

Fostering a thriving PETs ecosystem

October 2024



Contents

About	2
Executive summary	3
Background	4
Methods	6
The Consolidated Framework for Implementation Research (CFIR)	8
Adapting CFIR for PETs	10
Fostering a thriving PETs ecosystem	13
Tackling human trafficking with differentially private synthetic data	14
Background	14
Implementing the PET	17
Enabling granular health data research with the SAIL Databank	
trusted research environment (TRE)	19
Background	19
Implementing the PET	21
Using secure multi-party computation to measure wage-gap	
inequality in Greater Boston	23
Background	23
Implementing the PET	24
Cross-case analysis	29
Towards a PET ecosystem typology of actors	29
Development of a PETs implementation framework	31
The seven phases of the implementation framework and	
recommended activities	32
Scope	32
Engage with stakeholders to assess the problem/ opportunity	33
Map and analyse the data ecosystem	33
Define and prioritise use cases	34
Blueprint	34
Explore a range of approaches	34
Agree on the functions that different actors will need to perform	35
Co-design the four layers of the PET implementation	36
Engage with stakeholders to iterate and finalise the design	37
Secure funding for the further phases, if applicable	37
Roadmap	38
Deploy	39
Operate	40
Evaluate	40
Iterate/ retire	41
Conclusion	43
Acknowledgements	43

Annex	44
Limitations	44
PET case study selection	44
Future research	44
Future implementation framework based research	44

About

This report has been researched and produced by the Open Data Institute, and published in October 2024. Its authors were Neil Majithia, Elena Simperl, Claudine Tinsman, Elea Himmelsbach, Jared Keller and Calum Inverarity. If you want to share feedback by email or would like to get in touch, contact the privacy-enhancing technologies (PETs) programme team at pets@theodi.org.

Executive summary

Privacy-enhancing technologies (PETs) have the potential to enable the use of sensitive data that would otherwise need to be kept private. Yet their adoption remains somewhat limited. While literature exists on how these technologies function and where they have been successfully deployed, less research has been conducted on the factors that have contributed towards their successful implementation. This research addresses that gap by analysing the roles played by various actors in the ecosystems where PETs have been successfully deployed. Based on this research, we developed a ‘Version 0’ PETs implementation framework for prospective PET adopters.

In this report, we provide guidance on how prospective adopters of PETs could choose, develop and implement a PET to help them access and share sensitive data responsibly. The intended audiences for this guidance are smaller and medium enterprises, which might have fewer resources – be that financial or knowledge – but nonetheless could benefit from the adoption of a PET. This is in line with UK efforts to promote the use of these technologies.

We provide guidance to encourage innovation through responsible access to, and sharing of, data.¹ This report seeks to address a gap in knowledge and understanding of the actors that have contributed to the successful implementation of PETs, through examining three separate PET ecosystems. We adapted an existing framework from implementation science, which contributed towards our development of a [‘Version 0’](#) PETs implementation framework. This framework draws upon the lessons learned from the three examples of implemented PETs that are included in this report as case studies, as well as previous ODI research on facilitating safe access to sensitive data.

As a next step, we propose that this [‘Version 0’](#) PETs implementation framework should be subjected to user testing by prospective PET adopters to evaluate for suitability and modification.

¹ Department for Digital, Culture, Media and Sport (2020), [‘National Data Strategy’](#).

Background

Interest in and the use of privacy-enhancing technologies (PETs) has been growing in recent years, following the passing of data protection legislation such as the General Data Protection Regulation (GDPR)² and the California Consumer Protection Act.³ At the Open Data Institute (ODI), we consider these technologies as ‘tools and practices that can enable access to data that may otherwise be kept closed for reasons of privacy, as well as commercial sensitivity or national security’.⁴ Our definition is intentionally broad, as we appreciate that some of these tools and practices – such as passwords and encryption – are more established, understood and commonly used than others that have received increasing attention in recent years.

Some of the more novel PETs, such as federated learning and federated analytics,⁵ secure multi-party computation⁶ and synthetic data,⁷ have been well documented in how they function. However, fully-deployed examples of their practical application are much more limited. Similarly, while documentation of attempted and fully-deployed PETs exists and examples continue to be added,⁸ analysis of the success of these deployments could benefit from further examination. Understanding what contributes to a successful deployment of a PET requires an ecosystem approach, which is currently lacking in the vast majority of documented cases. We address this gap through considering the following questions:

- In contexts where PETs have been successfully deployed, who have been the actors involved?
- What roles have these actors played in successfully deploying PETs?

In exploring these research questions, we sought to establish whether there are identifiable characteristics of the ecosystems in which full-deployment of particular PETs has been successful. This included analysis of the types of actors involved, and the ways in which the relationships between these actors contributed to the implementation’s success. Through our research, we also identified an opportunity to consolidate our findings to inform the development of a ‘[Version 0](#)’ PETs implementation framework for prospective

² EUR-Lex (2016), ‘[General Data Protection Regulation](#)’.

³ State of California Department of Justice (2024), ‘[California Consumer Privacy Act \(CCPA\)](#)’.

⁴ The Open Data Institute (2023), ‘[Privacy enhancing technologies \(PETs\)](#)’.

⁵ Nair, A. and Inverarity, C. (2022), ‘[What is federated learning?](#)’.

⁶ Himmelsbach, E. et al. (2024), ‘[PETs in Practice](#)’.

⁷ Thereaux, O. (2019), ‘[Anonymisation and synthetic data: towards trustworthy data](#)’.

⁸ For example, see the Centre for Data Ethics and Innovation’s (2023), ‘[Repository of Use Cases](#)’ and the United Nations (2023), ‘[United Nations Guide on Privacy-Enhancing Technologies for Official Statistics](#)’.

adopters of PETs. This framework has been informed by existing ODI research, in combination with our consolidated findings. This is meant only as a '[Version 0](#)' work-in-progress that could benefit greatly from community testing. We encourage interested organisations and individuals to get in contact to explore this opportunity with us.

Methods

We first identified frameworks and methodologies for the adoption of new technologies. Through this process, we identified a number of promising frameworks that could be adapted and drawn upon to address the question: ‘What combination of roles and responsibilities makes for a thriving PETs ecosystem?’ We then sought to apply this to several examples of successful deployments of PETs.

We reviewed various frameworks, including the National Institute of Standards and Technology’s (NIST) National Initiative for Cybersecurity Education (NICE) Workforce Framework for Cybersecurity⁹ and the European Union Agency for Cybersecurity’s (ENISA) European Cybersecurity Skills Framework (ECSF).¹⁰ These guided our research and helped make sense of the different actors and roles that are engaged in developing a thriving PETs ecosystem. However, upon closer examination, we determined that neither framework aligned closely enough with our research objectives. NIST’s competency-based approach was too narrow for our purposes, whereas we deemed ENISA’s profiles approach too prescriptive. After broadening our research, we discovered The Consolidated Framework for Implementation Research ([CFIR](#)), a method used in implementation science.¹¹

We selected three different cases, rather than focus on three deployments of a single PET in different circumstances. The rationale behind this decision was based on our awareness that publicly documented examples of successful, fully-deployed cases of PETs still remain relatively limited. Furthermore, situational variables have been widely acknowledged as highly significant to the successful deployment of PETs, which warranted further interrogation. Therefore, we deemed it more appropriate to focus on a variety of examples that allow for consideration of contrasting factors, such as whether a PET was deployed with a large degree of autonomy by a single actor, or whether it was a collaborative effort between partners.

The first case that we focused on was the creation of a differentially private synthetic dataset, constructed by Microsoft in partnership with the International Organization for Migration (IOM).¹² This example was selected as it involved the use of two PETs – differential privacy and synthetic data – and collaboration between a company and an inter-governmental organisation.

⁹ NIST (2020), ‘[National Initiative for Cybersecurity Education \(NICE\) Cybersecurity Workforce Framework](#)’.

¹⁰ The European Union Agency for Cybersecurity (ENISA) (2024), ‘[European Cybersecurity Skills Framework \(ECSF\)](#)’.

¹¹ For additional information about implementation science, please see: Bauer, M. and Kirchner, J. (2020), ‘[Implementation science: What is it and why should I care](#)’.

¹² Microsoft (2022), ‘[IOM and Microsoft release first-ever differentially private synthetic dataset to counter human trafficking](#)’.

The second case study is the SAIL Databank Trusted Research Environment (TRE),¹³ which itself is a type of PET where informational architecture and governance processes are combined to create a secure environment for researchers to access data in. This case was selected partially as the functioning of this TRE required academic and public sector collaboration.

The third case study selected is the Boston Women's Workforce Council and Boston University's use of secure multi-party computation (MPC) to address gender and racial wage gaps in the Greater Boston area.¹⁴ This use of MPC proved a rich ecosystem case study as it involved cooperation between numerous actors that played different roles in providing the architecture and stewarding the use of the PET, and supplying the data required to undertake the analysis. Given the diversity of actors taking part, it could be argued that a greater degree of trust was required between those involved – as became apparent during interviews with those involved in the deployment of this PET.

In addition to the different numbers and types of actors involved in each of the PET deployments, rationale for selecting these three examples included the volume of publicly available documentation on each. This helped us construct a detailed timeline of the course of each PET's development and implementation.

The case studies within this report are the results of a combination of desk research and follow-up interviews and questions with relevant persons involved in the development and implementation of the chosen examples. They begin with a short background to provide necessary context to aid the understanding of the environment in which the PET was implemented, before moving on to discussion of the critical actors involved, which we have plotted against corresponding implementation timelines.

We then conducted cross-case analysis to identify the similarities and differences between developmental cycles of the three PETs. We consolidated the findings into a '[Version 0](#)' PETs implementation framework, included later in this report.

¹³ SAIL Databank (2021), '[What is a TRE?](#)'.

¹⁴ Trusted CI (2020), '[Transition to Practice success story: Boston University - Secure multiparty computation and the Boston Women's Workforce Council](#)'.

The Consolidated Framework for Implementation Research (CFIR)

Implementation science is a relatively new applied field of study that aims to bridge the gap between research and practice. Its aim is to create evidence-based thinking to understand how to best use specific interventions and strategies that have been proven to work in similar settings. Implementation science, and by extension CFIR, is predominantly used to evaluate and guide implementations in healthcare, but has also been successfully employed in related domains to evaluate non-healthcare-related interventions in fields such as agriculture and public health.

CFIR is a theory-based model designed to study the effectiveness of implementation strategies. It can be applied for two distinct purposes:

1. **Implementation design**¹⁵ helps tailor implementation strategies to mitigate barriers and leverage facilitators based on outcomes from an initial context assessment
2. **Implementation evaluation**¹⁶ guides how to collect the necessary data/insights to assess an implementation's success

Both approaches – the context assessment to enable implementation design and implementation evaluation – are relevant to our approach, given our aim to derive insights and learnings from analysing established case studies, and use them to model how to foster a thriving PETs ecosystem.

¹⁵ CFIR Research Team (2024), '[Strategy Design](#)'.

¹⁶ CFIR Research Team (2024), '[Evaluation Design](#)'.

The framework offers an overarching typology that consists of a list of ‘constructs’. These constructs promote theory development and verification on what innovation works where, and why, in practice across multiple contexts. Specifically, the framework offers a structured method to:

1. **Identify** the people and organisations with the power to effect change
2. **Align** disparate groups within a single organisation and/or across an ecosystem
3. **Communicate** needs and actions across an organisation and/or ecosystem
4. **Assess** implementation outcomes to understand how to make changes where necessary
5. **Understand** how to translate knowledge and experience from one context to another and make necessary adjustments to fit the new context

CFIR comprises 39 constructs organised into five domains. The constructs comprise considerations that adopters of the framework can bear in mind when designing their intervention, which can assist in the process of building a picture of where they might need to address potential obstacles or seek assistance. These five domains and illustrative constructs are:

1. **Innovation domain** – the ‘thing’ being implemented, such as a new clinical treatment, educational programme or city service.
 - Example constructs include: source, evidence-base, relative advantage and adaptability.
2. **Outer setting domain** – the setting in which the inner setting exists, such as hospital system, school district or state.
 - Example constructs include: critical incidents, local attitudes and conditions, financing and external pressures.
3. **Inner setting domain** – the setting in which the innovation is implemented, such as hospital, school or city.
 - Example constructs include: structural characteristics, physical infrastructure, IT infrastructure and culture.
4. **Individuals domain** – roles and characteristics of individuals.
 - Example constructs include: high, mid-level and opinion leaders; implementation facilitators, leads and team members; and innovation deliverers and recipients.

5. **Implementation process domain** – activities and strategies used to implement the innovation.

- Example constructs include: teaming, assessing needs and context, and assessing context.¹⁷

In a survey of 19 authors of research articles in which CFIR had been applied to low and middle-income countries, a substantial proportion of respondents found four out of the five constructs within the individuals' domain were either incompatible or irrelevant to research on the structure of healthcare systems in these countries.¹⁸ Furthermore, a substantial proportion of these authors cited that the notion of individuality within healthcare teams did not align with the prevailing organisational culture. They expressed a need for context-specific adaptations to the framework in situations where high-level factors played a greater role in a given system than individuals. Both of these instances demonstrate some of the limitations of applying a standardised framework to implementations that will inevitably require a reasonable degree of contextualisation. It is in appreciation of the variety of PETs that exist and the purposes that they may be used for that we determined that the in-built flexibility of CFIR lent itself as a suitable candidate framework when evaluating the deployment of these technologies. Further, while the studies in question were conducted in the context of healthcare settings, CFIR has also been adapted to evaluate non-healthcare-related interventions such as agriculture¹⁹ and public health²⁰, thus demonstrating contextual transferability.

Adapting CFIR for PETs

To our knowledge, CFIR has not been applied to PETs implementation. At the time of writing CFIR also had no prior guidance on the selection of relevant constructs. In the absence of concrete examples, we relied on the guidance from the CFIR's authors to select the appropriate constructs through a series of group discussions. We evaluated each construct with respect to its role in the PETs implementation to determine which constructs were most compatible with our unique research context.²¹

¹⁷ Summarised from: CFIR Research Team (2024), '[Updated CFIR Constructs](#)'.

¹⁸ Means, A.R. et al. (2020), '[Evaluating and optimizing the consolidated framework for implementation research \(CFIR\) for use in low- and middle-income countries: a systematic review](#)'. *Implementation Science* 15, 17.

¹⁹ Tinc, PJ et al. (2018), '[Applying the Consolidated Framework for implementation research to agricultural safety and health: Barriers, facilitators, and evaluation opportunities](#)'.

²⁰ Allen, M. (2021), '[Applying a Race\(ism\)-Conscious Adaptation of the CFIR Framework to Understand Implementation of a School-Based Equity-Oriented Intervention](#)'.

²¹ CFIR Research Team (2024), '[Evaluation Design](#)'.

We identified three actor-diverse use-cases that include a differentially private synthetic dataset to counter human trafficking²²; a trusted research environment in Wales to enable health and social research; and secure multi-party computation in the Greater Boston area²³ to measure income inequality. Across these three cases, we managed to capture a variety of PETs. Each has served to preserve privacy in different contexts, whether when data is collected, being shared or published. At the same time these are uses that benefit society, which we believe PETs should be directed towards.²⁴

We started by applying the CIFR framework to the PET deployments by first attempting to gather data that would be necessary in order to fill out the various fields within the framework. In the process, we identified several modifications:

1. **Constructs that were not relevant** – These modifications involved scrutiny of the relevance of certain constructs —namely ‘critical incidents’ and ‘market pressure’ in the outer domain setting and ‘space’ in the inner domain setting.
2. **Constructs that were relevant, but required adaptation** – There were substantially more constructs that we found relevant, but in need of adapting if they were then to be relevant to assessing PETs. Some adaptations included the consideration of metrics to measure constructs such as ‘innovation complexity’ and ‘innovation cost’ in the innovation domain. Specifically, consideration of both the technical and practical complexity of a PET benefits from being considered separately, in order to acknowledge the types of difficulty that prospective adopters might encounter. Similarly, we believe ‘innovation cost’ would benefit from explicit tailoring when applied to PETs applications. This would allow analysis to consider the cost of the PET architecture and the non-technical costs that should be anticipated from an early stage.

For the objective of this research – building comprehensive overviews of implemented PETs, based on successful examples of the types of actors that contributed to creating constructive conditions for the adoption of a specific PET – the ‘individuals domain’ and related constructs provided a solid foundation for analysis. For our purposes, we added the constructs ‘data contributors’ and ‘functional roles’, which we believed were sufficiently significant actors in the implementation of PETs and warranted discrete attention. Further, the ‘individuals domain’ of the framework was adapted slightly to first focus on the roles of organisational stakeholders

²² Microsoft (2022), [‘IOM and Microsoft release first-ever differentially private synthetic dataset to counter human trafficking’](#).

²³ Himmelsbach, E. et al. (2024), [‘PETs in Practice’](#).

²⁴ ODI (2023), [‘Privacy-enhancing technologies’](#).

rather than individuals. We took this decision as our initial research, as an external organisation, was desk-based and involved analysis of publicly available documents, from which it was not always possible to get a granular picture of the influence of individuals key to the development and implementation of the PET. When possible we added this layer of detail following research interviews in the second phase.

Our experience of adapting the ‘implementation process domain’ section of CFIR also informed the development of our own [‘Version 0’ implementation framework](#), covered later in this report.

Fostering a thriving PETs ecosystem

Domains I-III of CFIR (innovation, inner setting and outer setting) include constructs that can largely be answered through a process similar to a political, economic, sociological, technological, legal and environmental (PESTLE) analysis.²⁵ This is a widely-understood and used methodology that is straightforward for an individual or organisation to carry out. It enables organisations to get an understanding of the situational context and environment in which a PET is being deployed, which can help to inform early decision-making processes, such as which PET may be most suitable for the problem. Similarly, domain V – implementation process – provides prospective users of PETs with guidance that can inform the practical planning considerations when implementing an intervention.

However, domain IV – the ‘individual domain’ – deals primarily with relational dynamics between individuals and organisations that can influence and impact the rollout of an intervention. In our case, this includes the decisions behind the planning, development and implementation of a PET. From our research to date on PETs²⁶, we have observed that this component has received less attention compared with more comprehensive efforts in documenting how these technologies can work. This is why our research sought to explore the roles and relationships that contributed towards the successful adoption of specific PETs. This section contains our critical evaluation of these constructs in three contrasting, successful implementations.

For each case study, we created an implementation timeline to illustrate the development and adoption of the three PETs. Through an iterative process, we settled on a grid where the phases of the PET’s development are plotted along the table’s x-axis, while the types of actor we identified during these phases are plotted along the y-axis. We took inspiration from FinOps development processes.²⁷ While detailed explanations of the phases included in our implementation timeline are provided in our section on the [development of our PETs implementation framework](#), for present purposes these include:

²⁵ Government Communication Service (2024), ‘[GCS Knowledge Hub](#)’.

²⁶ ODI (2023), ‘[Privacy-enhancing technologies](#)’.

²⁷ FinOps Foundation (2024), ‘[Architecting VM-based Applications for Cost Efficiency](#)’.

- **Scoping** – consideration of the problem to be addressed, and how this could be achieved
- **Blueprinting** – agreeing on the functions of different actors and necessary stewardship of data and the PET
- **Roadmapping** – planning out activities such as the launch of the PET and necessary engagement between actors involved in the implementation
- **Deploying** – building of the implementation
- **Operating** – running the implementation of the PET in practice
- **Evaluating** – assessment of the implementation against agreed measurements of success
- **Iterating/retiring** – making the decision on whether to amend or adapt the PET, or to cease its operation

Tackling human trafficking with differentially private synthetic data

Background

Many anti-trafficking actors collect large quantities of sensitive data about the characteristics of victims and perpetrators. This raises a number of concerns about privacy and civil liberties, particularly where the possibility of identification is high, carrying significant risks to the victims whose data is included in the dataset.²⁸ These limitations have hindered relevant anti-trafficking actors from sharing information with each other or external stakeholders.

In the past, [The Counter Trafficking Data Collaborative](#) (CTDC), an initiative of the [International Organization for Migration](#) (IOM), addressed this problem by pooling data from its partner organisations and applied k-anonymity to anonymise human-trafficking data.²⁹

K-anonymity is a data anonymisation technique used to ensure that no single individual can be identified from a dataset containing potentially

²⁸ International Organization for Migration (2022), '[IOM-Microsoft Collaboration Enables Release of Largest Public Dataset to Bolster Fight Against Human Trafficking](#)'.

²⁹ Microsoft (2023), '[The Global Victim-Perpetrator Synthetic Explainer Presentation](#)'.













personally identifiable values.³⁰ In this context, ‘K’ refers to the number of times each set of attributes that could be used to identify individuals appearing in the dataset. These sets of attributes are known as quasi-identifiers.

A dataset achieves k-anonymity when there are at least ‘K-1’ other records that share the same quasi-identifiers. At this point, the data is no longer unique to specific individuals, making reidentification significantly more difficult for prospective attackers. This is achieved by clustering records with similar values of quasi-identifiers. For instance, consider a dataset containing the attribute; age, gender and postcode of a certain group of individuals, which would constitute a quasi-identifier.³¹ Achieving k-anonymity requires clustering and generalising records with similar values for their quasi-identifiers. For instance, this could be achieved by replacing the integer values for age (for example, Age=26) with a range (Age=20-39), and replacing postcodes (N1 9AG) with broader geographic areas (‘United Kingdom’). While k-anonymity is a method widely used to anonymise sensitive data, it has certain drawbacks: aggregating data leads to the redaction of outliers, which can result in the loss of valuable data. This loss can potentially lead to instances of misidentification, which in particularly sensitive cases – such as that of identifying instances of human trafficking – can be especially consequential.

In the following case study, the primary actors involved in the implementation of this example of differential private synthetic data were internal to Microsoft – within its Special Research Projects group. This presented a useful case to analyse, as development and implementation of the PET was largely centralised.

³⁰ Sweeney, L. (2002) ‘[k-anonymity: A model for protecting privacy](#)’. *International journal of uncertainty, fuzziness and knowledge-based systems*, 10(05), pp.557-570.

³¹ Ibid.

	Scope	Blueprint	Roadmap	Deploy	Operate	Evaluate	Iterate/Retire
LEADERS	 [IOM] identifies the need for safe, privacy-preserving sharing of human trafficking data [MSR-SP] matched with IOM to research solutions	[MSR-SP] researches potential options for safe data anonymisation and dissemination	[MSR-SP] identifies differential privacy as a viable option and sets the task of developing the synthesiser	[MSR-SP] tests synthetic dataset against aggregate data and develops web application for data aggregation/synthesis	[MSR-SP] and [IOM] launch dataset, dashboard, and web application data aggregator/synthesiser		[IOM] periodically update the dataset with new data, with technical support from [MSR-SP] as the synthesiser is being updated
TECHNICAL PROVIDERS			 [MAG] adapts existing differential privacy method to human trafficking data problem				
EXTERNAL SUPPORT	 Input provided by [TAT] network organisations on the requirements for human trafficking data sharing		 [HTCDS] standard consulted during the development process	 [OpenDP] provides support to make synthetic data and synthesiser open source on the Smart Noise Library	 [MSR-SP] support the development of data dashboard and visualisation of trafficking data on the CTDC website.		
ENGAGEMENT PARTNERS	  [BSR] organises the [TAT] Accelerator to match technology companies with anti-trafficking organisations					 [IOM] departments provide feedback to improve the user experience of data dashboard on the CTDC website.	
CONTRIBUTORS AND RECIPIENTS		 [CTDC], an IOM initiative, pools human trafficking data from multiple organisations: OM, Polaris, RecollectiV, A21, (OTSH)			 Anti-trafficking organisations and external partners have full access to the synthetic data and open-source synthesiser.		 [CTDC] periodically publish updated human trafficking synthetic data.

[IOM] - International Organization for Migration
 [CTDC] - Countertrafficking Data Collaborative
 [MSR-SP] - Microsoft Special Projects
 [MAG] - Microsoft Algorithms Group
 [Open DP] -
 [BSR] - Business for Social Research
 [OSTH] - Portuguese Observatory on Trafficking in Human Beings
 [HTCDS] - The Human Trafficking Case Data Standard (HTCDS)
 [TAT] - Tech Against Trafficking
 [ES] - External Stakeholders

Implementing the PET

In 2019, Tech Against Trafficking (TAT), in collaboration with Business for Social Responsibility (BSR),³² launched the TAT Accelerator Program,³³ an initiative that pairs anti-trafficking organisations with technology companies to develop solutions to humanitarian problems. CTDC participated in the accelerator and was matched with the Microsoft Research Special Projects (MSR-SP)³⁴ group to tackle the data anonymisation issue.

Researchers within the group quickly settled on an approach of creating synthetic data that enforced k-anonymity across all attributes of the dataset, while accurately preserving attribute counts. Differential privacy was not initially considered at the time, as most implementations were focused on creating differential privacy query mechanisms that added calibrated noise to query results in ways that confer privacy. This poses a problem when these fabricated combinations distort the real-world state of affairs and mislead subsequent decision-making processes, policy formulation or resource distribution, to the detriment of the human trafficking victims (for example, by directing law enforcement resources towards non-existent trafficking routes). Furthermore, this approach to differential privacy only permits queries until a prespecified privacy budget is reached, after which the utility of the query mechanism is zero.

To address this challenge, MSR-SP collaborated with Microsoft's Algorithms and Data Sciences group,³⁵ emphasising that the practice of fabricating unobserved combinations, though intended to preserve privacy, could hinder the understanding of real-world exploitation patterns. The solution was to adapt a differential privacy method developed to extract accurate counts of n-gram word combinations from a corpus of private text data to generate differentially private marginals.³⁶ In this context, 'marginals' are the counts of all possible short combinations of attributes within a dataset that can be made while preserving individual privacy, and short combinations of attributes are subsets of attributes within that dataset that are used to generate privacy-preserving synthetic data. Once the marginals are generated, the resulting aggregate counts can be used to derive synthetic records that retain differential privacy.

The new synthesiser controls the degree to which the synthesis of spurious attribute combinations is permitted and supplements the synthetic datasets with 'real'

³² Business for Social Responsibility (2024), '[Sustainable Business Network and Consultancy](#)'.

³³ Tech Against Trafficking (2020), '[About the Accelerator Program](#)'.

³⁴ Microsoft (2024), '[Microsoft Research Special Projects](#)'.

³⁵ Microsoft (2024), '[Algorithms and Data Sciences](#)'.

³⁶ Kim, K., Gopi, S., Kulkarni, J. and Yekhanin, S., (2021) "Differentially private n-gram extraction' Advances in neural information processing systems', 34, pp.5102-5111.

aggregate data. The aggregate data thereby supports both the evaluation of synthetic data quality and retrieval of accurate counts for official reporting. The solution enabled the IOM to use synthetic human-trafficking data for interactive exploration and machine learning, aggregate data for official reporting, as well as ensure differential privacy guarantees that safeguard privacy even across multiple overlapping data releases.³⁷

Between 2021 and 2024, the collaboration between IOM and MSR-SP produced several publicly available outputs:

- In 2021, the CTDC published the first Global Synthetic Dataset with k-anonymity, along with an interactive map providing visualisations of case data representing more than 156,000 victims and survivors of trafficking, as identified by IOM and partners across 189 countries and territories from 2002 to 2021.³⁸
 - In 2024, CTDC published an updated Global Dataset with differential privacy, representing more than 206,000 victims and survivors identified by IOM and partners across 190 countries and territories from 2002 to 2022.
- In 2022, the CTDC published the Global Victim-Perpetrator Synthetic Dataset with differential privacy, along with data dashboards showing the relationships between victims and perpetrators. The dataset includes IOM case data from more than 17,000 victims and survivors of trafficking identified across 123 countries and territories, and their accounts of more than 37,000 perpetrators who facilitated the trafficking process from 2005 to 2022.³⁹
- MSR-SP and [OpenDP](#) collaborated to produce the open-source [SmartNoise Library](#) on GitHub to ensure that the synthesiser and the dataset were publicly available.
- MSR-SP created a public utility web application that allows users to aggregate and synthesise data locally in a web browser so no data leaves the user's computer.

In 2024, MSR-SP and IOM updated the synthetic and aggregate datasets⁴⁰ for continued use by anti-trafficking organisations, governments, and external stakeholders.

In this example, collaboration between relatively few actors was required for the implementation of this PET. This serves to demonstrate that certain PETs – such as

³⁷ Microsoft (2022), '[IOM and Microsoft release first-ever differentially private synthetic dataset to counter human trafficking](#)'.

³⁸ Counter-Trafficking Data Collaborative (2024), '[Global Synthetic Dataset](#)'.

³⁹ Counter-Trafficking Data Collaborative (2024), '[The Global Victim-Perpetrator Synthetic Dataset](#)'.

⁴⁰ International Organization for Migration (IOM) (2024), '[IOM releases the Global Synthetic Dataset](#)'.

differentially private synthetic data – can be deployed in some cases without the need for extensive collaboration. However, it must be acknowledged in this specific example that the data on which the synthetic set was created had already been collated, thus minimising the need for extensive data-sharing agreements between numerous actors. Furthermore, much of the expertise required to deploy the PET in this example existed within the different teams at Microsoft and, to a large extent, within the MSR-SP team, thus limiting the dependencies on partners. As such, this example demonstrates that some PETs, such as synthetic data, can be developed and deployed with a high degree of autonomy, thus minimising the need for an extensive ecosystem of actors.

Enabling granular health data research with the SAIL Databank trusted research environment (TRE)

Background

A trusted research environment (TRE) is a privacy-enhancing infrastructure designed for the purpose of providing researchers with access to sensitive data whilst simultaneously ensuring no privacy loss. These infrastructures also serve to promote data protection and limit access to sensitive data by malicious or unintended actors. TREs such as the one in the following case study are often designed around the ‘Five Safes’ framework in the UK, in which TRE host organisations steward data to certify it is as anonymous and de-identified as possible, while also ensuring that researchers who want to use the data are compliant with rigorous authorisation protocols. Once authorised, researchers can use only hardware administered by the TRE hosts to access the anonymised data and analyse it, completing their research and exporting the results, but not the raw data. This export is audited by TRE administrators to certify that no data leaves the TRE, ensuring impactful research can be conducted without privacy loss.

SAIL (Secure Anonymised Information Linkage) Databank⁴¹ is an example of such a TRE. SAIL Databank provides a rich source of health and social data on individuals in Wales. It provides researchers access to sensitive information whilst simultaneously ensuring no privacy loss malicious use by the means of rigorous protocols in both data ingress and user authorisation.

⁴¹ SAIL Databank (2021), ‘[About us](#)’.

	Scope	Blueprint	Roadmap	Deploy	Operate	Evaluate		Iterate/Retire
LEADERS	[HIRU] scopes possible strategies to realise electronic, person-based data sharing for health research.	[HIRU] identifies 7 objectives a TRE must fulfil to be in accordance with desired data security and privacy.	[HIRU] decide timeline for development of infrastructure for a small scale pilot.	[HIRU] lead a pilot and fulfil their role within it. They gather feedback and iterate on the external audit by setting up the IGRP.	[HIRU] perform their role in stewarding and governance of the architecture.	[HIRU] evaluate successes and determine points for iteration and improvement.		[HIRU] govern expansion of SAIL to a remote-access model and facilitate the subsequent spin-out of SeRP.
TECHNICAL PROVIDERS		[HSW] define technical infrastructure and conceptualise processes.	[HSW] design and develop algorithms for data ingress, architecture for data storage, and means for data access.	[HSW] test their algorithms and work in the pilot in their role as a TTP.	[HSW] perform matching, encryption, deidentification, and standardisation, while supporting technical infrastructure.	[HSW] maintain their processes and evaluate them as the technological landscape changes over time.		Matching, encryption, de-identification, and standardisation algorithms are subject of [HSW]'s ongoing research.
EXTERNAL SUPPORT		Standards and regulations set up by [NHS Information Governance] and UK legislation consulted for blueprint.		[RSM Bentley-Jennison] perform a bespoke audit to ensure compliance to [NHS Information Governance].	[IGRP] review researchers' authorisation applications for access to SAIL Databank.	Certifications of adherence to [ISO] and [UK Statistics Authority] frameworks.		
ENGAGEMENT PARTNERS								
CONTRIBUTORS AND RECIPIENTS		[DPOs] provide their input on data sharing requirements, expectations, and concerns.		[DPOs] like the NHS and some social service institutions sign initial data sharing agreements.	More [DPOs] agree to share their data, signing bespoke DSAs that give them full control of projects their data is used for.	[DPOs] evaluate their DSAs and what their data is being used for.		

[HIRU] - Health Information Research Unit
 [HSW] - Health Solutions Wales
 [SeRP] - Secure eResearch Platform
 [NHS Information Governance] - provides framework for health data sharing
 RSM Bentley-Jennison] - audit services firm, now insolvent
 [IGRP] - the Information Governance Review Panel
 [ISO] and [UK Statistics Authority] - certifying organisations that endorse secure data practices
 [DPOs] - Data Providing Organisations

Implementing the PET

The overall purpose of the databank is to host de-identified information on both individuals' demographics and their clinical data. This allows researchers to find groundbreaking insights into the way someone's lifestyle impacts their health – for example, a researcher could compare the individual's residential location to their history of respiratory illness to determine whether the two could be associated with one another, and in what specific circumstances.⁴² Having this resource therefore enables innovative medical science for the public benefit. The SAIL data could however be used by malicious actors for re-identification attacks,⁴³ where individuals are identified from their demographics so that their clinical data can be used for blackmail, predatory advertisement, or raised insurance premiums. SAIL's architecture mitigates this possibility by taking stringent measures to de-identify the data while also being extremely rigorous when sanctioning access to the data for researchers.^{44 45}

The primary actors in the development of SAIL Databank were the Health Informatics Research Unit (HIRU, the parent organisation that sits as part of Swansea University), Health and Care Research Wales, Digital Health and Care Wales, and the Secure eResearch Platform (SeRP, a team once part of SAIL that is now a separate, accredited entity).

SAIL Databank was piloted and launched in 2007 as a physical TRE. Remote access via SAIL Gateway came in 2011, following collaboration with SeRP and facilitated by prior research authored by OCLC.⁴⁶ It remains operational.

In 2006, the HIRU was set up to develop ideas for health data sharing and to contribute as the Welsh component of the UK's health research initiative. The HIRU began surveying solutions that would enable clinical researchers to access individualised data without privacy loss, which led them to TRE architectures.

Once HIRU had scoped TRE architectures, its researchers identified the criteria that a bespoke TRE for health data sharing should comply with. These criteria were based on the NHS Information Governance (IG) framework and concerned how data entered, moved within, and left a TRE.

⁴² Ford, DV et al. (2009), '[The SAIL Databank: building a national architecture for e-health research and evaluation](#)'.

⁴³ Understanding Patients Data (2024), '[What are the risks around patient data?](#)'.

⁴⁴ SAIL Databank (2024), '[Privacy by Design](#)'.

⁴⁵ SAIL Databank (2024), '[Apply to work with the data](#)'.

⁴⁶ SAIL Databank (2021), '[Our History](#)'.

During development, Health Solutions Wales (HSW) built the technical means of data extraction, transportation, storage and analysis to conform with these criteria. HSW developed algorithms from the ground up to preserve security and privacy when datasets enter the TRE and are linked with other data.

Using data provided by NHS Wales and a few social services institutions, it ran a small-scale pilot to test the architecture and refine the technology, ingress protocol, and authorisation specifications. Following the pilot the Independent Review Process was established to enhance accountability and trust.

During and after the pilot, RSM Bentley-Jennison ran a bespoke compliance audit, designed from the Control Objectives for Information and Related Technologies (COBIT)⁴⁷ and following recommendations in the Health Insurance Portability and Accountability Act (HIPAA).⁴⁸ The audit included stakeholder interviews, control checks and documentation reviews. It resulted in a formal report of findings and recommendations for HIRU designed to improve SAIL Databank for large-scale use. Before deployment, a wide consultation was carried out with government, regulatory and professional agencies to scope further requirements surrounding suitability and ethics.

HIRU is responsible for all governance and stewardship roles, including but not limited to:

- Building partnerships with data providers
- Keeping the physical means of storage for the data
- Curating ‘data views’ for researchers (subsets of the data stored in the databank relevant for their use)
- Governing the researcher authorisation protocol
- Auditing use of the TRE and exports of results, and
- Coordinating initiatives (such as ‘One Wales’⁴⁹ during the Covid-19 pandemic).

Meanwhile, HSW works as a trusted third party, employing its algorithms in data ingress to anonymise and standardise data before it enters the Databank. The Independent Governance Review Panel⁵⁰ serves as an external party in the authorisation process, with members from academia, health services and the general public providing their opinions on applications by researchers to use SAIL Databank. All processes are audited internally.

⁴⁷ Brook, C. (2020), ‘[What is COBIT?](#)’.

⁴⁸ Centers for Disease Control and Prevention (2022), ‘[Health Insurance Portability and Accountability Act of 1996 \(HIPAA\)](#)’.

⁴⁹ Swansea University (2023), ‘[One Wales - Population Data Science](#)’.

⁵⁰ SAIL Databank (2021), ‘[Approvals & Public Engagement](#)’.

The largest iteration of the architecture came in 2011, when the Databank moved to a remote-access model. Previously, approved researchers could only access data on-site in a physical safe room in Swansea. The remote-access model provides researchers with a USB dongle and unique password that allows them to access a virtual environment hosted in Swansea via their own device, removing a significant barrier to access. This model was called SAIL Gateway, developed by part of HIRU (now SeRP, an organisation that can help with implementation of remote-access TRE models across the world).⁵¹

As the implementation timeline for the SAIL example shows, more actors have been involved in the implementation of this PET than in the previous example of the differentially private synthetic dataset. While much of the development of the TRE was undertaken by HIRU, relationships were required with NHS Wales to obtain the data that would be held within SAIL. This has since been expanded to further collaborations through additional data donations from more actors. This has been achieved through concerted efforts over time, through an iterative process of demonstrating utility of the service and the SAIL team's ability to uphold data security guarantees.

Using secure multi-party computation to measure wage-gap inequality in Greater Boston

Background

MPC is a cryptographic protocol that enables multiple stakeholders to carry out joint computations without revealing their individual data inputs. This protocol relies on employing a secret sharing algorithm to obscure and divide sensitive information, like a company's payroll records, into shares distributed among participants. Once the data is obfuscated and distributed, individual shares remain hidden, preventing any insights from being gleaned unless trusted parties cooperate to combine the data. MPC provided a solution for the safe collection and analysis of combined sensitive payroll data without exposing individual businesses to potentially costly legal action.

⁵¹ Swansea University (2020), '[About SeRP](#)'.

Implementing the PET

In 2013, Boston City Council sought to conduct a pay equity study to benchmark and address racial and gender pay gaps in the Greater Boston area. It commissioned the Boston Women's Workforce Council (BWWC) to lead the project. Though supportive of the initiative, prospective data stewards and local businesses were apprehensive of the legal and commercial risks associated with sharing sensitive payroll business data.

Details of BWWC's interest in performing this kind of challenging analysis permeated throughout the Boston area, which then led to individuals from within Boston University (BU), proposing the use of secure multi-party computation (MPC) to address the problem. This cryptographic protocol allows stakeholders to analyse data collectively without disclosing individual inputs, thus facilitating the safe collection and analysis of sensitive payroll data while addressing businesses' concerns about privacy and legal risks.⁵²

In this set-up, BU serves as the service provider and is the sole entity mandated to maintain an online presence throughout an MPC session, which may span a predetermined duration (such as two weeks).

The resulting pay equity report created through the use of MPC, which has been published biannually since 2016, is more accurate than other US wage gap measurements, which are based on self-reported census data. The report is publicly available,⁵³ and its principal beneficiaries are:

- Businesses from the Greater Boston area, which contribute their payroll data and use the report to benchmark their wage gap metrics against those of other businesses and to inform recruitment and remuneration decisions
- The Boston Mayor's Office, which commissioned the study to inform its public pay equity policies
- Researchers, who use the report as a source of accurate payroll measurement.

The use of MPC in the case of the Greater Boston Area materialised following the initial impetus of the Boston City Council, which in 2013 sought to establish a clearer picture of racial and gender-based wage gaps in the area than was possible at the time due to a lack of granular data. Until this point, the data used was uneven, given organisations' reluctance to share sensitive information for many reasons. Coupled with this, data

⁵² For a comprehensive overview, including an interactive walkthrough of how this PET works in practice, see: Himmelsbach, E. et al. (2024), '[PETs in Practice](#)'.

⁵³ Boston Women's Workforce Council (2023), '[Gender and racial wage gaps in Boston by the Numbers](#)'.

that was mandatorily provided existed at the federal level for census purposes in the US. The federal-level data is subject to aggregation that – while providing a degree of privacy – compromises its utility for establishing a more precise understanding of specific challenges.

As a result, BWWC were tasked with leading on the initiative to establish a clearer picture of the extent of these wage-gap disparities in the Greater Boston area. Initially, discussions took place between BWWC and individuals from Simmons University regarding working together to tackle this problem, however this did not materialise after it was deemed unfeasible.

In parallel, research had been taking place at the Rafik B. Hariri Institute for Computing and Computational Science & Engineering at Boston University⁵⁴ on a variety of privacy-enhancing technologies. Chief among the proponents of the potentially transformative impact of these technologies was the inaugural Associate Provost of the Hariri Institute, Professor Azer Bestavros,⁵⁵ who had been party to discussions on the challenge BWWC was facing. Based on these discussions, a partnership was formed between BWWC as the client and BU, which performed the role of primary technical provider to develop and operate the architecture required to carry out the MPC protocol.







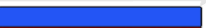
It is worth noting at this point that several other factors contributed to BU's decision to actively participate in this collaboration. This includes some historical efforts trialled in the Boston area, including by the City of Boston Department of Innovation and Technology, which had shown previous interest in the early 2010s in utilising data and analytics in combination with privacy-enhancing technologies to address administrative challenges.⁵⁶ While this did not lead to the trialling of PETs in practice, it did draw their potential applications to the attention of Professor Bestavros, who would later go on to propose this to BWWC as a possible solution. Additionally, had a more straightforward solution, such as using a trusted third party to steward the data – or even BWWC serving as one itself – been sufficient, it is unlikely that MPC would have been tried in this instance.⁵⁷

⁵⁴ Boston University (n.d.), '[About the Hariri Institute](#)'.

⁵⁵ Boston University (n.d.), '[Azer Bestavros](#)'.

⁵⁶ City of Boston (n.d.) '[Department of Innovation and Technology \(DoIT\)](#)'.

⁵⁷ Trusted CI (2020), '[Transition to Practice success story: Boston University - Secure multiparty computation and the Boston Women's Workforce Council](#)'.

	Scope	Blueprint	Roadmap	Deploy	Operate	Evaluate		Iterate/Retire
LEADERS	 [BC], interested in a pay equity study, commission [BWWC], who begin planning and outreach	[BWWC] collects their requirements for usability and their preferences regarding output	[Azer Bestavros] critical to promoting MPC as the proposed architectural solution for problem	[BWWC] schedule and initiate the MPC session		[BWWC] ensure their staff are trained ready for deployment as either data validators or support staff		[BWWC] note feedback after each operation and update [BC] decides whether to continue support for initiative depending on cost benefit analysis
TECHNICAL PROVIDERS	 [BWWC]'s difficulties with getting firms to share their data leads them to [BU], who have interest in demonstrating the functionality of MPC in practice	[BU] identifies constraints in the planned operation and their preference for deployment mechanisms	[BU] collect all requirements, preferences, constraints, and risks.	[BU] deploy the infrastructure and do last benchmarking tests	[BU] adapt their proposed architecture to use common infrastructure and change from peer-to-peer to client-server	 [BU] conduct small scale tests in-house, then with [BWWC], taking feedback and iterating		 [BU] team publishes academic papers on experience of developing solution and works with [BWWC], realising the improvements wanted in technical structure
EXTERNAL SUPPORT	 [NSF] transition to practice grant enabled the trialling of SMPC in a practical context [BU] colleagues from elsewhere in the university contributed expertise			 [AWS] provide increased server capacity when necessary	[AWS] provides back-end capabilities and a guarantee of security of data transit			
ENGAGEMENT PARTNERS		 [BWWC] serves as the engagement partner with participating organisations from Greater Boston Area		 [BWWC] outreach to their network and organise data contributors		 Use case featured in reports and explainers such as that from [RS] and [UN].		
CONTRIBUTORS AND RECIPIENTS	 [POs] express their hesitancy to make their payroll data open, citing legal and commercial risks	[POs] are asked what can be done to make a trustworthy environment for them		 [POs] are invited to submit their data	[POs] continue to provide data for future iterations of BWWC study, held every two years (some organisations drop out, while others join)	 [POs] volunteer for focus groups to learn more about the program		 [BWWC] gains greater insight into scale of pay-gaps in Greater Boston Area. [POs] learn their position relative to others based on benchmark

[BC] - Boston Council
 [BWWC] - Boston Women's Workforce Council
 [BU] - Boston University - solution developer
 [AZ] - Azer Bestavros - Associate Provost for the Faculty of Computing and Data Sciences at Boston University
 [NSF] - National Science Foundation - grant awarding organisation
 [AWS] - Amazon Web Services
 [RS] - Royal Society - publishes report highlighting work of BWWC and BU
 [UN] - United Nations PETs Lab - produces report publicising BWWC-BU example
 [POs] - Participating organisations from Greater Boston area

As noted in the implementation timeline, the Transition to Practice grant⁵⁸ provided by the National Science Foundation was pivotal in providing the financial support needed to apply MPC in the specific case of the BWWC/BU collaboration.⁵⁹

During the blueprinting phase, BWWC and BU worked collaboratively to define the questions and determine how this would be addressed through the use of MPC. Concurrently, BWWC served as the engagement partner in this collaboration, seeking to convince and onboard organisations in the Greater Boston area to contribute their data as part of the wage gap analysis initiative. This involved substantial effort on the part of BWWC, through running workshops and providing further resources that explained to prospective participating organisations how the protocol would work and how data privacy would be ensured throughout the process.

Once the specifics of the roadmap were settled on, primarily by BWWC and BU, the protocol was deployed through an initial session in 2015, with subsequent iterations in 2016 and 2017. Aside from BWWC and BU, Amazon Web Services were a critical actor – albeit one with limited scope to influence the development and deployment – as the provider of server capacity required to run the MPC protocol.

Continued engagement between BWWC and the participating organisations from within the Greater Boston area was necessary during sessions when the protocol was running in order to maintain buy-in. While some organisations inevitably dropped out during subsequent rounds, efforts were made to bring other organisations on board to ensure there were enough to run the MPC protocol. This was essential, given that – due to the nature of MPC – the greater the number of organisations that participate in a session, the lower the likelihood that participants can collude and undermine the process in any way.

On concluding each session and running the analysis, BWWC and BU worked collaboratively to measure the efficacy of the protocol and look for opportunities for improvement in future rounds. Furthermore, the reports created by BWWC served to demonstrate the utility of using MPC as a means of overcoming a challenge in which sensitive data was required. This helped BWWC to make the case to Boston City Council for further iterations of the wage gap study. The success of this initiative has since been promoted by organisations that have conducted independent

⁵⁸ Trusted CI (2020), '[Introduction to the Trusted CI Cybersecurity Technology Transition to Practice \(TTP\) program](#)'.

⁵⁹ Trusted CI (2020), '[Transition to Practice success story: Boston University - Secure multiparty computation and the Boston Women's Workforce Council](#)'.

research on the application of PETs, such as The Royal Society⁶⁰ and the United Nations PETs Lab.⁶¹ This has helped raise the profile of the initiative and the various actors involved in the process.

Following the successful implementation of MPC in the Boston case, this process has developed a somewhat cyclical pattern, where – post-evaluation – BWWC and Boston Council discuss the next iterations of the report. This has taken place every two years since, apart from in 2018. Due to the way in which the MPC protocol functions, it is essential to determine the research questions ahead of beginning a session. These are deliberated in advance by BWWC and Boston City Council, taking into account the most pressing issues facing the Greater Boston community at the time.

Cross-case analysis

In our analysis, we identified five types of actors critical to the success of efforts towards implementing PETs. The importance of these actors will depend on the context of the implementation, as well as the PET being implemented.

Towards a PET ecosystem typology of actors

We included the types of actors on the left-hand side of the case study implementation timelines. There are variations between the three cases, including distribution of actors within the timeline, the number of actors included, and the frequency and locations in which some of the actors appear. This is reflective of the variations in the circumstances under which the different cases were implemented, and illustrates the influence of taking a more centralised, as opposed to a more distributed and collegiate, approach. Similarly, these variations reflect the extent to which capabilities, resources and autonomy are available to those implementing PETs.

⁶⁰ Royal Society (2023), [‘From privacy to partnership’](#).

⁶¹ United Nations (2023), [‘United Nations Guide on Privacy-Enhancing Technologies for Official Statistics’](#).

We identified the following types of actors:

- **Leaders** – the key drivers of the development of the PET, whether in instigating the project or determining its direction - particularly at critical junctures.
- **Technical providers** – both internal and external providers who played direct roles in the development and deployment of the PET. It can be useful to consider where there have been collaborations between different teams (again, either internally or externally).
- **External support** – other actors who contributed to the development of the PET in an indirect manner. These could be service providers whose services were procured, but who had no direct influence on the development of the PET itself, or how it was used.
- **Engagement partners** – any stakeholders who played a supporting role in promoting or communicating the work done, and in the PET more broadly. This will likely be specific to the purpose of the PET and the communities that engage with it
- **Contributors and recipients** – includes those who have provided inputs (such as data) at the ingress stage that have been fundamental to the functioning of the PET. A relevant consideration for these actors is whether data was provided by individual participants, or by organisations that have stewarded data on their behalf. If the latter, it can be worth determining the operating terms for stewardship of that data.

As can be seen through the implementation timelines, they vary in terms of the volume of entries and the distribution of roles. This is reflective of the differences between the types of PETs that we looked into, as well as the contexts in which they were implemented and the organisations that took part in these efforts. As an example, the relatively centralised nature of the processes involved in the IOM/Microsoft and SAIL Databank cases resulted in noticeably less activity in the engagement partners portion of the implementation timeline. This reflects the fact that in these instances, there were pre-existing relationships between prospective data contributors and the coordinating organisation, which then needed to be able to prove that the PET would work as intended and to enable the analysis intended without increased risk. In the BWWC/BU case, however, the contributing organisations were much more diffuse and required more extensive engagement and reassurances in order to build trust in the proposed solution.

Development of a PETs implementation framework

By modifying the CFIR to the specifics of privacy-enhancing technologies, we have been able to begin documenting, understanding and modifying how PETs have been successfully implemented in the past. However, just as the CFIR can be used for evaluating existing implementations (see [above](#)), our research holds the potential to help organisations design successful PET implementations as well.

To that end, we have developed ‘Version 0’ of a PET implementation framework, which introduces recommended activities for each of the seven implementation phases introduced above. The framework draws on the structures contained within the ‘implementation process domain’ of the CFIR, on our own findings from interviews with those involved in the implementation case studies detailed above, and in previous work at the ODI.⁶² As with previous ODI resources and guidance, this framework contains information that can be of use to all types of organisations. However, this will likely be more relevant to smaller organisations with fewer resources, or less familiarity with the process of selecting, developing and implementing a PET. Through our research, we have gathered insights from those who have successfully implemented a PET and have incorporated the phases and related activities they undertook – including those they feel they should have taken but did not at the time – into the framework. The framework is still at an early stage and therefore would benefit greatly from community testing. Through constructing the implementation framework, we also identified areas for further research, which we have included in the [annex](#).

⁶² This includes previous research, including but not limited to: Keller, J.R. (2021), ‘[How do data institutions facilitate safe access to sensitive data?](#)’ and ODI (2019) and ‘[Data trusts: lessons from three pilots](#)’.

The seven phases of the implementation framework and recommended activities

The 'Version 0' PETs implementation framework includes seven phases:

- Scoping
- Blueprinting
- Roadmapping
- Deploying
- Operating
- Evaluating
- Iterating/ Retiring

Each phase includes recommended activities, drawn from our research, the CFIR and previous ODI work. Each implementation is different, so you might not need to work through every phase or activity. But from experience of designing and building data sharing initiatives, we have found it useful to at least discuss each phase and activity before deciding whether to skip it. This way, you can determine if it is either not applicable or that you have the correct knowledge or documentation to progress to the next phase.

Scope

Scoping a PET implementation should begin by aligning around an issue, challenge or need related to accessing data or insights, researching previous efforts to address similar challenges, and identifying existing organisations and infrastructure that could be adapted for, or included in, the implementation. This Scoping phase also involves working to build momentum and commitments that will be necessary for the Blueprinting and Roadmapping phases.

The recommended activities in the Scoping phase are:

- Engage with stakeholders to assess the problem/ opportunity
- Map and analyse the data ecosystem
- Define and prioritise use cases

Engage with stakeholders to assess the problem/ opportunity

Conduct outreach and research to better understand what the ecosystem needs and/or what the PET implementation is trying to achieve. Through engaging with potential users and stakeholders, work to understand what each group within the ecosystem needs - for example, the data or insights they need access to, who currently holds it, and what they want to do with it.

These engagement activities are crucial to understanding the needs and concerns of stakeholders across the ecosystem of a potential PETs implementation. Early efforts at engagement can save time by spotlighting potential points of disagreement or conflicting needs earlier in the implementation timeline. For instance, in the case of [Boston City Council and BWWC](#) mentioned above, after engaging with their ecosystem, they found that local businesses and prospective data stewards were generally supportive of efforts to study wage gap disparities, but were concerned about the potential legal and commercial risks associated with sharing payroll data. This helped Boston City Council and BWWC scope the work needed going forward in order to address these concerns while fulfilling the potential of the use case.

One way of bringing an ecosystem together and engaging with stakeholders is through accelerator programmes, challenges and hackathons.⁶³ For instance, as discussed above, in an effort to generate engagement between anti-trafficking organisations and technology companies, in 2019 Tech Against Trafficking (TAT), in collaboration with Business Social Responsibility (BSR), launched the TAT accelerator program.⁶⁴ This helped to pair CTDC with the Microsoft Research Special Projects (MSR-SP) team to tackle challenges around data anonymisation.

In the IOM/ Microsoft example above, CTDC hosted a kick-off session for member organisations, outlining the problem with attendees from all member organisations and TAT, including speakers from IOM, law enforcement, and victims of trafficking.

Map and analyse the data ecosystem

Using some combination of data ecosystem mapping,⁶⁵ data landscape review⁶⁶ and PESTLE analysis, strive to understand the environment and context within which the potential PET implementation will operate. Data ecosystem mapping in particular can be useful for documenting and understanding the key actors and data infrastructure within the data ecosystem, and how value and services flow across it.

⁶³ Recent examples include the joint [US-UK PETs Prize Challenges](#).

⁶⁴ Tech Against Trafficking (2020), '[About the Accelerator Program](#)'.

⁶⁵ D'Addario, J. (2022), '[Mapping data ecosystems: methodology](#)'.

⁶⁶ ODI (2023), '[Data Landscape Playbook](#)'.

Define and prioritise use cases

Once you understand the needs of each group within the ecosystem, create clear use cases for the available data/ insights, including rough sketches of the data ecosystem map of each. The Value of Data Canvas⁶⁷ can help to assess the potential social, economic and environmental value of the implementation and prioritise the use cases capable of incentivising stakeholders to proceed - those that members of that ecosystem are willing and able to pursue. This often comes down to how well the use case balances value and risk for the organisations involved.

Blueprint

When developing a blueprint for a PET implementation, the goal is to agree its purpose and begin to define how the implementation will function across the four layers of a PET implementation: the legal foundation; commercial/value model; technical infrastructure; and governance processes. This process should be as collaborative as possible, with the emphasis on engaging with stakeholders to co-design the implementation rather than designing it internally and then seeking validation or approval after the fact. During this phase, it is also important to challenge whether a PET is viable and is indeed the right way to meet the needs of the use case(s) identified during the previous phase.

The recommended activities in the Blueprinting phase are:

- Explore a range of stewardship options
- Agree on the functions that different actors will need to perform
- Co-design the four layers of the PET implementation
- Engage with stakeholders to iterate and finalise the design
- Secure funding for the further phases, if applicable

Explore a range of approaches

At this point in the process, it is worth pausing to explore whether the prioritised use cases are best served by a PET implementation or by another approach to enabling access to data, such as a data institution, API or open data portal. Using the ODI's data access map⁶⁸ and our collection of real-world examples of data sharing approaches can clarify your options and help determine the right approach:

⁶⁷ ODI Value of Data Canvas in Stiglich, L., Sharp, M. and Keller, J.R. (2023), '[Understanding the social and economic value of sharing data](#)'.

⁶⁸ Keller, J.R. (2019), '[Mapping the wide world of data sharing](#)'.

- PETs, such as:
 - Secure Multi-party Computation
 - Homomorphic Computation
 - Federated Learning/ Federated Analytics
 - Trusted Researcher Environment(s)
 - Differential privacy
 - Synthetic data
- Data governance mechanisms and infrastructures, such as:
 - Data institution
 - Open data platform
 - Technical help with APIs
 - Data dashboards
 - Co-designed data standards
 - Data marketplace

You can always decide based on further investigation to pursue a different approach, but it is important during this phase to ensure that your choice of approach is being led by needs of the use case(s), rather than by an interest in the technology itself or hype around the approach.

For instance, the [HIRU](#) began by surveying a range of solutions that might be capable of enabling clinical researchers to access individualised data without privacy loss. That survey ultimately led them to TRE architectures. Even then, once they had identified TREs as a solution to many privacy concerns, they had to conduct further research to ensure that the TRE approach would work for health data specifically.

Agree on the functions that different actors will need to perform

Making a PET implementation successful and sustainable usually requires many different organisations to perform a variety of functions within their ecosystem. These functions can be consolidated in a few organisations or distributed across the ecosystem. Some of the functions often required for a successful PET implementation include:

- Convening, prioritising, providing leadership, setting the agenda
- Empowering people to play a role in stewarding data
- Developing and/or managing identifiers, standards and other data infrastructure
- Collecting, linking and importing (could include hardware to collect and store and transmit, such as APIs)
- Managing, protecting, curating, validating datasets, assessing quality (metadata, enhancement?)
- Providing analytical, querying and search tools
- Developing or training algorithms or models
- Generating insights, conducting analyses, developing visualisations
- Continued monitoring, evaluation and learning

A crucial part of designing a PET implementation is identifying and documenting which functions are necessary and which actors are capable and willing to perform them. Not all of these functions will need to be performed for each implementation, and the right balance will depend on the specifics of the ecosystem, the capabilities of the actors involved, and how value will be distributed across the ecosystem.

In the case of HIRU, the design teams identified the need for a trusted third party that could perform technical processes. It recruited Health Solutions Wales to join the initiative.

Co-design the four layers of the PET implementation

Using the ODI's Facilitating Safe Access Framework⁶⁹ as a guide, work with stakeholders to define the four layers of the PET implementation:

1. Define the legal foundation of the implementation and its supporting policies and policy frameworks
2. Define the commercial/ value model by identifying use cases and investigating how the PET implementation could become financially sustainable
3. Define the technical infrastructure necessary to support the implementation and deliver on the aims of the use case(s)

⁶⁹ Keller, J.R. (2021), '[How do data institutions facilitate safe access to sensitive data?](#)'.

4. Define the governance processes necessary to ensure that the implementation is designed, built and operated in an ethical, equitable and responsible way

During this phase, it is not necessary to define the exact set-up of each layer. It is important, however, to identify desired structures, priorities and ‘must-haves’ or ‘red lines’ for different stakeholders. At this point, it should be possible to sketch the range of satisfactory set-ups of an implementation based on the results of this exercise. Further work will be required during successive phases to define the exact set-up.

For instance, once the HIRU had identified TRE architectures as a suitable approach, they made sure that the way that data entered, moved within, and left the TRE was aligned with regulations and governance frameworks around sharing health data, including the NHS Information Governance framework.⁷⁰

Engage with stakeholders to iterate and finalise the design

Present the range of satisfactory set-ups to a wider group of stakeholders and gather feedback on their preferences for different options within each of the four layers of the PET implementation. After gathering feedback, you should be able to narrow down the range of satisfactory set-ups. For instance, if actors in the ecosystem are uncomfortable adopting a specific technology or paying for a proposed service, it will be necessary to adjust the proposed set-up, potentially across all four layers of the PET implementation.

At this point, it is also useful to discuss and agree on short-term and long-term goals for the implementation, including options for how it might develop or scale and additional use cases that might be enabled down the road. able to assess the viability of delivering each prioritised use case

In the case of [SAIL](#), DPOs were consulted to understand their expectations, concerns and requirements, with their feedback incorporated into the final design.

Secure funding for the further phases, if applicable

Some PET implementations will have funding already identified or set aside for the remaining phases of work, whereas others will need to secure funding to move from design to implementation. The documentation produced to this point can be used to form the basis of an operating model or proposal that can be submitted to potential funding bodies to concisely demonstrate the use case(s), the potential impact and the systems and structures required to implement the PET in safe, responsible and sustainable ways.

⁷⁰ NHSX (2021), '[Information Governance Framework for Integrated Health and Care: Shared Care Records](#)'.

As we discovered in the case of the BWWC implementation, additional ‘bridging’ funding, such as that provided by the NSF, was critical to facilitating the trialling of MPC in a practical context. When speaking with those involved in the development of this solution, it quickly became apparent that these types of initiatives and funding mechanisms that help convert ideas into pilots are critical, yet few and far between. This appears to be a critically under-explored component of efforts towards greater adoption and experimentation with PETs, which could accompany statements and votes of confidence.⁷¹

Roadmap

Once you have designed how the PETs implementation will function (the Blueprint phase), you can start designing how to actually go about building, deploying and operating the implementation, including identifying potential challenges and means of addressing them and anticipating how the implementation will operate and iterate in the future.

As mentioned in the introduction to this section, the remaining five phases of the implementation framework contain less detail about the recommended activities. These recommendations are based on our early findings and on previous work at the ODI. We see potential for further research to develop guidance around these activities in order to match the detail provided for the first two phases.

The recommended activities in the Roadmapping phase are:

- Document the plans produced during the previous steps
- Co-design the remaining steps required to build and launch
- Create an engagement plan to test/ iterate the roadmap
- Create a communications plan to raise awareness about the PETs’ implementation

In the case of the [IOM/ Microsoft](#) example, once differential privacy was selected as a suitable approach, the special projects team identified the weaknesses of differential privacy and the potential implementation, consulted The Human Trafficking Case Data Standard,⁷² and developed a plan to address those weaknesses throughout launch and operation of the initiative.

⁷¹ Information Commissioner’s Office (2023), [‘ICO urges organisations to harness the power of data safely by using privacy enhancing technologies’](#).

⁷² International Organization for Migration (IOM) (2024). [‘Human Trafficking Case Data Standards. Toolkit and Guidance \(HTCDS\)’](#).

Regarding the need for engagement, while SAIL Databank was being developed, HIRU conducted a large consultation with government, regulatory and professional agencies to scope further requirements related to the suitability and ethicality of the proposed system before incorporating those requirements in further iterations.

Deploy

When launching a PET implementation, it is necessary to set up and register the implementation in line with its agreed purpose; transparently publish information about its processes, as agreed during the Blueprinting and Roadmapping phases; develop the technology to support data sharing, services and operations; make agreements with data holders for how the data will be provided; and communicate with stakeholders.

The recommended activities in the Deployment phase are:

- Enlist stakeholders and suppliers to build the implementation
- Build and set up the agreed components of the implementation
- Launch pilot (if necessary) and iterate where appropriate
- Enact plans

Before the full launch of SAIL, HIRU ran a small-scale pilot using data provided by NHS Wales and a few social services institutions. This helped to test the architecture, refine the technical implementation, ingress protocols, and authorisation specifications. In an effort to increase transparency and trust, HIRU brought in RSM Bentley Jennison to conduct a compliance audit during and after the pilot. Consisting of stakeholder interviews, control checks and documentation reviews, the audit resulted in a formal report of findings and recommendations for HIRU to act upon to improve SAIL Databank for large-scale use. Following the pilot, the need for independent review was identified, which led to the establishment of the Independent Review Process.⁷³

⁷³ SAIL Databank (2021), '[Approvals & Public Engagement](#)'.

Operate

To operate a PET implementation, it is crucial to ensure that the data, infrastructure and/or models are maintained and will continue to be available. Beyond the more technical aspects of the implementation, it is also crucial to ensure the continued operation of services and reporting systems, fundraising and business development activities, administration and distribution of benefits to data holders or beneficiaries, and any auditing or oversight functions.

The recommended activities in the Operation phase are:

- Perform chosen functions/ roles
- Conduct monitoring, evaluation and learning activities
- Perform audit, due diligence and compliance checks
- Continue engagement and communicate successes and/or case studies

For instance, SAIL continues to run the Independent Governance Review Panel that was established after the pilot. The panel works as an external party in the authorisation process, with members from academia, health services and the general public providing their opinions on applications by researchers to use SAIL Databank. Additionally, SAIL regularly conducts internal audits of its governance processes.

Evaluate

The goal of the Evaluation phase is to examine the operation of the PET implementation and identify areas where it needs to be updated, improved or altered. Conducting regular evaluation phases enables the original design to be iterated on. This is particularly important with PETs, since they are an emerging form of data infrastructure and we are still learning how they can and should operate. As far as we are aware, there is not currently a publicly available consolidated repository of PET implementation evaluations. We therefore propose that research should be carried out to inform how such a repository might look and what metrics should be considered for evaluation.

When evaluating a PET, it is important to assess whether the implementation is achieving the goals of the use case, and its positive and negative impact on stakeholders and the ecosystem. It is also necessary to examine how the implementation is being used in practice to ensure that it is in line with the agreed set-up and purpose. A proper evaluation should

also include engaging with stakeholders to understand where they would like to see improvements or refinements, and a thorough assessment of the finances and business model to gauge the ability of the implementation to remain sustainable. In certain cases, these assessments may need to be conducted by external parties, such as external auditors, regulators and governance panels. Using an adapted version of the CFIR evaluation framework can be a useful point of embarkation for this task, and implementers will want to proactively think through their specific circumstances and define what metrics are appropriate to evaluate the implementation against.

The recommended activities in the Evaluation phase are:

- Survey stakeholders to understand value, impact and potential harms
- Re-evaluate functions/ roles that each actor will play
- Re-evaluate the four layers of the PET implementation
- Determine whether to repeat or iterate

HIRU, for instance, regularly evaluates its successes and impacts, determining points for iteration and improvement as the technological landscape of both research and security changes. Similarly, HSW evaluates its own algorithms as new types of data come in. External evaluation comes in the form of compliance checks, with certifications awarded by ISO and the UK Statistics Authority.

Iterate/ retire

The result of the Evaluation phase could be to iterate based on feedback, or to retire the PET implementation altogether. If the decision is to iterate, the focus should return to the Scoping and Blueprinting phases to rescope and redesign the implementation where necessary. For instance, the decision may be to expand to additional use cases or shift the focus to a different use case, in which case the operating model, infrastructure and supporting roles may need to be reworked. Some iterations will be small, while others will require significant time and effort.

Some PET implementations will eventually decide to close down as the needs of stakeholders and ecosystems change, new technologies emerge or different sources of data become available, new regulations are passed, or because their funding models eventually fail. The goal of the Retirement phase is to minimise any harmful impacts from the PET implementation ceasing operation. In these cases, a timeline for closure will need to be agreed and communicated to stakeholders, and services and agreements will

need to be properly closed down, which may include archiving or deleting datasets or transferring intellectual property.

The recommended activities in the Iteration/ Retirement phase are:

- Communicate the timeline for iteration or closure
- Enact the iteration or closure timeline

These iterations could be simple, as in the case of MSR-SP and IOM's updates to the synthetic and aggregate datasets for continued use by anti-trafficking organisations, governments and external stakeholders. Or they could be more complex, as with SAIL Databank's transition from providing access to data via a physical TRE to remote access via the SAIL Gateway in 2011, a move made in collaboration with SeRP and facilitated by prior research authored by OCLC.

Conclusion

This research began as an effort to develop our understanding of the key ingredients that contribute towards fostering a thriving PETs ecosystem. We set out with the question of what actors, roles and responsibilities had been critical in specific instances of successful adoption of a particular PET.

In our examination of three successful implementations of PETs that span three different applications, we demonstrated that PETs can both be successfully implemented in ecosystems containing relatively few actors, as well as in larger, more complex ecosystems. This observation evidences, in part, many of the contributing factors that we have identified as important for prospective adopters to consider when exploring the possible adoption of a PET. These include factors such as the extent to which the data required for the PET to function is distributed across distrusting actors, or whether this is already held by one actor. We have reflected this and other factors in the implementation timeline template, which includes our corresponding PET ecosystem typology of actors.

Furthermore, through the examination of successful implementations, we have proposed our '[Version 0](#)' PETs implementation framework, to guide discussions within organisations. This resource is intended to be used in combination with the implementation timeline, as a '[Version 0](#)'. However, we appreciate that this will require thorough user testing. We would welcome the opportunity to work with individuals and organisations on this, with the intention of building upon it to create a more comprehensive framework.

Acknowledgements

This report would not have been possible without the help and insights provided by Darren Edge of Microsoft, Lorraine Wong, William Tomlinson of Boston University, and Andrei Lapets of Magnite. We are incredibly grateful for the time they contributed to answering our questions on their experiences of both developing and deploying the use cases included.

We would also like to express our thanks to our colleagues at the ODI who have contributed their time and assistance to the creation of this written report.

Annex

Limitations

PET case study selection

We are aware of the limitations of the approach that we took in selecting the three cases we focused on in this research. In choosing to focus on one example of each type in our research, our findings should be considered as illustrative. That said, we encourage fellow researchers to consider our methodological approach and use this in future evaluations of PET implementations, which can contribute towards a growing evidence base upon which to iteratively test our findings. We have included a suggestion of further research, to build on our findings through the analysis of wider examples of PET implementations.

Future research

Future implementation framework based research

Through developing the typology of actors, we identified that it would be helpful to be able to measure the competencies of the actors to carry out the tasks included in each stage of the implementation framework. Currently, we are unaware of any existing guidance to aid an organisation in evaluating this. This would be a helpful addition to the framework but would require research into existing tools or guidance that could be adapted, or the development of an entirely new tool.

Through the course of creating the [‘Version 0’](#) PETs implementation framework, we identified questions that require further research and activities that could be undertaken to address these.

As a first port of call, we suggest that more PET implementations should be monitored and evaluated into the further phases of their deployment in order to test and validate the implementation framework and provide more bespoke guidance for PETs implementations. At present, this appears to be taking place on an ad-hoc basis at the discretion of primary implementing partners. An opportunity exists to gather data during these stages for additional prospective adopters to learn from. This could be approached through

interviews, surveys and/or ethnographies. Undertaking this research would be particularly informative to improve the evaluate phase in the framework. As a starting point, we propose that research should be conducted to explore whether the CFIR framework can facilitate evaluations of PETs implementations as well, or if other approaches are needed.

As identified in our research limitations, we suggest that analysis of a wider range of novel PETs is necessary for testing the implementation framework. Through this process, further comparisons and contrasts can be factored into iteration of the framework. This would be especially useful for PETs that involve training of models and PETs where data is not ‘flowing’, to see if those systems are different enough from more traditional data sharing systems that they need their own bespoke guidance.

We also propose further investigation of the evaluate and iterate phases. Many initiatives start out with one revenue model or source of funding and then need to shift to another to remain sustainable. This raises questions over how this can be managed and how this affects the timeline of implementation. Further research on this should explore the common barriers and/or challenges to PETs implementations. Relatedly, are they similar/different to the barriers encountered by other implementations?