on the Cluster Filesystem
- Objective-specific data is copied to the Data Delivery DMZ, while the data still remains on the Cluster Filesystem according to contextual data retention needs

Management Symbol Legend



SequencingLab Managed



### LINDDUN Analysis\*

No.	Source	Dataflow Type	Data Action 1	Data Action 2	Destination	Context (Purpose)		e LINDDUN eats
1	Receiving Clerk (S1-PH)	Physical sample	Transfer		Lab Tech (S2-A)	Send physical sample to lab tech for research project	L2.2.1	Sending samples to wet lab known to be researching a specific disease at that time could link samples to that disease

### Privacy Activities (1/2) 05.01.01 Re-identification PA06.02 Unvetted data source PA01.03 Imprecise PA01.04 Absent PA03.04 Sniffing & affording sniffing 06.03.01 Bias of data not evaluated PA04.05 Misconfigured permissions **PANOPTIC** Analysis 06.09.02 Data processing out of scope PA03.13 Legal proceeding © 2024 THE MITRE CORPORATION. ALL RIGHTS RESERVED. APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED 23-3809.

PA07 Manageability	PA08 Aggregation	PA 09 Processing	PA 10 Sharing	PA11 Use	PA12 Retention & Destruction	PA 13 Deviations
PA07.01 No individual access to information	PA08.01 Profiling	PA09.01 Deriving new information	PA10.01 Affording revelations	PA11.01 Implication	PA12.01 Data not destroyed after use	PA13.01 Deviating from usage restrictions
PA07.02 No individual management of information content	08.01.01 Single source profiling	09.01.01 Deriving information about individuals	PA10.02 Exposure	PA11.02 Targeting	PA12.02 Data improperly destroyed	PA13.02 Deviating from stated policy or user agreement
PA07.03 No individual deletion of information	08.01.02 Multi-source profiling	09.01.02 Deriving aggregate information	10.02.01 Doxing	11.02.01 Tailored content		PA13.03 Deviating from claimed certification conformance
PA07.04 No individual control of information disclosure	PA08.02 Clustering	09.01.03 Deriving sensitive information	PA10.03 Mis- representation	PA11.03 Manipulation		PA13.04 Deviating from regulatory requirements
PA07.05 No individual control of information use	08.02.01 Single source clustering	09.01.04 Deriving derogatory information		11.03.01 Extortion		
PA07.06 Privacy configurations compromised by outside forces	08.02.02 Multi-source clustering	PA09.02 Behavioral analysis		PA11.04 Intrusion		
PA07.07 Confounded user controls		PA09.03 Introducing bias		PA11.05 Revenue		
PA07.08 Bypass of user controls		PA09.04 Trawling datasets for information		PA11.06 Reprisal		
PA07.09 Pre-emption of privacy settings		PA09.05 Internal appropriation				

### Validation\*

PA01	PA02	PA03	PA04	PA05	PA06	PA07	PA08	PA09	PA10	PA11 Use	PA12	PA13
Notice	Consent	Collection	Insecurity	Identifica- tion	Quality Assurance	Manage- ability	Aggrega- tion	Processing	Sharing		Retention & Destruction	Deviations

Attack	Scenario	Threat	LINDDUN Analysis	Implicated Privacy Engineering Objectives
PA03.09 PA03.11 PA08.01.01 PA10.01 PA11.01	S1.1	L2.2.1	Sending samples to wet lab known to be researching a specific disease at that time could link samples to that disease	Predictability Disassociability

PA03.09	PA03.11	PA08.01.01	PA10.01	PA11.01
Recording	Biological	Single source	Affording	Implication
	sample	profiling	revelations	

What are we going to do about it?

### Prioritization\*

Difficulty	Negligible	Minor	Moderate	Significant	Severe
Feasibility					
Plausible	1.0	0.8	0.6	0.4	0.2
Indeterminate	0.9	0.7	0.5	0.3	0.1
Implausible	0.8	0.6	0.4	0.2	0.0

LINDDUN Threat Type	Weight
Data Disclosure	1.0
Identifying	0.85
Linking	0.7
Non-compliance	0.5
Unawareness and Unintervenability	0.5
Detecting	0.3
Non-repudiation	0.2

No.	LINDDUN Threat	Feasibility	Difficulty	Ranking Value
55	DD.4.1.2	Plausible	Minor	0.80
5	L.2.1.2	Plausible	Minor	0.56
26	1.2.1.1	Plausible	Moderate	0.51
1	L.2.1.2	Plausible	Moderate	0.42
3	L.2.1.2	Plausible	Moderate	0.42
4	L2.1.2	Plausible	Moderate	Ø.42
14	L.2.2.1	Plausible	Moderate 1	0.42
15	L.2.2.1	Plausible	Moderate	0.42
65	U.1.1	Plausible	Minor	0.40
2	L.2.1.2	Plausible	Significant	0.28

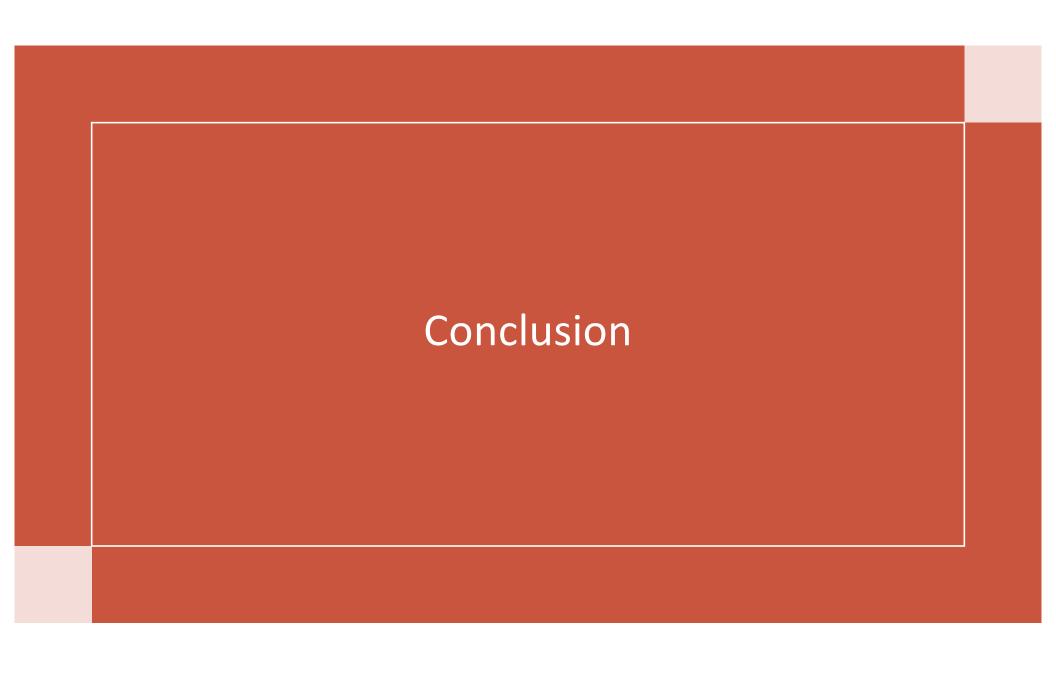
Physical samples or raw sequencing data

### Options

- Eliminate
- Disrupt
- Transfer responsibility
- Accept

### Potential Controls\*

PANOPTIC Threat Action	Privacy Framework Sub-Categories	SP 800-53 Controls
PA08.01.01 Aggregation: Profiling: Single source profiling	Disassociated Processing CT.DP-P2 Data are processed to limit the identification of individuals [1 2 1 2]	Identification & Authentication IA-4(8) Pairwise Pseudonymous Identifiers — Generate pairwise pseudonymous identifiers.  System & Information Integrity SI-12(1) Limit Personally Identifiable Information Elements — Limit personally identifiable information being processed in the information life cycle to the following elements of PII: [Assignment: organization-defined elements of personally identifiable information].
	Protective Technology PR.PT-P2 The principle of least functionality is incorporated by configuring systems to provide only essential capabilities [3 2 2 2]	Configuration Management CM-7 Least Functionality – Configure the system to provide only [Assignment: organization-defined mission essential capabilities].
	Disassociated Processing CT.DP-P3 Data are processed to limit the formulation of inferences about individuals' behavior or activities [2 3 2 2]	Audit & Accountability AU-16(3) Disassociability – Implement [Assignment: organization-defined measures] to disassociate individuals from audit information transmitted across organizational boundaries.



### **Takeaways**

- Privacy threat modeling identifies privacy threats to data subjects, including those posed by the system itself
- Potential risk is addressed at the earliest possible point: at the origin of the risk chain
- Privacy threat modeling complements cybersecurity threat modeling but is distinct

Ordinary mortals can do this, in a reasonable amount of time with a reasonable amount of effort

It's not a zero-sum game

### Resources

- NIST Special Publication 1800-43C (ipd): https://www.nccoe.nist.gov/sites/default/files/2025-08/nist-sp-1800-43c-draft.pdf
  - Appendices: https://pages.nist.gov/nccoe-genomic-data-threat-modeling/
- LINDDUN: https://linddun.org/
- MITRE PANOPTIC™: https://ptmworkshop.gitlab.io/#/panoptic
- NIST Privacy Risk Assessment Methodology (PRAM): https://www.nist.gov/document/nist-pramfeb2019zip
- NIST Privacy Framework v1.0: https://doi.org/10.6028/NIST.CSWP.01162020
- NIST Genomic Data Community Profile (2pd): https://doi.org/10.6028/NIST.IR.8467.2pd
- NIST Special Publication 800-53r5: https://doi.org/10.6028/NIST.SP.800-53r5
- NIST Privacy Framework SP 800-53r5 mapping: https://csrc.nist.gov/files/pubs/sp/800/53/r5/upd1/final/docs/csf-pf-to-sp800-53r5-mappings.xlsx

Stuart Shapiro

+1-603-365-6296

sshapiro@nemojr.com

www.nemojr.com

# Thank you