Standard Architecture for Trusted Research Environments

Release 0.0

The Contributors

Oct 12, 2023

OVERVIEW

CHAPTER

ONE

THE SATRE SPECIFICATION

The specification is presented in terms of the capabilities that a team running a TRE should aim for across all aspects of TRE provision.

This page explains what the specification is and how it's structured. It also describes how the importance of components is categorised. Consult the *FAQs* for more on what the specification *is not*.

Note: Throughout this document, we will use the term "*TRE operator*" to refer to the team running a particular TRE.

1.1 Structure

The SATRE specification comprises three key parts:

Fig. 1: SATRE Specification Architecture

Architectural Principles
: The *principles* that all *TRE operator*s looking to use the specification should hold themselves accountable to.

Specification Pillars
: The broad areas of TRE provisioning the specification covers. Each pillar is broken down into one or more *capabilities*. Each capability is broken down into one or more *components*.

Roles
: *Roles* that are necessary for the operation and use of a TRE.

Together, these provide a framework that *TRE operator*s can measure themselves against.

1.2 Architectural Principles

The SATRE specification has been developed based on the following principles:

- *Usability*
- *Maintaining public trust*
- *Observability*
- *Standardisation*

1.3 Specification Pillars and Capabilities

The SATRE specification contains three core pillars for a TRE, plus supporting capabilities:

Fig. 2: SATRE Pillars Capability Map

1. Information governance
What the *TRE operator*s do to ensure information risk is measured and managed to an acceptable level.

2. Computing technology
What the *TRE operator*s do to manage systems for storing, retrieving, and sending information.

3. Data management
What the *TRE operator*s do to manage data assets and ensure information remains secure.

4. Supporting capabilities
A *TRE operator* will need to possess various supporting capabilities, such as complying with legal requirements and managing relationships with stakeholders.

1.3.1 Importance

The TRE capabilities are broken down into components. Each component is a statement of a process, method or practice that the operators should have in place to ensure they fulfil the capability requirements. These components are each labelled with an importance:

Mandatory
This is required: if this component is not supported, then the capability, and therefore the specification, is not met.

Recommended
Most TREs should have this component, but it is not essential.

Optional
Many TREs would benefit from this component, however, we recognise there are reasons a *TRE operator* may actively choose not to implement it.

Some components are mandatory in some circumstances but not others. These are indicated by an asterisk `Mandatory*`, with details provided in the statement.

*TRE operator*s are able to demonstrate that they meet the specification by showing they can fulfil all **mandatory** components. Future versions of the specification may introduce more granular levels of evaluation, for instance tiered level of accreditation based on fulfilment of mandatory, recommended and optional components respectively.

Any particular TRE implementation should be able to *score itself against each capability*.

1.4 Roles

A TRE needs to consider many different stakeholders. SATRE provides specific roles which may or may not match titles used in your organisation. However each of these are important to the successful operation of a TRE. Roles are grouped into:

1. Project Roles
Roles for TRE end-users conducting research or analysing data in the TRE and others involved in managing this research.

2. Data Management Roles
: Roles for people managing data and databases used in a TRE

3. Infrastructure Management Roles
: The IT professionals and software engineers who will be responsible for developing, deploying and managing instances of a TRE conforming to the SATRE specification.

4. Governance Roles
: Roles that uphold the governance of TREs. Such governance responsibilities typically involve establishing policies and procedures to ensure the responsible use of data, protecting the privacy and confidentiality of research participants, and promoting transparency and accountability in research activities.

5. Public Roles
: Roles that concern members of the public with regard to TREs and TRE research.

CHAPTER

TWO

THE ARCHITECTURE

This Standard Architecture for Trusted Research Environments (TREs) provides a comprehensive high-level architecture for research organisations handling sensitive data safely. The architecture is documented using the ArchiMate modelling language with models created using the open source modelling tool Archi.

Capabilities an organisation requires to run a TRE are documented and deconstructed to show the elements (Roles, processes, applications and data) needed to realise those capabilities. Views are provided aligned to the capabilities and in version 1.0 an additional view is provided to show alignment to the Five Safes Framework.

The main document is available on Zenodo.

The architecture and SATRE standard maps to the architecture meta-model below.

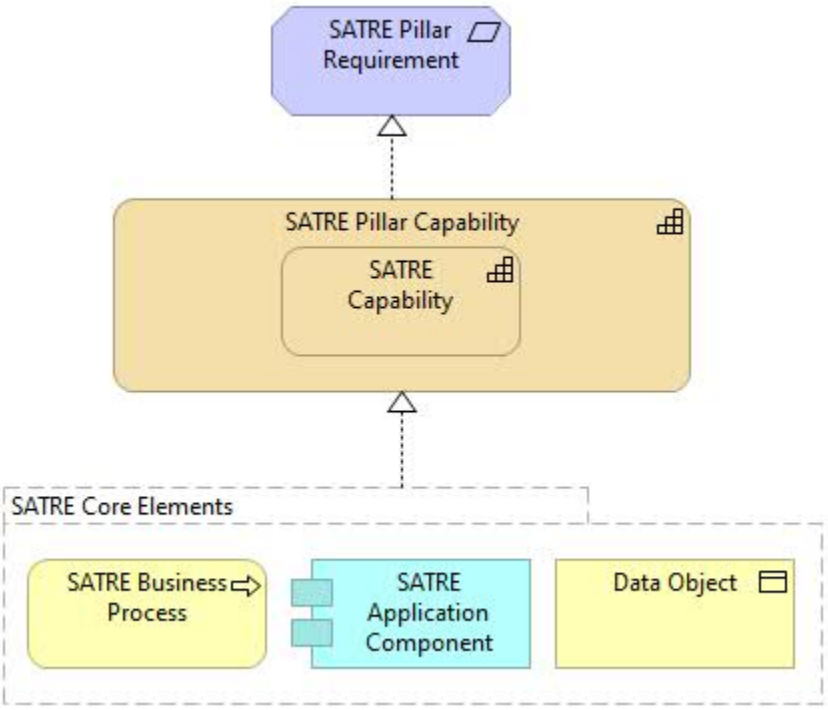

Fig. 1: SATRE Architecture meta-model

CHAPTER

THREE

ARCHITECTURAL PRINCIPLES

Architectural principles influence and shape the way you design and deliver a SATRE-aligned TRE. They are a set of guiding considerations that sit above any specific architectural requirement, and can be applied across the entire architecture.

They consist of the following parts:

Statement
: A singular sentence that summarises the principle

Rationale
: Justification as to why this principle is important for the specification

Implications
: Things you need to consider or do to practise this principle

3.1 Usability

3.1.1 Statement

A TRE instance that works for all users minimises barriers to use, and provides a productive and accessible analysis environment for research.

3.1.2 Rationale

There is often a trade-off between increased operational security and the usability of a TRE. In order to maintain productivity, a TRE must balance these two competing aims. The design and configuration of a TRE should allow all individuals involved with a TRE to effectively fulfil their roles.

3.1.3 Implications

- Robust TRE design and implementation should start by understanding users' diverse expectations, needs, existing skillsets and preferences and responsibilities.
- Design, configuration and testing of TREs must recognise a diversity of users. For instance, not all users are researchers and not all researchers are users. Other users include TRE operators, information governance officers, and TRE builders/developers.
- Because of diverse user needs, it is unlikely that a specific TRE instance will perfectly match the needs of all users.

- A TRE that is overly strict on tool and software provision may risk becoming unusable for users with different and varied backgrounds and skillsets.
- Working environments can differ significantly from users' preferred setups This has design and resource implications for supporting new users, and consideration should be given to resources and time required to help users get up to speed with new and unfamiliar TRE instances.
- Improving user experience takes time and resource, and will involve trade-offs between investing time in improved standards, better functional design, improving work and organisational culture, boosting users' skills and knowledge through training and making help more readily available at an organisational level. These trade-offs will need to be addressed at an organisational level, and teams may want to consider resourcing staff to focus specifically on these questions, for instance in the positions of product managers or service functions.

3.2 Maintaining public trust

3.2.1 Statement

TREs holding public data should build and maintain the trust of data subjects and any other impacted individuals, groups, communities and organisations by protecting privacy, keeping data secure and being transparent about their work.

3.2.2 Rationale

To ensure continued public support both for the use of data and TREs to facilitate research, it is vital to maintain public trust in the ways TREs hold and use data and to alleviate possible concerns. This could include maintaining the trust of members of the public whose data is held, those who are impacted by research conducted using TREs, and the trust of commercial data providers.

In the case of public sector data, public engagement work has indicated there is support for the use of regulated and ethical TREs working for the public benefit as long as conditions are met surrounding security and transparency. Consulting impacted parties, including the public, can help ensure TREs are being used for positive, impactful and agreed-upon purposes. Being transparent about the data held and the projects or organisations who access the data can also help maintain trust.

3.2.3 Implications

- Being as transparent as possible is key to building trust. TREs holding public data should practice transparency. For instance, accreditation from an external body, adherence to a specific design framework (such as the Five Safes), and details of the projects or organisations that access their data should be publicly available information. This should be provided in an accessible way.
- In addition to transparency, actively involving members of the public in oversight of TREs and their processes is important for accountability. Public involvement and engagement takes time and resources. TREs holding public data should consider allocating specific staffing and budget to public engagement activities.
- Where TREs have consulted impacted parties, they should be auditable by those parties, and consult them as part of any decision making processes. This may include the provision of documentation and educational resources for a diverse audience.
- Access to public sector data should be reviewed by an independent panel, which includes members of the public where possible, and follow agreed-upon governance to ensure projects using this data are in the public benefit, and provide clarity around any commercial access.

3.3 Observability

3.3.1 Statement

Human initiated and automated processes resulting in change within the TRE should be observable.

3.3.2 Rationale

System/process observability is key to understanding whether your policies and controls are actually doing what is intended.

It also allows operators to continuously improve their systems and processes, measure their effectiveness, and identify the causes of incidents. Data can also be made available to other parties such as auditors, data subjects and data providers as part of the assurance process.

This applies to both technical systems and policies/processes.

3.3.3 Implications

- In order to understand what is happening within the TRE, both automated and human initiated processes should generate sufficient data, for instance through audit logs. Any generated data should follow standards for provenance and transparency for audit trails.
- Different levels of observability may be needed for different users. Any data collected from an observability perspective should consider the needs of those who will use it, and minimise collection accordingly.
- There may be ethical and confidential issues to consider when implementing the observability principle.

3.4 Standardisation

3.4.1 Statement

TREs should adhere to standards or well-known patterns wherever possible.

3.4.2 Rationale

Standardisation makes it easier to design, operate, use and understand TREs, and reduces duplication of work. This includes making TREs easier to use, deploy, and audit.

TREs should be built in such a way that they do not restrict or prevent interoperability where this may be desirable in future, by identifying and avoiding or removing barriers to interoperability.

Standardisation is also linked to the public trust principle, as a standard approach to TRE provision will make it easier for impacted parties to understand how their data will be used within TREs.

3.4.3 Implications

- Owing to the broad definition of *TREs*, there are currently no technical or information governance standards focused on TREs. The SATRE specification has therefore been designed to help *Developers and Operators* with a variety of technical/policy requirements consider their options
- TRE *Developers and Operators* should be prepared identify the technical standards that are appropriate for them to work towards meeting whilst developing or maintaining their particular TRE(s).
- Standards that TREs adhere to may range in scope, including technical, operational and governance requirements.
- TRE *Developers and Operators* should ensure that when they aim to meet multiple standards, those standards align with one another to ensure there is no contradiction in requirements.

There might be good reasons why any particular TRE does not possess one or more of the capabilities listed in this specification, but most TREs should aspire to meet them in the long-term.

CHAPTER

FOUR

ROLES

A TRE conforming to the SATRE specification should provide a broadly similar experience for stakeholders operating in each of these defined roles. There is not necessarily a one-to-one mapping between roles and people. One person can have multiple roles.

4.1 Project Roles

Roles for TRE end-users conducting research or analysing data in the TRE and others involved in managing this research.

Role name	Role description
Data Consume	General term for any individuals who will be provided access to data via a TRE.
Data Analyst	Specific term for people provided access to data via a TRE, who intend to carry out analysis or conduct research using the data. These could be programmers and data scientists, but could also be scientists working in fields where deep computing expertise is less common. Analysts working with TREs that meet the SATRE standard should have a broadly similar user experience, at least where the type of analyst is consistent (e.g. data scientists). This includes both the user experience of the platforms themselves, and the associated documentation.
Proje Manager	The person in charge of coordinating other roles for the duration of a specific TRE project. See *Project and programme management*.
Proje Team	Refers to the team of data analysts and project manager(s) working on a specific project that uses a TRE.

4.2 Data Management Roles

Roles for people managing data and databases used in a TRE.

Role name	Role description
Data Steward	People who ensure data within a TRE is maintained and processed in ways useful to *data analysts*, including providing data extracts. May also be known as Data Wranglers, Data Engineers or Data Cleaners.
Database Administrator	People responsible for managing any databases included in the TRE. Where a database is used by multiple projects, this includes handling segregation of users and databases belonging to different projects. See *Advanced Computing Systems*.
Information Asset Owner	General term for custodians or owners of a datasets, projects or other information assets within a TRE. For example, the owner of a dataset who has liaised with a *TRE Operator* on *secure data ingress* to the TRE.
Output Checker	People responsible for checking the disclosure risk of project outputs, before egress, as part of the disclosure control process. See *Output management*.

4.3 Infrastructure Management Roles

The IT professionals and software engineers who will be responsible for developing, deploying and managing instances of a TRE conforming to the SATRE specification.

Role name	Role description
Operator	People responsible for the management of the TRE's IT infrastructure and general processes documented throughout the SATRE specification. Examples include carrying out data ingress/egress and managing user access. TRE operators should expect to have access to documentation regarding all processes they are required to carry out, developed by themselves or (in partnership with) the TRE Developers. This documentation should be comprehensive and include troubleshooting steps (see *Knowledge management*).
Developer	People responsible for building the software infrastructure that can be used as a TRE. These could be Research Software Engineers, whose job involves applying professional software engineering expertise to challenges in scientific research. Alternatively, these could be developers who are contracted to build a TRE for a given institution or project. TRE developers include people creating bespoke platforms catering to the specific requirements of a project or dataset, as well as developers building generalisable solutions to TRE provision that can be configured based on the research context.
Build	People responsible for carrying out the *Infrastructure Deployment Process*. This role could be taken on by either the TRE Operators or the TRE Developers.

4.4 Governance Roles

Roles that uphold the governance of TREs. Such governance responsibilities typically involve establishing policies and procedures to ensure the responsible use of data, protecting the privacy and confidentiality of research participants, and promoting transparency and accountability in research activities.

Role name	Role description
Information Governance Manager	People responsible for writing and/or compiling the correct operating procedures and policies for the TRE.
Quality Manager	People responsible for ensuring the TRE is operating correctly, and all procedures and policies are being followed by other roles and work effectively. See *Quality Management*.
Top Management	People who lead and control an organisation at the highest level. This definition is taken from *ISO 9000:2015* and in this context refers to the most senior governance official who own the risks associated with TRE research, can make decisions and allocate resources. See *Risk Ownership Process*.
Data Protection Manager	People responsible for *data protection* at the organisation hosting the TRE.
Auditor	General IT term for people who evaluate an organisation's IT systems on whether they meet technology or cybersecurity regulatory requirements. For TREs, this may include requirements around sensitive data handling and information security controls. Auditors can be internal, or external people working for a consulting firm.

4.5 Public Roles

Roles that concern members of the public with regard to TREs and TRE research.

Role name	Role description
Lay Panel	Members of the public charged with oversight of TRE operational decisions on behalf of parties affected by data usage.
Data Subject	People who are identifiable by data being used for research, e.g. patients in healthcare record data.

CHAPTER

FIVE

FREQUENTLY ASKED QUESTIONS

5.1 What is a TRE?

TRE stands for *Trusted Research Environment*. The simplest definition tends to be *any* kind of computing environment set up for research with sensitive data that has built-in security controls and user access management features. The definition of TRE relevant to SATRE encompasses the set of information governance and data management processes alongside the computing technology used to support research with sensitive data; the definition of sensitive data being broadly any data for which there may be considerations around disclosure control, for any reason.

We recognise that in the UK several other names such a *Secure Data Environment* or *Data Safe Haven* have been used in the literature on computing with sensitive data, and that these systems may go by other names elsewhere. For more information about TREs, visit the UK TRE Community website.

5.2 Why do we need a Standard Architecture for TREs?

A variety of approaches have been taken to building computing infrastructure and designing processes and policies for research with sensitive data. There are also a range of standards or frameworks that may apply to TREs such as ISO27001 or 5 Safes. However, they don't provide prescriptive guidance on how TREs comply or achieve accreditation. In recognition of this, SATRE aims to find the commonalities and compile a resource for TRE *Operators, Builders and Developers* to refer to and benefit from. See *What is SATRE?* for more information.

5.3 How has the SATRE specification been developed and why?

See the information on the homepage of *these docs*.

5.4 Who has developed the SATRE specification?

Take a look at our *Contributors* page.

5.5 Is SATRE an ISO technical standard?

No. The SATRE specification aims to provide a helpful guide for TRE *Operators, Builders and Developers*. It can be used to inform the development process of new TREs, or to evaluate existing TREs and inform how they could be improved. Evaluating a TRE with the SATRE specification may help to identify which technical standards (e.g. ISO 27001) are already met and which (if any) are desirable to work towards meeting.

5.6 Does SATRE provide everything I need to build a TRE?

No. The SATRE specification defines a set of stakeholder *roles* and feature *capabilities* for TREs, which were decided according to our *architectural principles*. It does not dictate which specific technologies could or should be used to build a TRE.

5.7 What do TRE Builders/Developers gain by reading the SATRE specification?

By reading through the SATRE specification, developers tasked by their institution with designing and/or building a TRE for sensitive data projects can avoid re-inventing the wheel. The specification does not dictate answers to the specific technology or policy choices that need to be made when developing a TRE, but it does provide a guide for thinking about which choices need to be made and what *capabilities* the TRE should possess.

5.8 What do TRE Operators gain by evaluating their TRE with SATRE?

See: *Why should I evaluate my institution's TRE?*

5.9 How do I build and run a SATRE compliant TRE?

We encourage TRE *Operators and Builders* to publicly evaluate their TREs against the SATRE specification; see *Evaluating TREs against SATRE*. TRE *Developers* can use the specification and published TRE evaluations as a starting point. Some of evaluated TREs such as the Alan Turing Institute's *Data Safe Haven* and the University of Dundee Health Informatics Centre's *TREEHOOSE* are deployed from open source infrastructure-as-code, and can be deployed by other institutions.

5.10 Is the SATRE specification set in stone?

Absolutely not. We know that TREs vary greatly in their design architecture, purposes for being built, the kinds of research they support and data they handle. We have tried to build a specification with as broadly useful a set of capabilities as possible, whilst acknowledging these different approaches. We won't have covered everything, and if you find SATRE valuable but think there is something we've missed, please consider *contributing*. Additionally, the best practices in TRE provision may evolve over time as technologies and regulations change. We hope that the SATRE specification will be maintained in the future and accommodate these changes as appropriate.

5.11 My TRE is designed for data that doesn't require this level of protection. Should I still follow SATRE?

Yes. At the moment, the SATRE specification contains a set of capabilities marked as "Mandatory" which we believe are essential for a system to be defined as a TRE, as well as many "Recommended" and "Optional" capabilities. Some of the non-mandatory capabilities will likely not be needed for TREs containing data that does not require all the possible protections, and there may well be tradeoffs to be made in terms of accessibility vs security that depend on the data the TRE holds. A future specification may include the idea of different archetypes of TREs, or data sensitivity tiers, with different requirements for each.

5.12 Does SATRE describe approaches to TRE federation or interoperability of TREs?

No. However, it's intended that SATRE could form the foundation for future standards and guidance on federation, interoperability, and related work.

CHAPTER

SIX

INFORMATION GOVERNANCE

Fig. 1: SATRE Pillars Capability Map

This pillar concerns actions taken by the *TRE operator* to ensure information risk is measured and managed to an acceptable level.

Each *TRE operator* will have its own information governance requirements. These will be informed by the context of the organisation, the work it performs and the nature of the data it processes. For example, some requirements will arise from national legislation such as GDPR, discipline specific regulation like GCP, or contractual requirements from a specific *information asset owner* such as a company or research partner organisation.

6.1 Governance Requirements

6.1.1 Requirements Gathering and Monitoring

This *business process* involves collecting, documenting, and managing the functional and non-functional requirements for the TRE based on the TRE organisation's goals and data assets.

	Statement	Guidance	Importance
1.1.1	You must gather and monitor the information governance requirements needed to fulfil any legal, regulatory and ethical standards.	Requirements will come from a variety of sources including legislation, contractual obligations and ethical standards. Requirements must be monitored to ensure the TRE controls remain appropriate.	Mandatory

6.1.2 Controls

This *business process* involves measures, safeguards, or mechanisms implemented to manage or mitigate risks associated with your organisational requirements.

	Statement	Guidance	Importance
1.1.2	You must ensure controls are implemented to ensure the requirements are met.	Control implementation should be systematic and directly aligned to the internal and stakeholder requirements.	Mandatory

6.1.3 Resource Allocation Process

This *business process* involves assigning, distributing, and managing resources (such as personnel, finances, equipment, or time) within the TRE organisation to meet information governance requirements.

	Statement	Guidance	Importance
1.1.3	You must ensure there are adequate resources to meet information governance requirements.	Ensuring information governance controls are suitable and enforced requires an investment of funding and people appropriate to the size of the TRE.	Mandatory

6.2 Quality Management

What the organisation does to measure and control quality of processes, documentation and outputs.

6.2.1 Document and SOP Management Process

This *business process* involves creating, organising, updating, and controlling documents and Standard Operating Procedures (SOPs) within the TRE organisation.

	Statement	Guidance	Importance
1.2.1	You must ensure that changes to policies and standard operating procedures can only be made by trusted individuals.	It is important to ensure that policies and SOPs are relevant, up-to-date and carefully controlled to maintain the integrity and security of your TRE organisation.	Mandatory
1.2.2	You must use versioning and a codified change procedure for all policies and standard operating procedures.	This includes recording dates of changes, person responsible for carrying out changes, and summary of changes.	Mandatory

6.2.2 Quality Management Process

This *business process* involves the generation and dissemination of reports or dashboards that provide insights and metrics on the performance and effectiveness of quality management processes and activities.

	Statement	Guidance	Importance
1.2.3	You should measure the performance of information governance within the TRE with regular reporting available to your TRE organisation's management team.	This may include reports and dashboards showing security incidents, quality management deviations and audit findings.	Recommended

6.2.3 Internal Audit Process

This *business process* involves an independent evaluation process within the TRE organisation that assesses and improves its internal controls, risk management, and governance.

	Statement	Guidance	Importance
1.2.4	You must audit your TRE organisation against relevant requirements and standards.	If you are publicly accredited against a standard, for instance ISO27001, DSPT, CE+ *etc.*, you must have processes in place to ensure you remain compliant.	Mandatory
1.2.5	You must report on and share outcomes of each audit of your TRE organisation with the required bodies.	This may include regulatory bodies or the organisations that manage accreditations you have.	Mandatory

6.2.4 Supplier Management and Monitoring Process

This *business process* involves a structured approach to managing and monitoring relationships with external suppliers, vendors and contractors, including selection, contract management and compliance oversight.

	Statement	Guidance	Importance
1.2.6	You must ensure that suppliers, contractors and sub-contractors with access to your TRE align with your security requirements.	These should be included as mandatory, non-functional requirements in during procurement and contracting. This will also include contractor staff contracts for example, legal liability and NDAs.	Mandatory
1.2.7	You must monitor compliance of your suppliers with the terms of the contracts.	This will include monitoring changes in the services and infrastructure being delivered and quality management within the contractor's organisation. This may be done through formal audit or by monitoring change and quality documentation provided by the supplier.	Mandatory

6.2.5 Asset Management Process

This *business process* involves a systematic approach to acquiring, operating, maintaining, and disposing of assets within an organization, aimed at maximizing their value and minimizing risks.

	Statement	Guidance	Importance
1.2.8	You must track and maintain any physical assets used by your TRE.	All physical assets should be maintained and covered by warranty if applicable. At the end of their lifetime, assets should be securely disposed of in such a way that data cannot be recovered from them.	Mandatory (where physical assets are in scope)

6.2.6 Issue Management Process

This *business process* involves a systematic approach to identifying, tracking, resolving, and managing issues or problems that arise within a TRE organisation, aiming to minimize their impact and ensure timely resolution.

	Statement	Guidance	Importance
1.2.9	You must log, track and resolve any issues resulting from deviations from processes, incidents and audit findings.	This process could, for example, be tracked through an electronic record and workflow system with records retained.	Mandatory
1.2.1(	You must use reported issues to inform changes, such as for process improvement and risk management.	All issues should be analysed for their root cause and improvements put in place to prevent further occurrence.	Mandatory

6.2.7 Quality Management Data

This *data object* consists of data, including training records and configuration data, collected and used to monitor, evaluate, and improve the quality of processes, or services within the TRE organisation.

	Statement	Guidance	Importance
1.2.1	You should collect and maintain quality management data for measuring the effectiveness of a TRE.	Large amounts of data will be produced by elements within the TRE. These data should be analysed with reports and dashboards provided to guide TRE implementer's improvements and provide re-assurance to *data consumers* and *data subjects*.	Recommended

6.2.8 Quality Management System Application

This *application component* is a software application or platform used to manage and automate quality management processes, including document control, corrective actions, audits, and performance tracking.

	Statement	Guidance	Importance
1.2.1(	You could use a QMS (Quality Management System) to standardise and automate quality management tasks and workflows, and to generate quality data and reports automatically.	A basic QMS could be a set of spreadsheets or documents held in a repository which are manually maintained. More mature applications will provide workflows and generate quality data through manual and automated actions.	Optional

6.3 Risk Management

What the organisation does to ensure information risk is measured and managed to an acceptable level.

6.3.1 Risk Assessment Process

This *business process* involves the systematic evaluation and analysis of potential risks, threats, or vulnerabilities, including their likelihood, potential impact, and the effectiveness of existing controls or mitigation measures.

	Statement	Guidance	Importance
1.3.1	You must have a way to score risk to understand the underlying severity.	You have a risk assessment methodology for scoring risks on multiple axes such as impact and likelihood.	Mandatory
1.3.2	You must carry out a data processing assessment for all projects requiring a TRE.	A data processing assessment is a process designed to identify risks arising out of the processing of sensitive data and to minimise these risks as far and as early as possible. This may take the form of an existing regulatory requirements such as Data Protection Impact Assessment.	Mandatory

6.3.2 Risk Treatment Process

This *business process* involves the selection and implementation of strategies, controls, or measures to manage or mitigate identified risks, such as risk avoidance, risk transfer, risk reduction, or risk acceptance.

	Statement	Guidance	Importance
1.3.3	You must have a process for designing, implementing and recording risk mitigations where indicated by a risk assessment.	Actions that are taken or not taken following a risk assessment must be recorded.	Mandatory

6.3.3 Risk Ownership Process

This *business process* involves the assignment of responsibility and accountability to individuals or entities for managing and mitigating specific risks within the TRE organisation.

	Statement	Guidance	Importance
1.3.4	You must have a clear set of roles and responsibilities relating to risk including who owns risks and how they are escalated and delegated.	The highest level of risk ownership is the Top Management of the TRE organisation (see *Governance Roles*). In order to ensure escalations to this level are rare, suitable structures should be put in place to own, mitigate and accept risk.	Mandatory
1.3.5	You must understand the risk appetite of your TRE organisation.	This includes understanding ownership of risk, and ability to accept risk which falls outside of the appetite should that become necessary.	Mandatory

6.4 Study Management

What the organisation does to create and maintain research projects and work packages within the TRE.

6.4.1 Study Onboarding Process

This *business process* involves onboarding or initiating a research study, including setting up necessary infrastructure, obtaining approvals, and defining protocols or methodologies.

	Statement	Guidance	Importance
1.4.1	You must have checks in place to ensure a project has the legal, financial and ethical requirements in place for the duration of the project.	This includes checks that contracts are in place where required, adequate funding is available for the duration of the project, and responsibilities concerning data handling are understood by all parties.	Mandatory

6.4.2 Compliance Checking Process

This *business process* involves verifying and ensuring adherence to applicable laws, regulations, standards, or internal policies within the TRE organisation.

	Statement	Guidance	Importance
1.4.2	You must have checks in place to ensure that any time limited compliance requirements are maintained.	This includes ensuring contracts remain in valid and action is promptly taken should they expire. Any changes in the status of responsible persons should also be monitored, for example a data owner leaving an organisation.	Mandatory
1.4.3	You must have checks in place to ensure that changes in regulations are met for a project.		Mandatory

6.4.3 Study Closure Process

This *business process* involves the formal conclusion of a research study or project, including final data analysis, reporting, documentation, and archiving.

	Statement	Guidance	Importance
1.4.4	You must have standard processes in place for the end of a project, that follow all legal requirements and data security best practice.	This includes the archiving of quality and log data along with the archiving or deletion of data sets.	Mandatory

6.4.4 Study Management Portal

This *application component* is an online platform that provides centralised access to manage research studies including onboarding studies, control of access and administration of compliance tasks.

	Statement	Guidance	Importance
1.4.5	You could implement a portal that can provide a workflow engine and database which automates the processes within this capability.	A portal should automate as much of the processes within the capability as possible. Where processes are automated, process maturity is easier to achieve, with more consistent completion and automatic production of quality control and monitoring data.	Optional

6.4.5 Data Asset Register

This *data object* is a database or other electronic record that documents and manages information about the TRE organisation's data assets, including their characteristics, ownership, usage, and other relevant details.

	Statement	Guidance	Importance
1.4.6	You must keep a complete record of all the data assets held within the system.	Details of all data assets (current and past) held by the system should be retained along with meta-data useful for ensuring compliance can be demonstrated. This would include ownership, data lifecycle, contracts, risk assessments and other quality data. This is likely to already exist within the wider organisation but may require augmenting for the TRE.	Mandatory

6.4.6 Study Register

This *data object* is a centralised record or database that tracks and manages information about research studies and projects.

	Statement	Guidance	Importance
1.4.7	You should keep a complete record of all the research studies and projects within the TRE current and past.	The study register should contain all data related to a study including a reference to data assets, *project team members*, *information asset owners* and any compliance activities required.	Recommended

6.5 Member Accreditation

Ability to ensure that people with access to data are correctly identified and they are suitably qualified.

6.5.1 Identity Verification Process

This *business process* involves confirming or authenticating the identity of individuals or entities, often through the verification of personal information, credentials, or biometric data.

	Statement	Guidance	Importance
1.5.1	You must have a robust method for identifying accredited members of your TRE organisation, prior to their accessing of sensitive data.	This may include ID checks or email/phone verification.	Mandatory

6.5.2 User Onboarding Process

This *business process* involves introducing and integrating *data consumers* onto a TRE's systems, processes, including training, access provisioning, and orientation.

	Statement	Guidance	Importance
1.5.2	You must have clear onboarding processes in place for all roles within your TRE organisation.	This may include all members signing role-specific terms of use or confirming that they have completed role specific training.	Mandatory

6.5.3 Identity and Access Management Services

This *application component* is a system to govern and control user identities, access privileges, authentication, and authorization within an organisation.

	Statement	Guidance	Importance
1.5.3	You must have a set of services to manage access to resources based on identity.	This will include a security model for role based access with technical controls to ensure the principle of least privilege is enforced.	Mandatory
1.5.4	You must not give anyone access to datasets without agreement from the Data Controller.	The Data Controller may choose to delegate this authority.	Mandatory

6.5.4 Authentication Application

This *application component* is a software system that verifies and validates the identities of users or entities accessing a system through multifactor authentication.

	Statement	Guidance	Importance
1.5.5	You must have robust and secure applications in place to authenticate users (and services) within the TRE.	The number of authentication applications should be kept to a minimum with common controls and standards applied across all such as MFA, password complexity *etc.*.	Mandatory

6.5.5 User Identity Attributes

This *data object* consists of characteristics or attributes associated with a user's identity, such as username, email address, role, permissions, or affiliations.

	Statement	Guidance	Importance
1.5.6	You must give each user of the TRE a unique logon with changes to any records strictly controlled.	The unique identifier and all associated records for a user should be traceable across the entire TRE. This will include training records, affiliations, contract agreements and ethics approvals where required.	Mandatory

6.6 Training Delivery and Management

Ability to deliver, track and maintain adequate training levels to ensure competence of all people within the TRE organisation.

6.6.1 Curriculum Creation and Management Process

This *business process* involves designing, developing, and managing educational curricula, courses through training needs analysis for required competency.

	Statement	Guidance	Importance
1.6.1	You must determine what training is relevant for all roles within the TRE organisation.	This may include, for instance, cyber security training, GDPR training, and higher level training for system operators. Specialised roles are likely to need more tailored training. Identification of these specialities should be done through a systematic training needs analysis. Specific training may also be required based on the data or *information asset owner* such as GCP.	Mandatory
1.6.2	You must ensure that relevant training is available for all roles within the TRE organisation.	All TRE organisation members need to complete all relevant training and keep their training current. You may need to provide help or guidance to enable them to do so. Details of what training is needed will have been determined above.	Mandatory
1.6.3	You must provide repeat or updated training where necessary to account for changes in competency requirements.	Training is not a one-off event. Electronic reminders for refresher training should be considered. Ideally, training should remain relevant and so policies and processes should enable people to demonstrate competency rather than unnecessarily repeating training.	Mandatory

6.6.2 Certification Management Process

This *business process* involves managing and overseeing certifications or qualifications held by individuals or entities, including tracking expiry dates, renewals, and compliance requirements.

	Statement	Guidance	Importance
1.6.4	You must maintain accurate training records that are directly tied to the role and access levels within the TRE.	Training records should be tied to a user record and carefully maintained. Maintaining training records enables you to ensure all people have completed the required training and that repeat training happens regularly.	Mandatory
1.6.5	You should accept proof of relevant training certifications from trusted third parties.	You might choose to trust certifications provided by known training providers or your institution's partner organisations.	Recommended

6.6.3 Learning Management System

This *application component* is a software platform or application that facilitates the administration, delivery, and tracking of educational or training programs, often including course materials, assessments, and learner progress tracking.

	Statement	Guidance	Importance
1.6.6	You could have a training platform capable of delivering online training in a variety of formats.	This could be a simple content delivery platform or a more comprehensive LMS platform. It could also include a range of multimedia delivery formats, and accessible training modules for those with access requirements.	Optional
1.6.7	You could implement a learning management system (LMS) to manage courses and deliver training as required.	Where possible an LMS should support a variety of course content and testing.	Optional

6.6.4 Courses Data

This *data object* consists of information or data associated with educational courses, including course materials and syllabi, assessments.

	Statement	Guidance	Importance
1.6.8	You could ensure that any courses you use are available in standard, transferable formats.	Support for standard formats such as SCORM allows courses to be shared between providers. This could help facilitate standardisation of training provision for TRE users across organisations.	Optional
1.6.9	You could keep historical copies of courses in order to demonstrate competency at a given point in time.	*Information asset owners* and regulators may be required to audit historical records, *e.g.* for clinical trials. It may be necessary to retain copies of superseded training along with versions of certifications within the training record.	Optional

CHAPTER

SEVEN

COMPUTING TECHNOLOGY AND INFORMATION SECURITY

Fig. 1: SATRE Pillars Capability Map

This pillar concerns actions taken by the *TRE operator* to manage TRE computing systems.

Each *TRE operator* will have its own computing technology requirements. The security controls needed by the computing infrastructure will depend on information governance requirements. Other computing requirements will be influenced by the technical knowledge and experience of those using the TRE, along with the work they need to perform within the system. For example, a data scientist will have very different requirements to a clinician. The required compute resources will vary according to the scale of data and computational techniques employed during research.

7.1 End user computing

The ability of the *TRE operator* to provide and manage devices, workspaces, interfaces and applications used by *data consumers* to interact with underlying systems and data.

7.1.1 End user computing interfaces

This group of *application components* is a collection of systems and software that allows people to interact with the TRE. This may include desktop, command-line and/or code-submission interfaces.

	Statement	Guidance	Importance
2.1.1	You must not allow users to copy data out of your TRE via the system clipboard.	A TRE user must not be able to copy sensitive data out of a workspace using the system clipboard. A TRE may allow user to paste text into a workspace. This might not be relevant to your TRE, for example if your user interface does not have a clipboard.	Mandatory
2.1.2	Your TRE workspace should provide an environment familiar to your users.	This may take the form of a virtual Windows or Linux desktops, non-desktop interfaces such as JupyterLab and other web applications, or a terminal. Bespoke TRE-specific software should be avoided when widely used alternatives already exist.	Recommended
2.1.3	A TRE could restrict data access from *data consumers* entirely and provide an interface for submitting code.	For example, you might use a system where users submit jobs that run over the data and return results without allowing direct data access.	Optional

7.1.2 End user software tools

This *application component* is the tools used by *data consumers* inside a TRE, such as programming languages, IDEs and desktop applications.

	Statement	Guidance	Importance
2.1.4	Your TRE should be accessed via a user interface accessible using commonly available applications.	TREs which allow users to connect from their own devices should not require the installation of any bespoke TRE application on the user's device. In practice a web browser is the most common way to achieve this.	Recommended
2.1.5	Your TRE must provide clear guidance on how to use software tools and work with data in the TRE.	TREs that provide a virtual desktop environment for *data consumers* to work in should provide documentation detailing the available tools. TREs where the analysis code is developed on the access machine (as oppose to within the TRE) should provide documentation detailing the mechanism by which code is submitted to the TRE.	Mandatory
2.1.6	Your TRE should, where possible, automatically apply security related updates for user software.	Reducing the risk of exploitable vulnerabilities in installed software will increase the security of your TRE.	Recommended
2.1.7	Your TRE could provide shared services that are accessible to users in the same project.	This may include shared file storage, databases, collaborative writing, and other web applications. This must only be shared amongst users within the same project.	Optional
2.1.8	Your TRE must ensure that any shared services are only available to users working on the same project.	Poorly designed shared services could enable the unintended mixing of data between projects. To prevent this it is necessary that each instance is only shared between users of a single project.	Mandatory
2.1.9	You must mitigate and record any risks introduced by the use in your TRE of software that requires telemetry to function.	For example, some licenced commercial software must contact an external licensing server at start-up. You must be confident that only licensing information is sent to this server and that any network connections are secure.	Mandatory
2.1.10	Your TRE must provide software applications that are relevant to	The tools provided will depend on the types of data in the TRE, and the expectations of users of the TRE. For users working in a TRE via a virtual desktop, this may include programming languages such as Python and R, integrated development environments, Jupyter notebooks, office type applications such as word processors and spreadsheets, command line tools, etc. TREs with non-desktop interfaces should	Mandatory

7.1.3 Code Version Control System

This *application component* is the systems and tools providing version control and collaboration features for code developed inside the TRE.

	Statement	Guidance	Importance
2.1.1	Your TRE should provide tools to encourage best-practice in reproducibly analysing data.	Reproducibility of analyses improves auditability and accountability of how data has been used, as well as being best-practice in research. This may include version control software, and tools for developing and running data analysis pipelines.	Recommended

7.1.4 Artefact Management Application

This *application component* is a service that manages and organises third-party software artefacts such as packaged code libraries or containers.

	Statement	Guidance	Importance
2.1.1	Your TRE could provide access to some public software repositories or container registries.	For example, a TRE may allow direct installation of packages from Python or R repositories, or provide an internal mirror.	Optional
2.1.1	Your TRE could tightly control which packages are available.	For example, a TRE may only allow installation of a pre-defined set of approved packages. You might also choose to scan for malicious packages and/or go through an approval process before allowing code into the technical environment.	Optional

7.1.5 Advanced Computing Systems

This *application component* involves the use of advanced, powerful computer resources to solve complex problems and process large amounts of data, possibly using specialised hardware.

	Statement	Guidance	Importance
2.1.1	Your TRE must maintain segregation of users and data from different projects when using non-standard compute.	High performance or specialist compute is often shared amongst multiple users. Users and data must remain segregated at all times. For example, when using physical compute resources, all sensitive data could be securely wiped before another user is given access to that same node. In a cloud hosted TRE virtual machines could be destroyed and recreated.	Mandatory
2.1.1	Your TRE should be able to provide access to high performance computing or other scalable compute resource if required by users.	If a TRE supports users conducting computationally intensive research it should provide access to dynamically scalable compute or the equivalent. For example this may be in the form of a batch scheduler on a HPC cluster, or a dynamically created compute nodes on a cloud platform.	Recommended
2.1.1	Your TRE should be able to provide access to accelerators such as GPUs if required by users.	GPUs and other accelerators are commonly used in machine learning and other computationally intensive research. TREs should make it clear to users whether GPUs and other resources are available whilst projects are being assessed.	Recommended
2.1.1	Your TRE could make data available to *data consumers* using common database systems such as PostgreSQL, MSSQL or MongoDB.	Databases must be secured and only accessible to users within the same project. If shared (multi-tenant) database servers are used, *database administrators* must ensure that the database server enforces segregation of users and databases belonging to different projects.	Optional
2.1.1	Your TRE could integrate with large-scale data analytics tools for working with large datasets.	For example, Spark and Hadoop can be used for distributed computing across a cluster. This may be an advantage where a TRE is using an amount of data that is too large for single-machine computing to be practical.	Optional

7.2 Infrastructure management

The ability of the *TRE Builder* to deploy, change or remove physical or virtual infrastructure.

7.2.1 Infrastructure Deployment Process

This *business process* involves setting up and configuring infrastructure components and resources to support applications or services. This requires development, installation, configuration, and validation.

	Statement	Guidance	Importance
2.2.1	You must have a documented procedure for deploying infrastructure.	This might, for instance, be a handbook that is followed or a set of automated scripts.	Mandatory
2.2.2	You should, where possible, automate any repeatable aspects of your deployment.	This might involve using infrastructure-as-code tools or a series of scripts.	Recommended
2.2.3	You must have a documented procedure for making changes to deployed infrastructure.	This refers both to changes that might be expected in the course of normal operation and emergency changes that might be needed. Your change management process may form part of a wider accreditation such as ISO 27001.	Mandatory
2.2.4	You must test changes before they are used in production.	This might involve a separate development environment or another system for testing.	Mandatory
2.2.5	You should have a development environment that mirrors your production environment which you use to test infrastructure changes before committing them to production.	If possible, you should automate application of changes between development and production environments. Consider the costs and practicality of whether this will work for your situation.	Recommended

7.2.2 Infrastructure Removal Process

This *business process* involves retiring or removing infrastructure assets that are no longer needed or outdated, ensuring proper data handling and disposal.

	Statement	Guidance	Importance
2.2.6	You must have a documented procedure for removing infrastructure when it is no longer needed.	Removing unused infrastructure not only reduces costs and management burden but also reduces the attack surface of a TRE and reduces the risk of unaddressed vulnerabilities.	Mandatory

7.2.3 Availability Management Process

This *business process* involves ensuring all IT infrastructure meets the agreed levels of availability.

	Statement	Guidance	Importance
2.2.7	You should understand the availability and uptime guarantees of any providers that you rely on.	For remote TREs this might include your cloud provider(s) and/or data centre operators. For on-premises TREs, it might be worth using an uninterruptable power supply (UPS) and planning how you would deal with internet outages.	Recommended
2.2.8	You should develop an availability target or statement and share this with your users.	Understanding how and when the TRE might be unavailable will help your projects in planning their work.	Recommended

7.2.4 Network Management Application

This *application component* is an application used to manage network infrastructure, ensuring proper functioning, security, and performance.

	Statement	Guidance	Importance
2.2.9	Your TRE must control and manage all of its network infrastructure in order to protect information in systems and applications.	Network infrastructure must prevent unauthorised access to resources on the network. This may include firewalls, network segmentation, and restricting connections to the network.	Mandatory
2.2.1	Your TRE must not allow connectivity between users in different projects, or with access to different datasets.	Connectivity between users in the same project may be allowed, for example to support shared network services within the project.	Mandatory
2.2.1	Your TRE must block outbound connections to the internet by default.	Limited outbound connectivity may be allowed for some services.	Mandatory
2.2.1	You should be able to monitor the network configuration of your TRE to check for misconfigurations and vulnerabilities.	This may include regular vulnerability scanning, and penetration testing.	Recommended
2.2.1	You should regularly monitor the network configuration of your TRE to check for misconfigurations and vulnerabilities.	This will involve following the monitoring procedure detailed above.	Recommended

7.2.5 Infrastructure analytics application

This *application component* is an application which enables the *TRE operator* to record and analyse data about the usage of the TRE.

	Statement	Guidance	Importance
2.2.1	Your TRE must record usage data.	This may include the number of users, number of projects, the amount of data stored, number of datasets, the number of workspaces, etc.	Mandatory
2.2.1	Your TRE should record which datasets are accessed, when and by who.	This helps maintain auditability of how sensitive data has been used.	Recommended
2.2.1	Your TRE should record computational resource usage at the user or aggregate level.	This is useful for optimising allocation of resources, and managing costs.	Recommended

7.3 Capacity management

The ability of the *TRE operator* to ensure the right amount of resources are available at the right time to provide a cost-effective service.

7.3.1 Capacity Planning Process

This *business process* involves forecasting and determining the resources required to meet the demands of an application or system, ensuring that adequate resources are available when needed.

	Statement	Guidance	Importance
2.3.1	You must ensure that all projects understand what resources are available and what the associated costs will be before the project starts.	For on-premises systems this might be related to the available hardware, for cloud-based systems there might be limits on how many instances of a particular resource (*e.g.* GPUs) can be used Projects should use this information to understand whether the available resources will be sufficient for their requirements.	Mandatory
2.3.2	You should ensure that the anticipated needs of projects can be satisfied using available resources.	Note that this does not require you to accept requests for additional resources, but rather that promises made about resource availability before a project starts should be honoured wherever possible.	Recommended
2.3.3	You must have a procedure for allocating available resources among projects.	For cloud-based TREs this may involve scaling resources, such as virtual machines or databases, or deploying additional resources. For on-premises TREs this may involve a procurement process to ensure that necessary resources are available. Not all requests for capacity increase must necessarily be granted, but having a clear process will help projects understand when/why/how they can make use of additional capacity.	Mandatory

7.3.2 Billing Process

This *business process* involves generating and managing invoices and bills for projects within the TRE. It involves calculation, issuance, and recording of payments and receipts.

	Statement	Guidance	Importance
2.3.4	You must ensure that the anticipated resource requirements will not result in overspending by the TRE.	For cloud-based TREs this may involve budgeting and/or restricting resource consumption on a project-by-project basis. For on-premises TREs this may involve managing expectations to match the available resource.	Mandatory

7.4 Configuration management

7.4.1 Configuration Management Process

This *business process* involves the *TRE operator* identifying, maintaining, and verifying information on IT assets and configurations in the TRE organisation.

	Statement	Guidance	Importance
2.4.1	You must have a documented procedure for configuring infrastructure.	This might, for instance, be a handbook that is followed or a set of automated scripts.	Mandatory
2.4.2	You should use configuration management tools to automate application of your configuration wherever possible.	This might involve configuration-as-code tools such as Ansible, Chef, Puppet or Windows Desired State Configuration or simply automated scripts.	Recommended
2.4.3	You should be able to verify whether the configuration is valid.	This might, for instance, involve running your configuration management tool in 'check' mode.	Recommended
2.4.4	You should regularly verify your TRE configuration.	This will limit the amount of time the TRE can spend in a non-compliant state.	Recommended
2.4.5	You must be able to replace a non-compliant TRE with a compliant system.	This might involve reconfiguring a running system or by replacing it with a compliant one.	Mandatory

7.5 Information security

What the *organisation* does to safeguard research to ensure the confidentiality, integrity and availability of research resources and data.

Measures taken to ensure information security can be further categorised into:

- *vulnerability management*: applying security updates or fixes for identified vulnerabilities
- *security testing*: proactive penetration testing of a deployed system
- *encryption*: ensuring that data is protected even if the TRE is compromised
- *physical security*: restricting TRE access to known secure locations

A TRE conforming to the SATRE standard should enact broadly similar measures to protect against the unauthorised use of information, especially electronic data. These measures include vulnerability management of TRE infrastructure (whether physical or virtual/cloud-based), carrying out compliance checks and security tests of the TRE, common approaches to data encryption, and (where appropriate) physical security measures to prevent unauthorised access to the TRE.

7.5.1 Vulnerability Management

The ability of the *TRE operator* to identify, assess, report on, manage and remediate technical vulnerabilities across endpoints, workloads, and systems.

7.5.2 Vulnerability Management: Resilience Processes

A set of processes which ensures the TRE infrastructure can withstand disruption from incidents that risk confidentiality, integrity or availability of data.

	Statement	Guidance	Importance
2.5.1	You should keep backups of data and research environments, provided that this is permitted by law.	Keeping backups could help reduce the impact of events like accidental deletion and data corruption on work in a TRE. *TRE developers* may want to consider how different elements such as sensitive input data or users' workspaces may be backed up, and whether they should be.	Recommended
2.5.2	You should build redundancy into infrastructure and storage.	Infrastructure should be as resilient as necessary to interruption. This could include redundant infrastructure in different physical locations, load balancing and replication of data between multiple storage locations.	Recommended
2.5.3	You should keep backups of infrastructure, applications and configurations.	This may include virtualised infrastructure snapshots which can restored as needed to recover from failure.	Recommended

7.5.3 Vulnerability Management: Response Process

A process which ensures the organisation can quickly deal with incidents that risk confidentiality, integrity or availability of data.

	Statement	Guidance	Importance
2.5.4	You must have procedures in place for rapid incident response.	There may be legal requirements to disclose details of any incidents, such as data breaches for organisations subject to GDPR. Having robust processes in place will ensure a swift and effective response when an incident occurs.	Mandatory
2.5.5	You should test your incident response through simulation.	During simulated incidents the TRE organisation can measure their effectiveness. This may involve people across the broader enterprise and/or external suppliers.	Recommended

7.5.4 Vulnerability Scanning

The automated process of scanning computer systems or networks to identify and assess potential security vulnerabilities.

	Statement	Guidance	Importance
2.5.6	You should have an application in place to scan for vulnerabilities across infrastructure.	Software used to identify vulnerabilities should also report and alert. Such an alert should be triaged, risk assessed and treated accordingly.	Recommended

7.5.5 Security Patching

The process of applying updates or patches to software and systems to address known security vulnerabilities and flaws.

	Statement	Guidance	Importance
2.5.7	You must have a process in place for applying security updates to all software that forms part of the TRE infrastructure.	This includes any software used for remote desktop portals, databases, webapps, creating and destroying compute infrastructure, configuration management, or software used for monitoring the TRE.	Mandatory
2.5.8	Infrastructure should be automatically patched for vulnerabilities.	Planning will be required across infrastructure and software systems to ensure security patches remain available from suppliers. Many systems may be isolated from the internet making TRE infrastructure more difficult to automatically patch.	Recommended

7.5.6 Security testing

Security testing enables the *TRE operator* to gain assurance in the security of a TRE by testing or attempting to breach some or all of that system's security.

	Statement	Guidance	Importance
2.5.9	You should carry out penetration tests on your TRE.	By intentionally attempting to breach their TRE, organisations can proactively discover unnoticed vulnerabilities before they are exploited maliciously. Tests can evaluate the effectiveness of security controls in preventing data breaches, unauthorised access, or other security incidents.	Recommended
2.5.1(	You should update the security controls of your TRE based on the results of security tests.	Security testing can reveal bugs and discrepancies in the TRE architecture which should be addressed in advance of sensitive data being uploaded, or with urgency in the case of an operational TRE. Regular testing will allow organisations to refine their TRE security controls and incident response capabilities. It enables them to adapt to any new security concerns that may arise as a result of changes in the underlying software.	Recommended
2.5.1	You should publish details of your security testing strategy and, where possible, the results of each test.	Knowledge that regular security testing occurs will help to ensure stakeholders, including *data consumers* and *information asset owners*, can trust that the data they work with or are responsible for is secure within a TRE. If security flaws are identified in a test, it may not be sensible to publicise these until a fix is in place.	Recommended

7.5.7 Encryption

The ability of the *TRE operator* to deploy and manage encryption to protect information assets, including data for TRE research projects.

Here we define 'project' data as the data brought in for work which is very likely to be sensitive and 'user' data, as the working files of a project which might hold copies of all or part of the project data or otherwise reveal sensitive data (*e.g.* through hard coded row/column names).

	Statement	Guidance	Importance
2.5.1	Your TRE must encrypt project and user data at rest.	This prevents unauthorised access to the data even if the storage media is compromised. This may involve encrypted filesystems or tools to encrypt and decrypt data on demand. The encryption keys may be managed by the *TRE operator* or by a trusted external actor, for example a cloud provider.	Mandatory
2.5.1	Your TRE must encrypt data when in transit between the TRE and external networks or computers.	Data encryption must be used to safeguard against interception or tampering during transmission. This includes both data ingress and egress and users accessing the TRE, for example over a remote desktop or shell session.	Mandatory
2.5.1	Your TRE should encrypt data when in transit inside the TRE.	If possible, data transfers between different components of a TRE should also be encrypted.	Recommended
2.5.1	You should use encryption algorithms and software that are widely accepted as secure.	Encryption algorithms widely accepted as secure today may become insecure in the future, for instance due to newly-identified flaws, or advances in compute capabilities. The latest security patches and updates should be applied to any encryption software being used by the TRE. This helps address any known vulnerabilities or weaknesses in the encryption implementation.	Recommended

7.5.8 Key Management Application

Software or tools dedicated to generating, storing, and managing encryption keys securely, ensuring their availability and protection.

	Statement	Guidance	Importance
2.5.1	Your TRE should use secure key management.	TREs should employ secure key management practices, including storing encryption keys separately from the encrypted data and implementing strong access controls (*e.g.* Single Sign On) for key management systems.	Recommended

7.5.9 Physical security

The ability of the *TRE operator* to manage and protect physical assets from unauthorised access, damage or destruction.

Physical security controls can provide TREs using highly sensitive data an extra layer of security, even if technical controls are already in place for less sensitive data:

	Statement	Guidance	Importance
2.5.1	Your TRE could offer physical protection measures against data leakage or theft via physical means.	Restricting access to research facilities containing computers logged into TREs can help prevent malicious actors from viewing or stealing sensitive data, for example by photographing a computer screen. Physical controls on access to a TRE could include surveillance systems, restricting physical access to authorised personnel only, visitor management systems and employee training.	Optional
2.5.1	Your TRE may need to comply with specific regulatory requirements due to the types of data it is hosting.	Regulatory frameworks often emphasise the need for security controls to protect sensitive data. Compliance with these regulations could require organisations to implement specific security measures to safeguard their TRE from unauthorised access.	Mandatory

CHAPTER

EIGHT

DATA MANAGEMENT

This capability concerns the ability of the *TRE operator* to manage data assets and ensure information remains secure.

Fig. 1: SATRE Pillars Capability Map

8.1 Data lifecycle management

The ability of the *TRE operator* to manage how and where data is stored, how it moves, changes and is removed.

	Statement	Guidance	Importance
3.1.1	You must have processes in place to assess the legal and regulatory implications of handling the data through its full lifecycle.	This involves considering your obligations to data controllers and subjects, and whether any security controls may be legally or contractually required. An assessment of the risks involved will also be needed. It may involve classifying the project into a predefined sensitivity category or defining bespoke controls.	Mandatory
3.1.2	You should keep records of data handling decisions.	Decisions that are made as part of the process discussed above should be recorded and made available for inspection by all stakeholders.	Recommended
3.1.3	*Information asset owners* must classify data sets according to a common process and data classification methodology.	To classify the data, information asset owners must have a good understanding of the datasets and the process of classification. Once classified, data can be stored in a TRE with an appropriate security controls (see *later section on security levels and tiering*), which can factor in the requirements for confidentiality, integrity and availability of the data.	Mandatory
3.1.4	You must have a data ingress process which enforces information governance rules/processes.	The data ingress process needs to ensure that information governance is correctly followed. In particular, it should require that an ingress request has been approved by all required parties.	Mandatory
3.1.5	You must have a data egress process which enforces information governance rules/processes.	The data egress process needs to ensure that information governance requirements are adhered to. In particular, it should require that an egress request has been approved by all required parties.	Mandatory
3.1.6	Egress must be limited to the *information asset owners* or their delegates.	Egress of data from a TRE must be a specific permission associated with individual users This permission must be given by information asset owners. Egress may still require further approval (see 3.1.5).	Mandatory
3.1.7	Your data egress process could sometimes require project-independent approval.	There may be cases where there are multiple stakeholders for a piece of analysis including *information asset owners*, data analysts, data subjects, the *TRE operator*. A data egress process may then require approval from people not on the *project team*, for example an external referee or *TRE operator* representative	Optional
3.1.8	You must keep a record of what data your TRE holds.	Good records are important for ensuring compliance with legislation, understanding risk and aiding good data hygiene. The record should include a description of the data, its source, contact details for the data owner, which projects use the data, the date it was received, when it is expected to no longer be needed.	Mandatory
3.1.9	You must have a policy on data deletion.	There should be a clear, published policy on when data will be retained or deleted. This may allow time for data owners to consider outputs they may want to extract from the TRE. Any sensitive data, including all backups, should be deleted when they are no longer needed. Having clear policies will help to avoid problems with data being kept longer than necessary or accidental deletion of outputs.	Mandatory
3.1.10	You should have a method of providing proof of deletion/removal	*Information asset owners* may require certification of the deletion of files. You should have a method of providing proof of deletion if challenged.	Recommended

8.2 Identity and access management

The ability of the *TRE operator* to ensure the right people (identities) can only access the tools and data they need.

	Statement	Guidance	Importance
3.2.1	You must not create user accounts for use by more than one person.	It is important that each user account should be used by one, and only one, person in order to facilitate the assignment of roles or permissions and to log the actions of individuals.	Mandatory
3.2.2	You must be reasonably convinced of the identity of each person being granted an account.	It is important to ensure an account has been given to the correct person. For example, multiple credentials may be used before account creation to verify identity or, when appropriate, photo ID checks may be required.	Mandatory
3.2.3	You must restrict a user's access to only data required in their work.	There is no need to grant an individual access to data they do not require. Access may be assigned in a manner appropriate to a TREs design, for example through roles granted to user accounts or through isolated project workspaces.	Mandatory
3.2.4	You must ensure that multi-factor authentication is enabled for all users.	Multi-factor authentication ensures that to successfully connect a user must have more than one piece of evidence in different categories. Categories include something the user knows (*e.g.* a password), something the user possesses (*e.g.* a TOTP key) or something the user is (*e.g.* biometric data). A TRE does not need to implement multi-factor authentication checks itself if it is provided by a third-party identity provider.	Mandatory
3.2.5	You could use federated authentication or single sign-on (SSO) for user login.	Institutions that use a SSO for other applications may wish to extend this login capability to a TRE. This will simplify the login process for *data consumers* using a TRE and prevent them having to remember or store multiple login credentials.	Optional
3.2.6	You could restrict access to particular networks or physical locations.	Restricting access to a set of known, static, personal or institutional IP addresses can help avoid speculative attacks. When appropriate, access could also be restricted to physical locations with security controls and access requirements.	Optional

8.3 Output management

The ability of the *TRE operator* to ensure outputs are safely published and shared.

	Statement	Guidance	Importance
3.3.1	You should have a system to help classify outputs.	Removing data from a TRE can be a difficult process as there is potential for sensitive data to be revealed. Having guidance, processes and methods will help ensure that outputs are correctly classified and, furthermore, that outputs due to be openly published are identified. Encouraging openly published outputs will enhance a TRE's impact and transparency.	Recommended
3.3.2	You should establish the intended outputs of each project from the outset.	Identifying the purpose of a piece of work is important for compliance with data protection legislation. Results will be produced which address the project's purpose, some of which may be outputs that are removed from the TRE. Understanding what these outputs are likely to be and their sensitivity as early as possible will help prepare for their processing and publication.	Recommended
3.3.3	You must have a documented process for disclosure control of outputs from the TRE.	This process should define expected risks and how to mitigate them. All TRE outputs must be subject to this process. You might choose to follow existing guidelines, for example around statistical disclosure.	Mandatory
3.3.4	You must have a process for assigning responsibility for output checking.	*Output checkers* should be given responsibility for checking outputs. They must follow your disclosure control process and will be responsible for any automated parts of this process. Output checking can help mitigate against unintentional data disclosure or leaks.	Mandatory
3.3.5	You must have a documented policy for handling disclosure risks associated with any outputs that cannot be manually checked.	Some categories of output, for instance binary files or very large numeric files, can be difficult to manually check. If egress of such files is permitted then the risks of inadvertent disclosure must be mitigated and documented. Refusing to allow egress of such files is also a valid policy decision.	Mandatory
3.3.6	You should have a statistical basis to guide the decisions of an output checker on the safety of outputs.	There should be a solid basis to allow decisions to be made about data based on risk factors such as re-identification of an individual or risk to commercial operations posed by outputs from the TRE.	Recommended
3.3.7	You could create a semi-automated system for checks on common research outputs.	Automation helps make decisions on outputs more consistent and reduces the overhead for output checkers. It's unlikely however that a fully automated output checking system (without humans in the loop) would be appropriate, given the risks associated with accidental data disclosure.	Optional
3.3.8	TRE outputs should be limited to the minimum required for sharing results of any analyses.	This decreases the risk of inadvertent disclosure, and makes it easier to comply with data protection legislation (e.g. GDPR).	Recommended

8.4 Information search and discovery

The ability to query and browse the data within an environment at various levels of abstraction.

	Statement	Guidance	Importance
3.4.1	You should provide a metadata catalogue of available datasets for users.	This is particularly relevant for TREs with population-level data collection of general interest. This may not be appropriate for TREs where each project has its own data sharing agreement with one or more data provider or very sensitive datasets.	Recommended

8.5 Security Levels and Tiering

The ability of the TRE deployment software (or active TRE) to configure security controls appropriate to the sensitivity of the data used in a project or workspace.

Security controls can add friction to the user experience and hinder work. A one-size-fits-all approach forces all projects to use the strictest security configuration even when that is unnecessary. Throughout the rest of this document, we will refer to each pre-defined security configuration supported by a particular TRE as a "security tier".

	Statement	Guidance	Importance
3.5.1	You must be able to specify what categories of data your TRE is able to support.	Your TRE must provide an explanation of the kinds of data it has been designed to hold, with reference to its security capabilities, that can be understood by all stakeholders. Relevant stakeholders may include *information asset owners* and *project teams* and they may have different levels of technical expertise.	Mandatory
3.5.2	Your TRE could support projects with differing security requirements through configurable security controls.	This allows projects with different security requirements to each be met with a suitable level of controls. It helps ensure that users can work effectively, with minimal barriers.	Optional
3.5.3	Your TRE could offer a pre-defined set of security control tiers.	Security control tiers can be designed to cover the types of project or data you expect to handle. Projects may be placed into the most suitable tier rather than having a bespoke design. This reduces the number of unique configurations that need to be supported.	Optional

8.6 Research Meta-Data

Descriptive information about research data, helping researchers understand and manage the data effectively.

	Statement	Guidance	Importance
3.6.1	You should have a consistent and easily accessible meta-data data model or similar to describe what a data asset contains.	Where possible, existing data models should be employed (and extended if necessary). More detailed information on the data schema for data assets should also be provided to assist researchers in understanding what data may be available without the need to see the underlying data.	Recommended
3.6.2	You could provide summary, abstracted or synthetic data to researchers without exposing the underlying data set.	To reduce the need for access to row level data researchers could be provided with non-sensitive versions of the data either as summary data or using synthetic versions of the data for activities such as code development and cohort planning.	Optional

8.7 Meta-Data Search and Discovery Application

Software designed to help users locate and retrieve specific metadata or information within a database or system.

	Statement	Guidance	Importance
3.7.1	You could provide an interface application for *data consumers* and *data subjects* to query elements of the data.	In order to make data findable, an application which queries the meta-data or elements of the research data could be made more easily accessible than the data itself.	Optional

8.8 Data Archiving

The practice of storing data that is no longer actively used but needs to be retained for historical or compliance reasons.

	Statement	Guidance	Importance
3.8.1	Archived data within the TRE should be read only.	Archived data by its very nature should not change and therefore be maintained as a read only store. If an update is required, it may be pulled from archive into a separate operational store.	Recommended
3.8.2	Long-term archives must be held in simple, standard formats to ensure accessibility.	Some data archives may be required by policy or legislation to be kept for very long periods within the scope of the TRE. Such data should be held in the simplest possible file format, conforming to international standards if available, to ensure they are platform and application agnostic.	Recommended

CHAPTER

NINE

SUPPORTING CAPABILITIES

Fig. 1: SATRE Pillars Capability Map

9.1 Business continuity management

What the *TRE operator* does to ensure the development, testing, and maintenance of business continuity plans.

	Statement	Guidance	Importance
4.1.1	You should have a business continuity plan that includes consideration of loss of service for deployed TREs.	This may be due to downtime from service providers, a breach, or loss of power. Your plan should detail your process for managing loss of service for deployed TREs, and evaluation of impact of such loss.	Recommended
4.1.2	You should regularly test the aspects of your business continuity plan concerning TREs, and have a process in place to iterate the plan if required.		Recommended

9.2 Project and programme management

What the *TRE operator* does to ensure effective management of programmes and projects.

	Statement	Guidance	Importance
4.2.1	You should ensure that all projects using your TRE have a named project manager.	The project manager has responsibility to ensure the smooth running of the project. Their responsibilities may include budget management, tracking TRE status, managing communications with the TRE operations team, and other project support tasks.	Recommended
4.2.2	You should not give project managers direct access to the TRE.	Doing so ensures a separation between those able to access sensitive data, and those overseeing access to sensitive data.	Recommended

9.3 Knowledge management

What the *TRE operator* does to acquire, enrich, share, store, publish and enhance expertise across their organisation.

	Statement	Guidance	Importance
4.3.1	You must document all features of your TRE implementation.	This includes ensuring all documentation is discoverable, clear, and able to be easily updated based on stakeholder feedback	Mandatory
4.3.2	You should have an education programme in place to upskill stakeholders in the use and management of your TRE.	This may include learning modules, workshops and other resources on how to effectively access and use a TRE, FAQ pages, and accessible pathways for additional support	Recommended
4.3.3	You should periodically carry out a training needs analysis (TNA) for all stakeholders included within your TRE provision.	At least once every 12 months you should assess the training needs of your stakeholders, and ensure they have easy access to all required training materials	Recommended

9.4 Financial management

All activities aimed at the efficient and effective management of money (funds) in such a manner as to allow the *TRE operator* to accomplish its objectives.

	Statement	Guidance	Importance
4.4.1	You must ensure that all projects using your TRE are aware of any associated costs and are able and willing to pay them.	Costs may include provision of the underlying TRE infrastructure, additional resources required in a specific TRE (for instance memory or additional compute), hardware including managed devices, and staff support costs	Mandatory
4.4.2	You should be able to track the costs associated with each TRE project.	This includes knowing which costs are associated with which project, and having an appropriate charging mechanism in place in line with your organisational policy.	Recommended
4.4.3	You should have a process in place to ensure your TRE provision remains financially sustainable.	This could include having a cost recovery process in place, or setting up a long-term funding mechanism to support projects with TREs. At any given time, you should have funds free to cover all potential foreseen TRE provision for at least 12 months.	Recommended
4.4.4	You should minimise the cost of your TRE infrastructure wherever possible	You should have regular reviews of your TRE provision and actively work to bring down costs, streamline provision, and optimise support.	Recommended

9.5 Procurement

What the *TRE operator* does to ensure the effective sourcing, purchasing and supply of the goods and services that enable them to operate.

	Statement	Guidance	Importance
4.5.1	You must identify any goods or services that will be needed to operate the TRE and ensure that a plan is in place to purchase them as needed.	These may include computing hardware, cloud credits or devices through which users access the TRE.	Mandatory

9.6 IT Service management

The implementation and management of quality IT services that meet the needs of the *TRE operator*.

	Statement	Guidance	Importance
4.6.1	Your TRE must have a team of *Operators* in place to support projects working with TREs.	This may be part of your organisation's IT support team, or separate. Responsibility should be clear and stakeholders should easily be able to access support appropriate to their needs.	Mandatory

9.7 Relationship management

All activities aimed at ensuring a continuous level of engagement is maintained between the *TRE operator* and its customers, stakeholders and other interested parties.

9.7.1 Stakeholder relationships

Activities aimed at engaging with TRE stakeholders.

	Statement	Guidance	Importance
4.7.1	You should have a clear process in place for stakeholders to feedback on your TRE infrastructure.	This may include a GitHub repository where people can open issues and discussions, communication streams like Slack or email, or forms stakeholders can fill in.	Recommended

9.8 Public Involvement and Engagement

How the *TRE operator* involves the public in its processes and work in order to maintain trust in its operations.

	Statement	Guidance	Importance
4.8.1	All public engagement activities must include a range of perspectives and be inclusive (*optional for TREs without personal data).	Any public engagement activity carried out by TREs should involve diverse participants and that activities are accessible. Recruitment plans should consider how to proactively reach a representative sample of people or target particular groups of people where relevant This could include following guidelines such as PEDRI.	Mandatory*
4.8.2	Details of TRE operations, data available and projects which have accessed the data should be publicly available (*optional for TREs without personal data).	TREs should be as transparent as possible by providing information online. Where information is made available online this should be written in clear language understandable to general public. A record of projects which have accessed data via the TRE should be kept and made available. Where possible it should include name, summaries, public benefit (if relevant) and organisations involved	Mandatory*
4.8.3	Members of the public should be included in TRE operations and/or oversight (*optional for TREs without personal data).	Members of the public can be involved via presence on steering groups or project approvals panels. Alternatively TRE's can establish separate public panels available for both researchers and TRE staff to consult.	Mandatory*
4.8.4	You should publicly share details of incidents, near misses, and mitigations in a timely fashion, in line with good practices for responsible disclosure.	This may be via the TRE website or annual reports. Sharing this information is particularly important when a TRE holds public sector data.	Recommended

9.9 Legal services

The ability of the *TRE operator* to access suitable and timely legal advice.

9.9.1 Legal advisory

The ability of the *TRE operator* to provide suitable and timely legal advice.

	Statement	Guidance	Importance
4.9.1	You should identify areas where legal advice may be required and ensure that you have ready access to it.	It is likely that legal advice will be necessary for several issues around the handling of sensitive data, and managing project contracts. *TRE operators* should have ready access to legal advice, including a way to solicit advice and carry out associated actions.	Recommended

9.9.2 Data protection

Ability to ensure data is used fairly, lawfully and transparently; for specified, explicit purposes; and in a way that is adequate, relevant and limited to only what is necessary.

	Statement	Guidance	Importance
4.9.2	You should identify areas where advice on data protection issues may be required and ensure that you have ready access to it.	It is likely that data protection advice will be necessary for several issues around the handling of sensitive data.	Recommended

9.9.3 Contract management

What the organisation does to ensure that all contracts are effectively managed within required frameworks.

	Statement	Guidance	Importance
4.9.3	You should identify who will be responsible for managing contracts related to the TRE.	These contracts may include data sharing agreements, secondments of personnel or limitations on how results obtained with the data can be distributed.	Recommended

CHAPTER

TEN

EVALUATING TRES AGAINST SATRE

This section details the method for evaluating a TRE against the SATRE specification.

This document also includes two example evaluations for *The Alan Turing Institute's Data Safe Haven* and *The University of Dundee/HIC's TREEHOOSE*. We hope that these examples will help you to write your own evaluation.

10.1 Who should evaluate a TRE against SATRE?

This section is aimed at *Operators* and *Information Governance Managers* of TREs at institutions hosting sensitive data research projects. The example evaluations provided may also be of use to TRE *Developers* who wish to review existing implementations as well as the specification.

10.2 Why should I evaluate my institution's TRE?

The SATRE specification has been compiled from the knowledge around successful TRE provision from a variety of institutions. This includes information governance procedures, computing technology, data management and other capabilities.

By scoring your institutions' TRE against the specification using the method below, you can:

1. Identify any technical oversights in the way your TRE is designed that could lead to unintended disclosure of sensitive data or inappropriate user access.
2. Identify any operating procedures that could be improved for your TRE and how to improve them, which will also minimise risks and ensure the smooth operation of TRE-based research projects.
3. Compile a wish list of capabilities that your TRE lacks (or could be improved). You could for example, cite the SATRE specification as evidence for resources (computational or human) needing to be allocated by your institution.

Note: SATRE is *not* a technical standard for which formal accreditation can be achieved. For more info see: *Is SATRE an ISO technical standard?*

10.3 Method

You should score your TRE against each statement in the SATRE specification using this scoring system:

0 Not met
: The TRE does not meet this requirement (if this is **Mandatory** this means the TRE is not SATRE compliant)

1 Sufficient
: The TRE meets this requirement met but there is substantial scope for improvement

2 Satisfied
: The TRE meets this requirement met but there may still be scope for improvement

N/A
: Not applicable: The statement is not relevant to a TRE, may apply to **Recommended** or **Optional** statements, and a very limited number of **Mandatory*** statements.

A score of **1** or above means you have met the requirement. Optionally you can use **1** and **2** to indicate potential areas of improvement in your TRE.

An evaluation may simply give your TRE scores for each statement. We recommend a more detailed evaluation, which includes a score, a justification and, where applicable, suggestions for improvement.

The example evaluations are detailed, including the supporting text as well as scores.

10.3.1 Combining scores

The scores for each statement can be easily combined at the capability, pillar or overall level. If all the **Mandatory** statements in a capability are met, either at level **1** or level **2**, then the capability is met. If all capabilities in a pillar are met then the pillar is met. If all pillars are met then the SATRE specification is met.

10.4 Evaluation spreadsheet

You can use this spreadsheet as a template for your evaluation.

CHAPTER

ELEVEN

ALAN TURING INSTITUTE DATA SAFE HAVEN

The Alan Turing Institute is the UK's national institute for data science and artificial intelligence. The Data Safe Haven project's goal is to remove barriers to working safely and effectively with sensitive data, by promoting and demonstrating a culture of open, community-led development of interoperable foundational infrastructure and governance. The project maintains an open-source TRE project which covers governance, documentation and programmatic deployment of a TRE. Data Safe Haven can be freely used and adapted to deploy a TRE to Microsoft Azure.

The Turing uses the Data Safe Haven TRE and governance to enable research on sensitive data. This includes work with external partners and our Data Study Group collaborative hackathons. The evaluation below has been carried out for the Turing's production TRE, using the Data Safe Haven technical implementation and institutional governance processes.

11.1 Governance Requirements

	Score	Response
1.1.1.	1	• We rely on the Alan Turing Institute legal and data protection teams to inform us of necessary legal requirements. • Our TRE (the Alan Turing Institute Data Safe Haven) is self-assessed against the NHS Data Security and Protection Toolkit. • The Institute has a centralised ethics approval procedure which all projects are required to follow. • Project specific requirements are discussed with *information asset owners* before start of each project. **Potential Improvements** • Be more proactive about trying to stay up-to-date with changes to requirements. • Hold an annual review with the Turing Legal and Data Protection teams to ensure all current requirements are still met by the TRE organisation.
1.1.2.	1	• We have documented our institutional risk appetite and understand what kinds of work we are prepared to support. • All projects handled in the TRE have gone through our institutional Data Protection Assessment Process and Ethics Approval Process. • All projects are tracked through a ticketing system with relevant documents and agreements stored on Sharepoint. **Potential Improvements** • Improve our document handling workflow by developing an Information Governance App through which all projects can be managed directly by approved users.
1.1.3	1	

11.2 Quality Management

	Score	Response
1.2.1.	1	• Our SOPs are held on a public website backed by a private GitHub repository with limited edit access. • Policies and other forms that need signatures as part of our SOPs are held in a private Sharepoint folder with limited access. • Acceptance of policies is recorded in a document held in a private Sharepoint folder. • We do not have an explicit process for determining who should have administrator access to these folders and repositories. **Potential Improvements** • Develop an induction process and/or mandatory training programme for potential administrators.
1.2.2.	1	• All SOPs are stored in a version controlled repository. • Policies, forms and project documentation are stored in a controlled Sharepoint folder. • We do not currently use explicit versioning for forms/documents that need to be signed. **Potential Improvements** • Use tags to refer to documents handled by git version control. • Use in-file versioning for files stored on Sharepoint.
1.2.3.	0	• We regularly discuss our information governance processes with impacted projects and look for ways to improve and streamline them. • We have SOPs for handling security incidents and running investigations. • We do not have a formal measurement process. • We do not regularly report our information governance per-

11.3 Risk Management

	Score	Response
1.3.1.	1	• We have a risk register which scores risks in a matrix based on their likelihood and potential consequences. **Potential Improvements** • Regularly update our risk register.
1.3.2.	2	• All Turing projects must carry out a Data Protection Assessment Process. • We also have a flowchart that *project teams* and *information asset owners* must follow to agree on the security tier of their project before it starts.
1.3.3.	1	• We decide on risk mitigations during our risk assessment process, but this tends to be an *ad-hoc* process rather than anything formalised.. **Potential Improvements** • Develop a formal risk mitigation process.
1.3.4.	1	• We have a documented set of roles together with responsibilities required for each role. • There is an implicit association of risks with particular roles, but we are not explicit about the relationship between risks and roles. **Potential Improvements** • Create explicit mapping of risks to roles.
1.3.5.	1	• Our TRE deployment is guided by the risk appetite of the wider Institute. • We have not yet encountered risks that fall outside our known risk appetite. **Potential Improvements** • Develop a procedure for han-

11.4 Study Management

	Score	Response
1.4.1.	2	• All Turing projects must carry out a Data Protection Assessment Process. • All projects must have an agreed security tier before starting. • We inform projects in advance of their estimated directly incurred infrastructure costs and require them to confirm that they will be able to pay for these. • Data sharing agreements must be in place before any data ingress.
1.4.2.	1	• As soon as we are informed of the need to revoke user access, we will do so. • Lists of responsible persons are established at the beginning of each project and these are kept up-to-date. • We tend to reactively respond to user removal requests rather than actively confirming that users are active. **Potential Improvements** • Develop a regular process for confirming that users are active.
1.4.3.	1	• Our updating process is passive, as we rely on our Data Protection and Legal teams to inform us of changes to relevant legislation. • We do not have formal checks in place. **Potential Improvements** • Regularly check for changed requirements with Legal and Data Protection teams.
1.4.4.	2	• We have processes in place to handle egressing results, removing access, securely deleting any data and destroying the infrastructure. • Data egress needs to be agreed

11.5 Member Accreditation

	Score	Response
1.5.1.	1	• Project managers are required to provide an email address and phone number for each user before we set up their account. • Their username is sent to their email address together with a link to a self-service password reset page. • Their password cannot be updated without providing a code that is sent to their registered phone number. **Potential Improvements** • Consider making more detailed checks on user ID, possibly delegating to a trusted third-party.
1.5.2.	2	• Onboarding documentation exists for both *TRE operators* and *project teams*. • Users must complete appropriate training and sign our terms of use before being granted access to the TRE.
1.5.3.	2	• We use Microsoft Entra to manage user accounts. • Access to resources and data is controlled by RBAC.
1.5.4.	2	• We have a process for agreeing which people are able to take which actions involving sensitive data. • Delegation of approval authority is also included here. • A document summarising these decisions must be signed by the *project manager*, *information asset owner* and referee before the project begins.
1.5.5.	2	• Initial log in is delegated to Microsoft Entra via OAuth. • This requires username, password and MFA. • Log in to computing resources within the TRE is controlled

11.6 Training Delivery and Management

	Score	Response
1.6.1.	1	• Part of the project initiation process involves agreeing on appropriate user training. • All TRE users carry out GDPR and cyber security training. • Depending on the project, they may also complete eLfH data security awareness training (level 1) and MRC Research, GDPR & Confidentiality Training • *TRE operators* must pass all user training requirements as well as data handling and cyber security training. **Potential Improvements** • Put a regular systematic training needs analysis into place.
1.6.2.	1	• Turing employees have access to internal GDPR and Cyber Security courses. • For non-Turing users we rely on their host organisation having sufficient training in these areas. • The eLfH and MRC courses are available to all. **Potential Improvements** • Develop data handling courses that will be available to all to close the current gap between Turing and non-Turing users.
1.6.3.	1	• We require all users and administrators to keep their training up-to-date. • All training certifications must be refreshed each year. **Potential Improvements** • Implement a system that creates alerts for users/admins when training is needed.
1.6.4.	1	• We have a register of people, their training require-

11.7 End user computing interfaces

	Scor	Response
2.1.1.	2	We do not allow data to move between the system clipboard and workspace in any instance.
2.1.2.	2	We provide both virtual desktop and command line interfaces to Linux virtual machines. Self-hosted web applications focused on collaborative work are accessible within the environment.
2.1.3.	2	We do not provide a job-submission interface, all users have direct access to the data they are working with.
Capability met?	**YES**	

11.8 End user software tools

	Score	Response
2.1.4.	2	We use a virtual Linux desktop, accessible via a web browser. We use standard, open-source tools, like Apache Guacamole, to support this.
2.1.5.	1	• We have a user guide that explains how to use the installed software, as well as how to configure your user account. **Potential Improvements** • We intend to iterate on the design of the user guide to make it easier to navigate, follow and understand - and separate it entirely from developer docs.
2.1.6.	2	We use the Azure platform-level automation tools to run weekly software updates on all virtual machines that make up the TRE. Any update failures are flagged by the automation software.
2.1.7.	2	Within each project environment we have a range of shared services. These include shared folders, user services such as GitLab, for collaborating on code, CodiMD, for collaborating on document writing and several database systems.
2.1.8.	2	These shared services are only available to users working within the same environment.
2.1.9.	2	User-facing software and tools are all open source. We do not allow any software to contact external licensing servers.
2.1.10.	2	We provide a wide range of tools and applications for data science, influenced by the needs of users. Our users are typically data scientists working with data directly. This data can only be accessed from inside the TRE, either via a database or a shared folder.
Capability met?	**YES**	

11.9 Code Version Control System

	Score	Response
2.1.11.	1	• Version control tools are provided to users, including an internal GitLab instance. • Users are encouraged to version control their code and we provide training for those who are unfamiliar with git. **Potential Improvements** • We do not provide specific tools to aid or encourage reproducibility or creating data analysis pipelines. • We do not support CI pipelines on our GitLab server. • We do not have a method to ensure that work done inside the environment can be reproduced outside such as containerisation.
Capability met?	**YES**	

11.10 Artefact Management Application

	Score	Response
2.1.12.	2	We provide proxied access to external software repositories, currently PyPI and CRAN, using Sonatype Nexus. For our highest sensitivity projects we instead provide a local mirror. In either case, we can support either access to every package in the remote repository or a pre-specified allowed list of approved packages.
2.1.13.	2	• For higher sensitivity environments, we restrict access to a pre-specified allowed list. These allowed lists are configurable on a per-project basis and, by default, include a minimal set of well-used and useful packages plus their dependencies.
Capability met?	**YES**	

11.11 Advanced Computing Systems

	Score	Response
2.1.14.	2	Non-standard resources are segregated in the same way as standard resources. We do not share any resources between projects.
2.1.15.	2	We are able to deploy high capacity virtual machines if required. These can have many cores and/or large amounts of RAM.
2.1.16.	2	We are able to deploy VM sizes featuring GPUs within the limits of what is available on Azure and compatible with our pre-built x64 image.
2.1.17.	2	We make Microsoft SQL server and/or PostgreSQL servers available to projects as needed. These databases are only accessible from inside a single project environment.
2.1.18.	1	We do not currently support large-scale data analytics tools. **Potential Improvements** • We would consider supporting Spark but it has not been requested by users.
Capability met?	**YES**	

11.12 Infrastructure Deployment Process

	Score	Response
2.2.1.	2	We have a detailed deployment guide which system managers follow to deploy a TRE instance.
2.2.2.	1	• We use reproducible Powershell scripts that are stored in GitHub to handle our deployments. • This code is regularly tested and new versions released. **Potential Improvements** • We do not currently use third-party infrastructure-as-code tools but we are in the process of moving to them.
2.2.3.	1	• We have documented procedures to add/remove users and to resize, add or remove infrastructure components such as GPU-enabled machines. • We do not make ad-hoc or unusual changes to deployed infrastructure in the course of normal operation. • In emergencies, we would deploy a fix that had been tested in development and then hold an incident report meeting. **Potential Improvements** • We do not currently have a formal process for making emergency changes to our production system.
2.2.4.	1	• We use separate development environments to test changes before they make it into a release. • Emergency fixes to our production environments are also tested on development environments before being deployed. • Production environments are created from known, tested releases of the codebase. **Potential Improvements** • We do not currently have a for-

11.13 Infrastructure Removal Process

	Score	Response
2.2.6.	2	We use Powershell scripts to automate the removal of unused infrastructure. We have documented procedures that detail when this should be done.
Capability met?	**YES**	

11.14 Availability Management Process

	Score	Response
2.2.7.	2	Azure publishes availability and uptime guarantees for relevant services. We have chosen replication levels which balance high availability while keeping data within a single region.
2.2.8.	0	We do not have an availability target. We do not make any availability guarantees to our users.
Capability met?	**YES**	

11.15 Network Management Application

	Score	Response
2.2.9.	2	We use Azure network security groups and firewalls to control network traffic between different parts of the TRE. Only the minimum necessary categories of traffic are permitted. The TRE gateway only permits connections from pre-approved IP addresses.
2.2.10.	2	Different projects are isolated at the virtual network level. Data sets belong to a single project only and are stored in storage accounts which only that project can access. Normal users have no way to directly connect to other project environments even if they have valid accounts for them.
2.2.11.	2	We block outbound connections to the internet unless these are required for functionality, such as system updates. All outbound connections are monitored by the Azure firewall.
2.2.12.	0	We do not actively monitor our TRE for misconfiguration. Unexpected connections would show up in our firewall logs. **Potential Improvements** • We are interested in hearing how other community members approach this.
2.2.13.	0	We do not actively monitor our TRE for misconfiguration. **Potential Improvements** • We are interested in hearing how other community members approach this.
Capability met?	**YES**	

11.16 Infrastructure analytics application

	Score	Response
2.2.14.	1	We keep track of users in Microsoft Entra, projects on a GitHub project board, datasets associated with each project in Sharepoint and workspaces associated with each project on GitHub issues. **Potential Improvements** • This data is not currently stored in one place, and the processes for tracking data are not clearly defined.
2.2.15.	1	Each dataset is associated with a single project. Only users associated with that project are able to access it. We do not keep track of instances of individual users accessing particular datasets. **Potential Improvements** • We cannot think think of a better way to do this now, but are interested in exploring options with the community.
2.2.16.	2	We record computational resource usage at the project level. We have no way to break down usage at the per-user level and do not think this would be useful for us since costs are managed at the project level.
Capability met?	**YES**	

11.17 Capacity Planning Process

	Score	Response
2.3.1.	2	At the planning stage, we make projects aware of possible resources, and associated costs. This information includes common configurations and requirements (such as GPUs), possible additional resources, and their costs. The costs of the shared aspects of the TRE and the TRE service (support, admin time) are also explained and broken down on a per-project basis.
2.3.2.	1	For our projects, we rely on the Azure availability guarantees about compute resources. We have limited control over the availability of Azure resources and sometimes there may not be available capacity.
2.3.3.	1	Our TRE is deployed on the Azure cloud. The availability of resources is therefore determined by the capacity of the cloud provider. Deciding on the distribution of resources between projects is not a large concern as the availability of resources, generally, greatly exceeds our need. **Potential Improvements** • Allocating resources to projects is currently done on an ad-hoc basis depending on project needs. We would like to make this process more formal and better documented.
Capability met?	**YES**	

11.18 Billing Process

	Score	Response
2.3.4.	2	We provide projects with estimates of their spend which are dependent on their requirements. We track spending on a per-project basis and allow the project manager to monitor spending. Spending alerts are sent out when spending reaches set thresholds: 50%, 90%, 100% of the pre-agreed limit. Overspend is possible but the additional spending must still be recovered from the project. **Potential Improvements** • We should be clearer about the consequences of overspending.
Capability met?	**YES**	

11.19 Configuration management

	Scor	Response
2.4.1.	1	We have a detailed deployment guide which system managers follow to deploy and configure a TRE instance. We have a limited set of documentation covering making common configuration changes after a TRE has been deployed.
2.4.2.	0	We do not use configuration management tools. We have a limited set of scripts to make some common configuration changes. Some changes involve manual steps which may be documented.
2.4.3.	0	There is no general, automated way to check the configuration of our TRE. A manual check would be time consuming and no process for doing so has been established. Security package update compliance for Ubuntu and Windows virtual machines can be confirmed.
2.4.4.	0	We are unable to verify configuration and so do not regularly check for compliance.
2.4.5.	1	We can replace non-compliant instances and/or components using out deployment processes and scripts. We are able to do this in a manner which avoids data loss. However, it will generally involve destruction and redeployment of infrastructure.
Capability met?	**YES**	

11.20 Information security

	Score	Response
2.5.1.	1	Some research environment data is backed up. This includes virtual disks and object storage accounts which contain users personal/configuration files and working data. Backups are distributed across data centres within a single region. Input data is only kept as a single, immutable copy which is not backed up (although users may make copies which would be). Because input data is always a copy, we are not concerned about the loss of input data. **Potential improvement** • We could ensure that non-file working data, such as database contents are also backed up.
2.5.2.	2	We use Microsoft Azure's features for geo-redundant storage for data, which can handle load balancing and replication of data between multiple storage locations. For TRE computng infrastructure, components are replicable via infrastructure as code.
2.5.3.	2	Infrastructure is defined by configuration files and replicable via infrastructure as code.
2.5.4	2	Our terms-of-use require users to report any potential data incident. We have a process in place for managing data incidents, whether raised by users or discovered independently, that ensures we meet our legal requirements and also implement any necessary changes, such as disabling access to a TRE if necessary.
2.5.5.	0	Although we have have a process for incidents, we don't have a incident response simulation process.
2.5.6.	0	Azure handles update automation, but not vulnerability scanning specifically.
2.5.7.	1	• Many cloud services, for example virtual networks, are kept up to date by the cloud provider. • All Windows and Ubuntu virtual machines have system package updates automat-

11.21 Data lifecycle management

	Score	Response
3.1.1.	2	Legal and regulatory implications are considered as part of the Data Protection Assessment Process (DPAP) when projects are first proposed. Each project is classified into one of five pre-defined security tiers before any work starts. Each tier has an associated set of security controls, although additional controls can be imposed on top of these if required.
3.1.2.	2	A signed approval form is required for each instance of data ingress or egress. A signed validation form must be filled out by the project team to confirm that any data moved in or out of the environment is as expected. A signed approval form for the security tier of each project is also required. These signed forms are kept in a private sharepoint folder, maintained by the TRE operators.
3.1.3.	2	*Information asset owners* must undergo a data classification process by following a flow chart to determine which of five sensitivity tiers data falls into. The data will then only be used within a TRE of an equivalent security tier (or higher).
3.1.4.	2	We implement data handling restrictions on data coming into the environment. These involve getting agreement from the *information asset owner*, *project manager* of the project and an independent representative from the Institute before any data or outputs are moved into the TRE. These stakeholders must sign a form detailing the requested ingress to confirm their agreement.
3.1.5.	2	We implement data handling restrictions on data coming out of the environment. These involve getting agreement from the *information asset owner*, *project manager* of the project and an independent representative from the Institute before any data or outputs are moved out of the TRE. These stakeholders must sign a form detailing the requested egress to confirm their agreement. These signed forms are kept in a private sharepoint folder, maintained by the TRE operators.

11.22 Identity and access management

	Score	Response
3.2.1.	2	Each user only has a single account. We assume that only the authorised user will have possession of the correct username, password and physical MFA device.
3.2.2.	2	Our user creation process involves multiple identification factors to reasonably convince us of the identity of the person holding access to an account. **Potential improvements** • We could perform more detailed ID checks, perhaps by requiring photo ID.
3.2.3.	2	Each project's data is held separately. It is not possible to mix data between projects, even if an individual is a member of multiple projects.
3.2.4.	2	MFA is enforced for all users through Microsoft Entra. The second factor can be either push notification or a phone call.
3.2.5.	2	We use dedicated credentials for our TRE that are separate from any other accounts. A user who is working on multiple projects will use the same credentials for each of them.
3.2.6.	2	We are able to restrict access to known IP addresses. Where appropriate, IP addresses are restricted to the static institutional or personal IP addresses of the users allowed to connect to the environment. Sometimes, users are required to only access the TRE from inside the Institute's office space.
Capability met?	**YES**	

11.23 Output management

	Score	Response
3.3.1.	1	All outputs from a TRE go through our security classification process, carried out by the project investigator, *information asset owner* representative and an independent referee at the Turing. Different egress processes are required according to the sensitivity of the outputs. **Potential improvements** • We would like to create better guidance and documentation for classification, or possibly build tools to classify/create classification reports. • We would also like to better document the different methods available for outputs, depending on the security level of the classification.
3.3.2.	1	We require all projects to classify work packages, which considers all input data and the work to be done within the project. This process does not require a detailed description of the outputs and it does not restrict what outputs may be suggested for egress. **Potential improvements** Ensure we more precisely define the expected outputs for projects before they begin.
3.3.3.	1	We rely on the project stakeholders to reach a consensus on output disclosure risks. They must classify all outputs and, depending on the classification, the outputs might be made publicly available, available to named parties or available only inside another TRE. We do not feel that existing statistical disclosure processes are sufficient for the types of data we encounter, for example, unlabelled image files. **Potential improvements** • We should improve documentation of this process
3.3.4.	2	The *project manager*, *information asset owner* representative and ref-

11.24 Information search and discovery

	Score	Response
3.4.1.	2	As each project brings its own data we do not have a catalogue of available datasets.
Capability met?	**YES**	

11.25 Security Levels and Tiering

	Scor	Response
3.5.1.	2	We categorise projects into one of five security tiers. These are clearly defined in our documentation. We are able to support four of those tiers and would reject any projects falling into the most sensitive tier.
3.5.2.	2	We support projects with differing security requirements through security controls that are predefined for each tier.
3.5.3.	2	We support a documented set of security control tiers that projects can choose from at the outset.
Capability met?	**YES**	

11.26 Research Meta-Data

	Score	Response
3.6.1.	0	We do not hold a catalogue of data in this format for this purpose. The data is provided to us by the *information asset owner* for a specific purpose. Researchers do not apply to us to access specific datasets and thus do not need to have access to a description of the data.
3.6.2.	0	This is something we would expect the *information asset owner* to do, rather than implement ourselves. For example, they could use a high-security tier 3 TRE to summarise or produce a synthetic version of a sensitive dataset for use in a lower security, tier 2 TRE.
Capability met?	**YES (no mandatory statements)**	

11.27 Meta-Data Search and Discovery Application

	Score	Response
3.7.1.	0	We do not provide such an application. We do not maintain meta-data for sets of available datasets, since we do not maintain a corpus of datasets for people to apply for access to.
Capability met?	**YES (no mandatory statements)**	

11.28 Data Archiving

	Score	Response
3.8.1.	1	We don't have a particular method for data archiving in the TRE, though administrators do have the ability to move data to a read-only location if needed.
3.8.2.	0	We don't have a particular method for archiving data in the TRE, though it is possible to keep data in the Azure Storage Accounts whilst restricting access to users. We don't handle formatting or maintaining of datasets, which is up to project teams using the TRE.
Capability met?	**YES (no mandatory statements)**	

11.29 Business continuity management

	Scor	Response
4.1.1.	1	We rely on redundancy options provided by Azure, such as load-balancing and geo-redundancy, to maximise the uptime of the TRE. If there is a catastrophic failure of Azure, access to TREs will be lost until service is resumed. We believe this is an acceptable risk that does not need further mitigation.
4.1.2.	1	No part of our business continuity plan depends on actions that we can take, so we are not able to test it.
Capability met?	**YES**	

11.30 Project and programme management

	Score	Response
4.2.1.	2	All projects have a project manager named in a Sharepoint document that is signed by all stakeholders. This individual is responsible for all project support tasks. They will liaise with the TRE operations team as necessary.
4.2.2.	2	Only a list of named *data consumers* get access to to the TRE. The project manager is not currently forbidden from being a project team member but this situation has never arisen. **Potential improvements** • We could create a policy that the project manager is not allowed to be part of the project team.
Capability met?	**YES**	

11.31 Knowledge management

	Score	Response
4.3.1.	2	Our documentation is hosted across two public web sites. One site contains documentation for the TRE implementation, including deployment and user guides. The other describes the Institutes particular processes for TRE operations. These sites are generated from GitHub repositories which can be easily updated in response to feedback as needed.
4.3.2.	2	We have documentation in place for using and managing our TRE. **Potential improvements** • We could offer a consistent training programme for all projects.
4.3.3.	1	We have identified training needs for stakeholders and made plans to address this, but we do not currently have plans for reviewing these. **Potential improvements** • We should perform a regular training needs analysis review.
Capability met?	**YES**	

11.32 Financial management

	Score	Response
4.4.1.	2	We make estimates of our infrastructure costs publicly available in advance. The project manager for each project is also able to see real-time infrastructure spending. We do not currently charge for person-time, although we plan to do so in future. Our charging structure is simple, publicly available, and discussed with the project manager before each project starts.
4.4.2.	2	We make use of Azure's spending calculators and other bespoke tools developed by the Institute to manage this. We have dedicated management professionals for charging costs back to projects.
4.4.3.	1	We recover the infrastructure costs each project. Infrastructure common to all projects is centrally funded on a year-by-year basis. We do not currently recover person-time costs. We do not have a process in place for ensuring funding in the long-term. **Potential improvements** • We could look to secure longer-term commitments for ongoing funding.
4.4.4.	2	We use the Data Safe Haven codebase which is under active development and which considers cost-effectiveness as part of its update process. We start by deploying cheaper resources and resize them to more powerful (and expensive) versions only when requested by end users, for instance, GPU-enabled machines are available only on request. We turn off infrastructure components when not in use.
Capability met?	**YES**	

11.33 Procurement

	Scor	Response
4.5.1.	2	We have systems in place for ensuring cloud credits required for TRE provision can be purchased by projects requiring a TRE. We also have systems in place for providing Chromebooks to TREs users who require access to higher-security TREs where managed devices are required.
Capability met?	**YES**	

11.34 IT Service management

	Scor	Response
4.6.1.	2	We have a dedicated service team for deploying TREs and supporting processes. This is well documented and made available to *data consumers* via the company intranet. The documents themselves are publicly available online.
Capability met?	**YES**	

11.35 Relationship management

	Scor	Response
4.7.1.	2	The code that deploys our TRE infrastructure is open-source and open to contributions from anyone. We also have a dedicated Slack channel and email address for stakeholders to engage with the project team.
Capability met?	**YES**	

11.36 Public Involvement and Engagement

	Score	Response
4.8.1.	1	All public engagement activities we have undertaken as a project have been led by public engagement professionals and have followed best practice as outlined in PEDRI guidelines. **Potential improvements** • We should develop a public engagement strategy for the Turing DSH project in collaboration with the Institute's public engagement specialists.
4.8.2.	0	We do not currently share the details of projects using our TRE. **Potential improvements** • We might consider doing this after discussion our legal team and other stakeholders.
4.8.3.	0	We do not include members of the public in our approval process. We do not think this is appropriate in the case of commercially-sensitive data and we already have an Institute-wide ethics approvals process.
4.8.4.	0	There is a clear process in place for internal incident reporting. There is no process for publicly sharing this information.
Capability met?	**YES**	

11.37 Legal services

	Scor	Response
4.9.1.	2	The Institute has a legal team who can be contacted with matters relating to the handling of sensitive data, which includes TRE projects. The TRE operators can get legal advice from this team as required.
4.9.2.	2	The Institute has a data protection team who can be contacted with matters relating to the handling of sensitive data, which includes TRE projects. The TRE operators can get legal advice from this team as required.
4.9.3.	2	The project manager has responsibility for managing contracts related to data sharing and secondment agreements. The TRE operations team together with the project manager have responsibility for ensuring that user-access terms-of-use are signed.
Capability met?	**YES**	

CHAPTER

TWELVE

HEALTH INFORMATICS CENTRE TRUSTED RESEARCH ENVIRONMENT (HIC-TRE), UNIVERSITY OF DUNDEE

Health Informatics Centre (HIC) supports high impact research through the collection and management of population based data. HIC runs a cloud based TRE based on an older fork of the open-source TREEHOOSE platform. This evaluation applies to the HIC-TRE, but it should be possible to satisfy all **Mandatory** SATRE technical requirements using TREEHOOSE.

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.1.1.	You must gather and monitor the information governance requirements needed to fulfil any legal, regulatory and ethical standards.	Requirements will come from a variety of sources including legislation, contractual obligations and ethical standards. Requirements must be monitored to ensure the TRE controls remain appropriate.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.1.2.	You must ensure controls are implemented to ensure the requirements are met.	Control implementation should be systematic and directly aligned to the internal and stakeholder requirements.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Information governance	1.1.3.	You must ensure there are adequate resources to meet information governance requirements.	Ensuring information governance controls are suitable and enforced requires an investment of funding and people appropriate to the size of the TRE.	Mandatory	1	ISO 27001, Scottish Safe Haven charter, DSPT	One dedicated person isn't sufficient, especially regarding dealing with new technologies like AI
Information governance	1.2.1.	You must ensure that changes to policies and standard operating procedures can only be made by trusted individuals.	It is important to ensure that policies and SOPs are relevant, up-to-date and carefully controlled to maintain the integrity and security of your TRE organisation.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.2.2.	You must use versioning and a codified change procedure for all policies and standard operating procedures.	This includes recording dates of changes, person responsible for carrying out changes, and summary of changes.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Information governance	1.2.3.	You should measure the performance of information governance within the TRE with regular reporting available to your TRE organisation's management team.	This may include reports and dashboards showing security incidents, quality management deviations and audit findings.	Recommended	1		More resources needed for regular assessment/reporting
Information governance	1.2.4.	You must audit your TRE organisation against relevant requirements and standards.	If you are publicly accredited against a standard, for instance ISO27001, DSPT, CE+ *etc.*, you must have processes in place to ensure you remain compliant.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.2.5.	You must report on and share outcomes of each audit of your TRE organisation with the required bodies.	This may include regulatory bodies or the organisations that manage accreditations you have.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Information governance	1.2.6.	You must ensure that suppliers, contractors and subcontractors with access to your TRE align with your security requirements.	These should be included as mandatory, non-functional requirements in during procurement and contracting. This will also include contractor staff contracts for example, legal liability and NDAs.	Mandatory	1		Should be easier to find the information

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.2.7.	You must monitor compliance of your suppliers with the terms of the contracts.	This will include monitoring changes in the services and infrastructure being delivered and quality management within the contractor's organisation. This may be done through formal audit or by monitoring change and quality documentation provided by the supplier.	Mandatory	1		Should be easier to find the information

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.2.8.	You must track and maintain any physical assets used by your TRE.	All physical assets should be maintained and covered by warranty if applicable. At the end of their lifetime, assets should be securely disposed of in such a way that data cannot be recovered from them.	Mandatory (where physical assets are in scope)	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Information governance	1.2.9.	You must log, track and resolve any issues resulting from deviations from processes, incidents and audit findings.	This process could, for example, be tracked through an electronic record and workflow system with records retained.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Information governance	1.2.10.	You must use reported issues to inform changes, such as for process improvement and risk management.	All issues should be analysed for their root cause and improvements put in place to prevent further occurrence.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.2.11.	You should collect and maintain quality management data for measuring the effectiveness of a TRE.	Large amounts of data will be produced by elements within the TRE. These data should be analysed with reports and dashboards provided to guide TRE implementer's improvements and provide reassurance to data consumers and data subjects.	Recommended	1	Regularly ask users for feedback. Monitor technical performance.	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.2.12.	You could use a QMS (Quality Management System) to standardise and automate quality management tasks and workflows, and to generate quality data and reports automatically.	A basic QMS could be a set of spreadsheets or documents held in a repository which are manually maintained. More mature applications will provide workflows and generate quality data through manual and automated actions.	Optional	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Information governance	1.3.1.	You must have a way to score risk to understand the underlying severity.	You have a risk assessment methodology for scoring risks on multiple axes such as impact and likelihood.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.3.2.	You must carry out a data processing assessment for all projects requiring a TRE.	A data processing assessment is a process designed to identify risks arising out of the processing of sensitive data and to minimise these risks as far and as early as possible. This may take the form of an existing regulatory requirements such as Data Protection Impact Assessment.	Mandatory	2	DPIA, etc	
Information governance	1.3.3.	You must have a process for designing, implementing and recording risk mitigations where indicated by a risk assessment.	Actions that are taken or not taken following a risk assessment must be recorded.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.3.4.	You must have a clear set of roles and responsibilities relating to risk including who owns risks and how they are escalated and delegated.	The highest level of risk ownership is the Top Management of the TRE organisation (see Governance Roles). In order to ensure escalations to this level are rare, suitable structures should be put in place to own, mitigate and accept risk.	Mandatory	2		
Information governance	1.3.5.	You must understand the risk appetite of your TRE organisation.	This includes understanding ownership of risk, and ability to accept risk which falls outside of the appetite should that become necessary.	Mandatory	2		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.4.1.	You must have checks in place to ensure a project has the legal, financial and ethical requirements in place for the duration of the project.	This includes checks that contracts are in place where required, adequate funding is available for the duration of the project, and responsibilities concerning data handling are understood by all parties.	Mandatory	2		
Information governance	1.4.2.	You must have checks in place to ensure that any time limited compliance requirements are maintained.	This includes ensuring contracts remain in valid and action is promptly taken should they expire. Any changes in the status of responsible persons should also be monitored, for example a data owner leaving an organisation.	Mandatory	2	Managed through JIRA assets	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.4.3.	You must have checks in place to ensure that changes in regulations are met for a project.		Mandatory	1	Yes for legal regulations	Could be extended to non-legal regulations, currently too dependent on researchers staying up to date with regulations that apply to them
Information governance	1.4.4.	You must have standard processes in place for the end of a project, that follow all legal requirements and data security best practice.	This includes the archiving of quality and log data along with the archiving or deletion of data sets.	Mandatory	1	Have processes	Not always followed exactly

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.4.5.	You could implement a portal that can provide a workflow engine and database which automates the processes within this capability.	A portal should automate as much of the processes within the capability as possible. Where processes are automated, process maturity is easier to achieve, with more consistent completion and automatic production of quality control and monitoring data.	Optional	1	Implemented ISMS that abides by the above. E.g. forms to create new project, governance, JIRA workflows, etc	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.4.6.	You must keep a complete record of all the data assets held within the system.	Details of all data assets (current and past) held by the system should be retained along with meta-data useful for ensuring compliance can be demonstrated. This would include ownership, data lifecycle, contracts, risk assessments and other quality data. This is likely to already exist within the wider organisation but may require augmenting for the TRE.	Mandatory	1	ISO 27001, Scottish Safe Haven charter, DSPT	Quality needs improving

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.4.7.	You should keep a complete record of all the research studies and projects within the TRE current and past.	The study register should contain all data related to a study including a reference to data assets, project team members, information asset owners and any compliance activities required.	Recommended	2	JIRA, sharepoint/folios	
Information governance	1.5.1.	You must have a robust method for identifying accredited members of your TRE organisation, prior to their accessing of sensitive data.	This may include ID checks or email/phone verification.	Mandatory	2	Data use declaration, confidentiality agreements, MRC training	
Information governance	1.5.2.	You must have clear onboarding processes in place for all roles within your TRE organisation.	This may include all members signing role-specific terms of use or confirming that they have completed role specific training.	Mandatory	1	Have processes	Needs improving, e.g. formally link SOPS to roles

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.5.3.	You must have a set of services to manage access to resources based on identity.	This will include a security model for role based access with technical controls to ensure the principle of least privilege is enforced.	Mandatory	2	Identity management, Active Directory, Keycloak	
Information governance	1.5.4.	You must not give anyone access to datasets without agreement from the Data Controller.	The Data Controller may choose to delegate this authority.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Information governance	1.5.5.	You must have robust and secure applications in place to authenticate users (and services) within the TRE.	The number of authentication applications should be kept to a minimum with common controls and standards applied across all such as MFA, password complexity *etc.*.	Mandatory	2	Identity management, Active Directory, Keycloak	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.5.6.	You must give each user of the TRE a unique logon with changes to any records strictly controlled.	The unique identifier and all associated records for a user should be traceable across the entire TRE. This will include training records, affiliations, contract agreements and ethics approvals where required.	Mandatory	2	Identity management, Active Directory, Keycloak	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.6.1.	You must determine what training is relevant for all roles within the TRE organisation.	This may include, for instance, cyber security training, GDPR training, and higher level training for system operators. Specialised roles are likely to need more tailored training. Identification of these specialities should be done through a systematic training needs analysis. Specific training may also be required based on the data or information asset owner such as GCP.	Mandatory	1	MRC training, in-house cyber security training	Skill-based training

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.6.2.	You must ensure that relevant training is available for all roles within the TRE organisation.	All TRE organisation members need to complete all relevant training and keep their training current. You may need to provide help or guidance to enable them to do so. Details of what training is needed will have been determined above.	Mandatory	1	MRC training, in-house cyber security training	Skill-based training

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.6.3.	You must provide repeat or updated training where necessary to account for changes in competency requirements.	Training is not a one-off event. Electronic reminders for refresher training should be considered. Ideally, training should remain relevant and so policies and processes should enable people to demonstrate competency rather than unnecessarily repeating training.	Mandatory	2	Annual	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.6.4.	You must maintain accurate training records that are directly tied to the role and access levels within the TRE.	Training records should be tied to a user record and carefully maintained. Maintaining training records enables you to ensure all people have completed the required training and that repeat training happens regularly.	Mandatory	2	JIRA Asset management	
Information governance	1.6.5.	You should accept proof of relevant training certifications from trusted third parties.	You might choose to trust certifications provided by known training providers or your institution's partner organisations.	Recommended	1	Accept some (e.g. MRC) but not ONS	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.6.6.	You could have a training platform capable of delivering online training in a variety of formats.	This could be a simple content delivery platform or a more comprehensive LMS platform. It could also include a range of multimedia delivery formats, and accessible training modules for those with access requirements.	Optional	0		Nice to have
Information governance	1.6.7.	You could implement a learning management system (LMS) to manage courses and deliver training as required.	Where possible an LMS should support a variety of course content and testing.	Optional	0		Nice to have

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.6.8.	You could ensure that any courses you use are available in standard, transferable formats.	Support for standard formats such as SCORM allows courses to be shared between providers. This could help facilitate standardisation of training provision for TRE users across organisations.	Optional	0		Nice to have

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Information governance	1.6.9.	You could keep historical copies of courses in order to demonstrate competency at a given point in time.	Information asset owners and regulators may be required to audit historical records, *e.g.* for clinical trials. It may be necessary to retain copies of superseded training along with versions of certifications within the training record.	Optional	0		Nice to have

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.1.	You must not allow users to copy data out of your TRE via the system clipboard.	A TRE user must not be able to copy sensitive data out of a workspace using the system clipboard. A TRE may allow user to paste text into a workspace. This might not be relevant to your TRE, for example if your user interface does not have a clipboard.	Mandatory	2	Blocked by TRE	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.2.	Your TRE workspace should provide an environment familiar to your users.	This may take the form of a virtual Windows or Linux desktops, non-desktop interfaces such as JupyterLab and other web applications, or a terminal. Bespoke TRE-specific software should be avoided when widely used alternatives already exist.	Recommended	2	Windows and Linux desktops, typical software or equivalent available	
Computing technology and Information Security	2.1.3.	A TRE could restrict data access from data consumers entirely and provide an interface for submitting code.	For example, you might use a system where users submit jobs that run over the data and return results without allowing direct data access.	Optional	0	Desktop TRE, we're not OpenSAFELY	Not planned

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.4.	Your TRE should be accessed via a user interface accessible using commonly available applications.	TREs which allow users to connect from their own devices should not require the installation of any bespoke TRE application on the user's device. In practice a web browser is the most common way to achieve this.	Recommended	2	Web browser	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.5.	Your TRE must provide clear guidance on how to use software tools and work with data in the TRE.	TREs that provide a virtual desktop environment for data consumers to work in should provide documentation detailing the available tools. TREs where the analysis code is developed on the access machine (as oppose to within the TRE) should provide documentation detailing the mechanism by which code is submitted to the TRE.	Mandatory	1		Improvements Needed

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.6.	Your TRE should, where possible, automatically apply security related updates for user software.	Reducing the risk of exploitable vulnerabilities in installed software will increase the security of your TRE.	Recommended	0	Currently don't do it, TRE workspaces are firewalled	
Computing technology and Information Security	2.1.7.	Your TRE could provide shared services that are accessible to users in the same project.	This may include shared file storage, databases, collaborative writing, and other web applications. This must only be shared amongst users within the same project.	Optional	1	We have some shared services e.g. MSSQL server	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.8	Your TRE must ensure that any shared services are only available to users working on the same project.	Poorly designed shared services could enable the unintended mixing of data between projects. To prevent this it is necessary that each instance is only shared between users of a single project.	Mandatory	2	User access controls on shared services	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.9.	You must mitigate and record any risks introduced by the use in your TRE of software that requires telemetry to function.	For example, some licenced commercial software must contact an external licensing server at start-up. You must be confident that only licensing information is sent to this server and that any network connections are secure.	Mandatory	1	Improvement in recording required	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.10.	Your TRE must provide software applications that are relevant to working with the data in the TRE.	The tools provided will depend on the types of data in the TRE, and the expectations of users of the TRE. For users working in a TRE via a virtual desktop, this may include programming languages such as Python and R, integrated development environments, Jupyter notebooks, office type applications such as word processors and spreadsheets, command line tools, etc. TREs with non-desktop interfaces should similarly consider carefully which applications are best suited for the data consumers	Mandatory	2	We provide requested open-source packages, and commercial applications where licensed	

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.11.	Your TRE should provide tools to encourage best-practice in reproducibly analysing data.	Reproducibil of analyses improves auditability and accountability of how data has been used, as well as being best-practice in research. This may include version control software, and tools for developing and running data analysis pipelines.	Recommended	2	R, Python, and standard libraries are available	
Computing technology and Information Security	2.1.12.	Your TRE could provide access to some public software repositories or container registries.	For example, a TRE may allow direct installation of packages from Python or R repositories, or provide an internal mirror.	Optional	1	We provide limited access to some package repositories	Should improve allow/deny-listing capabilities

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.13.	Your TRE could tightly control which packages are available.	For example, a TRE may only allow installation of a pre-defined set of approved packages. You might also choose to scan for malicious packages and/or go through an approval process before allowing code into the technical environment.	Optional	1	We limit which package repositories can be accessed	Should improve allow/deny-listing capabilities

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.14.	Your TRE must maintain segregation of users and data from different projects when using non-standard compute.	High performance or specialist compute is often shared amongst multiple users. Users and data must remain segregated at all times. For example, when using physical compute resources, all sensitive data could be securely wiped before another user is given access to that same node. In a cloud hosted TRE virtual machines could be destroyed and recreated.	Mandatory	2	Flexibility of cloud compute means non-standard compute resources aren't shared	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.15.	Your TRE should be able to provide access to high performance computing or other scalable compute resource if required by users.	If a TRE supports users conducting computationally intensive research it should provide access to dynamically scalable compute or the equivalent. For example this may be in the form of a batch scheduler on a HPC cluster, or a dynamically created compute nodes on a cloud platform.	Recommended	2	Available where required and funded	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.16.	Your TRE should be able to provide access to accelerators such as GPUs if required by users.	GPUs and other accelerators are commonly used in machine learning and other computationally intensive research. TREs should make it clear to users whether GPUs and other resources are available whilst projects are being assessed.	Recommended	2	Available where required and funded	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.17.	Your TRE could make data available to data consumers using common database systems such as PostgreSQL, MSSQL or MongoDB.	Databases must be secured and only accessible to users within the same project. If shared (multi-tenant) database servers are used, database administrators must ensure that the database server enforces segregation of users and databases belonging to different projects.	Optional	2	MSSQL is required by many users	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.1.18.	Your TRE could integrate with large-scale data analytics tools for working with large datasets.	For example, Spark and Hadoop can be used for distributed computing across a cluster. This may be an advantage where a TRE is using an amount of data that is too large for single-machine computing to be practical.	Optional	1	Offer HPC	
Computing technology and Information Security	2.2.1.	You must have a documented procedure for deploying infrastructure.	This might, for instance, be a handbook that is followed or a set of automated scripts.	Mandatory	2	GitHub workflows, ISO documentation	
Computing technology and Information Security	2.2.2.	You should, where possible, automate any repeatable aspects of your deployment.	This might involve using infrastructure as-code tools or a series of scripts.	Recommended	2	GitHub workflows	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.2.3.	You must have a documented procedure for making changes to deployed infrastructure.	This refers both to changes that might be expected in the course of normal operation and emergency changes that might be needed. Your change management process may form part of a wider accreditation such as ISO 27001.	Mandatory	2	ISO /ISMS change management	
Computing technology and Information Security	2.2.4.	You must test changes before they are used in production.	This might involve a separate development environment or another system for testing.	Mandatory	2	We have a staging TRE	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.2.5.	You should have a development environment that mirrors your production environment which you use to test infrastructure changes before committing them to production.	If possible, you should automate application of changes between development and production environments. Consider the costs and practicality of whether this will work for your situation.	Recommended	2	We have a staging TRE, and can deploy additional dev TREs when required	
Computing technology and Information Security	2.2.6.	You must have a documented procedure for removing infrastructure when it is no longer needed.	Removing unused infrastructure not only reduces costs and management burden but also reduces the attack surface of a TRE and reduces the risk of unaddressed vulnerabilities.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.2.7.	You should understand the availability and uptime guarantees of any providers that you rely on.	For remote TREs this might include your cloud provider(s) and/or data centre operators. For on-premises TREs, it might be worth using an uninterruptable power supply (UPS) and planning how you would deal with internet outages.	Recommended	1	AWS Enterprise Agreement	
Computing technology and Information Security	2.2.8.	You should develop an availability target or statement and share this with your users.	Understanding how and when the TRE might be unavailable will help your projects in planning their work.	Recommended	1		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.2.9.	Your TRE must control and manage all of its network infrastructure in order to protect information in systems and applications.	Network infrastructure must prevent unauthorised access to resources on the network. This may include firewalls, network segmentation, and restricting connections to the network.	Mandatory	1	University of Dundee has it's own AWS Organisation	
Computing technology and Information Security	2.2.10.	Your TRE must not allow connectivity between users in different projects, or with access to different datasets.	Connectivity between users in the same project may be allowed, for example to support shared network services within the project.	Mandatory	2	Enforced by TRE	
Computing technology and Information Security	2.2.11.	Your TRE must block outbound connections to the internet by default.	Limited outbound connectivity may be allowed for some services.	Mandatory	2	Enforced by TRE	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.2.12.	You should be able to monitor the network configuration of your TRE to check for misconfigurations and vulnerabilities.	This may include regular vulnerability scanning, and penetration testing.	Recommended	2		
Computing technology and Information Security	2.2.13.	You should regularly monitor the network configuration of your TRE to check for misconfigurations and vulnerabilities.	This will involve following the monitoring procedure detailed above.	Recommended	1	Regular Pen Test, need to increase more vulnerability scanning	
Computing technology and Information Security	2.2.14.	Your TRE must record usage data.	This may include the number of users, number of projects, the amount of data stored, number of datasets, the number of workspaces, etc.	Mandatory	2	Asset Management - ISO27001	
Computing technology and Information Security	2.2.15.	Your TRE should record which datasets are accessed, when and by who.	This helps maintain auditability of how sensitive data has been used.	Recommended	1		Can be improved with realtime logging

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.2.16.	Your TRE should record computational resource usage at the user or aggregate level.	This is useful for optimising allocation of resources, and managing costs.	Recommended	1		Looking to implement dashboarding to improve
Computing technology and Information Security	2.3.1.	You must ensure that all projects understand what resources are available and what the associated costs will be before the project starts.	For on-premises systems this might be related to the available hardware, for cloud-based systems there might be limits on how many instances of a particular resource (*e.g.* GPUs) can be used Projects should use this information to understand whether the available resources will be sufficient for their requirements.	Mandatory	1	We provide quotes based on requirements, but many project don't understand their requirements	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.3.2.	You should ensure that the anticipated needs of projects can be satisfied using available resources.	Note that this does not require you to accept requests for additional resources, but rather that promises made about resource availability before a project starts should be honoured wherever possible.	Recommended	2	Cloud compute means we can scale as much as we want, but in practice this is limited by researcher's funding	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.3.3.	You must have a procedure for allocating available resources among projects.	For cloud-based TREs this may involve scaling resources, such as virtual machines or databases, or deploying additional resources. For on-premises TREs this may involve a procurement process to ensure that necessary resources are available. Not all requests for capacity increase must necessarily be granted, but having a clear process will help projects understand when/why/hc they can make use of additional capacity.	Mandatory	2	Part of the requirements gathering and quoting for a project	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.3.4.	You must ensure that the anticipated resource requirements will not result in overspending by the TRE.	For cloud-based TREs this may involve budgeting and/or restricting resource consumption on a project-by-project basis. For on-premises TREs this may involve managing expectations to match the available resource.	Mandatory	1	We don't anticipate overspend, however we don't have a technical solution and it's managed manually	
Computing technology and Information Security	2.4.1.	You must have a documented procedure for configuring infrastructure.	This might, for instance, be a handbook that is followed or a set of automated scripts.	Mandatory	2	ISO /ISMS change management	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.4.2.	You should use configuration management tools to automate application of your configuration wherever possible.	This might involve configuration as-code tools such as Ansible, Chef, Puppet or Windows Desired State Configuration or simply automated scripts.	Recommended	2	GitHub workflows	
Computing technology and Information Security	2.4.3.	You should be able to verify whether the configuration is valid.	This might, for instance, involve running your configuration management tool in 'check' mode.	Recommended	1	GitHub workflows, manual testing	
Computing technology and Information Security	2.4.4.	You should regularly verify your TRE configuration.	This will limit the amount of time the TRE can spend in a non-compliant state.	Recommended	1	GitHub workflows	
Computing technology and Information Security	2.4.5.	You must be able to replace a non-compliant TRE with a compliant system.	This might involve reconfiguring a running system or by replacing it with a compliant one.	Mandatory	2	Redeploy TRE	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.1.	You should keep backups of data and research environments, provided that this is permitted by law.	Keeping backups could help reduce the impact of events like accidental deletion and data corruption on work in a TRE. TRE developers may want to consider how different elements such as sensitive input data or users' workspaces may be backed up, and whether they should be.	Recommended	1	Research data is backed up. Workspaces are currently treated as ephemeral for backup purposes	Cost/benefit tradeoff when backing up ephemeral VMs

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.2.	You should build redundancy into infrastructure and storage.	Infrastructure should be as resilient as necessary to interruption. This could include redundant infrastructure in different physical locations, load balancing and replication of data between multiple storage locations.	Recommended	2	Using cloud native storage and execution where possible	
Computing technology and Information Security	2.5.3.	You should keep backups of infrastructure, applications and configurations.	This may include virtualised infrastructure snapshots which can restored as needed to recover from failure.	Recommended	2	Infrastructure as code, stored in GitHub	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.4.	You must have procedures in place for rapid incident response.	There may be legal requirements to disclose details of any incidents, such as data breaches for organisations subject to GDPR. Having robust processes in place will ensure a swift and effective response when an incident occurs.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	
Computing technology and Information Security	2.5.5.	You should test your incident response through simulation.	During simulated incidents the TRE organisation can measure their effectiveness. This may involve people across the broader enterprise and/or external suppliers.	Recommended	0		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.6.	You should have an application in place to scan for vulnerabilities across infrastructure.	Software used to identify vulnerabilities should also report and alert. Such an alert should be triaged, risk assessed and treated accordingly.	Recommended	1		Not 100% reliable
Computing technology and Information Security	2.5.7.	You must have a process in place for applying security updates to all software that forms part of the TRE infrastructure.	This includes any software used for remote desktop portals, databases, webapps, creating and destroying compute infrastructure, configuration management, or software used for monitoring the TRE.	Mandatory	0.5		Needs to be automated

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.8.	Infrastructure should be automatically patched for vulnerabilities.	Planning will be required across infrastructure and software systems to ensure security patches remain available from suppliers. Many systems may be isolated from the internet making TRE infrastructure more difficult to automatically patch.	Recommended	0.5		Needs to be automated

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.9.	You should carry out penetration tests on your TRE.	By intentionally attempting to breach their TRE, organisations can proactively discover unnoticed vulnerabilities before they are exploited maliciously. Tests can evaluate the effectiveness of security controls in preventing data breaches, unauthorised access, or other security incidents.	Recommended	2	Annual tests conducted by an external company	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.10.	You should update the security controls of your TRE based on the results of security tests.	Security testing can reveal bugs and discrepancies in the TRE architecture which should be addressed in advance of sensitive data being uploaded, or with urgency in the case of an operational TRE. Regular testing will allow organisations to refine their TRE security controls and incident response capabilities. It enables them to adapt to any new security concerns that may arise as a result of changes in the underlying software.	Recommended	2	We review security test results	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.11.	You should publish details of your security testing strategy and, where possible, the results of each test.	Knowledge that regular security testing occurs will help to ensure stakeholders, including data consumers and information asset owners, can trust that the data they work with or are responsible for is secure within a TRE. If security flaws are identified in a test, it may not be sensible to publicise these until a fix is in place.	Recommended	1	Not published publically but available on request	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.12.	Your TRE must encrypt project and user data at rest.	This prevents unauthorised access to the data even if the storage media is compromised. This may involve encrypted filesystems or tools to encrypt and decrypt data on demand. The encryption keys may be managed by the TRE operator or by a trusted external actor, for example a cloud provider.	Mandatory	2	Built in to TRE	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.13.	Your TRE must encrypt data when in transit between the TRE and external networks or computers.	Data encryption must be used to safeguard against interception or tampering during transmission. This includes both data ingress and egress and users accessing the TRE, for example over a remote desktop or shell session.	Mandatory	2	Built-in to file transfer protocols and documented processes	
Computing technology and Information Security	2.5.14.	Your TRE should encrypt data when in transit inside the TRE.	If possible, data transfers between different components of a TRE should also be encrypted.	Recommended	2	Default for AWS infrastructure	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.15.	You should use encryption algorithms and software that are widely accepted as secure.	Encryption algorithms widely accepted as secure today may become insecure in the future, for instance due to newly-identified flaws, or advances in compute capabilities. The latest security patches and updates should be applied to any encryption software being used by the TRE. This helps address any known vulnerabilities or weaknesses in the encryption implementation.	Recommended	2	We use standard encryption algorithms	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.16.	Your TRE should use secure key management.	TREs should employ secure key management practices, including storing encryption keys separately from the encrypted data and implementing strong access controls (*e.g.* Single Sign On) for key management systems.	Recommended	2	AWS KMS	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.17.	Your TRE could offer physical protection measures against data leakage or theft via physical means.	Restricting access to research facilities containing computers logged into TREs can help prevent malicious actors from viewing or stealing sensitive data, for example by photographing a computer screen. Physical controls on access to a TRE could include surveillance systems, restricting physical access to authorised personnel only, visitor management systems and employee training.	Optional	2	https://aws.amazon.com/compliance/data-center/controls/	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Computing technology and Information Security	2.5.18.	Your TRE may need to comply with specific regulatory requirements due to the types of data it is hosting.	Regulatory frameworks often emphasise the need for security controls to protect sensitive data. Compliance with these regulations could require organisations to implement specific security measures to safeguard their TRE from unauthorised access.	Mandatory	2	We comply with necessary requirements, e.g. for NHS data	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.1.	You must have processes in place to assess the legal and regulatory implications of handling the data through its full lifecycle.	This involves considering your obligations to data controllers and subjects, and whether any security controls may be legally or contractually required. An assessment of the risks involved will also be needed. It may involve classifying the project into a predefined sensitivity category or defining bespoke controls.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.2.	You should keep records of data handling decisions.	Decisions that are made as part of the process discussed above should be recorded and made available for inspection by all stakeholders.	Recommended	1	Everything is in project management system	Could make it easier to search old decisions

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.3	Information asset owners must classify data sets according to a common process and data classification methodology.	To classify the data, information asset owners must have a good understanding of the data sets and the process of classification. Once classified, data can be stored in a TRE with an appropriate security controls (see later section on security levels and tiering), which can factor in the requirements for confidentiality, integrity and availability of the data.	Mandatory	1		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.4.	You must have a data ingress process which enforces information governance rules/process	The data ingress process needs to ensure that information governance is correctly followed. In particular, it should require that an ingress request has been approved by all required parties.	Mandatory	1	Data ingress process	Manual, could be improved
Data management	3.1.5.	You must have a data egress process which enforces information governance rules/process	The data egress process needs to ensure that information governance requirements are adhered to. In particular, it should require that an egress request has been approved by all required parties.	Mandatory	2	Data egress process, managed with data egress application	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.6	Egress must be limited to the information asset owners or their delegates.	Egress of data from a TRE must be a specific permission associated with individual users This permission must be given by information asset owners. Egress may still require further approval (see 3.1.5).	Mandatory	2		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.7.	Your data egress process could sometimes require project-independent approval.	There may be cases where there are multiple stakeholders for a piece of analysis including information asset owners, data analysts, data subjects, the TRE operator. A data egress process may then require approval from people not on the project team, for example an external referee or TRE operator representative	Optional	2	Egress for some projects is managed by an external body	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.8.	You must keep a record of what data your TRE holds.	Good records are important for ensuring compliance with legislation, understanding risk and aiding good data hygiene. The record should include a description of the data, its source, contact details for the data owner, which projects use the data, the date it was received, when it is expected to no longer be needed.	Mandatory	2	Asset Management - ISO27001	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.9.	You must have a policy on data deletion.	There should be a clear, published policy on when data will be retained or deleted. This may allow time for data owners to consider outputs they may want to extract from the TRE. Any sensitive data, including all backups, should be deleted when they are no longer needed. Having clear policies will help to avoid problems with data being kept longer than necessary or accidental deletion of outputs.	Mandatory	2	ISO 27001	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.10.	You should have a method of providing proof of deletion/removal of files.	information asset owners may require certification of the deletion of files. You should have a method of providing proof of deletion if challenged.	Recommended	1		
Data management	3.1.11.	You should log how input data is modified.	If the input data is mutable a TRE should keep records of its modification. For example, when the data was modified and by who.	Recommended	0		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.12.	You must, to a reasonable extent, prevent unauthorised data ingress or egress.	Movement of data which has not been subject to information governance processes risks breaking rules and is more likely to result in a data breach. However, it is difficult to control for every possibility. For example, a user may take pictures of their computer screen to remove data, or use a device presenting as a USB HID keyboard to input large amounts of text. An example of a reasonable measure would be for a remote desktop based TRE to prevent data being copied from a local machine's	Mandatory	2	Data ingress/egres process	

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.1.13.	Data held within the TRE should be the minimum required for analysis or research.	Data stored and processed within the TRE should be limited to the amount required for that purpose. This increases the level of protection for data subjects, makes it easier to comply with data protection legislation and could reduce the overhead of storage and processing.	Recommended	2	Only a subset of raw data is made available to researchers	
Data management	3.2.1.	You must not create user accounts for use by more than one person.	It is important that each user account should be used by one, and only one, person in order to facilitate the assignment of roles or permissions and to log the actions of individuals.	Mandatory	2	Each user has their own account. User agreement requires users to not share access.	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.2.2.	You must be reasonably convinced of the identity of each person being granted an account.	It is important to ensure an account has been given to the correct person. For example, multiple credentials may be used before account creation to verify identity or, when appropriate, photo ID checks may be required.	Mandatory	2	ISO 27001, Scottish Safe Haven charter, DSPT	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.2.3.	You must restrict a user's access to only data required in their work.	There is no need to grant an individual access to data they do not require. Access may be assigned in a manner appropriate to a TREs design, for example through roles granted to user accounts or through isolated project workspaces.	Mandatory	2	Enforced by TRE	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.2.4.	You must ensure that multi-factor authentication is enabled for all users.	Multi-factor authentication ensures that to successfully connect a user must have more than one piece of evidence in different categories. Categories include something the user knows (*e.g.* a password), something the user possesses (*e.g.* a TOTP key) or something the user is (*e.g.* biometric data). A TRE does not need to implement multi-factor authentication checks itself if it is provided by a third-party identity provider.	Mandatory	2	Enforced by SSO provider	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.2.5.	You could use federated authentication or single sign-on (SSO) for user login.	Institutions that use a SSO for other applications may wish to extend this login capability to a TRE. This will simplify the login process for data consumers using a TRE and prevent them having to remember or store multiple login credentials.	Optional	2	SSO used	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.2.6.	You could restrict access to particular networks or physical locations.	Restricting access to a set of known, static, personal or institutional IP addresses can help avoid speculative attacks. When appropriate, access could also be restricted to physical locations with security controls and access requirements.	Optional	2	https://aws.amazon.com/compliance/data-center/controls/	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.3.1.	You should have a system to help classify outputs.	Removing data from a TRE can be a difficult process as there is potential for sensitive data to be revealed. Having guidance, processes and methods will help ensure that outputs are correctly classified and, furthermore, that outputs due to be openly published are identified. Encouraging openly published outputs will enhance a TRE's impact and transparency.	Recommended	0		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.3.2.	You should establish the intended outputs of each project from the outset.	Identifying the purpose of a piece of work is important for compliance with data protection legislation. Results will be produced which address the project's purpose, some of which may be outputs that are removed from the TRE. Understandin what these outputs are likely to be and their sensitivity as early as possible will help prepare for their processing and publication.	Recommended	1	AI/ML model process	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.3.3.	You must have a documented process for disclosure control of outputs from the TRE.	This process should define expected risks and how to mitigate them. All TRE outputs must be subject to this process. You might choose to follow existing guidelines, for example around statistical disclosure.	Mandatory	2	Data egress process	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.3.4.	You must have a process for assigning responsibility for output checking.	Output checkers should be given responsibility for checking outputs. They must follow your disclosure control process and will be responsible for any automated parts of this process. Output checking can help mitigate against unintentional data disclosure or leaks.	Mandatory	2	Named people are responsible	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.3.5.	You must have a documented policy for handling disclosure risks associated with any outputs that cannot be manually checked.	Some categories of output, for instance binary files or very large numeric files, can be difficult to manually check. If egress of such files is permitted then the risks of inadvertent disclosure must be mitigated and documented. Refusing to allow egress of such files is also a valid policy decision.	Mandatory	2	Default is to reject. If necessary further due diligence, risk assessment, and looking at mitigations, will be done	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.3.6	You should have a statistical basis to guide the decisions of an output checker on the safety of outputs.	There should be a solid basis to allow decisions to be made about data based on risk factors such as re-identification of an individual or risk to commercial operations posed by outputs from the TRE.	Recommended	1		Could be more thorough

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.3.7	You could create a semi-automated system for checks on common research outputs.	Automation helps make decisions on outputs more consistent and reduces the overhead for output checkers. It's unlikely however that a fully automated output checking system (without humans in the loop) would be appropriate, given the risks associated with accidental data disclosure.	Optional	0		
Data management	3.3.8.	TRE outputs should be limited to the minimum required for sharing results of any analyses.	This decreases the risk of inadvertent disclosure, and makes it easier to comply with data protection legislation (e.g. GDPR).	Recommended	2		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.4.1.	You should provide a metadata catalogue of available datasets for users.	This is particularly relevant for TREs with population-level data collection of general interest. This may not be appropriate for TREs where each project has its own data sharing agreement with one or more data provider or very sensitive datasets.	Recommended	2	Data Catalogue	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.5.1.	You must be able to specify what categories of data your TRE is able to support.	Your TRE must provide an explanation of the kinds of data it has been designed to hold, with reference to its security capabilities, that can be understood by all stakeholders. Relevant stakeholders may include information asset owners and project teams and they may have different levels of technical expertise.	Mandatory	2		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.5.2.	Your TRE could support projects with differing security requirements through configurable security controls.	This allows projects with different security requirements to each be met with a suitable level of controls. It helps ensure that users can work effectively, with minimal barriers.	Optional	1	We have the ability to spin up a non-default TRE configuration for a project if funding is available	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.5.3.	Your TRE could offer a predefined set of security control tiers.	Security control tiers can be designed to cover the types of project or data you expect to handle. Projects may be placed into the most suitable tier rather than having a bespoke design. This reduces the number of unique configurations that need to be supported.	Optional	0		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.6.1.	You should have a consistent and easily accessible meta-data data model or similar to describe what a data asset contains.	Where possible, existing data models should be employed (and extended if necessary). More detailed information on the data schema for data assets should also be provided to assist researchers in understanding what data may be available without the need to see the underlying data.	Recommended	1		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.6.2.	You could provide summary, abstracted or synthetic data to researchers without exposing the underlying data set.	To reduce the need for access to row level data researchers could be provided with non-sensitive versions of the data either as summary data or using synthetic versions of the data for activities such as code development and cohort planning.	Optional	1	We only provide the necessary data	
Data management	3.7.1.	You could provide an interface application for data consumers and data subjects to query elements of the data.	In order to make data findable, an application which queries the meta-data or elements of the research data could be made more easily accessible than the data itself.	Optional	0	Coming soon via HDR UK Cohort Discovery Tool	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.8.1.	Archived data within the TRE should be read only.	Archived data by its very nature should not change and therefore be maintained as a read only store. If an update is required, it may be pulled from archive into a separate operational store.	Recommended	2		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Data management	3.8.2.	Long-term archives must be held in simple, standard formats to ensure accessibility.	Some data archives may be required by policy or legislation to be kept for very long periods within the scope of the TRE. Such data should be held in the simplest possible file format, conforming to international standards if available, to ensure they are platform and application agnostic.	Recommended	1	Data is stored in the original format	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.1.1.	You should have a business continuity plan that includes consideration of loss of service for deployed TREs.	This may be due to downtime from service providers, a breach, or loss of power. Your plan should detail your process for managing loss of service for deployed TREs, and evaluation of impact of such loss.	Recommended	2	ISO 27001	
Supporting Capabilities	4.1.2.	You should regularly test the aspects of your business continuity plan concerning TREs, and have a process in place to iterate the plan if required.		Recommended	2	Internal and external audit	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.2.1.	You should ensure that all projects using your TRE have a named project manager.	The project manager has responsibility to ensure the smooth running of the project. Their responsibilities may include budget management, tracking TRE status, managing communications with the TRE operations team, and other project support tasks.	Recommended	2	Named lead on every project	
Supporting Capabilities	4.2.2.	You should not give project managers direct access to the TRE.	Doing so ensures a separation between those able to access sensitive data, and those overseeing access to sensitive data.	Recommended	1	Small projects just have a lead, usually the PI. Depends on requirements. Role based access process.	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.3.1.	You must document all features of your TRE implementation.	This includes ensuring all documentation is discoverable, clear, and able to be easily updated based on stakeholder feedback	Mandatory	1		Improve document management
Supporting Capabilities	4.3.2.	You should have an education programme in place to upskill stakeholders in the use and management of your TRE.	This may include learning modules, workshops and other resources on how to effectively access and use a TRE, FAQ pages, and accessible pathways for additional support	Recommended	1	Education is adhoc currently, resource required to build an educational programme	
Supporting Capabilities	4.3.3.	You should periodically carry out a training needs analysis (TNA) for all stakeholders included within your TRE provision.	At least once every 12 months you should assess the training needs of your stakeholders, and ensure they have easy access to all required training materials	Recommended	0	Resource required to build an educational programme	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.4.1.	You must ensure that all projects using your TRE are aware of any associated costs and are able and willing to pay them.	Costs may include provision of the underlying TRE infrastructure, additional resources required in a specific TRE (for instance memory or additional compute), hardware including managed devices, and staff support costs	Mandatory	2	Part of the requirements gathering and quoting for a project	
Supporting Capabilities	4.4.2.	You should be able to track the costs associated with each TRE project.	This includes knowing which costs are associated with which project, and having an appropriate charging mechanism in place in line with your organisational policy.	Recommended	1	Area for improvement	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.4.3.	You should have a process in place to ensure your TRE provision remains financially sustainable.	This could include having a cost recovery process in place, or setting up a long-term funding mechanism to support projects with TREs. At any given time, you should have funds free to cover all potential foreseen TRE provision for at least 12 months.	Recommended	2	Area for improvement	
Supporting Capabilities	4.4.4.	You should minimise the cost of your TRE infrastructure wherever possible	You should have regular reviews of your TRE provision and actively work to bring down costs, streamline provision, and optimise support.	Recommended	1	Area for improvement	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.5.1.	You must identify any goods or services that will be needed to operate the TRE and ensure that a plan is in place to purchase them as needed.	These may include computing hardware, cloud credits or devices through which users access the TRE.	Mandatory	2	AWS resources can be used on-demand. Requests for e.g. licensed software goes through procurement	
Supporting Capabilities	4.6.1.	Your TRE must have a team of Operators in place to support projects working with TREs.	This may be part of your organisation's IT support team, or separate. Responsibilit should be clear and stakeholders should easily be able to access support appropriate to their needs.	Mandatory	1	We have a support process	Could always do with more people!

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.7.1.	You should have a clear process in place for stakeholders to feedback on your TRE infrastructure.	This may include a GitHub repository where people can open issues and discussions, communication streams like Slack or email, or forms stakeholders can fill in.	Recommended	1	Annual User Feedback Questionnaire - User Community	Users need to engage
Supporting Capabilities	4.8.1.	You should ensure that all public engagement activities are representative and inclusive.	Any public engagement activity carried out by TREs should make sure they are involving a representative sample where possible and that activities are accessible and open. This could include following guidelines such as PEDRI.	Recommended	1	Project based PPI - Improvement required to do this at HIC level	Resource required

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.8.2.	You could publicly share the details of any projects which use the TRE.	This may be via the TRE website or annual reports.	Optional	0		
Supporting Capabilities	4.8.3.	You could include members of the public in your approvals process.	This may be carried out via a separate public panel or by including members of the public on an approvals panel.	Optional	2	Data Custodians provide approvals which have public involved.	
Supporting Capabilities	4.8.4.	You should publicly share details of incidents, near misses, and mitigations in a timely fashion, in line with good practices for responsible disclosure.	This may be via the TRE website or annual reports. Sharing this information is particularly important when a TRE holds public sector data.	Recommended	0		

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.9.1.	You should have identify areas where legal advice may be required and ensure that you have ready access to it.	It is likely that legal advice will be necessary for several issues around the handling of sensitive data, and managing project contracts. TRE operators should have ready access to legal advice, including a way to solicit advice and carry out associated actions.	Recommended	2	Legal Department	
Supporting Capabilities	4.9.2.	You should identify areas where advice on data protection issues may be required and ensure that you have ready access to it.	It is likely that data protection advice will be necessary for several issues around the handling of sensitive data.	Recommended	2	IG Manager and IG Department, DPO	

continues on next page

Table 1 – continued from previous page

Section	Item	Statement	Guidance	Importance	Score	Response	Improvements
Supporting Capabilities	4.9.3.	You should identify who will be responsible for managing contracts related to the TRE.	These contracts may include data sharing agreements, secondments of personnel or limitations on how results obtained with the data can be distributed.	Recommended	2	Operations Director and University Legal team	
Version	branch-latest-cdcb7c5						

satre-uod-evaluation-20231011.csv

CHAPTER

THIRTEEN

COMMONLY USED TERMS

Actor
: A person, organization, or system that has one or more *roles* that initiates or interacts with activities. Example: *The SATRE architecture needs actors such as* data analysts *and internal auditors.*

Application component
: An encapsulation of application functionality which is modular and replaceable. Example: *To perform work within a TRE a* data analyst *might need access to a Desktop or command line interface application component.*

Architectural principle
: Fundamental guidelines that inform the design, decision making and implementation of a TRE. These principles provide a framework to ensure that the design of the underlying components of a TRE are aligned to consistent goals, values and best practices.

Business process
: A set of actions which produce a specific desired outcome. Example: *to access the TRE a* data consumer *needs to complete an onboarding business process.*

Capability
: An ability that a system possesses. Capabilities are typically expressed in general and high-level terms. Achieving a capability typically requires a combination of organisation, people, processes, and technology.

Capability decomposition
: A set of *components* that realise a capability. These components will vary depending on the nature of the capability. Business-focused capabilities will be realised by *business processes*, *roles* and services. Technology-focused capabilities will be realised by *applications*, services and interfaces. In addition to the components realising the capability, a catalogue of standards, frameworks and controls linked to the capabilities will provide guidance on how to implement the capabilities safely.

Component
: The statements concerning processes, controls, practices and applications that make up a *capability*, together with an importance label.

Data Object
: A store of data or information. For example: to know what data is stored within the TRE a study database data object is needed. This contains information on the data assets within the TRE, who owns them and other compliance information.

Role
: A role is a set of connected behaviors, rights, obligations and norms within a TRE system. *Roles* are occupied by individuals, who are called *actors*.

Specification pillar
: A specification pillar is a group of related capabilities. SATRE has four specification pillars: Information governance, Computing technology, Data management and Supporting Capabilities.

Trusted Research Environment (TRE)
A Trusted Research Environment. See our *FAQs page*.

TRE organisation
A TRE organisation is the set of people, processes and technology that operate and use a particular TRE.

CHAPTER

FOURTEEN

CONTRIBUTING TO THE SATRE SPECIFICATION

We're excited that you want to contribute

Some ways to immediately get involved are:

- **Join a Collaboration Cafe**: These are online events where we discuss the specification and other TRE topics. They are a great way to meet other members of the community, find out more about the project, and are open to everyone.
- **Sign-up for email updates** from the SATRE project
- **Read the current SATRE specification**
- **Provide feedback and suggestions** on the specification:
  - If you are a GitHub user please open or comment on an issue
  - Alternatively, you can fill in this form (no login needed)

We want to ensure that every user and contributor feels welcome, included and supported to participate in the SATRE project and community. We hope that the information provided in this document will make it as easy as possible for you to get involved.

We welcome contributions to this project via GitHub issues and pull requests. Please follow these guidelines to make sure your contributions can be easily integrated into the project. As you start contributing, don't forget that your ideas are more important than perfectly formatted contributions :heart:.

If you have any questions that aren't discussed below, please let us know through one of the many ways to *get in touch*.

Jump straight to our contribution walkthrough

14.1 Code of Conduct

SATRE is a community-led and collaboratively-developed project. Therefore, we require that all our contributors and their contributions **adhere to our Code of Conduct (CoC)**. Please familiarise yourself with our CoC and ensure your contributions and engagement with this project follow it!

14.2 Contributing through GitHub

Git is a really useful tool for version control. GitHub sits on top of Git and supports collaborative and distributed working.

We know that it can be daunting to start using Git and GitHub if you haven't worked with them in the past. We are here to help you figure out any of the jargon or confusing instructions you encounter !

In order to contribute via GitHub, you'll need to set up a free account and sign in. Here are some instructions to help you get going. Remember that you can ask us any questions you need to along the way.

14.3 Alternative ways to comment

In line with our open principles we recommend commenting and *contributing to SATRE using GitHub*. If this is not possible you can also leave a comment via this form. Please read the form carefully. We would like to acknowledge your contributions in public, but if you prefer you can remain anonymous.

14.4 Contribution Model

We have designed our contribution model to be as accessible as possible, while utilising the full power of GitHub's collaboration and version-control tools.

This specification, including its governance procedures and contribution models, are open for the community to evaluate, challenge, amend and discuss. If you have improvements we can make to anything we are doing, please suggest them!

The community can suggest any amendments to the specification at any point. If you see a part of the specification you don't like, open an issue about it and start a conversation with the community .

Important: We are in a bootstrapping phase to get an initial specification written. As part of this initial work, we will propose a more formal governance model for the specification going forwards.

We have chosen to keep all discussion to issues for now, so contributors have a single place to engage in conversation. Pull requests are used when a specific change is ready to be proposed. This can be without discussion, however, it is best for substantial or significant changes to be discussed first in an issue. We have opted to not use GitHub Discussions at this point.

The community can suggest governance changes at any point. This includes the SATRE team, and any decision must be openly documented in the repository.

14.4.1 Specification Format

The specification source is kept in the specification repository. It is written as a Sphinx document in Markdown format.

The most up-to-date 'source of truth' will be the specification on the `main` branch of the specification repository. The community can decide when to 'tag' a new version of the specification. They may also decide to where to publish the specification.

The specification repository is self-contained and relates only to the specification specifically, or its governance. Any contributions to the wider SATRE project should be made through a different medium via the SATRE GitHub organisation, or by contacting the SATRE team at satre-contact@dundee.ac.uk.

14.4.2 Contribution Process

Issues should be used to discuss ideas, potential changes and to ask questions. Issue templates have been designed for common issue types to help collect the most important information and present it in a clear, consistent way. It is possible, however, to open a blank issue if none of the templates are suitable.

While we encourage opening issues, we understand that some may be more comfortable contributing ideas in other ways. We support other methods of contribution such as discussions and notes taken at SATRE Collaboration Cafés. The SATRE Team will aim to collate ideas and draft issues that welcome further discussion and attribute those involved in initial discussions. The SATRE Team will try to capture the ideas as accurately as possible, in good faith, and be guided by the SATRE Community to correct any misconceptions.

When ready, changes will be proposed in pull requests. Similarly to issues, there is a pull request template. This template prompts contributors to include important details in order to explain the contribution and make triage and review easier.

Pull requests will be used to review changes. During the review process, the pull request will be used for discussion, to suggest amendments and ultimately accept or reject the change.

We use this process to ensure that as much as possible of the discussion and decision-making process can be public. This is to provide as open and accessible as possible an environment for all contributors to engage in the conversation.

14.4.3 Consensus Mechanism

Approval from the SATRE team is based on lazy consensus. After a pull request has been open for at least 7 days, unless there are outstanding objections, the change will be presumed as accepted. SATRE team members are then free to merge the pull request at any point.

If any of the SATRE team objects to a particular issue, this will be raised by individuals using pull request reviews. The SATRE team will only merge pull requests that have no outstanding objections.

14.5 Writing in Markdown

The Myst Parser documentation has a guide on the Markdown format used in the specification source files. GitHub also has a helpful page on getting started with writing and formatting on GitHub, which will be useful when writing Markdown for GitHub (for example in issue or pull request comments).

You can think of Markdown as a few little symbols around your text that instruct how to render the text. For example, you could write words in **bold** (`**bold**`), in *italics* (`_italics_`), or as a link (`[link](https://medium.com/satre)`) to another web page.

Also, when writing in Markdown, please start each new sentence on a new line. Having each sentence on a new line will make no difference to how the text is displayed. A blank line is needed to start a new paragraph. However, it makes the source and diffs produced during the pull request review easier to read !

14.5.1 Linting and auto-formatting

We take advantage of `pre-commit` and related tools to help maintain consistent formatting within a repository, which improves review efficiency, and readability. `pre-commit` can be installed using pip:

```
pip install pre-commit
```

When you make some changes, you should run

```
pre-commit run -a
```

before committing any changes. See the pre-commit documentation for more advanced usage, including automatically running it as part of a commit.

14.6 SATRE Team Contributions

SATRE team members are free to contribute to the repository in the same way as any contributor, following the process above. The SATRE team is also doing ongoing work to identify the key features of this specification. Some contributions by SATRE team members may represent the output of this work. Any contribution that represents this work will be explicitly mentioned in the contribution.

This work is taking on two main forms:

1. Identifying what features the community feels are important for a TRE via the features survey. We will synthesise responses from this survey to suggest features here.
2. Evaluating the TREs used in production as part of the Alan Turing Institute Data Safe Haven, Microsoft's Azure TRE, and the TREEHOOSE TRE. The SATRE team will make recommendations for features of the specification based on similarities/differences across these three TRE provisions.

14.7 Get in touch

To get in touch with the SATRE team, please email satre-contact@dundee.ac.uk.

To report Code-of-Conduct violations, please use the contact specified in the Code of Conduct.

14.8 Recognising Contributions

We welcome and recognise all kinds of contributions, from discussing ideas, suggesting features, improving governance, maintaining the project, and more.

This project follows the All Contributors specification. The All Contributors bot usage is described here.

To add yourself or someone else as a contributor, comment on the relevant Issue or Pull Request with the following:

```
@all-contributors please add <username> for <contributions>
```

You can see the Emoji Key (Contribution Types Reference) for a list of valid `<contribution>` types. The bot will then create a Pull Request to add the contributor and reply with the pull request details.

Hint: Please only add one contributor with the bot at a time!

It is best to add each contributor in turn and merge the pull request before adding another one. Otherwise, you can end up with merge conflicts. Please check the open pull requests first to make sure there aren't any open requests from the bot before adding another.

What happens if you accidentally run the bot before the previous run was merged and you get those pesky merge conflicts? (Don't feel bad, we have all done it!) Simply close the pull request and delete the branch (`all-contributors/add-<username>`). If you are unable to do this for any reason, please let us know by opening an issue, and SATRE team members will be very happy to help!

CHAPTER

FIFTEEN

NEW CONTRIBUTORS

15.1 Overview

This guide aims to help new contributors get involved with the SATRE (Standard Architecture for Trusted Research Environments) project. SATRE is a collaborative effort between various universities and research institutions to develop a reference architecture for Trusted Research Environments in the UK. Your contributions, regardless of your experience level, are highly welcomed and appreciated.

15.2 Understanding the SATRE Project

Before getting involved, please read SATRE's Kick Off Blog Post (a 6-minute read) to understand the motivations of the project.

15.3 Join the Community

15.3.1 Join the Mailing List

To get the latest SATRE Newsletter and communications, please sign up to our Mailing List

15.3.2 Sign up to our Collaboration Cafés

We hold one hour online Zoom Collaboration Cafés to facilitate sharing ideas for what should be in the SATRE Specification. Collaboration Cafés happen at 3pm on the 1st Tuesday and 3rd Thursday of the month. Please fill in this form to register your attendance in advance.

Collaboration Cafés are run using HackMD documents that will contain the meeting agenda and a place to share notes. If you are unfamiliar with HackMD, please see this guide on how to use it. Please also see a sample HackMD from our Collaboration Café on the 18th May 2023. Each Collaboration Café uses Breakout Rooms where participants will work on themes, e.g. Risk Management or Training Requirements. Typically, each Breakout Room is linked to an Issue on the SATRE GitHub Issues Board.

Important: The HackMD for each Collaboration Café is made available in your calendar invites. We encourage you to propose a Breakout Room in advanced of the Collaboration Café, linking to a SATRE GitHub Issue if possible. This will help participants come prepared to talk about certain topics. We also welcome Breakout Room proposals on the day!

15.4 Review the SATRE Specification Document

The latest version of the SATRE Specification Document can be found *here*. It is a living document, and we intend to have a complete draft by October 2023.

15.5 Contribute to the SATRE Specification Document

As a collaborative project driven by community needs, we'd love for you to contribute directly to the specification.

There are two ways to do this:

1. Directly to the specification on GitHub (recommended)
2. *Submitting comments via a form*

Why the difference? GitHub is the live version of the specification and is great for things like community discussions, version control, editing changes and more. This is why we recommend contributing via GitHub — it is the most direct way to work with the SATRE team on the specification.

However, we realise this may be difficult if you aren't already familiar with GitHub's ways of working. Therefore, we also have a way to *make comments without a GitHub account*.

15.5.1 Contribute directly through GitHub

Below is a walkthrough of all the steps required to contribute via GitHub, from the very beginning.

This walkthrough is designed to give you the critical path steps to contributing directly to the Specification repository. For a broader, deeper introduction to GitHub, check out the Turing Way's Introduction to GitHub Workshop, run at CarpentryCon2022.

1. Go to the GitHub homepage
2. In the top right corner, click either 'Sign In' (if you have an account already) or 'Sign Up'. You will need an email, password, username and one or two other things.

1. Once you are logged in, navigate to the SATRE Specification page
2. This is the SATRE specification **repository**. You can think of this like a directory containing all files to do with the SATRE specification. The main things to know about are the link in the `About` Section, and the `Code`, `Issues` and `Pull requests` tabs. Let's look at these in turn.

About link

This link will take you to a readable version of the specification. You can access the current live version of the specification by clicking on `A Standard Architecture for TREs` from the left-hand navigation bar. You can navigate directly to it *here*.

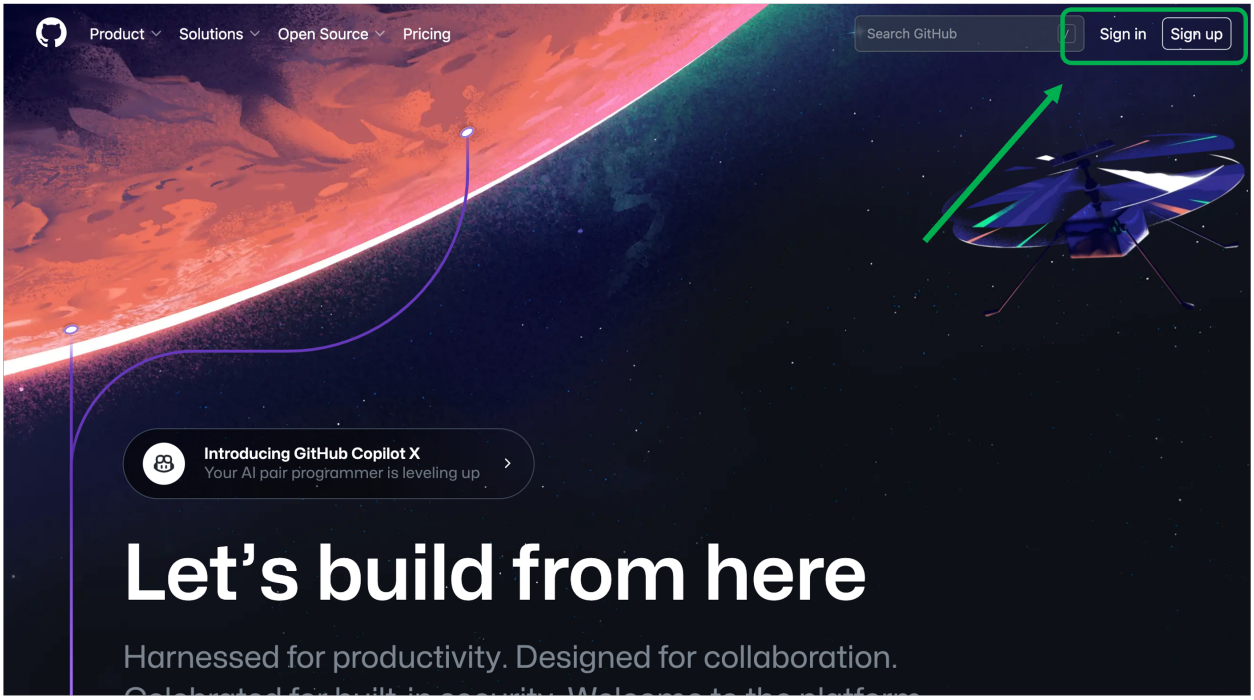

Fig. 1: GitHub homepage

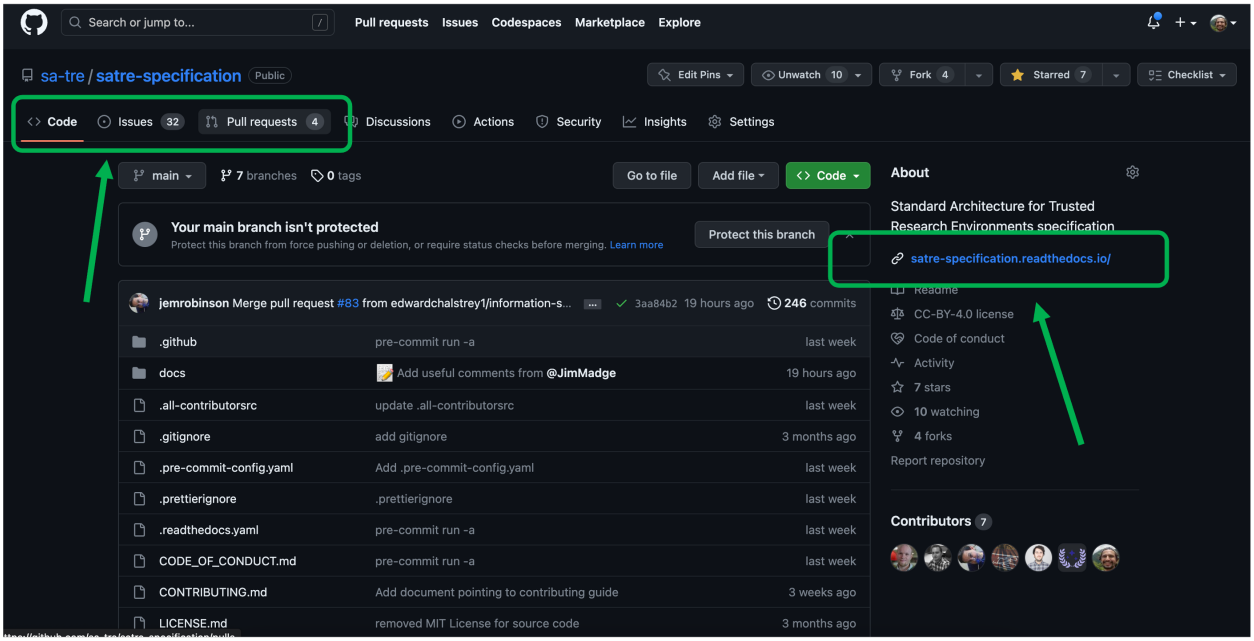

Fig. 2: SATRE specification repository

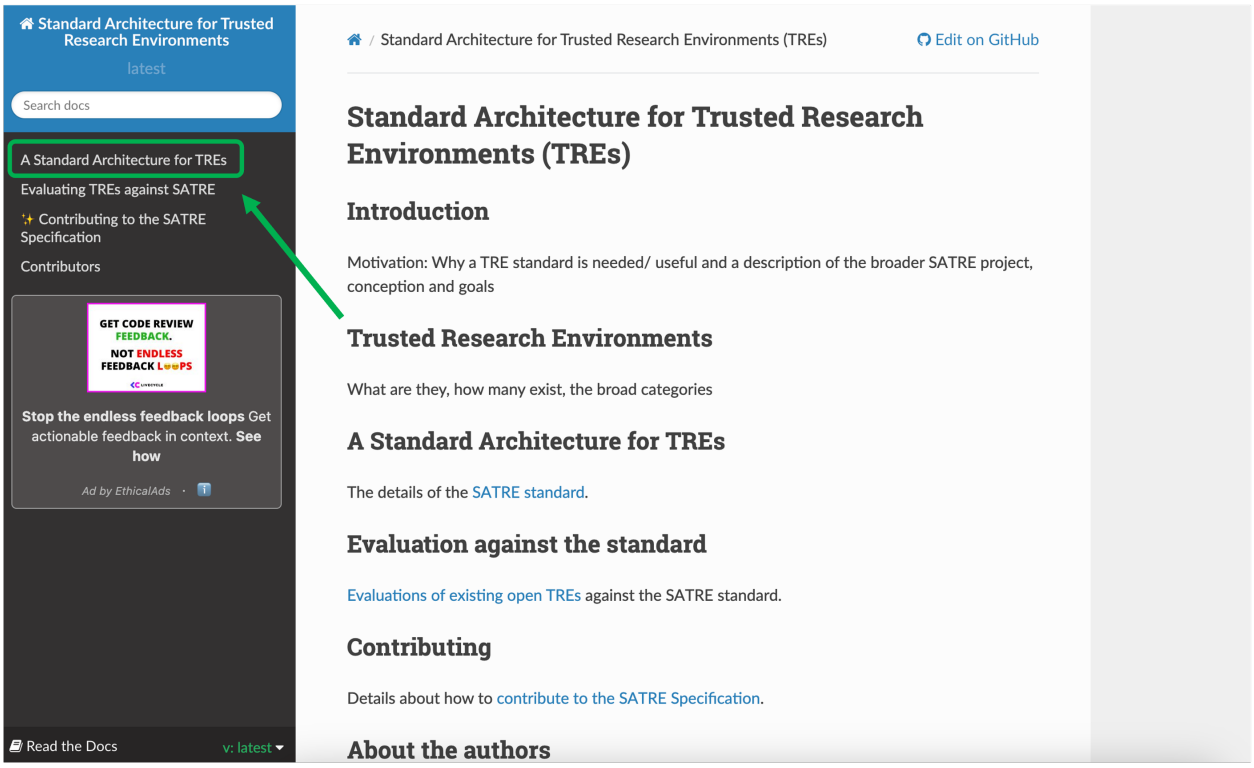

Fig. 3: SATRE specification website

Code

This is the default landing page of the repository, and contains all the code associated with the specification. The most important one for you to know about is where the actual specification is being written.

You can find this by navigating to `docs` → `source` → `specification.md`. This is the specification, written in Markdown, that renders into a website (which you can access from the `About` section).

Most `Issues` and `Pull requests` will be related to this file specifically - and this file is the official specification. You can read more about `Issues` and `Pull requests` below.

Issues

Navigate to the `Issues` tab.

This is where the community is having discussions about ideas for the specification. You can think of it like an online forum where you start discussions and comment on pre-existing discussions.

There are two main things you can do:

- Comment on a pre-existing issue
- Open a new issue

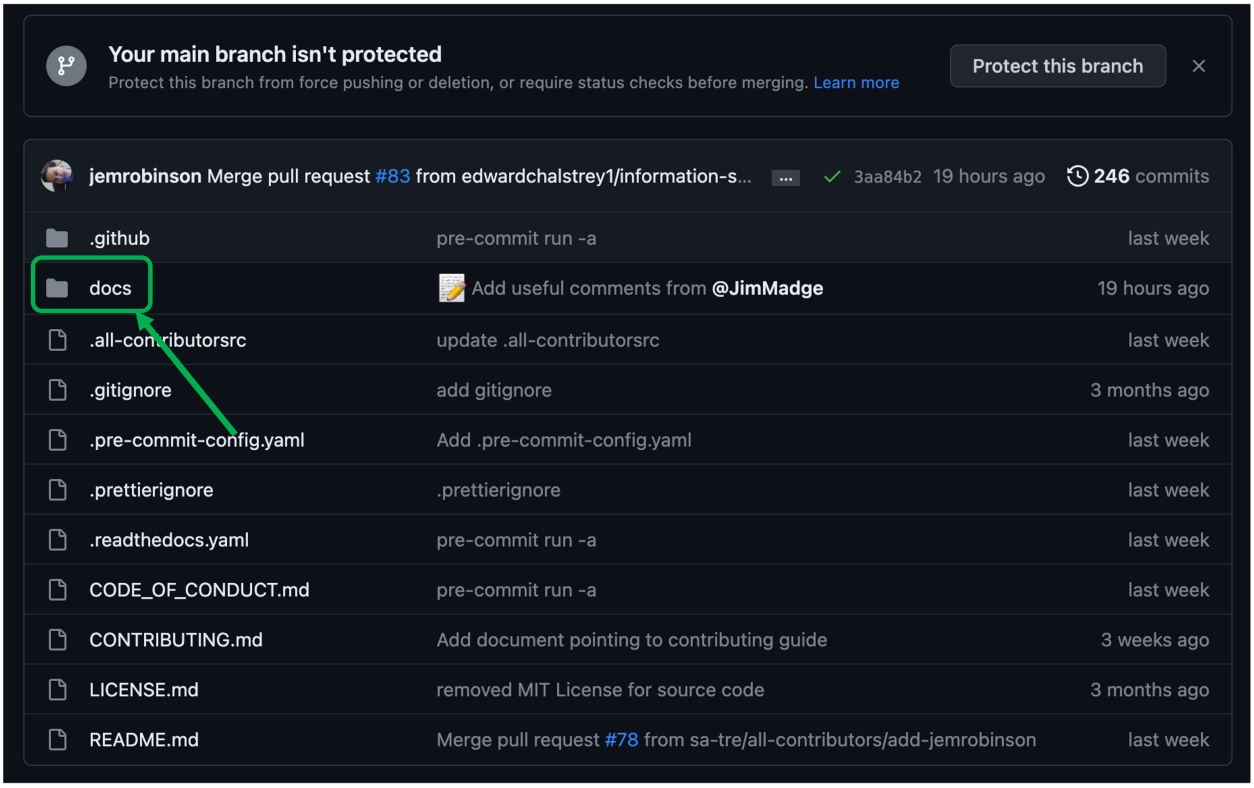

Fig. 4: SATRE repository docs directory

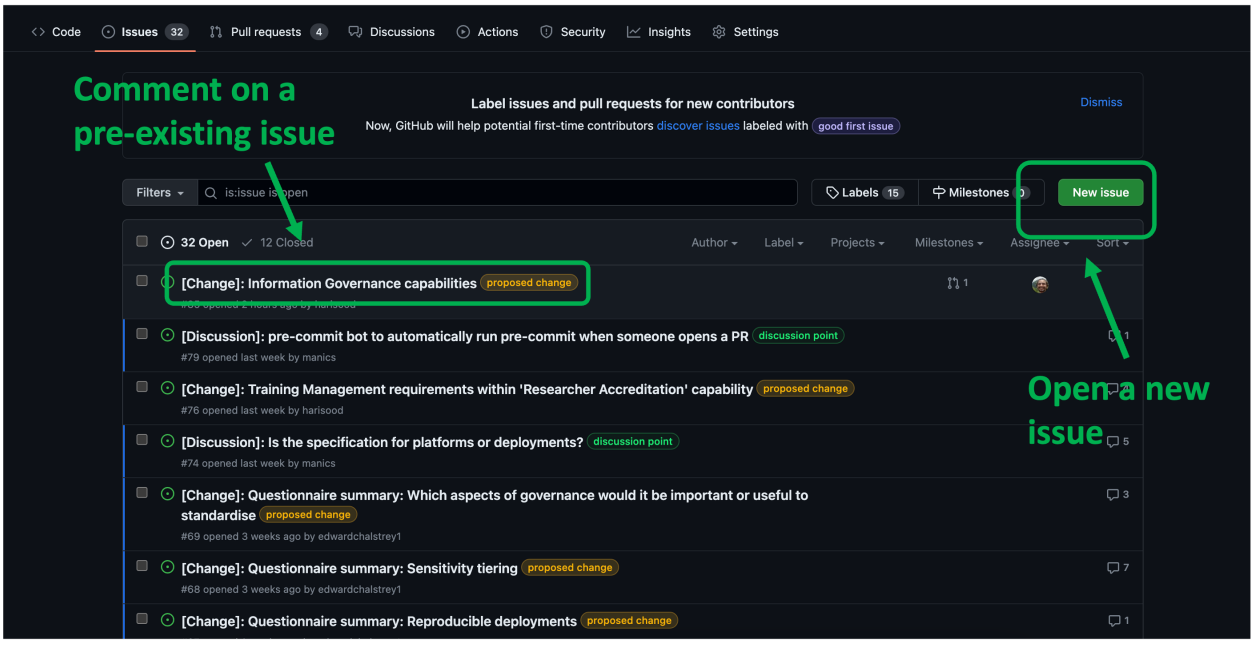

Fig. 5: SATRE repository issues

Commenting on a pre-existing issue

1. Click on the title of the Issue - in the screenshot above, you can click on `[Change]: Information Governance capabilities`.
2. This should load the issue. You should be able to see the title of the issue, the summary of it (provided by the person who created the issue), and a place you can add comments at the bottom.

Fig. 6: Comment on SATRE specification issues

1. Commenting on an issue is really easy! Just write your comment in the space provided, and click `Comment`. Anyone following the issue will get notified that you've commented.

Creating a new issue

1. From the `Issues` page, click the green button `New issue`
2. You should be able to select a template for the issue you want to create - you can choose from the available options, or `Open a blank issue` if no template is right.
3. Fill in the template and select `Submit new issue`.

For more information on how Issues are being governed, see our *Contribution Process*.

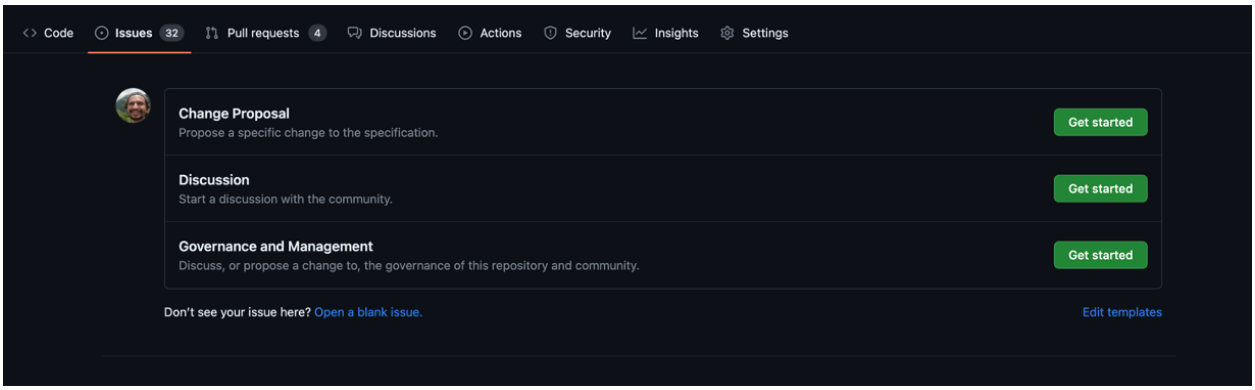

Fig. 7: Create new SATRE specification issue

Pull requests

Navigate to the `Pull requests` tab.

Pull Requests (PRs) are where the community is making specific change proposals to the wording of the specification. You can think of it like making suggested/track changes to a Google Doc or Word document. These need to be reviewed by the SATRE team before they are accepted.

PRs are a little more difficult to wrap your head around! This guide will show you how to comment on open PR. For a deeper dive, including how to create a PR, make changes and more, we recommend the Turing Way's Introduction to GitHub Workshop, run at CarpentryCon22

Comment on open Pull Requests

This process will be similar to the one for Issues above

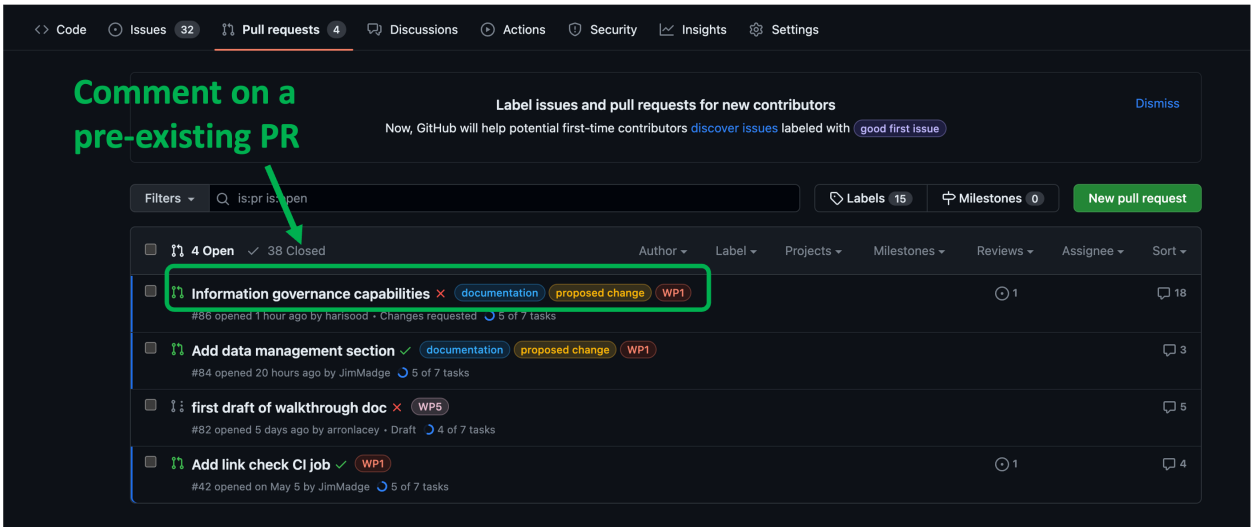

Fig. 8: Comment on a SATRE specification pull request

1. Click on the Title of a PR - in the screenshot above, you could click on `Information Governance capabilities`.

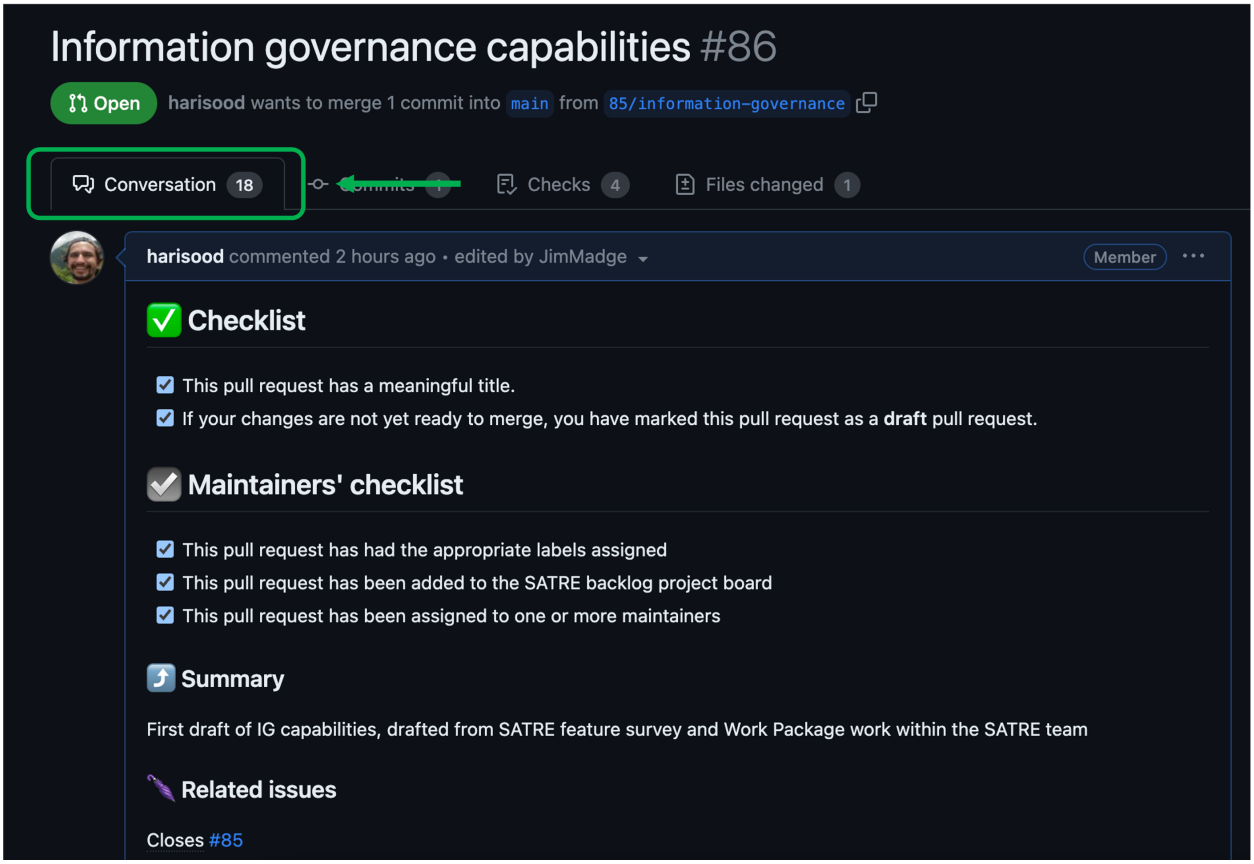

Fig. 9: View discussion on a SATRE specification pull request

2. There are many things you can do on a PR - make changes to a file, comment on proposed changes, and more! For the purposes of this guide, we will just focus on comments. Make sure you are on the `Conversation` tab.

3. From here, you can comment in the same way as you did with `Issues` (add your comment at the bottom, and click `Comment`).

Create your own Pull Request

If you would like to directly author a change to the specification yourself, then you can create a Pull Request. Currently, you will need to make a `Fork` of the SATRE repository to make a Pull Request. Here's how it works:

1. **Create a Fork of the SATRE specification repository:**

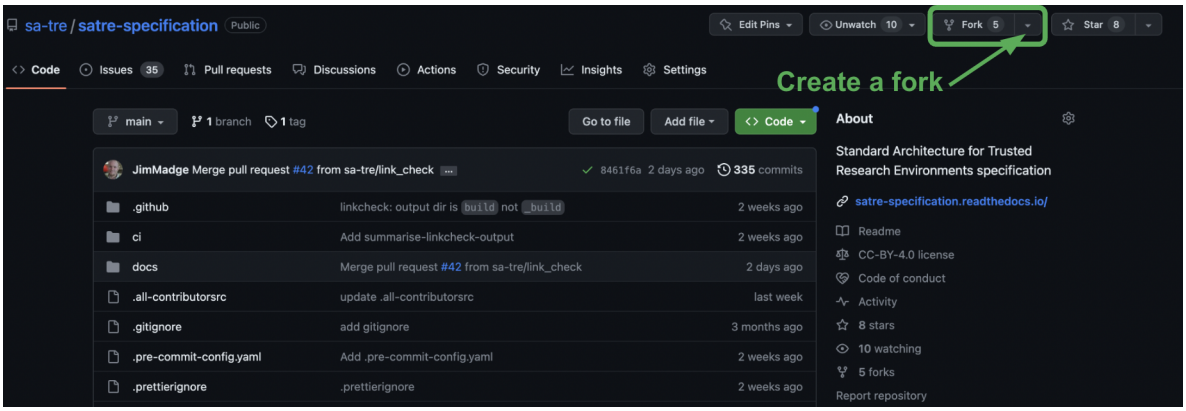

Fig. 10: Create a fork of the SATRE repository

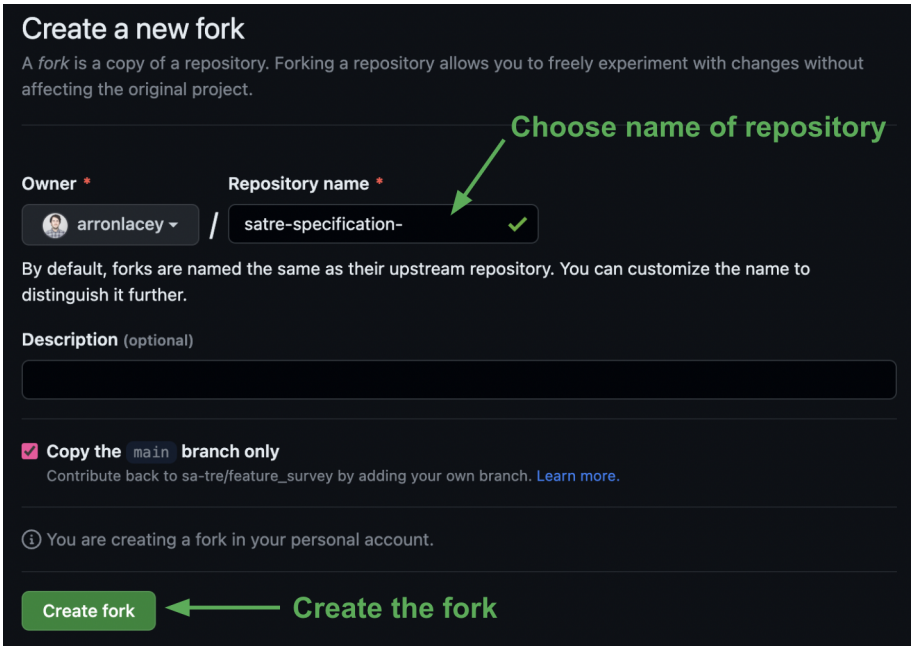

Fig. 11: Name your fork of the SATRE repository

2. **Keep the fork up to date.** It is common for work to continue on the original repository while you are working with your forked version of the repository. This means work on the original repository will not be reflected in

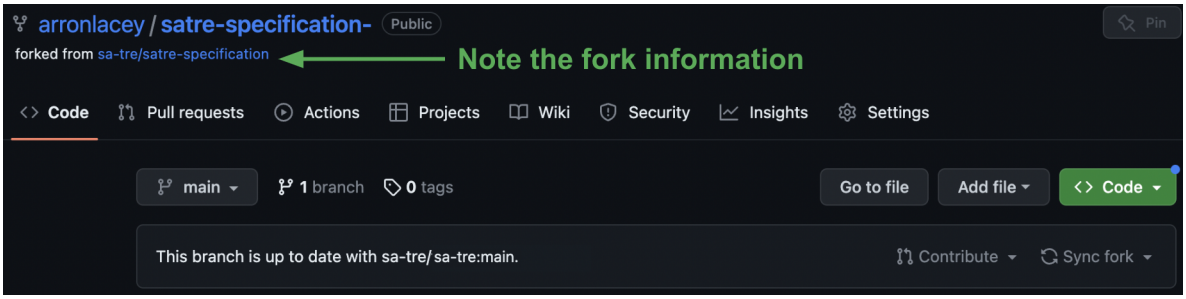

Fig. 12: Link to the main SATRE repository

your forked repository. You can keep your forked repository up to date by pressing the sync button (note that we are assuming there aren't any merge conflicts):

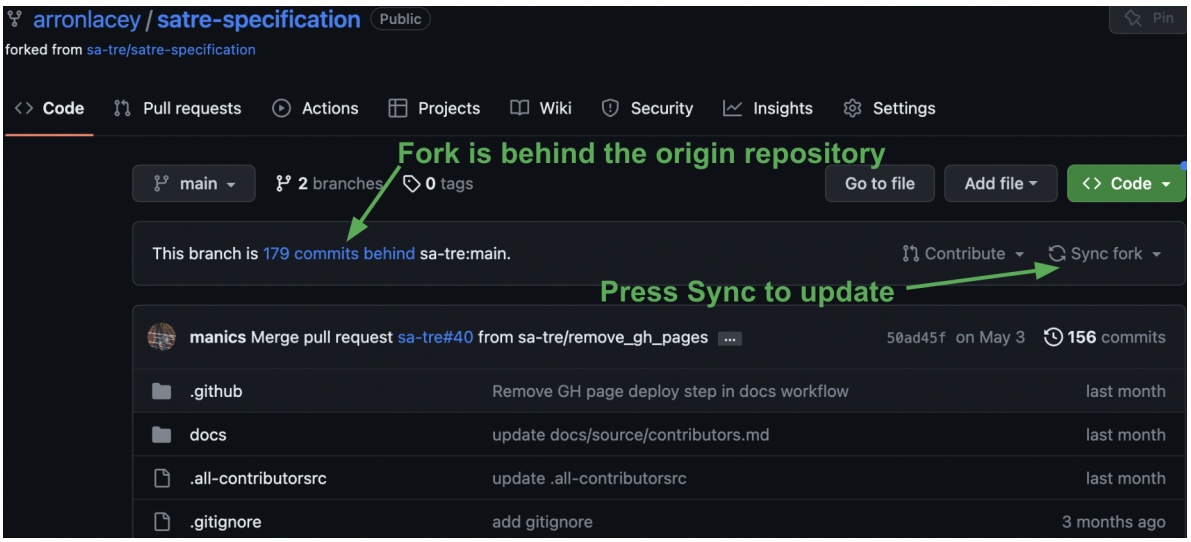

Fig. 13: Update your fork of the SATRE repository

1. **Edit the section you want to change.** Click on the `specification.md` file in the repository, where you will be able to edit it:

Note that although you are only changing your Forked repository (not the original repository) it is still recommended that you use Pull Requests and branches within your own Fork rather than committing directly to the `main` branch

4. **Create the Pull Request to the original repository.** To merge the changes on this new branch from the forked repository into the main branch of the original repository:

You have now just opened a Pull Request intended to merge the changes on your Fork to the original SATRE Repository! Don't forget to sync your Fork when the Pull Request gets accepted into the original repository.

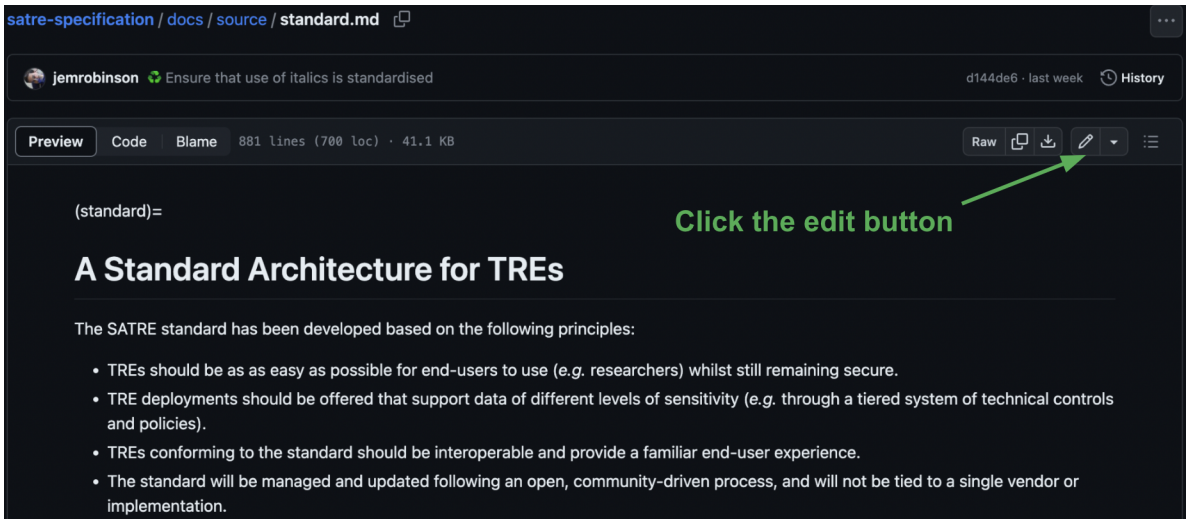

Fig. 14: Edit a file

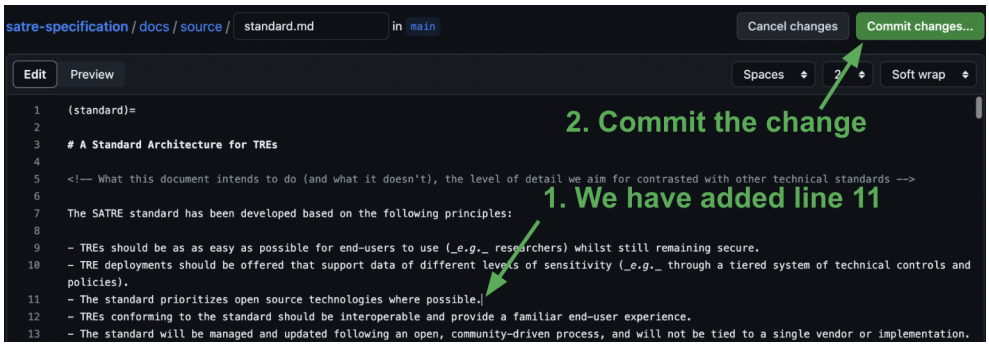

Fig. 15: Submit your changes

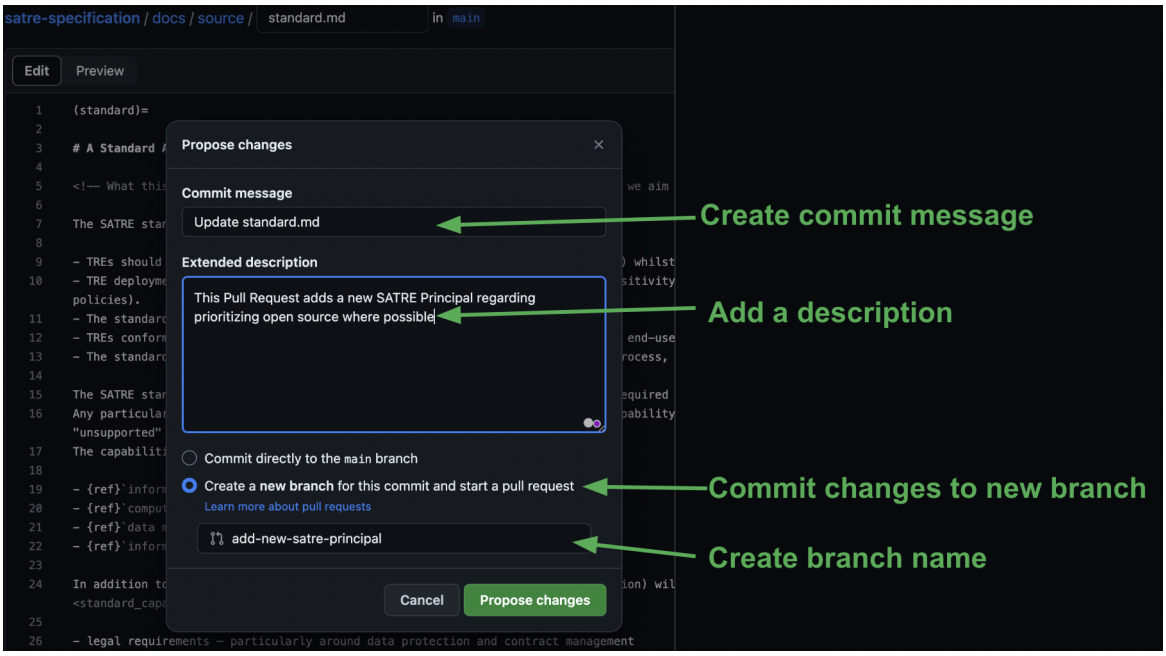

Fig. 16: Write a message to explain your changes

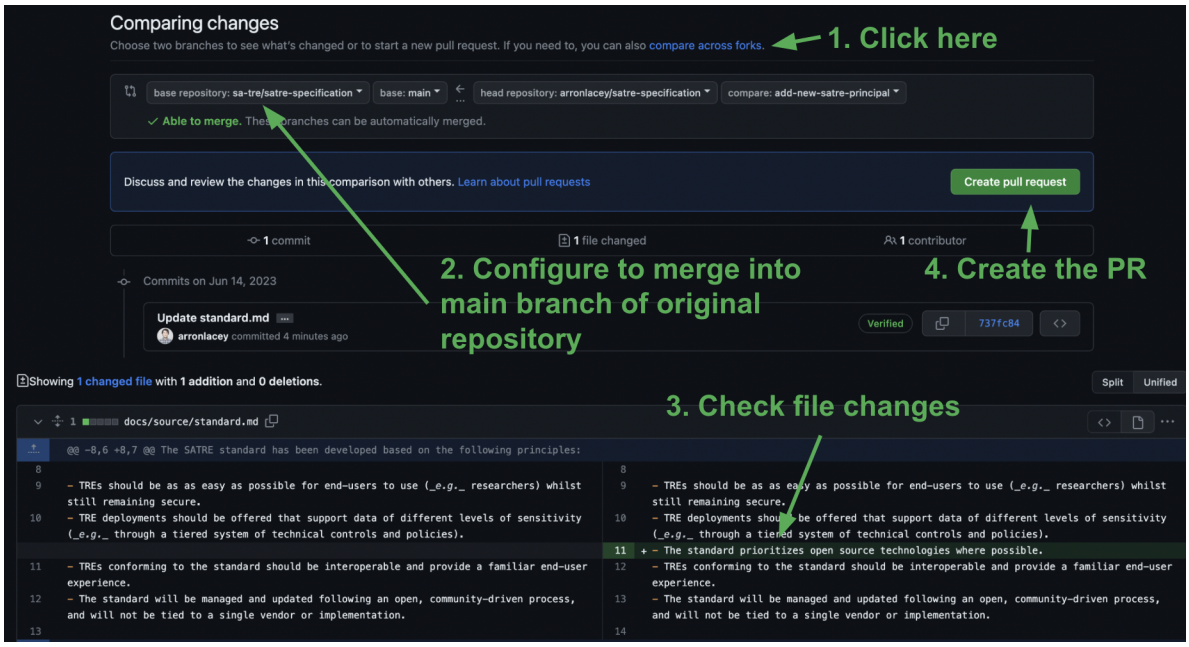

Fig. 17: Set the target to the main SATRE repository

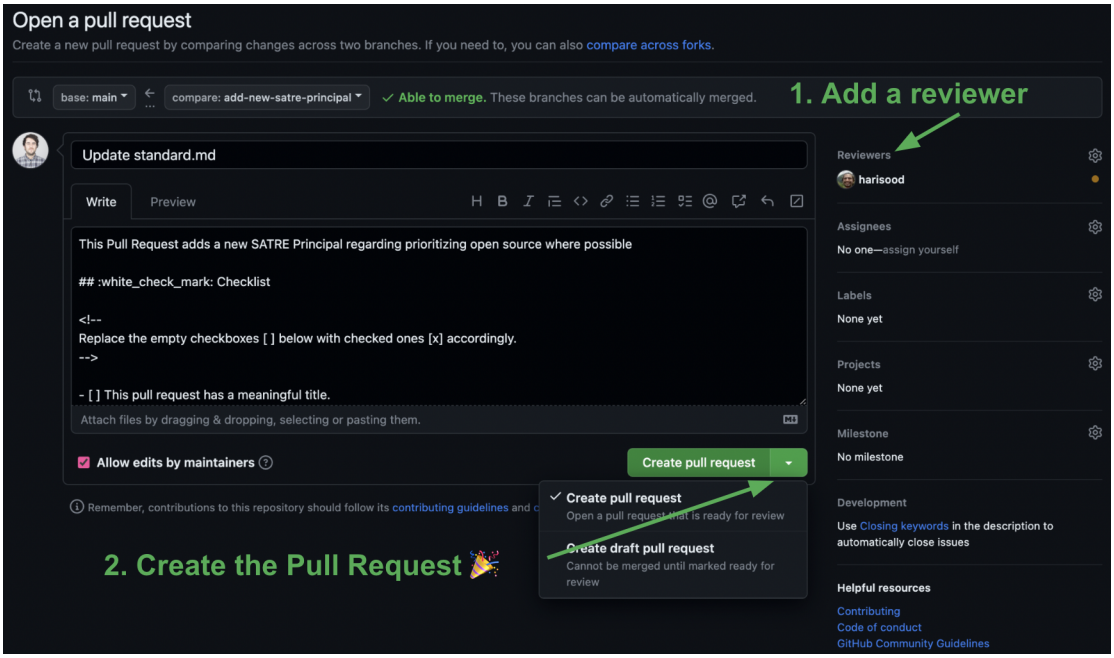

Fig. 18: Create your pull request!

15.5.2 Additional considerations for GitHub

Notifications

In order to make sure you stay informed of conversations you have joined, you need to make sure your notification settings are switched on

Notifications from the repository

From the `Code` tab, click `Watch` → `Participating and @mentions`. This will ensure you get notified to any conversations where you are already taking part, or are mentioned by someone else. If you are really keen, you can turn on notifications for the whole repository, to be notified of any new issues or pull requests people open!

You will know you have the right setting when a tick appears next to your chosen notification level.

Notification streams

You can also decide how you receive notifications - whether just on GitHub, or also via email. We recommend receiving `Participating and @mentions` notifications by email too, to ensure you don't miss any conversation you're involved in!

1. Click your profile in the top right corner of GitHub and select `Settings`.
2. On left-hand navigation bar, click `Notifications`. Ensure your email is the right one, and then in `Participating and @mentions` check both `GitHub` and `Email`, and click save.

You are done! You should now get emails for issues/PRs you are directly involved in across GitHub :rocket: If this becomes annoying, you can always come back to this page to turn them off.

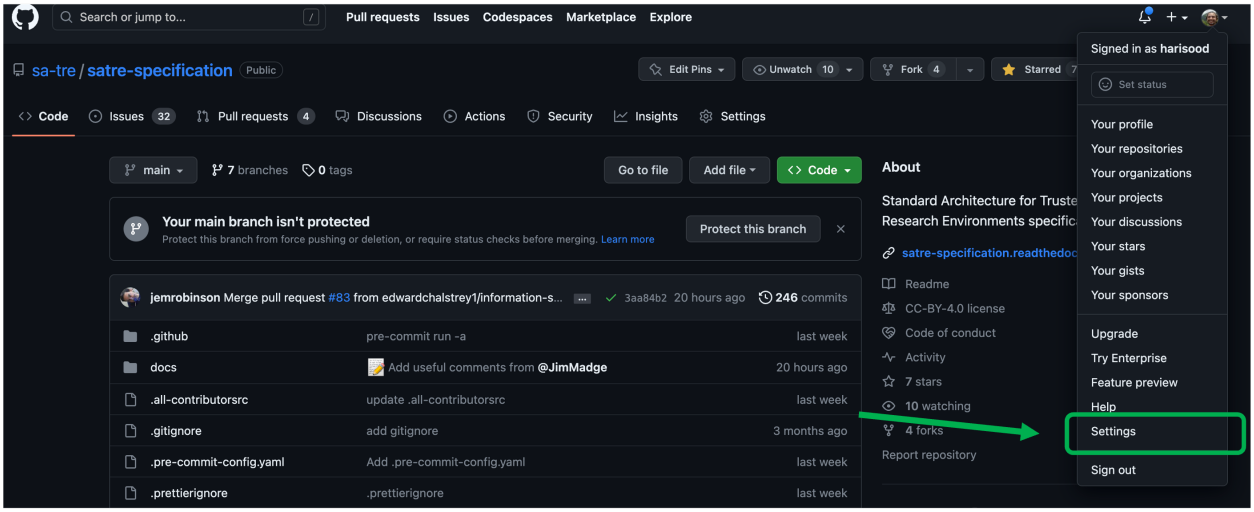

Fig. 19: GitHub settings

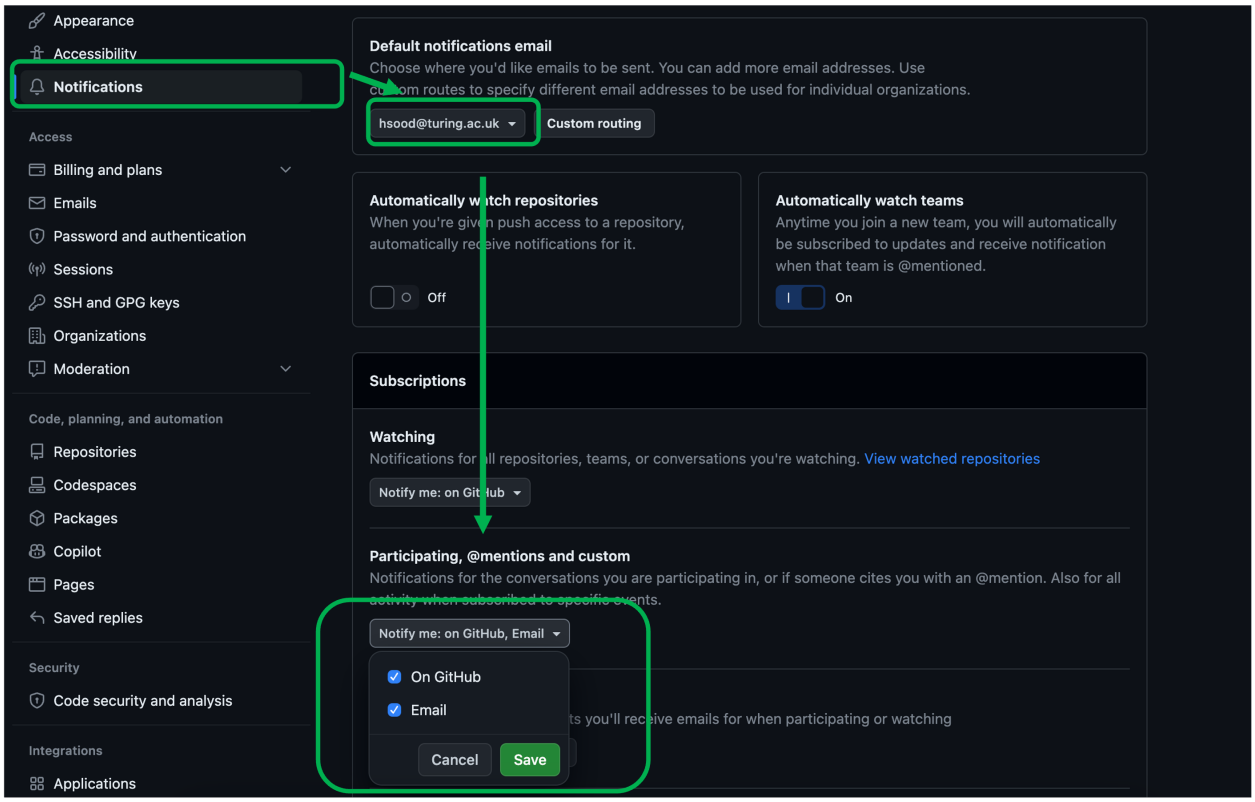

Fig. 20: GitHub notification settings

Helpful Markdown/GitHub tools

When you are commenting on Issues/PRs, there are a couple of handy things to know:

- You can mention others by tagging them with `@` followed by their GitHub username. For instance, to tag and notify Hari (GitHub username `harisood`), you can write `@harisood`
- All Issues and PR have a number associated with them. For instance, the number associated with the below issue is `85`. In any of your comments, you can reference an issue or PR by typing `#` followed by the number of the issue/PR you want to reference. For instance, to reference the below, you can type `#85`, and GitHub will magically create a link to it for you!

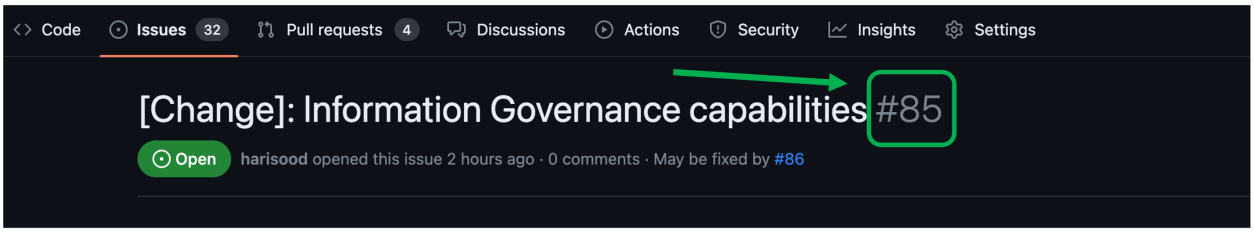

Fig. 21: Pull request number

There will always be a dedicated Breakout Room in the Collaboration Cafés where one of the SATRE Team will be on hand to answer any questions, guide you through the GitHub Repository, and help you get set up.

15.5.3 Contribute via alternate streams

We're working on other ways to contribute and will update this document when they are developed.

15.6 Building the specification website locally

Instructions for building the specification website locally can be found in the `README` of the project.

15.7 Code of Conduct

The SATRE project is dedicated to providing an inclusive and respectful environment for all participants. Please review the project's Code of Conduct before starting your contribution.

15.8 Contact

If you have any questions or concerns, reach out to SATRE project team member Hari Sood (@harisood, hsood@turing.ac.uk).

CHAPTER

SIXTEEN

CONTRIBUTORS

CHAPTER

SEVENTEEN

WHAT IS SATRE?

The SATRE project provides a Standard Architecture for *Trusted Research Environments (TREs)*. It incorporates knowledge and best practices from multiple institutions and sectors across the UK. This includes all aspects of TRE provision such as information governance procedures, computing technology, data management and other capabilities.

It aims to standardise the capabilities of TREs, making it easier for users, operators, and developers to work with sensitive data, and making the operation of TREs more transparent to data owners and the general public.

This specification should be useful if you are:

- a *TRE Operator* wanting to evaluate or improve their TRE with the suggested capabilities
- a *Developer or Builder* of new TREs looking for guidance in their thinking and decision making

We encourage all TREs in the UK to *evaluate* themselves against the SATRE specification, and to *contribute* to the project.

CHAPTER

EIGHTEEN

GETTING STARTED

If you are familiar with SATRE and want to evaluate your own TRE you can jump straight to the *evaluation section* which includes an *Excel spreadsheet* you can use for your evaluation.

If this is your first time here we recommend reading the rest of this page to understand the background behind SATRE, followed by:

- *Frequently asked questions*
- *The specification*
- *How to evaluate your TRE*

CHAPTER

NINETEEN

WHY DO WE NEED TRES?

Personal or sensitive data which have been collected for operational, commercial or governmental reasons need to be managed securely and safely. A TRE enables researchers to access the data in a secure environment following best practice. This should ensure that research projects and *data consumers* are properly authorised and that researchers only access the data they need, whilst minimising risk of data release or exposure.

CHAPTER

TWENTY

WHY ARE WE DOING THIS NOW?

The need for trusted research environments (TREs) is clear. Influential reports including the UK Government's Goldacre review and 'Data Saves Lives' policy paper, have highlighted the need for change in how sensitive data are handled. These papers set out a vision for the potential impact of research enabled by TREs.

At present operators have to interpret a range of frameworks, legislation and guidance when building and running a TRE. These include:

- Office for National Statistics: 5 Safes
- UK Health Data Research Alliance: TRE Green Paper
- UK Health Data Research Alliance: TRE Principles and Best Practices
- Design choices for productive, secure, data-intensive research at scale in the cloud
- ISO27001
- Digital Economy Act
- Handbook on Statistical Disclosure Control for Outputs

This makes for inconsistent governance standards and makes it hard for researchers to work consistently in different environments.

A common specification for TREs will improve governance and practice across the sector, simplify researcher and operator journeys. Furthermore, it will lay a foundation for interoperability that is required to maximise the impact of research by providing a trusted ecosystem for working with currently disparate and siloed data.

CHAPTER

TWENTYONE

WHO ARE WE?

The SATRE team contains representatives from several existing UK TREs, which host many different types of sensitive data. We will use the reference architecture specified here to bring these into closer alignment and make it easy for others to do the same. This supports DARE UK's aim of developing a coordinated national data research infrastructure.

CHAPTER

TWENTYTWO

CONTRIBUTING

We welcome contributions from anyone who is interested in the project. There are lots of ways to contribute, not just writing code! Find out more about how to *contribute to the SATRE Specification*.

CHAPTER

TWENTYTHREE

ACKNOWLEDGEMENTS

We are grateful for the following support for this project:

- UKRI via the DARE UK Phase 1 driver projects programme (SATRE)

INDEX