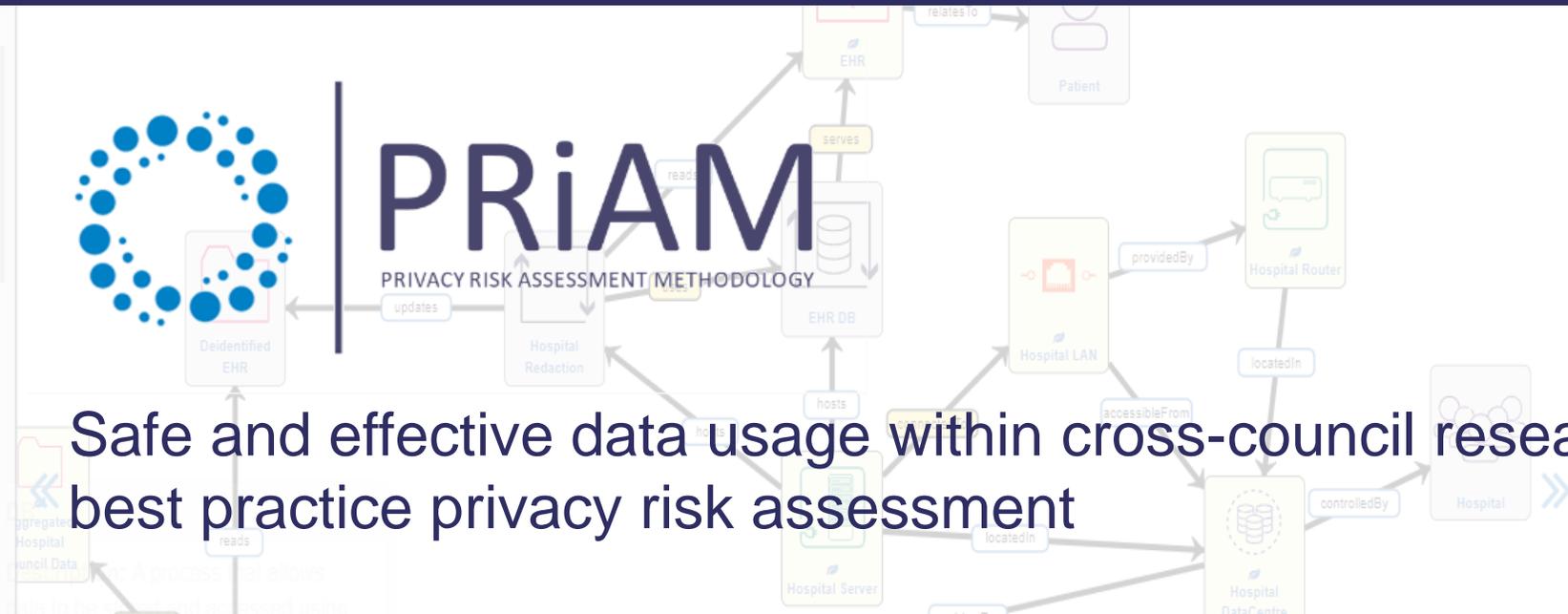


# PRiAM

PRIVACY RISK ASSESSMENT METHODOLOGY



Assets (92) ?

Controls (67) ?

Adverse Effects and their Impact (195) ?

Show filters Reset sort

Name	Asset	Impact	Likelihood	Risk
Loss of Confidentiality	EHR	High	Very High	Very High
Loss of Confidentiality	Benefits Record	High	Very High	Very High
Loss of Confidentiality	Benefits Record	High	Very High	Very High
Loss of Integrity	Benefits Record	High	Very High	Very High
Loss of Availability	EHR	Medium	Very High	High
Infected by Malware	Council Router	Very Low	Very High	Low

Safe and effective data usage within cross-council research networks through best practice privacy risk assessment

Presented by Professor Michael Boniface ([m.j.boniface@soton.ac.uk](mailto:m.j.boniface@soton.ac.uk))  
 University of Southampton, IT Innovation Centre  
 Interim Progress - April 2022

# Motivation

- Research to improve health and wellbeing increasingly depends on **combing diverse data from multiple organisations**
- However, “..the **use of data presents risks**; those risks need to be fully understood and taken into account“,  
*UK’s National Data Sharing Strategy, DCMS*
- Even with shared principles for safe data usage, **privacy risk management is still vague**
  - no consistent guidance for risk assessment, mitigation and management
  - resulting in different implementations of Trusted Research Environments
- A **common way to assess privacy risk** is needed

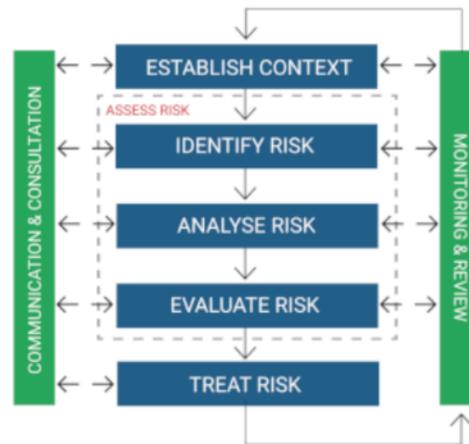


# Approach

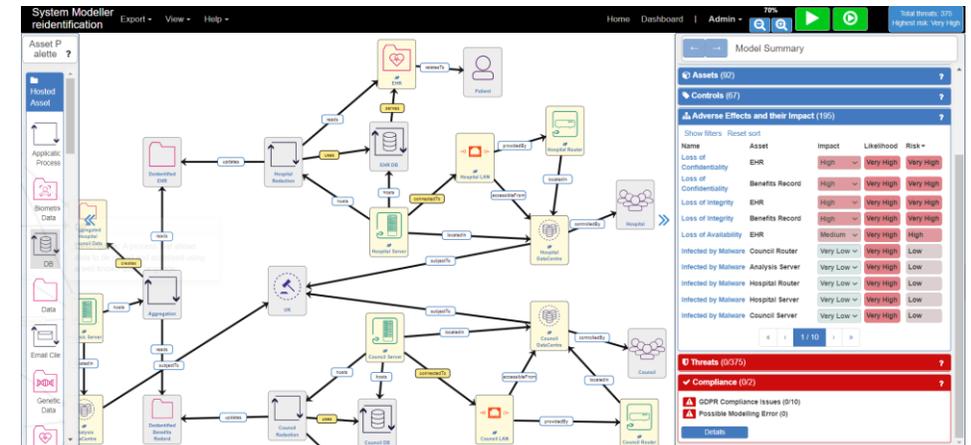
- We aim to published a best-practice **privacy risk assessment framework** that can describe and assess privacy risk for **safe data usage** in research networks
- We will bring together well-known principles for safe research - the **Five Safes** with methodology for information security risk management (**ISO 27005**) to enable consistent, efficient and usable privacy assessment



principles



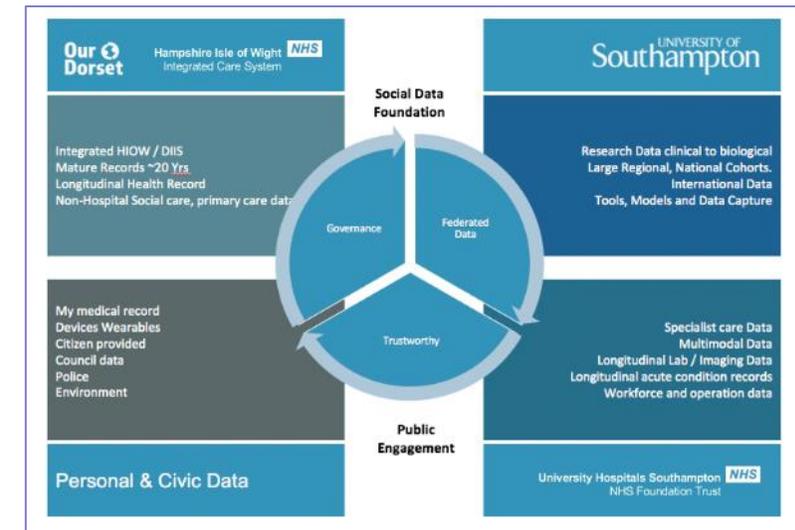
risk assessment and management



risk modelling of systems

# Objectives

- Analyse **driver use cases** in public health prevention and integrated care
- Identify **factors contributing to privacy risks** within the Five Safes
- Define a **framework** to provide a consistent methodology for privacy risk assessment
- Assess privacy risks for use cases using a cyber security **risk modelling and simulation** platform
- Codesign and evaluate the framework, modelling and simulation through **engagement with the public and multidisciplinary stakeholders**



Source: Wessex Trusted Research Environment (NHSx)

# Privacy Requirements for Safe Federations

- Explore context of privacy risks for federated research networks
  - address multiple **interpretations** of principles
  - consider multiple **perceptions** of risk
  - elaborate **harms** related to federation
  - focus on **information privacy**
  - define **privacy goals** including CIA, acceptability, intervenability, transparency and unlinkability
  - identify of **privacy controls**
- Introduce the principle of ‘**safe federation**’
  - Protocols for commitment from parties over goals, standards, success measures, costs, benefits and value creation
  - Benefits -> *local control, risk mitigation, large data, potential reduction in costs, cross border working*
  - Challenges -> *decision making complexity, new risks from infomediaries, new approaches to federated controls (e.g. intervenability)*
- Define of operational/functional privacy requirements for safe federations

Table 1: Different interpretations of Five Safes

Five Safes Framework	Original Five Safes	HDRUK [7] Interpretation	AHW (2021) [18] Interpretation	UK Data Service, SecureLab (2022) [19] Interpretation	Arbuckle and Ritchie (2019) [20] Interpretation
<b>Safe projects</b>	“Is this use of the data appropriate?”	“Data is only used for ethical, approved research with the potential for clear public benefit”	“Use of the data is legal, ethical and the project is expected to deliver public benefit”	“research projects are approved by data owners for the public good”	“Will personal data be anonymized? What are the legal/ethical boundaries?”
<b>Safe people</b>	“Can the researchers be trusted to use it in an appropriate manner?”	“Only trained and specifically accredited researchers can access the data”	“Researchers have the knowledge, skills and incentives to act in accordance with required standards of behaviour”	“researchers are trained and authorised to use data safely”	“Evaluate recipient trust, and manage their motives”
<b>Safe data</b>	“Is there a disclosure risk in the data itself?”	“Researchers only use data that have been de-identified to protect privacy”	“Data has been treated appropriately to minimise the potential for identification of individuals or organisations”	“data is treated to protect any confidentiality concerns”	“To determine the data transformations necessary to deal with residual risk, we need to understand the risk from the data”
<b>Safe settings</b>	“Does the access facility limit unauthorised use?”	“Access to data is only possible using secure technology systems – the data never leaves the TRE”	“There are practical controls on the way the data is accessed – both from a technology perspective and considering the physical environment”	“a SecureLab environment prevents unauthorised use”	“Assess security and privacy controls of the recipient”
<b>Safe outputs</b>	“Are the statistical results non-disclosive?”	“All research outputs are checked to ensure they cannot be used to identify subjects”	“A final check can be required to minimise risks leaving the project”	“screened and”	“Evaluate context”
<b>Key purpose</b>	To be used as discussion points about data access	Presented as key principles of TREs	Presented dimension with “pot to be miti		



+ Acceptability

Table 2: Identified Privacy Requirements of 'Safe federation' (O:operational, F:functional)

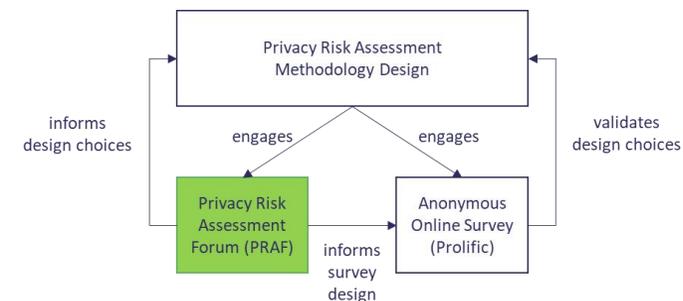
ID	Privacy Requirement for Federated Research Network	Related to Extended Five Safes Framework	Reference
O.1	<b>Standardised procedures for assessing outputs</b> , including (but not limited to): 1. Proposed statistical outputs 2. Proposed qualitative outputs 3. Other proposed outputs, such as (but not limited to) metadata, algorithms, workflows, models and software	Safe outputs	[36][40][41]
O.2	<b>Standardised appeals procedure for rejected outputs</b>	Safe outputs	[41]
O.3	<b>Standardised procedure to block and/ or embargo proposed outputs</b>	Safe outputs	[40]
O.4	<b>Standardised procedures to measure and evidence the benefit of approved outputs</b> (on release from a TRE) for individuals, communities and society by those appointed responsible by the TRE and/or the federated TRE network to which it belongs	Safe outputs	[27]
O.5	<b>Standardised procedures for archival</b> , including (but not limited to): 1. One or more workspaces related to a completed project 2. Datasets related to a completed project (including those linked to publications) 3. Tools related to a completed project	Safe outputs	[27]
F.1	<b>Standardised procedures to identify and manage conflicting standards across a federated network of TREs</b> , such as (but not limited to): 1. Screening, training, guidance and/or support 2. Assessing outputs and handling appeals	Safe federation, also related to Safe people, and Safe outputs	Interpretation
F.2	<b>Standardised procedures for intervenability</b> across a federated network of TREs, such as (but not limited to): 1. “Single Point of Contact (SPoC)” for a TRE and/or specified federated network of TREs has been established for data subjects to exercise their data-related rights 2. “disabling options for individual functionalities without affecting the whole system”	Safe federation	[42]
	<b>Standardised procedures for change management</b>		

# Public Involvement and Engagement – Privacy Risk Assessment Forum

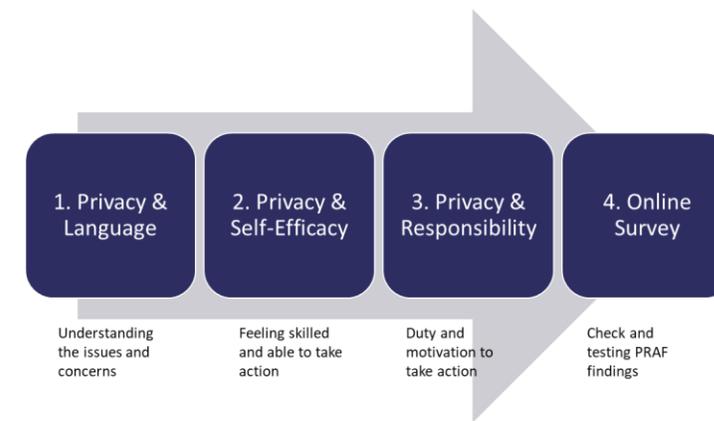
- Find ways to involve members of the public in data sharing decisions
- Approach
  - 12 members of the public
  - Participant journey
    - **1. Privacy attitudes and language (Done)**
    - 2. Privacy and self-efficacy
    - 3. Privacy and responsibilities
    - 4. Check and test findings for online survey
- Emerging themes (1<sup>st</sup> workshop analysis in progress)
  - Education and support
  - Communication of decisions
  - Polarities in the debate (you signed so your responsibility vs people don't have understanding)
  - Concerns for custodianship incl. data retention beyond business lifecycles
  - Concerns regarding business vs plain language



**Public Involvement**  
**James McMahon**  
*J.P.Mcmahon@southampton.ac.uk*



- Scenario 1 – Online Shopping
- Scenario 2 – Activity Tracking
- Scenario 3 – COVID Track and Trace
- Scenario 4 – Research Project



What do you understand by....

- Safety
- Privacy Harm
- Feared Event
- Data Stewardship
- Trusted Research Environment

....would you use different words?

What do you understand by....

- Privacy risk, likelihood and impact
- Asset, threat and vulnerability
- Security and privacy control
- Loss of confidentiality
- Identifier, quasi-identifier, and reidentification

....would you use different words?

# Advisory Group

22 experts including:

- Information governance practitioners
- Practitioners running or developing secure research facilities
- Legal professionals
- Oversight bodies
- Academic experts

**Semi-structured interviews** to understand the risk factors to consider when research projects request data, the controls available and the decisions tied to privacy risk assessment



The University of Manchester

Centre for Epidemiology Versus Arthritis



National Data  
Guardian



University Hospital Southampton  
NHS Foundation Trust



The  
Alan Turing  
Institute



UNIVERSITY OF  
CAMBRIDGE

Cambridge Health Informatics Limited

Imperial College  
London



University of  
Southampton



Bristows



Medical  
Research  
Council



Hampshire and Isle of Wight



## Early findings from the Advisory Board

- Decisions by committees to determine **functional anonymisation** guarantees can be subjective and lack transparency
- In data sharing contracts, **institutions** that the researcher requesting data is affiliated with matters a lot
  - problems for people who do not have affiliations with a stronger/well established institution
  - bottleneck for researchers to navigate IG inside their own organisation, especially if they are risk averse
- Controls on one safe can **compensate for risks** on the other in certain cases (e.g., people and settings) but not in others (e.g., project)

# Risk Tiers Framework

Develop a framework to help decision makers:

- **Document** level of risk along each axis of the five safes
- Establish a **shared view** that stakeholders can understand and reason about
- **Evaluate** risk and the actions to reduce risk for each data sharing scenario
- Respond to risk **consistently**

Project	Level 0				
Setting + People + Outputs	Level 0	Level 1	Level 2	Level 3	Level 4
Data	Level 0	Level 1	Level 2	Level 3	

For example:  
 + All activity logged  
 + Contractual agreement  
 + Trained researcher  
 + Differentially private outputs

Tier 1	Sum of risk levels = 0 or 1
Tier 2	Sum of risk levels = 2 or 3
Tier 3	Sum of risk levels > 3

Overall risk tier for project mapped to decisions. For example:

- Tier 1 = Fast track approval
- Tier 2 = Increased monitoring of project
- Tier 3 = Rejection

# Privacy and Security Risk Modelling – Example TRE system model

689 Threats

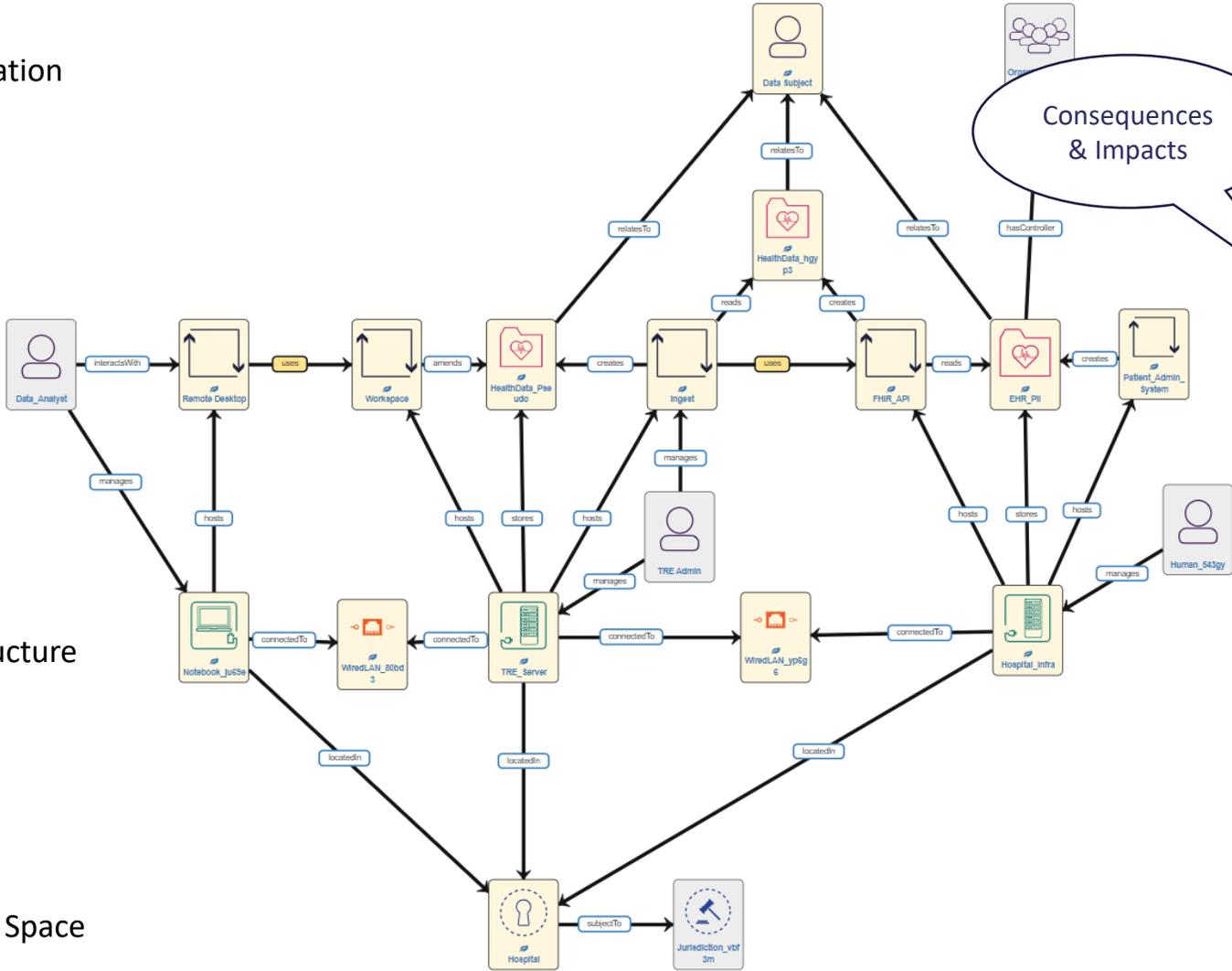
System Modeller demo-tre    Export    View    Help    Home    Dashboard    mjb    75%    Total threats: 689    Highest risk: Low

Organisation

Application (Dataflow)

Infrastructure

Physical Space



Consequences & Impacts

Residual risk after applying controls

Model Summary

**Details**

Name: demo-tre  
 Knowledgebase: NETWORK  
 Description:  
 Assets: 95  
 Relations: 490  
 Threats: 689

Buttons: Edit Details, Open Report

✓ Possible Modelling Errors (3/3)

Assets (95)

Controls (113)

Adverse Effects and their Impact (196)

Show filters    Reset sort

Name	Asset	Impact	Likelihood	Risk
LossOfAuthenticity	HealthData_hgyp3	High	Very Low	Low
LossOfAuthenticity	HealthData_Pseudo	High	Very Low	Low
LossOfAuthenticity	EHR_PII	High	Very Low	Low
LossOfAvailability	Ingest	Very Low	Very Low	Very Low
LossOfAvailability	FHIR_API	Very Low	Very Low	Very Low
LossOfAvailability	TRE_Server	Very Low	Very Low	Very Low
LossOfAvailability	TRE Admin	Medium	Very Low	Very Low
LossOfAvailability	WiredLAN_yp6g6	Very Low	Very Low	Very Low
LossOfAvailability	HealthData_hgyp3	Medium	Very Low	Very Low
LossOfAvailability	Hospital_Infra	Very Low	Very Low	Very Low

Threats (230/689)

Compliance (0/1)

GDPR (10/20)

Buttons: Details



# GDPR Compliance Explorer

System Modeller  
demo-tre

Export View Help

Home Dashboard | mjb

75%

Total threats: 689  
Highest risk: Low

Personal data  
"HealthData\_Pseudo" has no data controller: under the GDPR, a data controller (a person or in most cases, an organisation) who alone or jointly with others, determines the purposes and means of the processing of personal data, and has certain responsibilities defined by Article 24. Data "HealthData\_Pseudo" relates to the data subject "Data Subject" and is subject to the GDPR, so there should be a controller. Add an organisation or user if not already included, and a hasController relationship from "HealthData\_Pseudo" to its data controller.

### Compliance Explorer

GDPR (10/20)

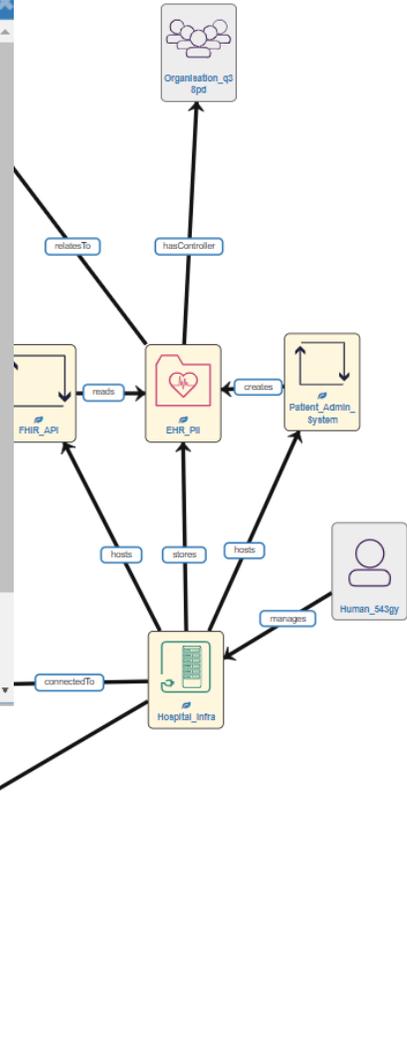
Description: GDPR compliance threats.

Compliant: false

Show filters Reset sort

Compliance Threat

Asset	Compliance Threat
HealthData_Pseudo	Personal data "HealthData_Pseudo" has no data controller (a275)
HealthData_Pseudo	Personal data "HealthData_Pseudo" has no data controller (faf6)
HealthData_Pseudo	Personal data "HealthData_Pseudo" has no data controller (c426)
HealthData_hgyp3	Personal data "HealthData_hgyp3" has no data controller (2a7c)
HealthData_hgyp3	Personal data "HealthData_hgyp3" has no data controller (9d88)
Hospital_Infra	Personal data "EHR_PII" should be stored in a suitable location (1d2d)
TRE_Server	Personal data "HealthData_Pseudo" should be stored in a suitable location (b245)
Data Subject	Condition for consent for processing of personal data "EHR_PII" related to "Data Subject" (ab8a)
Data Subject	Condition for consent for processing of personal data "HealthData_Pseudo" related to "Data Subject" (5283)
Data Subject	Condition for consent for processing of personal data "HealthData_hgyp3" related to "Data Subject" (51b0)
FHIR_API	Personal data "EHR_PII" related to "Data Subject" should be processed in a suitable location (6f00)
Patient_Admin_System	Personal data "EHR_PII" related to "Data Subject" should be processed in a suitable location (ad9b)
Ingest	Personal data "HealthData_Pseudo" related to "Data Subject" should be processed in a suitable location (bc91)
Workspace	Personal data "HealthData_Pseudo" related to "Data Subject" should be processed in a suitable location (7108)
Workspace	Personal data "HealthData_hgyp3" related to "Data Subject" should be processed in a suitable location (7108)



### Model Summary

**Details**

Name: demo-tre  
Knowledgebase: NETWORK  
Description:  
Assets: 95  
Relations: 490  
Threats: 689

Edit Details Open Report

**Possible Modelling Errors (3/3)**

**Assets (95)**

**Controls (113)**

**Adverse Effects and their Impact (196)**

Show filters Reset sort

Name	Asset	Impact	Likelihood	Risk
LossOfAuthenticity	HealthData_hgyp3	High	Very Low	Low
LossOfAuthenticity	HealthData_Pseudo	High	Very Low	Low
LossOfAuthenticity	EHR_PII	High	Very Low	Low
LossOfAvailability	Ingest	Very Low	Very Low	Very Low
LossOfAvailability	FHIR_API	Very Low	Very Low	Very Low
LossOfAvailability	TRE_Server	Very Low	Very Low	Very Low
LossOfAvailability	TRE Admin	Medium	Very Low	Very Low
LossOfAvailability	WiredLAN_yp6g6	Very Low	Very Low	Very Low
LossOfAvailability	HealthData_hgyp3	Medium	Very Low	Very Low
LossOfAvailability	Hospital_Infra	Very Low	Very Low	Very Low

Threats (230/689)

Compliance (0/1)

GDPR (10/20)

Details

## Conclusions

- Privacy requirements for safe federations and use cases analysed
  - D1 report to be published end-May
- Approach codesigned with stakeholder engagement through Advisory Board and the public Privacy Risk Assessment Forum
- Risk Tiers framework outlined and aligned with security and privacy risk modelling tools
- Extensions to privacy domain knowledge for system modelling based on privacy requirements started
- Plans for open community of privacy and security domain experts supported by open methodologies and tools



UK Research  
and Innovation



Thank you for listening

