From privacy to partnership

The role of privacy enhancing technologies in data governance and collaborative analysis

THE ROYAL SOCIETY

From privacy to partnership
Issued: January 2023 DES7924
ISBN: 978-1-78252-627-8

This report can be viewed online at:
royalsociety.org/privacy-enhancing-technologies

Cover image: Visualisation of the Internet 1997 – 2021, by Barret Lyon as part of the the Opte Project.

Contents

Foreword

Alison Noble OBE FREng FRS

The widespread collection and use of data is transforming all facets of society, from scientific research to communication and commerce. The benefits of using data in decision making are increasingly evident in tackling societal problems and understanding the world around us. At the same time, there are inherent vulnerabilities when sensitive data is stored, used or shared.

From privacy to partnership sets out how an emerging set of privacy enhancing technologies (PETs) might help to balance the risks and rewards of data use, leading to wider social benefit. It follows the Royal Society's *Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis*, which gave a snapshot of this rapidly developing field in 2019. This new publication offers a refreshed perspective on PETs, not only as security tools, but as novel means to establish collaborative analysis and data partnerships that are ethical, legal and responsible.

We have three objectives for this report. Our first objective is that the use cases inspire those collecting and using data to consider the potential benefits of PETs for their own work, or in new collaborations with others. Second, for the evidence we present on barriers to adoption and standardisation to help inform policy decisions to encourage a marketplace for PETs. Finally, through our recommendations, we hope the UK will maximise the opportunity to be a global leader in PETs – both for data security and collaborative analysis – alongside emerging, coordinated efforts to implement PETs in other countries.

Our report arrives at a time of rapid innovation in PETs, as well as data protection legislation reform in the United Kingdom. The intention is not to provide a comprehensive view of all technologies under the broad umbrella of PETs; rather, we have chosen to focus on a subset of promising and emerging tools with demonstrable potential in data governance. In demonstrating this value, we cite examples from the UK and international contexts. Realising the full potential of PETs across national borders will require further harmonisation, including consideration of data protection laws in various jurisdictions.

Artificial intelligence and machine learning are transforming our capacity to assess and confront our greatest challenges, but these tools require data to 'fuel' them. As a biomedical engineer using AI-assistive technologies to detect disease, I recognise that the greatest research problems of our time – from cancer diagnostics to the climate crisis – are, in a sense, data problems.

The value of data is most fully realised through aggregation and collaboration, whether between individuals or institutions. I hope this report will inspire new approaches to data protection and collaboration, encouraging further research in – and testing of – PETs in various scenarios. PETs are not a silver bullet, but they could play a key role in unlocking the value of data without compromising privacy. By enabling new data partnerships, PETs could spark a research transformation: a new paradigm for information sharing and data analysis with real promise for tackling future challenges.

Professor Alison Noble OBE FREng FRS, Chair of the Royal Society Privacy Enhancing Technologies Working Group

Executive summary

Privacy Enhancing Technologies (PETs) are a suite of tools that can help maximise the use of data by reducing risks inherent to data use. Some PETs provide new techniques for anonymisation, while others enable collaborative analysis on privately-held datasets, allowing data to be used without disclosing copies of data. PETs are multi-purpose: they can reinforce data governance choices, serve as tools for data collaboration or enable greater accountability through audit. For these reasons, PETs have also been described as 'Partnership Enhancing Technologies'[1] or 'Trust Technologies'[2].

This report builds on the Royal Society's 2019 publication *Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis*[3], which presented a high-level overview of PETs and identified how these technologies could play a role in addressing privacy in applied data science research, digital strategies and data-driven business.

This new report, developed in close collaboration with the Alan Turing Institute, considers how PETs could play a significant role in responsible data use by enhancing data protection and collaborative data analysis. It is divided into three chapters covering the emerging marketplace for PETs, the state of standards and assurance and use cases for PETs.

Scope

From privacy to partnership outlines the current PETs landscape and considers the role of these technologies in addressing data governance issues beyond data security. The aim of this report is to address the following questions:

- How can PETs support data governance and enable new, innovative, uses of data for public benefit?
- What are the primary barriers and enabling factors around the adoption of PETs in data governance, and how might these be addressed or amplified?
- How might PETs be factored into frameworks for assessing and balancing risks, harms and benefits when working with personal data?

Methodology

This work was steered by an expert Working Group as well as two closed contact group sessions with senior civil servants and regulators in April and October 2021 (on the scope and remit of the report, and on the use case topics and emerging themes, respectively).

1 Trask A. in Lunar Ventures (Lundy-Bryan L.) 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

2 Infocomm Media Development Authority (Singapore grows trust in the digital environment). See https://www.imda.gov.sg/news-and-events/Media-Room/Media-Releases/2022/Singapore-grows-trust-in-the-digital-environment (accessed 5 June 2022).

3 The Royal Society. 2019 Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis. See https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/privacy-enhancing-technologies-report.pdf (accessed 30 June 2022).

The findings in this report are the result of consultations with a wide range of data and privacy stakeholders from academia, government, third sector, and industry, as well as three commissioned research projects on the role of assurance in enabling the uptake of PETs[4], PETs market readiness in the public sector[5], and a survey of *synthetic data*: data that is artificially generated based on real-world data, but which produces new data points[6]. The use cases were drafted with input from domain specialists, and the report was reviewed by expert readers as well as invited reviewers. The details of contributors, Working Group members, expert readers and reviewers are provided in the Appendix.

Key findings

General knowledge and awareness of PETs remains low amongst many potential PETs users[7, 8], with inherent risk of using new and poorly understood technologies acting as a disincentive to adoption. Few organisations, particularly in the public sector, are prepared to experiment with data protection[9]. Without in-house expertise, external assurance mechanisms or standards, organisations are unable to assess privacy trade-offs for a given PET or application. As a result, the PETs value proposition remains abstract and the business case for adopting PETs is unclear for potential users.

Standardisation for PETs, including data standards, is lacking and is cited as a hindrance to adoption by potential users in the UK public sector[10]. Technical standards are required to ensure the underpinning technologies work as intended, while process standards are needed to ensure users know how and when to deploy them. While few PETs-specific standards exist to date, standards in adjacent fields (such as cybersecurity and AI) will be relevant. In the future, PETs-specific standards could provide the basis for assurance schemes to bolster user confidence.

4 Hattusia 2022 The current state of assurance in establishing trust in PETs. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

5 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/. This project was partly funded by a grant from CDEI.

6 Jordon J *et al*. 2022 Synthetic data: What, why and how? See https://arxiv.org/pdf/2205.03257.pdf (accessed 2 September 2022).

7 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/.

8 Lunar Ventures, Lundy-Bryan L. 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

9 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/.

10 *Ibid.*

A significant barrier to the widespread use of PETs is a lack of clear use cases for wider public benefit. To address this, Chapter 4 illustrates the potential benefit of PETs in the contexts of:

- Using biometric data for health research and diagnostics;
- Enhancing privacy in the Internet of Things and in digital twins;
- Increasing safe access to social media data and accountability on social media platforms;
- Generating population-level insights using synthesised national data;
- Collective intelligence, crime detection and voting in digital governance; and
- PETs in crisis situations and in analysis of humanitarian data:

The use cases demonstrate how PETs might maximise the value of data without compromising privacy.

A core question for potential PETs users is: What will PETs enable an analyst to do with data that could not be accomplished otherwise? Alternatively: What will PETs prevent an adversary from achieving? As the use cases illustrate, PETs are not a 'silver bullet' solution to data protection problems. However, they may be able to provide novel building blocks for constructing responsible data governance systems. For example, in some cases, PETs could be the best tools for reaching legal obligations, such as anonymity.

Data protection is only one aspect of the right to privacy. In most cases, PETs address this one aspect but do not address how data or the output of data analysis is used, although this could change as PETs mature. Some recent applications utilise PETs as tools for accountability and transparency, or to distribute decision-making power over a dataset across multiple collaborators[11], suggesting their potential in addressing elements of privacy beyond data security.

The field of PETs continues to develop rapidly. This report aims to consolidate and direct these efforts toward using data for public good. Through novel modes of data protection, PETs are already enhancing the responsible use of personal data in tackling significant contemporary challenges. The emerging role of PETs as tools for partnership, enhancing transparency and accountability may entail greater benefits still.

11 For example, Meta recently conducted a survey collecting personal data, which was encrypted and split into shares between third-party facilitators, namely universities. Analyses can be run using secure multi-party computation; requests for analysis must be approved by all third-party shareholders. See https://ai.facebook.com/blog/assessing-fairness-of-our-products-while-protecting-peoples-privacy/ (accessed 10 October 2022).

Recommendations

AREA FOR ACTION: COORDINATED INTERNATIONAL ACTION TO ENSURE THE RESPONSIBLE DEVELOPMENT OF PETS FOR PUBLIC BENEFIT

RECOMMENDATION 1

National and supernational organisations, including standards development organisations (SDOs) should establish protocols and standards for PETs, and their technical components, as a priority.

PETs have been developed by experts in different fields and with little coordination between them to date. The greatest potential for PETs – whether used in isolation or combination – are as components of data governance systems. Open standards (available for use by anyone) are likely to help drive the development, accessibility and uptake of PETs for data governance. Furthermore, standards will be necessary for audit and assurance, encouraging a marketplace of confident PETs users with effective regulation and quality assurance marks where appropriate.

SDOs such as the British Standards Institute (BSI) (UK), National Physical Laboratory (UK), Institute of Electrical and Electronics Engineers (IEEE) (US), the National Cyber Security Centre (UK) and National Institute of Standards and Technology (NIST) (US) should identify and convene international expert groups to address gaps in PETs technical standards. These should build on existing standards in cryptography and information security (Chapter 3). Open standards will be especially important in PETs that enable information networks, such as secure multi-party computation or federated learning (similar to how HTTP[12] provided a common set of rules that enabled communication over the Internet).

Alongside technical standards, process standards should guide best practice in the application of PETs in data governance. Privacy best practice guides, codes of conduct and process standards (such as the draft Institute of Privacy Design Process Standard[13]) could be used to integrate PETs into a privacy-by-design approach to data governance systems. Whereas technical standards will be essential for technical interoperability, codes of conduct for PETs in data management and use will be critical for 'social interoperability' and acceptance in partnerships and digital collaborations on new scales (such as international or cross-sector partnerships).

12 Hypertext Transfer Protocol.

13 Institute of Privacy Design (The DRAFT Design Process Standard). See https://instituteofprivacydesign.org/2022/02/11/the-draft-design-process-standard/ (accessed 2 September 2022).

RECOMMENDATION 2

Science funders, including governments and intergovernmental bodies, should accelerate and incentivise the development and maturation of PETs by funding prize challenges, pathfinder projects (such as topic guides or resource lists) and cross-border, collaborative test environments (such as an international PETs sandbox).

Science funders should foster a network of independent researchers and universities working on PETs challenges that address PETs in security, partnerships and transparency applications. They could involve the private sector (for example cloud providers and social media platforms) in designing challenges and through international cooperation on standards, guidance and regulation. To date, exemplary programmes include the UK-US PETs Prize Challenge led by the UK's Centre for Data Ethics and Innovation (CDEI) and the US White House Office of Science and Technology Policy; the Digital Security by Design Challenge[14] funded through UK Research and Innovation; the Data.org Epiverse Challenge funding call; and the French data protection authority sandbox on digital health and GDPR[15].

Intragovernmental bodies such as the United Nations and Global Partnership for Artificial Intelligence should lead by creating test environments and providing data for demonstrations to test the security, privacy, and utility potentials of specific PETs, as well as test configurations of PETs. An international PETs sandbox would allow national regulators to collaborate and evaluate PETs solutions for cross-border data use according to common data governance principles.

14 UK Research and Innovation (Digital security by design challenge). See https://www.ukri.org/what-we-offer/our-main-funds/industrial-strategy-challenge-fund/artificial-intelligence-and-data-economy/digital-security-by-design-challenge/ (accessed 20 September 2022).

15 Commission Nationale de l'Informatique et des Libertés *(Un «bac à sable» RGPD pour accompagner des projets innovants dans le domaine de la santé numérique*). See https://www.cnil.fr/fr/un-bac-sable-rgpd-pour-accompagner-des-projets-innovants-dans-le-domaine-de-la-sante-numerique (accessed 15 September 2022).

RECOMMENDATION 3

Researchers, regulators and enforcement authorities should investigate the wider social and economic implications of PETs, for example, how PETs might be used in novel harms (such as fraud or linking datasets for increased surveillance) or how PETs might affect competition in digitised markets (such as monopolies through new network effects).

The potential follow-on effects of PETs adoption are not well understood, particularly whether and how they might amplify data monopolies, or what oversight mechanisms are required to prevent the type of collaborative analysis that might be considered state surveillance[16]. For example, the Arts and Humanities Research Council could consider the ethical, social and economic implications of PETs within their program on AI (particularly where PETs could be dual use or surveillance technologies)[17, 18].

Regulators, such as the Information Commissioner's Office (ICO) and the Competition and Markets Authority (CMA), could investigate the wider economic implications of PETs, particularly where they could enable competition through greater interoperability (as with open banking, for example). It is not understood how the adoption of PETs aligns with FAIR[19] principles, particularly where PETs (such as privacy-preserving synthetic data) are used as an alternative to open data. In collaborative analysis, the ability to audit data that is not shared should be better understood by those who might use PETs (to identify potential for biased outcomes, for example). The relationship between PETs and data trusts also remains ambiguous.

16 Liberty Human Rights (Challenge hostile environment data-sharing). See https://www.libertyhumanrights.org.uk/campaign/challenge-hostile-environment-data-sharing/ (accessed 20 September 2022).

17 Ongoing research highlights the negative consequences of data sharing in dual-use or otherwise sensitive contexts. For example: Papageogiou V, Wharton-Smith A, Campos-Matos I, Ward H. 2020 Patient data-sharing for immigration enforcement: a qualitative study of healthcare providers in England. *BMJ Open*. (https://doi.org/10.1136/bmjopen-2019-033202)

18 Liberty Human Rights (Liberty and Southall Black Sisters' Super-complain on data-sharing between the police and home office regarding victims and witnesses to crime). See https://www.libertyhumanrights.org.uk/issue/liberty-and-southall-black-sisters-super-complaint-on-data-sharing-between-the-police-and-home-office-regarding-victims-and-witnesses-to-crime/ (accessed 20 September 2022).

19 Go FAIR (FAIR principles). See https://www.go-fair.org/fair-principles/ (accessed 20 September 2022).

AREA FOR ACTION: A STRATEGIC AND PRAGMATIC APPROACH TO PETS ADOPTION IN THE UK, LED BY THE PUBLIC SECTOR THROUGH PUBLIC-PRIVATE PARTNERSHIPS, DEMONSTRATION OF USE CASES AND COMMUNICATION OF BENEFITS

RECOMMENDATION 4

The UK Government should develop a national PETs strategy to promote the responsible use of PETs in data governance: as tools for data protection and security, for collaboration and partnership (both domestically and cross-border) and for advancing scientific research.

PETs could reform the way data is used domestically and across borders, offering potential solutions to longstanding problems of siloed and underutilised data across sectors. To ensure the use of PETs for public good, PETs-driven information networks should be stewarded by public sector and civil society organisations using data infrastructure for public good. A coordinated national strategy for the development and adoption of PETs for public good will ensure the timely and responsible deployment of these technologies, with the public sector leading by example.

PETs have a role to play in achieving the objectives outlined in Mission 2 of the National Data Strategy, securing a 'pro-growth and trusted data regime,' positioning the UK internationally as a trusted data partner, with wider implications for national security. This recommendation reflects emerging, coordinated PETs work in foreign governments (such as that led in the US by the White House Office for Science and Technology Policy)[20].

The PETs strategy should offer a vision that complements the Government's National Data Strategy[21] and National AI Strategy[22]. The PETs strategy should prioritise a roadmap for public sector PETs adoption, addressing public awareness and the PETs marketplace (Chapter 2), technological maturity, appropriate regulatory mechanisms and responsibilities, alongside standards and codes of conduct for PETs users (Chapter 3).

20 US Office for Science and Technology Policy (Request for Information on Advancing Privacy-Enhancing Technologies). https://public-inspection.federalregister.gov/2022-12432.pdf (accessed 17 July 2022).

21 HM Government (National Data Strategy). See https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy (accessed 9 September 2022).

22 HM Government (National AI Strategy). See https://www.gov.uk/government/publications/national-ai-strategy (accessed 9 September 2022).

RECOMMENDATION 5

Local, devolved and national governments across the UK should lead by example in the adoption of PETs for data sharing and use across government and in public-private partnerships, improving awareness by communicating PETs-enabled projects and their results.

Public sector organisations could partner with small and medium-sized enterprises (SMEs) developing PETs to identify use cases, which could then be tested through low-cost, low-risk pilot projects. Legal experts and interdisciplinary policy professionals should be involved from project inception, ensuring PETs meet data protection requirements and that outcomes and implications are properly communicated to non-technical decision-makers.

Use cases illustrated in Chapter 5 highlight areas of significant potential public benefit in healthcare and medical research, for reaching net zero through national digital twins and for population-level data collaboration.

Communication of PETs and their appropriate use in various contexts will be key to building trust with potential users[23], encouraging the PETs marketplace (Chapter 2). The ICO should continue its work on using PETs for wider good and communicating the implications – including barriers and potential benefits. The CDEI should continue to provide practical examples that will help organisations understand and build a business case for PETs' adoption. Proof of concept and pilot studies should be communicated to the wider public to demonstrate the value of PETs, foster trust in public sector data use and demonstrate value-for-money[24].

23 The Royal Society. Creating trusted and resilient data systems: The public perspective. (to be published online in 2023)

24 This is in line with the Digital Economy Act 2017. See: The Information Commissioner's Office (Data sharing across the public sector: the Digital Economy Act codes). See https://ico.org.uk/for-organisations/guide-to-data-protection/ico-codes-of-practice/data-sharing-a-code-of-practice/data-sharing-across-the-public-sector-the-digital-economy-act-codes/ (accessed 2 September 2022).

RECOMMENDATION 6

The UK Government should ensure that new data protection reforms account for the new systems of data governance enabled by emerging technologies such as PETs and ensure any new regulations are supported by clear, scenario-specific guidance and assessment tools.

While data protection legislation should remain technology neutral so as to be adaptable, current plans to review UK data protection laws provide an opportunity to consider the novel and multipurpose nature of these emerging technologies, particularly as they provide the technical means for new types of collaborative analysis. The ICO should continue its work to provide clarity around PETs and data protection law, encouraging the use of PETs for wider public good[25] and drawing from parallel work on AI guidance where relevant (such as privacy-preserving machine learning).

Further interpretation may be required to help users understand how PETs might serve as tools for meeting data protection requirements. For example, it may be required to clarify data protection obligations where machine learning models are trained on personal data in federated learning scenarios[26] or the degree to which differentially private or homomorphically encrypted data meets anonymisation requirements[27]. Where PETs enable information networks and international data collaborations, the ICO might anticipate clarification questions specific to international and collaborative analysis use cases. Regulatory sandboxes (as in Recommendation 2) will be useful for testing scenarios, particularly for experimentation with PETs in structured transparency[28] (such as in open research, credit scoring systems) and as accountability tools[29].

25 The Information Commissioner's Office (ICO consults health organisation to shape thinking on privacy-enhancing technologies). See https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2022/02/ico-consults-health-organisations-to-shape-thinking-on-privacy-enhancing-technologies/ (accessed 20 March 2022).

26 Nguyen T, Sun K, Wang S, Guitton F, Guo Y. 2021. Privacy preservation in federate learning An insightful survey from the GDPR perspective. *Computers & Security* 110. (https://doi.org/10.1016/j.cose.2021.102402)

27 See for example: Koerner K. 2021 Legal perspectives on PETs: Homomorphic encryption. *Medium*. 20 July 2021. See https://medium.com/golden-data/legal-perspectives-on-pets-homomorphic-encryption-9ccfb9a334f (accessed 30 June 2022).

28 Trask A, Bluemke E, Garfinkel B, Cuervas-Mons CG, Dafoe A. 2020 Beyond Privacy Trade-offs with Structured Transparency. See https://arxiv.org/ftp/arxiv/papers/2012/2012.08347.pdf (accessed 6 February 2022).

29 See for example: Meta AI (Assessing fairness of our products while protecting peoples privacy). See https://ai.facebook.com/blog/assessing-fairness-of-our-products-while-protecting-peoples-privacy/ (accessed 15 August 2022).

RECOMMENDATION 6 (CONTINUED)

The ICO could expand on its PETs guidance, for example, through developing self-assessment guides. Data ethics organisations, such as the CDEI, might also develop impact assessment tools, for example, a PETs impact assessment protocol that considers downstream implications on human rights. The Alliance for Data Science Professionals certification scheme[30], which defines standards for ethical and well-governed approaches to data use, could specifically consider the role of PETs in evidencing Skill Areas A (Data Privacy and Stewardship) and E (Evaluation and Reflection).

30 Alliance for Data Science Professionals (Homepage). See https://allianceforadatascienceprofessionals.co.uk/ (accessed 20 September 2022).

AREA FOR ACTION: FOUNDATIONAL SCHOLARSHIP AND PROFESSIONALISATION TO ENCOURAGE MATURATION OF PETS, FOSTER TRUST AND DRIVE UPTAKE OF PETS IN DATA-USING ORGANISATIONS

RECOMMENDATION 7

Universities, businesses and science funders should fund foundational scholarship in PETs-related fields, such as cryptography and statistics.

Foundational training and fellowships in PETs fundamentals (such as cryptography) for graduate level study will create the skilled workforce required for widespread development and implementation of PETs. Critical future-proofing questions could be addressed through fellowships and research posts (for example, evaluating the security guarantees of PETs in a post-quantum context, or the energy proportionality, sustainability and scalability of energy-intensive, cryptography-based PETs). Internships and work placement programmes in organisations developing PETs could assist new graduates in moving from academic fields into applied PETs research and development.

RECOMMENDATION 8

Organisations providing certifications and continuing professional development courses in data science, cybersecurity and related fields should incorporate PETs modules to raise awareness among data professionals.

Professional certifications and Continuing Professional Development opportunities (including British Computer Society Professional Certifications such as the Alliance for Data Science Professionals certification, Data Science Professional Certificates offered by Microsoft or IBM, or $(ISC)^2$ Certifications[31]) should include a primer on PETs to raise awareness and encourage baseline knowledge of PETs amongst in-house data professionals. For example, the International Association of Privacy Professionals now includes a module on PETs in their Certified Information Privacy Technologist Certification[32].

31 $(ISC)^2$ ($(ISC)^2$ Information Security Certifications). See https://www.isc2.org/Certifications# (accessed 13 May 2022).

32 International Association of Privacy Professionals (Privacy Technology Certification). See https://iapp.org/media/pdf/certification/CIPT_BOK_v.3.0.0.pdf (accessed 30 June 2022).

TABLE 1

Summary table of PETs explored in this report

	Trusted execution environment	Homomorphic encryption	Secure multi-party computation (PSI / PIR)
Context of data use	Securely outsourcing to a server, or cloud, computations on sensitive data	Securely outsourcing specific operations on sensitive data; Safely providing access to sensitive data	Enabling joint analysis on sensitive data held by several organisations
Privacy risk addressed	Revealing sensitive attributes present in a dataset during computation	Revealing sensitive attributes present in a dataset during computation	Revealing sensitive attributes present in a dataset during computation
Data protected	✓ In storage ✓ During computation X On release	✓ In storage ✓ During computation X On release*	X In storage ✓ During computation X On release
Benefits	Commercial solutions widely available; Zero loss of information; efficient computation of any operations	Can allow zero loss of information; FHE can support the computation of any operation	No need for a trusted third party--sensitive information is not revealed to anyone; The parties obtain only the resulting analysis or model
Current limitations	Many side-channel attacks possible; current commercial solutions limited with regard to distributed computation on big datasets	FHE, SHE and PHE are usable Highly computationally intensive Bandwidth and latency issues Running time PHE and SHE support the computation of limited functions Standardisation in progress Possibility for side channel attacks (current understanding is limited)	Highly compute and communication intensive; requires expertise in design that meets compute requirements and security models
Readiness level	Product	PHE / SHE / FHE in use (FHE on a smaller scale)	PSI / PIR / Product, Proof of concept--Pilot
Qualification criteria	Could be exclusive to established research groups	Specialist skills Custom protocols Computing resources	Specialist skills Custom protocols Computing resources

KEY

FHE: Fully Homomorphic Encryption **SHE:** Somewhat Homomorphic Encryption **PHE:** Partial Homomorphic Encryption
PIR: Private Information Retrieval **PSI:** Private Set Intersection

* If the client encrypts their data and sends it to a server for homomorphic computation, only the client is able to access the results (by using their secret decryption key).

	Federated learning / federated machine learning	Differential privacy	Privacy-preserving synthetic data
	Enables the use of remote data for training algorithms; data is not centralised	Prevents disclosure about individuals when releasing statistics or derived information	Prevents disclosure about individuals when releasing statistics or derived information
	Revealing sensitive information, including an individual's presence in a dataset	Revealing sensitive information, including an individual's presence in a dataset; Dataset or output disclosing sensitive information about an entity included in the dataset	Revealing sensitive attributes or presence in a dataset
	X In storage ✓ During computation X On release	✓ In storage (and at point of data collection) ✓ During computation (with limitations) ✓ On release (with limitations)	X In storage ✓ During computation (with limitations) ✓ On release (with limitations)
	Very little loss of information	Formal mathematical proof / privacy guarantee. Level of privacy protection may be quantifiable. Relative to other PETs, it is computationally inexpensive.	Applications beyond privacy Level of privacy protection may be quantifiable (eg, with differentially private synthetic data)
	Model inversion and membership inference attacks may be vulnerabilities	Noise and loss of information, unless datasets are large enough Setting the level of protection requires expertise Precision of analysis limited inversely to level of protection	Noise and loss of information Setting the level of protection requires expertise Privacy enhancement unclear
	Product, in use	Proof of concept, in use	Proof of concept, in use
	May require scale of data within each dataset (cross-silo federated learning) Distributed systems are complex and difficult to manage	Specialist skills Custom protocols Very large datasets As yet, no standards for setting privacy parameters	Specialist skills required As yet, no standards for generation or setting privacy parameters

Introduction

Background

Data about individuals, their unique characteristics, preferences and behaviours, is ubiquitous and the power to deliver data-driven insights using this information is rapidly accelerating[33, 34]. This unprecedented availability of data, coupled with new capabilities to use data, drives the frontiers of research and innovation – addressing challenges from the climate crisis to the COVID-19 pandemic[35, 36]. However, the greater collection, transfer and use of data – particularly data which is personal, commercially sensitive or otherwise confidential – also entails increased risks. The tension between maximising data utility (where data is used) and managing risk (where data is hidden) poses a significant challenge to anyone using data to make decisions.

This report, undertaken in close collaboration with the Alan Turing Institute, considers the potential for tools and approaches collectively known as Privacy Enhancing Technologies (PETs) to revolutionise the safe and rapid use of sensitive data for wider public benefit. It examines the possibilities and limitations for PETs in responsible data governance and identifies steps required to realise their benefits.

This work follows the Royal Society's 2019 report *Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis*[37], which highlighted the role of PETs in enabling the derivation of useful results from data without providing wider access to datasets. *Protecting privacy in practice* presented a high-level overview of PETs and identified how these potentially disruptive technologies could play a role in addressing tensions around privacy and utility.

The 2019 report made several observations for how the UK could realise the potential of PETs, including:

- The research and development of PETs can be accelerated through collaborative, cross-sector research challenges developed by government, industry and the third sector, alongside fundamental research support for advancing PETs;
- Government can be an important influencer in the adoption of PETs by demonstrating their use and sharing their experience around how PETs unlock new opportunities for data analysis. At the same time, public sector organisations should be given the level of expertise and assurance required to utilise new technological solutions;

33 The Royal Society. 2017 Machine learning: the power and promise of computers that learn by example. See https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf (accessed 30 May 2022).

34 The British Academy and the Royal Society. 2017 Data management and use: Governance in the 21st century. See https://royalsociety.org/-/media/policy/projects/data-governance/data-management-governance.pdf (accessed 28 July 2022).

35 Alsunaidi A J *et al*. 2021 Applications of big data analytics to control COVID-19 pandemic. *Sensors (Basel)* 21, 2282. (https://doi.org/10.3390/s21072282 s21072282)

36 The Royal Society. 2020 Digital technology and the planet: Harnessing computing to achieve net zero. See https://royalsociety.org/-/media/policy/projects/digital-technology-and-the-planet/digital-technology-and-the-planet-report.pdf (accessed 20 September 2022).

37 The Royal Society. 2019 Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis. See https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/privacy-enhancing-technologies-report.pdf (accessed 30 June 2022).

- PETs can promote human flourishing through enabling new and innovative ways of governing data, as well as promoting safe and secure data use. The Department for Digital, Culture, Media and Sport (DCMS), the Centre for Data Ethics and Innovation (CDEI), Office for AI, regulators and civil society can consider how PETs play a role in wider data governance structures, including how they operate alongside new data governance models such as 'data trusts'.

Key terms and definitions

This report draws on multidisciplinary concepts from cryptography, business, cybersecurity, ethics and analytics. Included here is a quick reference glossary of key terms used throughout.

Differential privacy: security definition which means that, when a statistic is released, it should not give much more information about a particular individual than if that individual had not been included in the dataset. See also 'privacy budget'.

Distributed Ledger Technology (DLT): an open, distributed database that can record transactions between several parties efficiently and in a verifiable and permanent way. DLTs are not considered PETs, though they can be used (as some PETs) to promote tra nsparency by documenting data provenance.

Epsilon (Ɛ): see 'privacy budget'.

Homomorphic encryption (HE): a property that some encryption schemes have, so that it is possible to compute on encrypted data without deciphering it.

Metadata: data that describes or provides information about other data, such as time and location of a message (rather than the content of the message).

Noise: noise refers to a random alteration of data/values in a dataset so that the true data points (such as personal identifiers) are not as easy to identify.

Privacy budget (also differential privacy budget, or epsilon): a quantitative measure of the change in confidence of an individual having a given attribute.

Privacy-preserving synthetic data (PPSD): synthetic data generated from real-world data to a degree of privacy that is deemed acceptable for a given application.

Private Set Intersection (PSI): secure multiparty computation protocol where two parties compare datasets without revealing them in an unencrypted form. At the conclusion of the computation, each party knows which items they have in common with the other. There are some scalable open-source implementations of PSI available.

Secure multi-party computation (SMPC or MPC): a subfield of cryptography concerned with enabling private distributed computations. MPC protocols allow computation or analysis on combined data without the different parties revealing their own private inputs to the computation.

Synthetic data: data that is modelled to represent the statistical properties of original data; new data values are created which, taken as a whole, reproduce the statistical properties of the 'real' dataset.

Trusted Execution Environment (TEE): secure area of a processor that allows code and data to be isolated and protected from the rest of the system such that it cannot be accessed or modified even by the operating system or admin users. Trusted execution environments are also known as secure enclaves.

Chapter one

The role of technology in privacy-preserving data flows

Left

The role of technology in privacy-preserving data flows

Data security relates to protecting data as an asset, whereas data privacy is more concerned with protecting people: ensuring the rights of data subjects follow their data.

The ever-growing quantity of data collected in contemporary life, coupled with increasing power to compute, is opening new possibilities for data-driven solutions[38]. At the same time, there is unprecedented potential for the misuse of data – whether intentional or unintentional – leading to downstream harms at individual, community, corporate and national scales[39, 40].

The Royal Society's 2019 report focused on the role of PETs in addressing data privacy. Acknowledging that privacy is a term with multiple meanings[41, 42], it referenced Daniel Solove's taxonomy of privacy. Solove's approach considers privacy violation as resulting from problematic data actions pertaining to personal data, including:

- Aggregation: the gathering together of information about an individual, which could be used to generate insights for reidentification or profiling[43];
- Identification: the linking of data (which may otherwise be anonymised) to a specific individual;
- Insecurity: the potential for data to be accessed by an intruder due to glitches, cybersecurity breach or intentional misuse of information;
- Exclusion: the use of personal data without notice to individuals;
- Disclosure: the revelation of personal data to others;
- Exposure: the revelation of an individual's physical or emotional attributes to others;
- Intrusion: invasive acts that interfere with an individual's physical or virtual life (such as junk mail).

Data privacy tools can include technologies, legal instruments or physical components (such as hardware keys) that mitigate the risk of problematic data actions. However, data privacy can mean many things, and can be subjective or contextual[44]. Broadly, privacy may be considered the right of individuals to selectively express themselves or be known. Data privacy entails a degree of control and influence over personal data, including its use. It may therefore be described as 'the authorized, fair, and legitimate processing of personal information'[45].

38 The British Academy and the Royal Society. 2017 Data management and use: Governance in the 21st century. See https://royalsociety.org/-/media/policy/projects/data-governance/data-management-governance.pdf (accessed 28 July 2022).

39 Wolf LE 2018. Risks and Legal Protections in the World of Big-Data. Asia Pac *J Health Law Ethics*. 11, 1-15. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6863510/

40 Jain P, Gyanchandani M, Khare N. 2016 Big data privacy: a technological perspective and review. *Journal of Big Data* 3, 25.

41 The British Academy and the Royal Society. 2017 Data management and use: Governance in the 21st century. See https://royalsociety.org/-/media/policy/projects/data-governance/data-management-governance.pdf (accessed 28 July 2022).

42 The Israel Academy of Sciences and Humanities and The Royal Society. 2017 Israel-UK privacy and technology workshop note of discussions. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/ (accessed 20 September).

43 This is distinct from aggregation across a population or group.

44 Nissenbaum H. 2010 *Privacy In Context: Technology, Policy, and the Integrity of Social Life*. Stanford: Stanford Law Books.

45 Bhajaria N. 2022 *Data privacy: A runbook for engineers*. Shelter Island: Manning.

A specific definition of privacy may be less useful than considering what privacy is for[46] and what is at stake by examining potential downstream harms. The loss of privacy may also be considered intrinsically harmful to an individual.

Data privacy, data protection and information security

Data privacy is related to information security, but there are important differences. Information security focuses on external adversaries and the prevention of undesired access to information[47]. Security is a necessary condition for data privacy, but privacy also entails the legitimate and fair use of (secure) data. Data security relates to protecting data as an asset, whereas data privacy is more concerned with protecting people: ensuring the rights of data subjects follow their data.

The unauthorised use of data shared for a given purpose is loss of privacy (a violation of intention). This suggests that data privacy tools should address accountability and transparency in data collection and use, in addition to helping meet security requirements. Data protection, on the other hand, refers to the legal safeguards in place to ensure data rights are upheld while data is collected, stored or processed.

What are privacy enhancing technologies (PETs)?

PETs are an emerging set of technologies and approaches that enable the derivation of useful results from data without providing full access to the data. In many cases, they are tools for controlling the likelihood of breach or disclosure. This potentially disruptive suite of tools could create new opportunities where the risks of using data currently outweigh the benefits. PETs can reduce the threats typically associated with collaboration[48], motivating new partnerships – for example, between otherwise competing organisations. For this reason, PETs have more recently been described as Partnership Enhancing Technologies[49] and Trust Technologies[50].

PETs are an emerging set of technologies and approaches that enable the derivation of useful results from data without providing full access to the data.

46 Zimmermann C. 2022 Part 1: What is Privacy Engineering? *The Privacy Blog*. 10 May 2022. See https://the-privacy-blog.eu/2022/05/10/part1-what-is-privacy-engineering/ (accessed 20 September 2022).

47 According to NIST, security is '[t]he protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability.' National Institute of Standards and Technology (Computer security resource center). See https://csrc.nist.gov/glossary/term/is (accessed 20 September 2022).

48 World Economic Forum. 2019 The next generation of data-sharing in financial services: Using privacy enhancing technologies to unlock new value. See https://www3.weforum.org/docs/WEF_Next_Gen_Data_Sharing_Financial_Services.pdf (accessed 20 September 2022).

49 Lunar Ventures, Lundy-Bryan L. 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

50 Infocomm Media Development Authority (Singapore grows trust in the digital environment). See https://www.imda.gov.sg/news-and-events/Media-Room/Media-Releases/2022/Singapore-grows-trust-in-the-digital-environment (accessed 5 June 2022).

The term Privacy Enhancing Technologies originates in a 1995 report co-authored by the Information and Privacy Commissioner of Ontario and the Dutch Data Protection Authority, which described technologies that allowed online transactions to remain anonymous[51]. Since then, PETs have evolved in different fields with limited coordination, and there is no consensus around a single definition of PETs. This report follows the European Union Agency for Cybersecurity (ENISA) definition of PETs: a group of technologies that support data minimisation, anonymisation and pseudonymisation as well as other privacy and security principles central to data protection[52].

A downstream harms-based approach: Taxonomy of harms

This report considers PETs beyond data security mitigation. However, a framework for data protection and risk is useful in understanding the drivers of data governance decisions (including reluctance to partner or share data).

PETs can help prevent downstream harms through bolstering data protection practices. A taxonomy of harms (Figure 1) provides a conceptual overview of how data might be used or shared, alongside the harms that may follow problematic data actions. It classifies harms into domains (individual, organisation, societal, national) and types (physical/psychological, relational, reputational, personal, economic, security).

To demonstrate the interconnectedness of risk factors and harms, the model shows both practical elements that may result in harm, as well as downstream effects – including damage that can occur far outside the perceived system[53]. It is important to note that, while there are general trade-offs between privacy and utility, the relationship is rarely a simple or linear one.

Threats to privacy are not always external to a data-holding institution. Internal actors may intentionally or unwittingly disclose personal data or other sensitive information. Additionally, there is no simple one-to-one mapping between an attack and the target (type of information release) or an outcome. Multiple attacks may be used in a sequence to reveal information.

The taxonomy is not an exhaustive list of all potential attacks and harms, but provides an illustrative tool designed to encourage a harms-based approach to data protection risks.

51 Information and Privacy Commissioner of Ontario and Registratiekamer (Netherlands) 2008. Privacy-Enhancing Technologies: The Path to Anonymity. Volume 1.

52 European Union Agency for Cybersecurity (Data Protection: Privacy enhancing technologies). See https://www.enisa.europa.eu/topics/data-protection/privacy-enhancing-technologies (accessed 20 September 2022).

53 National Institute of Standards and Technology (NIST Privacy Engineering Objectives and Risk Model Discussion Draft). See https://www.nist.gov/system/files/documents/itl/csd/nist_privacy_engr_objectives_risk_model_discussion_draft.pdf (accessed 20 September 2022).

Recent international developments in PETs

Beyond data security applications, PETs are gaining attention for their role in facilitating data use across national borders. In 2019, the World Economic Forum published a comprehensive review of PETs in financial services, a sector that is among the most cited in emerging PETs uptake[54]. In 2020 The Organisation for Economic Cooperation and Development (OECD) recommended data sharing arrangements that use technological access controls, such as PETs, in guidance on cross-border data flows and international trade. For international data use, they suggest PETs may be complemented with 'legally binding and enforceable obligations to protect the rights and interests of data subjects and other stakeholders[55].

In January 2022, the United Nations Committee of Experts on Big Data and Data Science for Official Statistics launched a pilot PET lab programme, which aims to enhance international data use with PETs[56]. The UN PET Lab is currently working with four National Statistical Offices (NSOs) and collaborating with PETs providers to safely experiment with PETs and identify barriers to their implementation.

In June 2022 Singapore's Minister for Communications and Information launched the new Digital Trust Centre, which will lead research and development in 'Trust Technologies', including PETs and explainable artificial intelligence[57].

Also in June 2022, the US Office for Science and Technology Policy and DCMS in the UK launched a joint PETs prize challenge to accelerate the adoption of PETs as tools for democracy[58]. Both governments are working closely with NIST (US) and the US National Science Foundation in developing the challenge. The transatlantic initiative is deemed an 'expression of our shared vision: a world where our technologies reflect our values and innovation opens the door to solutions that make us more secure'[59].

Beyond data security applications, PETs are gaining attention for their role in facilitating data use across national borders.

54 World Economic Forum. 2019 The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value). See https://www3.weforum.org/docs/WEF_Next_Gen_Data_Sharing_Financial_Services.pdf (accessed 20 September 2022).

55 Organisation for Economic Co-operation and Development (Recommendation of the Council on Enhancing Access to and Sharing of Data). See https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463 (accessed 20 September 2022).

56 Hurst A. 2022 UN launches privacy lab pilot to unlock cross-border data sharing benefits. Information Age. 25 January 2022. See https://www.information-age.com/un-launches-privacy-lab-pilot-to-unlock-cross-border-data-sharing-benefits-19414/ (accessed 20 March 2022).

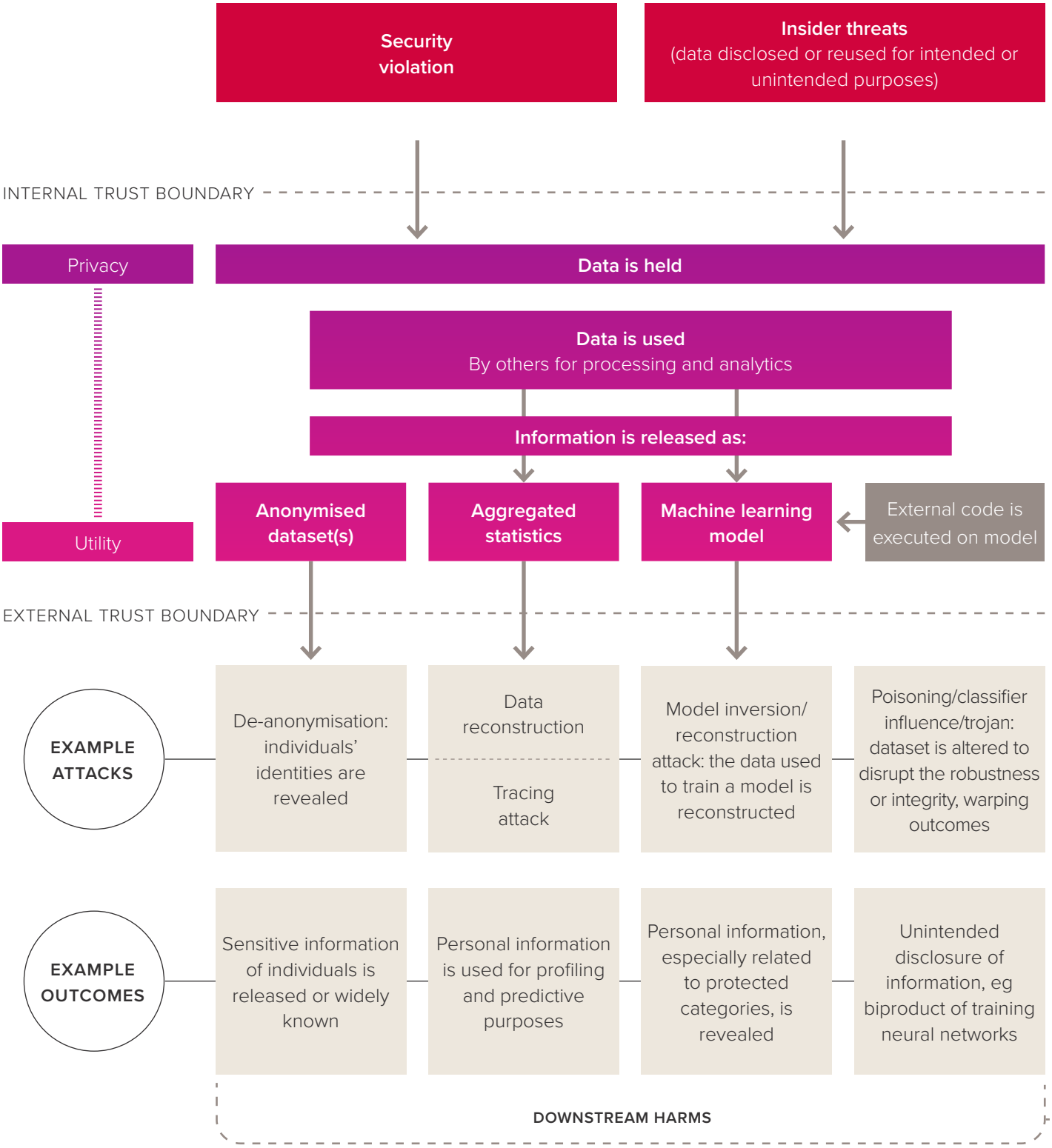
57 Infocomm Media Development Authority (Singapore grows trust in the digital environment). See https://www.imda.gov.sg/news-and-events/Media-Room/Media-Releases/2022/Singapore-grows-trust-in-the-digital-environment (accessed 5 June 2022).

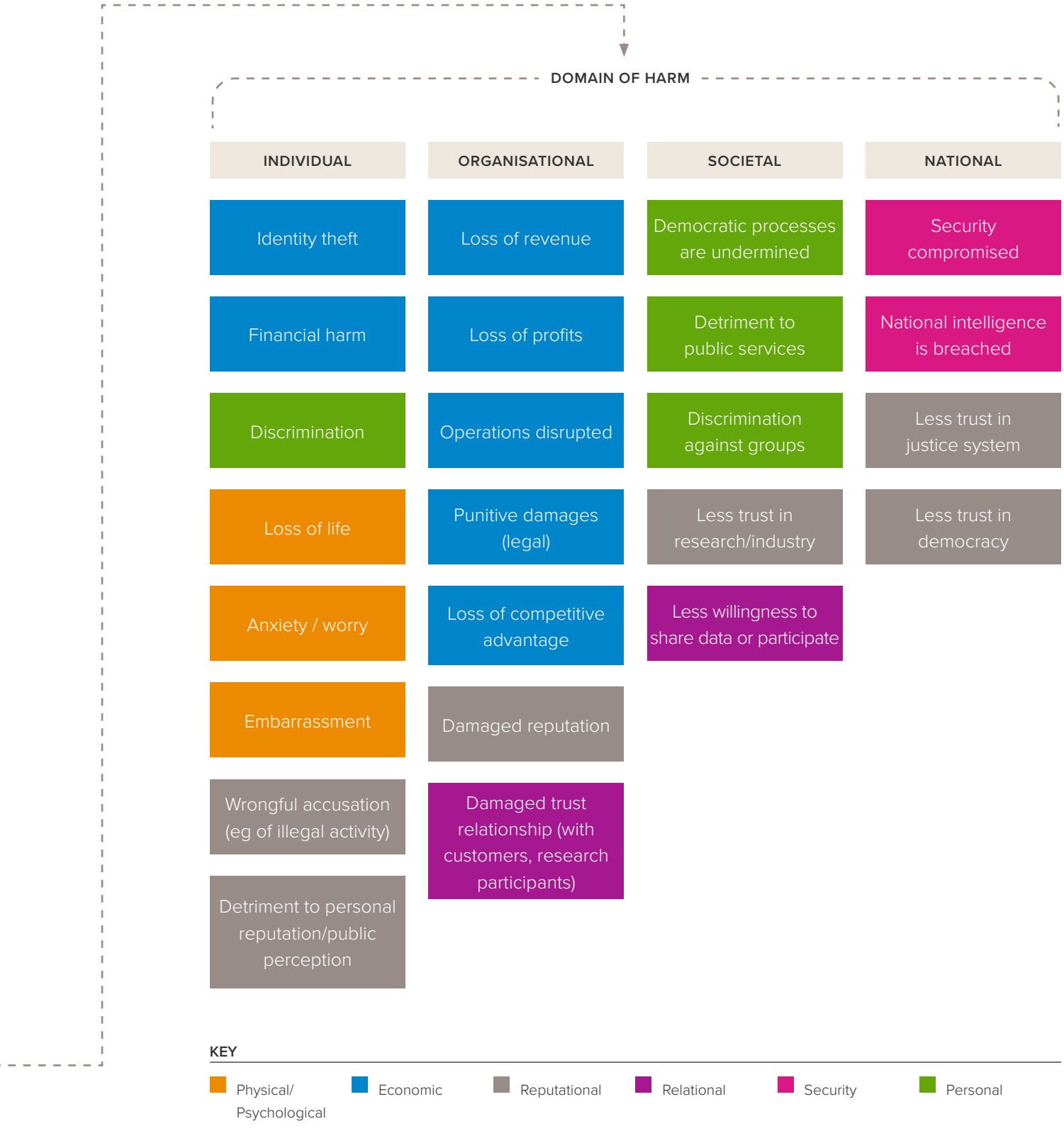
58 HM Government (U.K. and U.S. governments collaborate on prize challenges to accelerate development and adoption of privacy-enhancing technologies). See https://www.gov.uk/government/news/uk-and-us-governments-collaborate-on-prize-challenges-to-accelerate-development-and-adoption-of-privacy-enhancing-technologies (accessed 13 June 2022).

59 *Ibid*.

FIGURE 1

Taxonomy of harms





Source: Royal Society meetings with Working Group for Privacy Enhancing Technologies, November 2021 and April 2022.

Interest in PETs for international data transfer and use

A fragmented array of legal requirements covers data use across the globe. As of March 2022, there are 157 countries with data protection laws, entailing various stipulations for data transfer and use[60]. PETs can provide means for secure collaboration across borders, preventing unauthorised access to datasets; however, data use is still subject to local legal requirements. PETs do not provide 'loopholes' to data protection laws in the UK. Rather, PETs can be used as tools to help data users comply with regulatory requirements, such as anonymisation. While this report refers primarily to current UK GDPR, it restricts legal commentary to high-level observations, noting ongoing data reform in the UK and international relevance of PETs in other jurisdictions.

Accelerating PETs development: Sprints, challenges and international collaboration

Other PETs development initiatives include the PRIViLEDGE project, funded by Horizon Europe between 2017 and 2021. The project aimed to develop cryptographic protocols in support of privacy, anonymity and efficient decentralised consensus using distributed ledger technologies (DLTs).

As well as online voting (see Use case 5.3, page 95), PRIViLEDGE developed a number of toolkits and prototypes[61], including privacy-preserving data storage using ledgers (data residing on a blockchain) and secure multi-party computation (SMPC) on distributed ledgers, which allows two or more parties to compute using a ledger as a communication channel. Many of these resources have been opened further development.

State-level collaborations to accelerate PETs include the Digital Trust Centre (DTC), launched in 2022 in Singapore[62, 63]. The DTC is set to lead Singapore's efforts in research and development for 'trust technologies', such as PETs, which provide solutions for data sharing and evaluation of trustworthy AI systems. This national effort includes sandbox environments, academic-enterprise partnerships and national and international collaborations between research institutes. As a founding member of the Global Partnership for AI (GPAI), Singapore intends to use this platform to enhance its contributions to GPAI.

These initiatives have the potential to drive innovation and are raising the profile of PETs for privacy, partnership and trust. This will be key in motivating new users and creating a wider marketplace for PETs. The following section focuses on the UK public sector, describing enabling factors and barriers in the adoption of PETs.

60 Greenleaf G. 2022 Now 157 Countries: Twelve Data Privacy Laws in 2021/22. *Privacy Laws & Business International Report* 1, 3—8. See https://ssrn.com/abstract=4137418 (accessed 24 May 2022).

61 Livin L. 2021 Achievements of the priviledge project. *Priviledge blog*. 30 June 2021. See https://priviledge-project.eu/news/achievements-of-the-priviledge-project (accessed 30 June 2022).

62 Infocomm Media Development Authority (Singapore grows trust in the digital environment). See https://www.imda.gov.sg/news-and-events/Media-Room/Media-Releases/2022/Singapore-grows-trust-in-the-digital-environment (accessed 5 June 2022).

63 The DTC will serve as implementation partner for an international collaboration between the Centre of Expertise of Montreal for the Advancement of Artificial Intelligence (CEIMIA) and the Infocomm Media Development Authority (IMDA) in Singapore. This partnership seeks to develop solutions to demonstrate how PETs can help organisations leverage cross-institution and cross-border data.

BOX 1

PETs in financial services

A series of challenges, technology sprints and collaborative projects have propelled the development of PETs in financial services. The World Economic Forum has outlined potential uses for PETs in determining creditworthiness, identifying collusion, or flagging fraudulent transactions between multiple banks[64]. Financial information sharing is key in tackling financial crime, which amounts to around $1.6 trillion annually (between 2-5% of the global GDP). This requires collaboration and data sharing in a way that safeguards client data, adheres to legal requirements and does not compromise competitive advantage of banking institutions.

In the UK, the Financial Conduct Authority (FCA) explored potential use cases for PETs such as secure multi-party computation in enabling data-based financial crime detection and prevention, launching a TechSprint on Global Anti-Money Laundering and Financial Crime in July 2019[65, 66].

This event included over 140 active participants, and concluded with ten proofs of concept, including:

- Using homomorphic encryption to enable banks to share and analyse sensitive information in order to uncover money-laundering networks, or to support the identification of existing and new financial crime typologies, or to allow banks to distinguish good from bad actors through question-and-answer when onboarding new clients;
- Using secure multi-party computation to uncover patterns of suspicious transactions across networks involving multiple banking institutions, or to highlight transactional mismatches in risky categories, such as account names;
- Using federated learning to improve risk assessment between multiple banks by enabling sharing of typologies;
- Using pseudonymised and hashed customer data to enable sharing and cross-referencing, to highlight potential areas of concern or for further investigation.

These demonstrations illustrate how PETs can be used for a particular end goal: to identify criminal behaviour in order to target enforcement action. While this use case is applauded by those working to tackle financial crime, it is worth considering how the same methods might be used for surveillance of other behaviours (for example, to profile customers for targeted advertisements, or for enhanced credit scoring).

64 World Economic Forum. 2019 The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value). See https://www3.weforum.org/docs/WEF_Next_Gen_Data_Sharing_Financial_Services.pdf (accessed 20 September 2022).

65 Financial Conduct Authority (2019 Global AML and Financial Crime TechSprint). See https://www.fca.org.uk/events/techsprints/2019-global-aml-and-financial-crime-techsprint (accessed 20 September 2022).

66 Cook N. 2019 It takes a network to defeat a network: tech in the fight against financial crime. *Royal Society blog*. 19 September 2022. See https://royalsociety.org/blog/2019/09/it-takes-a-network-to-defeat-a-network/ (accessed 16 February 2022).


ACCESS
CARD

Chapter two

Building the PETs marketplace

Left
© iStock / PeopleImages.

Building the PETs marketplace

As highlighted in market research commissioned by the Royal Society and CDEI, the market for PETs is nascent[67]. However, a growing number of documented examples demonstrate PETs already being used in a range of contexts[68], with a substantial number of large organisations expected to use one or more privacy-enhanced computation techniques by 2025, particularly in secure cloud infrastructures[69]. In addition to safeguarding personal data (which is required by data protection legislation), PETs are increasingly used wherever data is sufficiently valuable (for example, where data is tied to intellectual property or natural resource management).

PETs are rapidly evolving through private enterprise, as well as significant third sector and open initiatives. The development of the technology is thus greater than might be expected, given the modest size of the PETs market[70]. While this chapter explores the UK public sector market for PETs, it does not fully consider how PETs might shape future digital and data markets at large. In some cases, PETs negate the need to make copies of datasets, allowing data holders to provide insights as-a-service and potentially disincentivising open data approaches. Considering the potentially disruptive nature of PETs in this way, further research is required to understand the full implications of PETs in digital and data markets.

PETs for compliance and privacy

Neither EU nor UK data protection regulation explicitly mention PETs (nor 'privacy'). However, compliance with data protection law is a substantial motivating factor for organisations using data protection approaches. One investment firm contends that the EU GDPR has 'created the enterprise privacy market'[71]. Data processors want to understand how PETs can help them in compliance (particularly where data analysis is a weakness in the data lifecycle).

While privacy challenges are risk-related, they are not always assessed as commercial problems[72], particularly where the use of data is not commercially motivated (or where data use is altogether optional). Many data-holding organisations already use secure cloud services and analytics by default, and PETs are unlikely to be more cost-effective security tools in the near-term. In the wider marketplace, collaborative analysis may provide the most compelling business case for these technologies.

67 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

68 Centre for Data Ethics and Innovation (Privacy Enhancing Technologies Adoption Guide). See https://cdeiuk.github.io/pets-adoption-guide/ (accessed 20 September 2022).

69 Gartner (Gartner identifies the top strategic technology trends for 2022). See https://www.gartner.com/en/newsroom/press-releases/2021-10-18-gartner-identifies-the-top-strategic-technology-trends-for-2022 (accessed 20 September 2022). Note that in Gartner's analysis PETs are defined similarly to this report.

70 Lunar Ventures (Lundy-Bryan L). 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

71 *Ibid.*

72 *Ibid.*

PETs in collaborative analysis

Collaborative analysis (including collaborative computing[73] and collaborative learning[74]) is a growing area of interest in PETs applications. Researchers requiring data to generate insights, or to 'fuel' machine learning and other AI applications, can leverage PETs to establish data partnerships – effectively augmenting the data available to them. For example, organisations with a mandate to use data for public good are using PETs to make in-house data usable for external analysts[75]; cross-sector partnerships between crime agencies and human rights NGOs involves the pooling of datasets for analysis without revealing their contents to one another[76], enabling efficient, collective intelligence between analysts who do not see the original data.

Data availability and access is a priority for public sector bodies with remit to use data for public benefit, provision of services or to provide digital functions. For example, the Greater London Authority's London Datastore is designed to proactively link data assets to generate insights. Likewise, DataLoch – a service developed between the University of Edinburgh and NHS Lothian – aims to encourage 'non-typical researchers', such as charitable organisations, to use in-house health and social care data for the region of South-East Scotland. In interviews, PETs for collaborative analysis were seen by such public sector bodies as possible methods for reaching these aims; however, no examples of this application of PETs were identified by the UK organisations interviewed[77].

73 *Ibid.*

74 Melis L, Song C, De Cristofaro E, Shmatikov V. 2018 Inference attacks against collaborative learning. Preprint. See https://www.researchgate.net/publication/325074745_Inference_Attacks_Against_Collaborative_Learning (accessed 20 September 2022).

75 See Use case 1.1, page 57.

76 See Use case 6, page 97.

77 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

Legal and technical friction points prevent timely and straightforward access to public sector data, limiting its value as a public resource. PETs that allow the sending or processing of datasets internationally could be key to realising the value of data use across institutions and borders, which has been estimated to be between $3-5 trillion USD annually[78]. Governments and data-holding organisations are beginning to understand this value in terms of both economic and social benefits, and are seeking technology-based tools to enable collaboration[79]. The same PETs could also enhance data use across departments *within* an organisation, whether for reuse or when subject to further restrictions (as with International Traffic in Arms Regulations compliance in the US).

For these reasons, collaborative analysis has been predicted by one firm as the largest new technology market to develop in the current decade[80]. Cloud services are one substantial market already being impacted through the widespread use of Trusted Execution Environments (TEEs), which allow for data processing and analysis in a secure environment with restricted access[81]. TEEs can provide an application domain for SMPC, enabling collaborative analysis of confidential datasets[82]. Given its role in secure and collaborative analysis, confidential cloud could be an area of significant market growth in the near future[83, 84].

78 McKinsey. 2013 Collaborating for the common good: Navigating public-private data partnerships. See https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/collaborating-for-the-common-good#:~:text=Overall%2C%20McKinsey%20estimates%20that%20connecting (accessed 18 July 2022).

79 World Economic Forum. 2019 The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value). See https://www3.weforum.org/docs/WEF_Next_Gen_Data_Sharing_Financial_Services.pdf (accessed 20 September 2022).

80 Lunar Ventures (Lundy-Bryan L.) 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

81 Gartner (Gartner Top Strategic Technology Trends for 2021). See https://www.gartner.com/smarterwithgartner/gartner-top-strategic-technology-trends-for-2021 (accessed 26 September 2022).

82 Geppert T, Deml S, Sturzenegger D, Ebert N. 2022 Trusted Execution Environments: Applications and Organizational Challenges. Front. Comput. Sci. 4 (https://doi.org/10.3389/fcomp.2022.930741)

83 Gartner (Gartner Top Strategic Technology Trends for 2021). See https://www.gartner.com/smarterwithgartner/gartner-top-strategic-technology-trends-for-2021 (accessed 26 September 2022).

84 The Confidential Computing Consortium, which is run by the Linux Foundation, is promoting the use of TEEs in cloud services internationally. The Consortium includes every large cloud provider (Alibaba, Baidu, Google Clous, Microsoft, Tencent), demonstrating confidential computing as a priority to leaders in digital technology. Confidential Computing Consortium Defining and Enabling Confidential Computing (Overview). See https://confidentialcomputing.io/wp-content/uploads/sites/85/2019/12/CCC_Overview.pdf (accessed 15 March 2022).

Barriers to PETs adoption: User awareness and understanding in the UK public sector

A number of barriers prevent the widespread use of PETs for data protection and collaborative data analysis in the UK public sector. The first obstacle is general knowledge and awareness of PETs, their benefits and potential use cases[85, 86]. Researchers and analysts are often familiar with traditional privacy techniques (such as anonymisation, pseudonymisation, encryption and data minimisation); for some, it is unclear what PETs can add to these approaches.

PETs that enable collaborative analysis include some of the most technically complex and least used to date (such as secure multi-party computation and federated learning). While PETs may be some of the most promising, the risk inherent to using new and poorly understood technologies is a strong disincentive to adoption: few organisations, particularly in the public sector, are prepared to experiment with privacy[87].

A lack of understanding around PETs within wider data protection requirements means stakeholders are hesitant to adopt them[88]. For example, anonymised personal data is not subject to the principles of data protection requirements detailed in the UK GDPR or EU GDPR[89, 90]; however, in the UK, there is no universal test of anonymity. Technology-specific guidance may be useful in interpreting requirements and best practices in emerging technologies, for example, how archived synthetic data should be handled[91]. Currently, organisations must turn to assessments by internal or external parties for guidance. These uncertainties lead to a culture of risk-aversion described by some UK public bodies[92]. Without assurance or technical standards, some question the genuine security PETs offer, particularly where privacy threats and adversaries are undefined or hypothetical[93].

85 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

86 Lunar Ventures (Lundy-Bryan L.) 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

87 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

88 *Ibid*.

89 Information Commissioner's Office (What is personal data?). See https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/what-is-personal-data/#:~:text=If%20personal%20data%20can%20be,subject%20to%20the%20UK%20GDPR (accessed 20 September 2022).

90 GDPR Info (EU GDPR Recital 26). See https://gdpr-info.eu/recitals/no-26/ (accessed 20 September 2022).

91 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

92 *Ibid*.

93 *Ibid*.

Where organisations are unable to assess privacy trade-offs for a given PET or application, cost-benefit analysis becomes impractical.

Where organisations are unable to assess privacy trade-offs for a given PET or application, cost-benefit analysis becomes impractical. As a result, the PETs value proposition remains speculative and the business case for adopting PETs is unclear. Demonstrations are needed to establish the potential benefit of PETs, for example, through case studies that include cost-benefit analyses[94]. The use cases and examples in Chapter Four (page 56) provide a starting point for such an approach.

According to those interviewed, market confidence could be enhanced through better data readiness and the development of standards (Chapter Three)[95]. PETs are subject to relevant legal frameworks and existing regulators, such as the ICO in the UK. However, they are not specifically regulated as technologies, and their efficacy is 'illegible' to non-experts. Standards could be followed by assurance and certifications. Implementation frameworks for PETs would allow some elements of decision-making to be outsourced, although additional expertise will likely be required in practice[96].

Other barriers are institutional in nature. For example, where technical expertise does exist in-house, these individuals are often organisationally removed from decision-makers[97]. Foundational data governance issues, such as data quality and interoperability, are primary concerns for many organisations and, as such new, unknown technologies are deprioritised. Compute power is also a practical limiting factor, particularly with energy-intensive approaches such as homomorphic encryption[98].

Barriers to PETs adoption: Vendors and expertise

The development of PETs requires a deep understanding of cryptography. However, unlike other computing-related fields (such as software engineering), the cutting edge of cryptography remains largely in academia. This leads to a gap between cryptography expertise and market drivers, such as cost and convenience. As a result, theoretical cryptography 'risks over-serving the market on security'[99]. Bridging the gap between cryptography talent and entrepreneurs could create viable PETs vendors.

94 *Ibid.*

95 *Ibid.*

96 *Ibid.*

97 *Ibid.*

98 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

99 Lunar Ventures (Lundy-Bryan L.) 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

Professional certifications and online courses for privacy professionals could integrate a PETs primer into existing courses to raise awareness and expertise in the profession. For example, the Alliance for Data Science Professionals[100], which defines standards to ensure ethical and well-governed data use, could consider PETs in designing standards around data stewardship and analysis.

Modules on general and specific PETs are appearing in university syllabuses, particularly at the postgraduate study level. Several of the universities within the Academic Centres of Excellence in Cyber Security Research have a focus on privacy, and PETs and privacy is a remit of the doctoral training. In more informal education, online courses are starting to appear such as OpenMined's 'Our Privacy Opportunity' 'Foundations of Private Computation' and 'Introduction to Remote Data Science'[101]. These can go a long way in raising general awareness and inspiring use cases.

Conclusions

A flourishing PETs market will require both trust in the technology and users' ability to discern appropriate applications. PETs vendors can help address scepticism by integrating PETs in wider data governance approaches, rather than promoting one-size-fits-all solutions. Where public sentiment around the use of PETs is unknown, further research – including focus groups or public dialogues – could be used toward ensuring end-user acceptance of (and demand for) the technologies[102].

Today, businesses are incentivised to accumulate data for exclusive use. PETs may engender new business models, for example data or analytics as-a-service. This could entail a data-holding organisation allowing clients to query or run analyses on in-house datasets. This could be done using PETs that do not reveal the data, only the insights or solutions gathered from the query or analysis. Data is not transferred and remains unseen by the external client.

In this way, PETs may enable a shift from data sharing (through agreements or otherwise) to a dynamic data processing and analytics market[103], such as through 'commissioned analyses'[104]. It will be important to consider this potential shift and incentivise organisations to utilise PETs for collaboration, rather than data gatekeeping.

100 British Computing Society (The Alliance for Data Science Professionals: Memorandum of Understanding July 2021). See https://www.bcs.org/media/7536/alliance-data-science-mou.pdf (accessed 2 September 2022).

101 OpenMined (The Private AI Series). See https://courses.openmined.org/ (accessed 7 October 2022).

102 The Royal Society. Creating trusted and resilient data systems: The public perspective. (to be published online in 2023)

103 Lunar Ventures (Lundy-Bryan L.) 2021 Privacy Enhancing Technologies: Part 2—the coming age of collaborative computing. See https://docsend.com/view/db577xmkswv9ujap?submissionGuid=650e684f-93eb-4cee-99e8-12a92d5d88a0 (accessed 20 September 2022).

104 London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/

TABLE 2

PETs described in this report and their function with regard to security and collaborative analysis[105].

	Homomorphic encryption	Trusted Execution Environments (TEEs)
What does this PET do?	Allows the use, or analysis, of encrypted data without decrypting it.	Allows data to be used or analysed within a secure, isolated environment.
In what circumstances would it be used?	To create meaningful insights in computation without revealing the contents of a dataset to those running the analysis (which could be done by a trusted third-party).	When data needs to be stored securely, or to generate insights from data without revealing the dataset to party running the analysis or hosting the TEE.
Whose data is being protected and from whom?	The data held by the institution running the computation is being protected from whoever runs the analysis, whether a third-party or the institution themselves. If the third-party were to act in bad faith, they would not have access to the data in question.	The data held by the institution running the research can only be decrypted and used within the TEE, and only used by approved code. The TEE is protected from outside environment, including the operating system and admin users.
Whose interests are being protected and what are they?	**The data controller** They have an interest to carry out their computation in the safest and most effective way possible. **The data subjects** Those who the data is about have an interest in making sure their data is not accessed by bad actors.	**The data controller** They have an interest to carry out their research in the safest and most effective way possible. **The data subjects** Those who the data is about have an interest in making sure their data is not accessed by bad actors.
Relevance to security and collaborative analysis	**Security** Data is protected from unauthorised access.	**Security** Data is protected from unauthorised access.

105 Modified from Hattusia 2022 The current state of assurance in establishing trust in PETs. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/ (accessed 20 September 2022).

106 A type of HE called 'multi-key FHE' can perform a similar function: several parties each have a secret key and can encrypt their own data, which is sent to a trusted third party who for computation. The result can be decrypted by all parties who contributed data to the process.

	Secure multi-party computation (SMPC)	Differential privacy	Federated l earning
	This allows multiple parties to run analysis on their combined data, without revealing the contents of the data to each other[106].	Mostly for use with large data sets, DP allows institutions to reveal data or derived information to others *without* revealing sensitive information about the groups or individuals represented in the data set.	This allows for the training of an algorithm across multiple devices or datasets held on servers.
	Removes the need for a trusted central authority that would have access to everyone's data. Rather, multiple organisations can keep their data sets private from each other, but still run joint analysis on the combined data.	An institution may want to share analytical insights that they have derived from their data with another group or with the public, but their data set contains sensitive information which should be kept private.	An organisation wants to train a machine learning model, but has limited training data available. They 'send' the model to remote datasets for training; the model returns having benefitted from those datasets.
	Each collaborating organisation holds data about individuals (or other sensitive data), and that data is protected from those collaborating on analysis. The data also is protected from any potential misconduct or incompetence from any of the parties.	Sensitive information about the groups or individuals present in the dataset is being protected from whoever the data is being shared with or analysed by, whether that's a trusted third-party, the general public, or the institution themselves.	Each collaborating organisation holds data about individuals (or other sensitive data) and that data is protected from those collaborating on analysis. Only the trained model is exchanged.
	The collaborating organisations They have an interest to carry out their research in the safest and most effective way possible. **The data subjects** Those who the data is about have an interest in making sure their data is not accessed by bad actors.	**The data controller** They have an interest to carry out their research and share data in the safest and most effective way possible. **The data subjects** Those who the data is about have an interest in making sure their data is not accessed by bad actors.	**The collaborating organisations** They have an interest to carry out their research in the safest and most effective way possible. **The data subjects** Those who the data is about have an interest in making sure their data is not accessed by bad actors.
	Security Data is protected from unauthorised access. **Collaborative analysis** Multiple parties can work on datasets held by parties of 'mutual distrust'; the data remains safe from unwarranted interference.	**Security** Data is protected from unauthorised access. **Collaborative analysis** There is potential for open access to the data without revealing the presence or attributes of individuals.	**Security** Data is protected from unauthorised access. **Collaborative analysis** Federated learning is also called collaborative learning; multiple parties are required.

Chapter three

Standards, assessments and assurance in PETs

Left
© iStock / olaser.

Standards, assessments and assurance in PETs

PETs are generally best used in a systems approach to data privacy by addressing the twin goals of compliance and trust.

The Royal Society's 2019 report, *Protecting privacy in practice* suggested a system of standards and certification for PETs may provide a pathway for assurance, leading to wider adoption of the technologies. Similar initiatives have shaped the development and uptake of emerging technologies (such as cybersecurity products) and global information sharing platforms (as with the protocols that continue to enable the internet). However, PETs are unlike cybersecurity in that they address highly contextual, often intersectional, privacy concerns[107].

This chapter reviews the role of trust and assurance in PETs implementation[108]. The review finds that, given their current state of maturation, PETs are generally best used in a systems approach to data privacy by addressing the twin goals of compliance and trust[109]. Compliance is adherence to legal and statutory obligations (such as the UK GDPR) to avoid penalties, while trust enables data flows and collaboration.

The 2020 Edelman Trust Barometer[110] identified two types of trust:

- Moral – the trustor believes the trustee can articulate and act on the best interests of the trustor and;
- Competence – the trustor believes the trustee has the ability to deliver on what has been agreed.

Trust in privacy systems is similarly twofold (see Table 3):

- Trust that the PET will be used in a way that protects the rights of the data subject (moral) and;
- Trust in the technical ability of PET as a security tool (competence).

Currently, only technical standards exist for PETs (and these are few). These pertain to the technical capabilities of PETs in achieving security (trust in competency). The following sections explore data privacy frameworks, technical standards and assurances in fostering the rapid and responsible use of PETs.

PETs and assurance: The role of standards

Assurance in new technologies takes many forms. Certifications, Kitemarks and other formal guarantees for products are perhaps most well-known. These official marks of assurance require external audit based on formal *standards*, which set out requirements for a product or system.

Global standards have been effective in cybersecurity and privacy; likewise, encryption-based PETs may rely on encryption standards. Similar approaches may be feasible where risk of disclosure is quantifiable, such as with differential privacy.

107 Hattusia 2022 The current state of assurance in establishing trust in PETs. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/ See also Table 3.

108 *Ibid*.

109 Zimmermann C. 2022 Part 1: What is Privacy Engineering? *The Privacy Blog*. 10 May 2022. See https://the-privacy-blog.eu/2022/05/10/part1-what-is-privacy-engineering/ (accessed 20 September 2022).

110 Edelman (2020 Trust Barometer). See https://www.edelman.com/trust/2020-trust-barometer (accessed 15 February 2022).

TABLE 3

Assurances and trust relationships in the use of PETs in privacy-preserving data governance systems.

Trustors	Trustees	Moral trustworthiness	Trust in competence	Assurances needed
PETs users (eg, engineers or data scientists)	The technology itself; collaborators; external actors; organisation's executives (decision-makers) or PETs vendors (if using).	Have the executives or PETs vendors prescribed the right PET for the application, such that it functions in a privacy-preserving way?	Will the PET fulfil its expected technical function? Will the data remain secure from outside actors who want access to it?	**Technical assurance** Technological specifications demonstrating the PET will function as intended. **Assurance in the application** The use of the PET is appropriate for the given use case; the PET is part of wider responsible data governance.
Executives and PETs vendors (those 'diagnosing' use cases and deploying PETs)	PETs users; PETs vendors; PETs developers; the technology itself.	N/A	Are the developers competent in delivering a fit-for-purpose technology? Will the PET fulfil its expected function?	**Technical assurance** Professional qualifications detailing the PET user's ability. **Technical assurance** Technological specifications demonstrating the PET will function as intended.
Data subjects (the people whom the data is about)	The data governance ecosystem of organisations that collect and use their data	Will personal data be used in accordance with intent, and not lead to increased surveillance and exploitation?	Will data remain safe from interference from unauthorised users?	**Assurance in the application** The PET is used as part of wider responsible data governance.

A standardisation of approach to PETs will be essential in:

- Developing higher-level guidance for 'best practice' and codes of conduct;
- facilitating the early phases of PETs adoption;
- incorporating PETs into privacy frameworks and impact assessments in an informed and responsible manner.

The National Institute of Standards and Technology (NIST) also highlight the need for technical standards[111]. NIST promotes the standardisation of technologies that underpin PETs (such as secret-sharing and encryption regimes), alongside a guidance-based approach to the standardised use of PETs themselves.

Process standards for data protection

Process standards can be used to assist in compliance with data protection law and general privacy protection. Privacy frameworks are one example; these are built around a set of questions or controls: points that must be considered and addressed in building an effective system. This structure allows frameworks to specifically address data protection laws, such as the UK GDPR.

A popular privacy framework approach entails:

1. mapping of information flows.
2. conducting a privacy risk assessment (or 'privacy impact assessment').
3. strategising to manage identified risks.

Frameworks do not prescribe methods or technologies for implementation; rather, the implementer may decide to use classic and emerging PETs to fulfil the framework requirements.

Existing standards, guidance and frameworks that address privacy systems are highlighted in Table 4.

The pathway to PETs standards

Standards for PETs are being developed through a range of international, national and sector-specific SDOs. In addition, there is an emergence of open standards initiatives. These initiatives seek to make standards on PETs accessible by anyone and can entail a collaborative approach to standards development, involving community-led groups and stakeholders from government, industry and academia. There is a growing movement for this standardisation approach, particularly within emerging technologies. An example of this is the UK's AI Standards Hub which aims to create practical tools and standards to improve the governance of AI[112].

111 National Institute of Standards and Technology (*Roadmap to the Privacy Framework*). See https://www.nist.gov/privacy-framework/roadmap (accessed 15 March 2022).

112 The AI Standards Hub is led by the Alan Turing Institute with support from the British Standards Institution and the National Physical Laboratory. HM Government (New UK initiative to shape global standards for Artificial Intelligence). See https://www.gov.uk/government/news/new-uk-initiative-to-shape-global-standards-for-artificial-intelligence (accessed 19 March 2022).

BOX 2

Lessons from standardisation: Open standards and the internet

The internet operates smoothly thanks to consensus-driven protocols that continue to be developed by a vast community of technologists. The Internet Engineering Task Force (IETF) is an informal, volunteer-led group that serves as the standards body for the Internet. The IETF has played a critical role in the development of the internet without a formal, centralised standards body. They developed such inter-domain standards such as HTTP (HyperText Transfer Protocol) and TCP (Transmission Control Protocol), allowing users to access the same internet and transfer data around the world.

Open standards can be led by technologists, who know what is technically possible and can propose standards to adapt and meet new legal or other requirements. They may also benefit from additional inputs from other stakeholders. In being 'open', standards are made available for anyone who wishes to use them. Innovators can then use these protocols in the development of new technology; assurance against such standards becomes a marketable added value to such organisations.

The development of open standards in PETs will be crucial in ensuring PETs work for everyone by allowing for the global and interoperable use of data.

TABLE 4

Example standards and guidance relevant to data privacy

Name	Number	Standards development organisation	Date published
Information technology – Security techniques – Privacy framework	ISO/IEC 29100:2011/ AMD 1:2018	ISO and IEC	June 2018
Information technology – Security techniques – Privacy architecture framework	ISO/IEC 29101:2018	ISO and IEC	Nov 2018
Information technology – Security techniques – Information security management systems – Requirements	ISO/IEC 27001:2013	ISO and IEC	Oct 2013, will be replaced by ISO/IEC FDIS 27001 (under development)
Information security, cybersecurity and privacy protection – Information security controls	ISO/IEC 27002:2022	ISO and IEC	Feb 2022
Security techniques – Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management – Requirements and guidelines	ISO/IEC 27701:2019	ISO and IEC	Aug 2019
Information technology – Security techniques – Privacy capability assessment model	ISO/IEC 29190:2015	ISO and IEC	Aug 2015, reviewed 2021
Data protection – Specification for a personal information management system	BS 10012:2017+A1:2018	BSI	Jul 2018
IEEE Standard for Data Privacy Process	IEEE 7002-2022	IEEE	Apr 2022
Privacy enhancing data de-identification terminology and classification of techniques	ISO/IEC 20889:2018	ISO and IEC	Nov 2018
Privacy enhancing data de-identification framework	ISO/IEC DIS 27559	ISO and IEC	TBC
Anonymisation, pseudonymisation and privacy enhancing technologies guidance		ICO	TBC
Information technology – Security techniques – Code of practice for personally identifiable information protection	ISO/IEC 29151:2017	ISO and IEC	Aug 2017
De-Identification of Personal Information	NISTIR 8053	NIST	Oct 2015

113 See for example the Professional Evaluation and Certification board training courses https://pecb.com/en/education-and-certification-for-individuals.

	Training available	Description	Reference to PETs
	Certificate	Privacy framework for Personal Identifiable Information (PII) use	
		Focus on ICT systems for PII	PETs used as privacy controls; refers to PETs 'such as pseudonymization, anonymization or secret sharing'. Briefly mentions HE in regards to encryption.
	Certificate	Cyber security focussed standard with related standards that include guidance for auditing.	
	Courses available[113]	Includes reference materials for security controls and implementation guidance, used regularly in conjunction with ISO/IEC 27001.	
	Certificate	Guidance for Privacy Information Management Systems (PIMS), building on ISO 27001	
		Provides high-level guidance for organisations to assess their management of privacy-related processes.	
	Yes, no certificate	Guidance for PIMS with specific application to UK law (also a mapping to ISO/IEC 27701 exists). Training covered under GDPR implementer/Self Assessor training.	
		Requirements for a systems/software engineering for privacy	'Organizations should also put in place policies on the following: Privacy enhancing technologies and techniques: Which technologies the organization uses, and how and when these technologies are used.'
		Description of privacy enhancing data de-identification techniques and measures to be used in accordance with ISO/IEC 29100.	Content on homomorphic encryption, differential privacy and synthetic data.
		Framework for identifying and mitigating re-identification risks, building on ISO/IEC 20889.	
		Upcoming guidance on anonymisation and PETs, suggests motivated intruder tests.	Forthcoming.
		Information security guidelines specifically for PII.	Recommends to 'consider whether, and which, privacy enhancing technologies (PETs) may be used.'
		Guidance on de-identification, suggests motivated intruder tests.	Suggestions of use of differential privacy and synthetic data.

TABLE 4 (continued)

Example standards and guidance relevant to data privacy

Name	Number	Standards development organisation	Date published
The Anonymisation Decision-Making Framework: European Practitioners' Guide		UK Anonymisation Network	Jul 2012
The NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management		NIST	Jan 2020
Roadmap for Advancing the NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management		NIST	Jan 2020
PETs Adoption Guide		CDEI	Jul 2021

	Training available	Description	Reference to PETs
		Framework for anonymisation by an open group, led by academics at the University of Manchester.	Suggestions of use of differential privacy and synthetic data.
		Privacy framework based on NIST's successful cybersecurity framework.	-
		Roadmap for NIST's Privacy framework highlighting challenges	Passing reference to differential privacy
		Guidance including flowchart for identifying appropriate PETs	

TABLE 5

Existing and forthcoming initiatives related to PETs standards development

PET	Name	Number	SDO
HE	IT Security techniques – Encryption algorithms – Part 6: Homomorphic encryption	ISO/IEC 18033-6:2019	ISO/IEC
HE	Information security – Encryption algorithms – Part 8: Fully Homomorphic Encryption	ISO/IEC AWI 18033-8	ISO/IEC
HE	Homomorphic Encryption Security Standard		Open
TEEs	Advanced Trusted Environment	OMTP TR1	OMTP
TEEs	TEE Trusted User Interface Low-level API	GPD_SPE_055	GlobalPlatform
TEEs	PSA Certified IoT Security Framework		PSA Certified
TEEs	IEEE Standard for Technical Framework and Requirements of Trusted Execution Environment based Shared Machine Learning	IEEE 2830-2021	IEEE
TEEs	Standard for Secure Computing Based on Trusted Execution Environment	P2952	IEEE
TEEs	Information technology – Trusted platform module library	ISO/IEC 11889-1:2015	ISO/IEC
DP	Privacy enhancing data de-identification terminology and classification of techniques	ISO/IEC 20889:2018	ISO/IEC
DP	NIST blog series		NIST
DP	ε KTELO: A Framework for Defining Differentially-Private Computations		Academic
SMPC	Information technology – Security techniques – Secret sharing – Part 1: General	ISO/IEC 19592-1:2016	ISO/IEC
SMPC	Information technology – Security techniques – Secret sharing – Part 2: Fundamental mechanisms	ISO/IEC 19592-2:2017	
SMPC	Information security – Secure multi-party computation – Part 1: General	ISO/IEC CD 4922-1.2	ISO/IEC
SMPC	Information security – Secure multi-party computation – Part 2: Mechanisms based on secret sharing.	ISO/IEC WD 4922-2.3	ISO/IEC
SMPC	IEEE Recommended Practice for Secure Multi-Party Computation	IEEE 2842-2021	IEEE
SD	Synthetic Data – Industry Connections	IC21-013-01	IEEE
SD	Synthetic Data – what, why and how?		The Alan Turing Institute

	Date	Type	Comment
	May 2019	Standard	Looks at two PHE algorithms, appropriate parameters and the process of homomorphically operating on the encrypted data.
		Standard	Continuation of ISO/IEC 18033-6:2019 for FHE
	Mar 2018	Standard	Standard produced by an open consortium of industry, government and academia.
	May 2009	Standard	Originally made for mobile phone TEEs, but applicable more generally, setting out core requirements, best practice and examples.
	Oct 2018	Standard	Highly technical standard used extensively in industry products.
		Standard	Internet of Things (IoT) certification for hardware, software and devices. This is used In the standardisation of TEE hardware (e.g. ARM TrustZone).
	Oct 2021	Standard	Standard on the applied use of TEEs in privacy preserving machine learning done using third parties and MPC.
		Project	Standard on cyber security application of TEEs.
	Aug 2015	Standard	A four-part standard on trusted platform modules, a related technology, developed by an industry collaboration and later adopted by ISO/IEC.
	Nov 2018	Guidance	Discusses differential privacy as a metric and also related noise addition methods.
	Dec 2021	Project	General explainer on DP in 12 parts, concluding with a statement that they have plans to use it as a foundation on which to develop technical guidelines.
	May 2018	Guidance	Example academic paper sharing a framework for developing DP algorithms.
	Nov 2016	Guidance	Sets out terminology.
	Oct 2017	Standard	Covers five secret sharing algorithms that meet requirements of message confidentiality and recoverability.
		Standard	Incoming standard on SPMC.
		Standard	Incoming standard on SPMC specifically where it uses secret sharing.
	Nov 2021	Standard	A ‘technical framework’ for SMPC including security levels and use cases.
		Project	Industry (and academic) collaboration, sets out goals to produce best practice and terminology guidance for a standard project authorization request for a synthetic data privacy and accuracy standard.
	May 2022	Guidance	An academic review of synthetic data as a technology highlighting some of the challenges.

Measuring privacy and utility in PETs

One potential barrier in developing PETs standards is achieving consensus on metrics for privacy and utility. There are many different metrics that can be used for privacy; one review categorises over 80 privacy metrics and suggests a method of how to choose them[114].

The cybersecurity community uses security metrics. Encryption, for example, has security metrics such as key length, which estimate the computing power it would take to break encryption and therefore the degree of security provided. SDOs are also interested in privacy metrics, as in *Privacy enhancing data de-identification terminology and classification of techniques* (ISO/IEC 20889), which concerns differential privacy and its use as a measure. However, privacy-utility trade-offs vary according to context, making metrics and thresholds difficult to generalise[115, 116].

Using a single privacy metric also risks over-simplification, failing to adequately address all relevant harms (as privacy metrics can only account for one harm at a time).

Threat modelling can be used to identify potential risks, attacks or vulnerabilities in a data governance system. Threat models are constantly evolving as attacks reach new levels of sophistication. For example, *anonymisation* originally meant zero risk of reidentification. However, increasingly sophisticated reidentification techniques, such as those that make use of statistical approaches and publicly available datasets, are changing the requirements of adequate anonymisation[117].

Considering these constraints, the best approach may be technical standards and metrics where feasible (as with encryption or noise addition algorithms), complemented by scenario-based guidance, assessment protocols and codes of conduct.

114 Wagner I, Eckhoff D. 2018 Technical Privacy Metrics: A Systematic Survey. See https://arxiv.org/abs/1512.00327 (accessed 20 September 2022). Note that more general mathematical approaches also exist, which aim for a definition of privacy more like that of epsilon in differential privacy. One example of this is Pufferfish, a self-professed framework for mathematical privacy definitions, which can be used in the context of PETs: Kifer D, Machanavajjhala A. 2014 Pufferfish: a framework for mathematical privacy definitions. ACM Transactions on Database Systems 39, 1—36. (https://doi.org/10.1145/2514689).

115 Lee J, Clifton C. 2011 How Much Is Enough? Choosing ε for Differential Privacy (conference paper). See https://link.springer.com/chapter/10.1007/978-3-642-24861-0_22 (accessed 23 April 2022).

116 Abowd JM, Schmutte IM. 2019 An Economic Analysis of Privacy Protection and Statistical Accuracy as Social Choices. Am Econ Rev. 109, 171—202.

117 A useful analysis of the UK's approach to anonymisation in data protection regulation can be found in: Bristow and Privitar. 2021 Introduction to Anonymisation. See https://www.bristows.com/app/uploads/2021/07/Introduction-to-Anonymisation-Privitar-and-Bristows.pdf (accessed 20 September 2022).

BOX 3

Motivated intruder tests

A type of attack-based risk assessment, this is an established method for assessing the efficacy of a privacy regime (or PET). This requires anticipation of 1) the technologies and methods that might be used to attack the data; 2) the vulnerabilities of a given PET to various attacks; 3) the kinds of knowledge that could enable the attack; 4) the goals of potential attacks and how they might cause harms. An exhaustive list and test of every attack is not feasible. Rather, it is important to know what kind of attacks are most possible and most likely. Because it is impossible to anticipate every scenario, even motivated intruder testing does not provide a guarantee of privacy. Nonetheless, it has been the primary legal test in determining whether data is identifiable or not.

Motivated intruder testing can provide a degree of assurance. However, this approach cannot provide a quantitative *measure* of assurance. In the past, a motivated intruder has been defined as someone without specialist skills or computing power, which may not be a realistic adversary for some data sets (such as highly desirable datasets). More explicit guidance on testing, including choosing what and how to test a PET, could be included either in process standards or PETs guidance.

Testing does not remove the need for expert users and developers. Social and educational infrastructure must be in place to educate data scientists (and privacy professionals) on PETs and risk assessment.


STOP

Chapter four

Use cases for PETs

Left
Elizabeth Line, London, UK. © Unsplash / Kevin Grieve.

Use cases for PETs

This chapter is comprised of a set of use cases highlighting the various roles PETs are playing – or could play – in real-world data governance scenarios. The use cases are intended to represent broad scenarios where PETs could help reach a wider data objective.

As these examples demonstrate, the role of PETs is not exclusively one of protecting privacy – rather, they can serve to enhance transparency, increase collaboration and strengthen data partnerships.

The efficacy and appropriateness of a PET in data governance is highly dependent on context. Therefore, the aim of these use cases is not to prescribe reproduceable solutions, but rather:

1. To inspire discussions between UK government, regulators and organisations that use data to consider how technology may play an enabling role in data governance, along with allowing faster and safer ways of partnering to find data-driven solutions to multidisciplinary challenges;
2. To illustrate the importance of context-based solutions and a privacy by design approach by including various types of data and circumstantial sensitivities (individual, commercial, national);
3. To showcase where PETs could play a critical difference in data-driven problem solving, allowing for data use that would otherwise be legally, technically or socially prohibitive.

Considerations and approach

These use cases were chosen for their relevance to significant real-world data-driven challenges. The choice of scenarios was informed by two workshops with a PETs Contact Group and validated through further discussions with stakeholders. They were developed with technical and legal input of the report's Working Group, as well as invited external experts and desk-based review.

The intention is for these cases to be an aide for anyone who relies on information flows to imagine how PETs could enhance a systems approach to data governance. The use cases are meant to explain PETs in various scenarios; they are not intended to be an endorsement or recommendation for action.

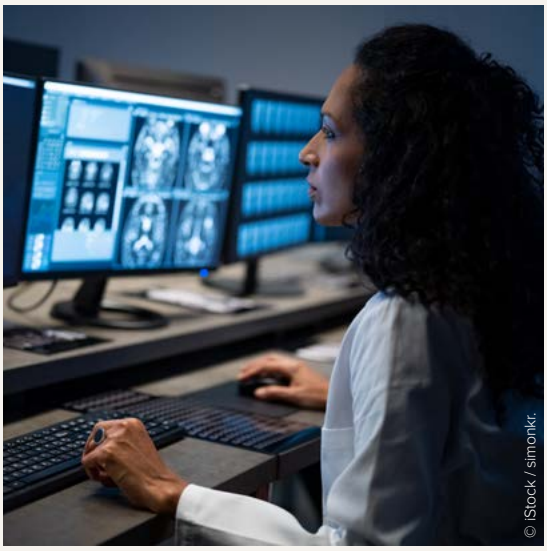
USE CASE 1.1

Privacy in biometric data for health research and diagnostics

The challenge

Recent advances in medical imaging, audio and AI have led to unprecedented possibilities in healthcare and research. This is especially true of the UK, where the public health system is replete with population-scale electronic patient records. These conditions, coupled with strong academic and research programmes, mean that the UK is well positioned to deliver timely and impactful health research and its translation to offer more effective treatments, track and prevent public health risks, utilising health data to improve and save lives[118].

Internationally, hospitals produce an estimated 50 petabytes[119] of data annually[120], though only 20% is structured for digitisation[121], let alone further research or analysis. The public benefit of utilising this joint resource is substantial, and AI-assisted analytics are essential for realising the value of big health data. Because patient-level health data is inherently personal, there is potential for public distrust if health data is misused and privacy is compromised.



© iStock / simonkr.

Anonymous data is not covered by current data protection law in the UK and EU. However, it is difficult to be certain that health data is anonymous, particularly in biometric and other non-textual data. Health data is subject to specific legal requirements in the UK, as well as the common law duty of confidentiality. The following three examples illustrate how PETs could help in meeting best practice standards in non-textual health data use, while making data more readily available for researchers and innovators[122].

118 HM Government (Life sciences industrial strategy update). See https://www.gov.uk/government/publications/life-sciences-industrial-strategy-update (accessed 15 March 2022).

119 One petabyte is roughly the equivalent of 500 billion pages of standard printed text.

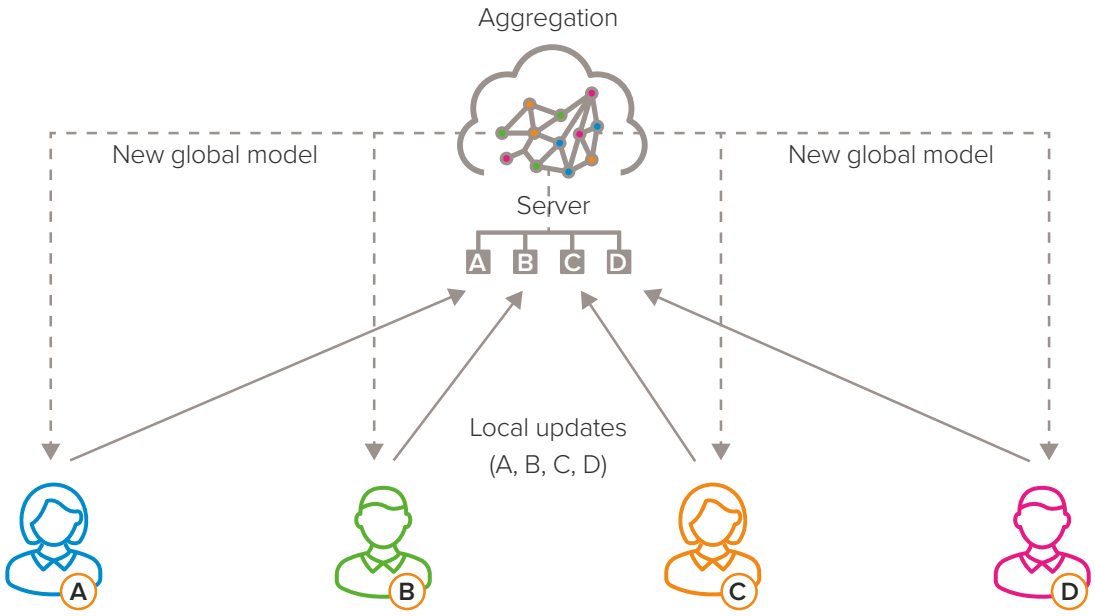
120 InfoDocket (How Large is the Digital Universe? How Fast is It Growing?). See https://www.infodocket.com/2014/04/16/how-large-is-the-digital-universe-how-fast-is-it-growing-2014-emc-digital-universe-study-now-available/ (accessed 20 September 2022).

121 HIT Consultant (Why unstructured data holds the key to intelligent healthcare systems). See https://hitconsultant.net/2015/03/31/tapping-unstructured-data-healthcares-biggest-hurdle-realized/#.XFvZ1lwvOUk (accessed 20 September 2022).

122 HM Government (National Data Strategy). See https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy (accessed 9 September 2022).

FIGURE 2

Federated machine learning



Preserving privacy in medical imaging for research and diagnostics

Magnetic Resonance Imaging (MRI) is a type of scan that produces detailed images of the inside of the body and internal organs by using strong magnetic fields and radio waves. The images produced by MRI scanning provide critical information in the diagnosis and staging of disease progression. Sets of MRI images can be used to train machine learning algorithms to detect certain features or abnormalities in images. This technology can be deployed to screen large numbers of images for research purposes: identifying patterns that link variables like patient behaviour, genetics, or environmental factors with brain function.

MRI imaging and metadata can reveal sensitive information about a patient. Indeed, even an individual's presence in a dataset may be sensitive. While the images themselves may be de-identified through removal of names, addresses and scan date, neuroimages can sometimes be reidentified (as demonstrated in a 2019 Mayo Clinic study)[123].

123 Schwarz C G *et al*. 2019 Identification of Anonymous MRI Research Participants with Face-Recognition Software. *N Engl J Med*. 381, 1684—1686. (https://doi.org/10.1056/nejmc1908881

Privacy solutions that enable collaboration

Federated learning is a type of remote execution in which models are 'sent' to remote data-holding machines (eg, servers) for local training. This can allow researchers to use data at other sites for training models without accessing those data sets. For example, if researchers at different universities hold neuroimaging data, a federated approach would allow them to train models on all participants' imaging data, even as that data remains 'invisible' to analysts. This is an example of Federated Machine Learning (see Figure 2).

There are two approaches to accomplishing Federated Machine Learning in this case:

- In one approach, each site analyses its own data and builds a model; the model is then shared to a remote, centralised location (a node) common to all researchers involved. This node then combines all models into one 'global' model and shares it back to each site, where researchers can use the new, improved model[124];
- In a second approach, the model is built iteratively, where the remote node and local nodes take turns sending and returning information[125].

In either approach, all users' models are improved by 'learning' from remote datasets, which are themselves never revealed. By using federated learning, raw data is not shared, which rules out the most common issues associated with data protection. At the same time, federated learning does not offer perfect privacy; models are still vulnerable to some advanced attacks. These attacks may be of a sufficiently low risk to be acceptable to the parties such that they can proceed. Other safeguards may also be put in place. These could include detecting when repeated queries are made of an MRI dataset, which could be cross-referenced with public data to reidentify subjects.

124 In this approach, a single-shot algorithm can be used.

125 Each participant sends a gradient on its data set until the algorithm converges. Iterations use an optimisation routine (such as stochastic gradient descent). In this approach, a multi-shot algorithm can be used.

BOX 4

Collaborative Informatics and Neuroimaging Suite Toolkit for Anonymous Computation (COINSTAC)

COINSTAC[126], an open-source, cross-platform application created by the Center for Translational Research in Neuroimaging and Data Science (TReNDS) in Atlanta, Georgia, is one example illustrating how to overcome data access barriers in neuroimaging through federated learning and privacy preserving algorithms.

COINSTAC allows users who cannot directly share their data to collaboratively run open, reproduceable federated learning and coordinated pre-processing using software packages that can run in any environment (such as personal devices, private data centres, or public clouds). It uses containerised software (software which runs all necessary code within one environment that is executable regardless of host operating system and is therefore consistent across platforms). This software is available on GitHub under an MIT license[127].

COINSTAC developers have documented several case studies. In one study, a federated analysis using datasets from Europe and India found structural changes in brain grey matter linked to age, smoking, and body mass index (BMI) in adolescents[128]. Another case study uses a federated neural network classifier to differentiate smokers from non-smokers in resting-state functional MRI (fMRI) data. The federated models typically achieve results similar to those using pooled data and better than those drawing data only from isolated sites.

Additionally, TReNDS researchers are developing optimised algorithms for deep learning to reduce transmission bandwidth without sacrificing accuracy. In a third example, brain age estimation algorithms were trained to predict actual subject age using neuroimaging; this was then applied to estimate the biological brain age of new subjects[129]. This is useful because large gaps between estimation of biological brain age and actual age are potential biomarkers of brain disorders such as Alzheimer's disease. This model achieved results that were statistically equivalent to centralised models.

TReNDS is also currently developing a network of COINSTAC vaults, which will allow researchers to perform federated analysis with multiple large, curated datasets. This open science infrastructure will enable rapid data reuse, create more generalisable models on diverse datasets, and democratise research by removing barriers to entry for small or under-resourced groups.

126 Coinstac (Homepage). See https://coinstac.org/ (accessed 30 March 2022).

127 Github (Coinstac release v6.5.3). See https://github.com/trendscenter/coinstac (accessed 20 September 2022).

128 Gazula H *et al*. 2021 Decentralized Multisite VBM Analysis During Adolescence Shows Structural Changes Linked to Age, Body Mass Index, and Smoking: a COINSTAC Analysis. *Neuroinformatics*. 19, 553—566. (https://doi.org/10.1007/s12021-020-09502-7)

129 Basodi S *et al*. 2022 Decentralized Brain Age Estimation using MRI Data. Neuroinform 20, 981–990. (https://doi.org/10.1007/s12021-022-09570-x)

Differential privacy can also be applied to prevent reidentification of neuroimages. Differential privacy entails the addition of 'noise' (irrelevant or unwanted data items, features, or records) to the results; this makes the task of cross-referencing with public data more difficult. Differential privacy also allows for risk to be quantified as the *probability of reidentification*, allowing the controller to 'dial up or down' and adjust for performance-privacy trade-offs by referring to a set 'privacy budget', or how much data is determined acceptable to be leaked from the site[130].

Conclusions

Large, robust, international neuroimaging datasets are required for training machine learning models. These datasets exist around the world in various institutions. Securely using remote datasets to train machine learning models could transform research in this field. Further, safeguarding the privacy of imaging subjects could increase participation in research, enhancing the diverse, large-scale data required to make future strides in neuroscience.

130 Differential privacy and federated learning can be combined in two ways: output perturbation (where noise is added to the output of an optimisation algorithm) and objective perturbation (noise is added at every step of the optimisation algorithm). The latter may hold more functionality but requires identical pre-processing across sites and good local feature mapping.

BOX 5

PETs for machine learning with medical images: Emerging challenges

Radiology uses medical imaging to diagnose, treat disease and monitor in clinical and research settings. High-quality machine learning-based models can provide a second reading of images, acting as a 'digital peer' to medical researchers and clinicians. Once the model can identify patterns of disease, it can be exported for use by other clinicians and researchers (if the model is transferable). The potential public benefit of using these trained models is significant and currently being investigated, for instance, by Health Data Research UK and other stakeholders[131].

Once the model is exported, the original creators relinquish control. While a model is not 'raw data', there are potential vulnerabilities. Over-trained models may remain so faithful to the training dataset that they risk revealing granular details about the training data. Linkage attacks could harvest information derived from the model which, when linked with third-party data, result in the exposure of personal data[132].

Lastly, model inversion or reconstruction attacks may allow an attacker to reverse engineer the training dataset from a model[133]. As a relatively new possibility[134], risk-benefit assessment in model inversion is relatively immature.

Data protection regulation (such as the UK GDPR) can lack clarity regarding models trained on sensitive data. Traditionally, models have been treated as intellectual property or trade secrets, rather than personal data. However, 'trained models can transform seemingly non-sensitive data, such as gait or social media use, into sensitive data, such as information on an individual's fitness or medical conditions.' Legally, the possibility of revealing training data 'might render models as personal data in the sense of European protection law [...]'[135]. Recent publications demonstrate how inference attacks present real threats for collaborative analysis[136].

131 Health Data Research UK (HDR UK Strategic Delivery Plan 2021/22). See https://www.hdruk.ac.uk/wp-content/uploads/2021/02/Strategic-Delivery-Plan-2021_22.pdf (accessed 7 October 2022).

132 White T, Blok E, Calhoun V D. 2020 Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed. *Hum Brain Mapp*. 43, 278—291. (https://doi.org/10.1002/hbm.25120)

133 Veale M, Binns R, Edwards L. 2018 Algorithms that remember: model inversion attacks and data protection law. *Philos T R Soc A*. 376. (https://doi.org/10.1098/rsta.2018.0083).

134 First proposed by Fredrikson M, Jha S, Ristenpart T. 2015 Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures. See https://rist.tech.cornell.edu/papers/mi-ccs.pdf (accessed 6 September 2022).

135 Veale M, Binns R, Edwards L. 2018 Algorithms that remember: model inversion attacks and data protection law. *Philos T R Soc A*. 376. (https://doi.org/10.1098/rsta.2018.0083)

136 Melis L, Song C, De Cristofaro E, Shmatikov V. 2018 Exploiting Unintended Feature Leakage in Collaborative Learning. See https://arxiv.org/abs/1805.04049 (accessed 10 October 2022).

BOX 6

Privacy and compliance concerns around trained models

This should be addressed by considering:

- accessibility (who will use, or have access to, the trained model);
- identification of adversaries and their incentives (in the case of model inversion, the 'honest but curious' persona and deliberate reverse engineering of models for commercial gain)[137];
- legality (what contractual obligations or data sharing regimes model users are subject to);
- public acceptability of proposed model usage (for example, public health application versus commercial enterprise; UK implementation versus international humanitarian applications);
- potential for model reuse or repurposing for other tasks beyond original intentions[138] and;
- commercial sensitivities inherent to the model (such as the risk of a private actor improving upon and re-selling the model back to a public entity).

ICO guidance on model inversion and model inferencing attacks entails a series of actions to be documented. These include reviewing trade-offs on a regular basis, establishing clear lines of accountability with a risk-based approval process, and consideration of available technical approaches that minimise trade-offs[139]. Data management could include best practices for deidentification and removal of metadata. Legal instruments, such as Data Transfer Agreements (DTAs) or contracts provide further safeguarding. For example, NHS Digital has implemented, in collaboration with Privitar, a de-identification tool using a variety of pseudonymisation techniques and a form of homomorphic encryption to ensure safer linkage of data[140, 141].

137 The honest-but-curious adversary is 'a legitimate participant in a communication protocol who will not deviate from the defined protocol but will attempt to learn all possible information from legitimately received messages', as defined in Paverd A, Martin A, Brown I. Modelling and Automatically Analysing Privacy Properties for Honest-but-Curious Adversaries. See https://www.cs.ox.ac.uk/people/andrew.paverd/casper/casper-privacy-report.pdf (accessed 10 September 2022).

138 Melis L, Song C, De Cristofaro E, Shmatikov V. 2018 Exploiting Unintended Feature Leakage in Collaborative Learning. See https://arxiv.org/abs/1805.04049 (accessed 10 October 2022).

139 The Information Commissioner's Office. Guidance on the AI auditing framework: Draft guidance for consultation. See https://ico.org.uk/media/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf (accessed 20 September 2022).

140 The Royal Society. 2019 Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies in data analysis. See https://royalsociety.org/-/media/policy/projects/privacy-enhancing-technologies/privacy-enhancing-technologies-report.pdf (accessed 30 June 2022).

141 National Health Service Digital (Improving our Data Processing Services). See https://digital.nhs.uk/data-and-information/data-insights-and-statistics/improving-our-data-processing-services (accessed 15 May 2022).

BOX 6 (continued)

Training models could be submitted to robust motivated intruder testing, where potential attacks are simulated. For especially sensitive data, stronger attacker profiles (such as criminal groups) may be preferable. Libraries of attacks, which detail potential attacks and relative risks, provide resources for simulating and testing against potential attacks[142].

Biometric and imaging data shared within the usual research and clinical settings is handled by professionals who are incentivised to protect patient confidentiality. With the right data sharing practices in place, the risk of patient data being compromised is greatly reduced. Similar protocols could be followed in model training and use. For example, the model could be restricted to sharing in a trusted research environment[143] with access limited to approved researchers. Perhaps the most secure option, the researcher may retain control of the model and provides analysis as a service.

142 Github (Privacy Trust Lab Privacy Meter). See https://github.com/privacytrustlab/ml_privacy_meter (accessed 10 September 2022).

143 Trusted research environments can vary significantly in scope and security guarantee.

USE CASE 1.2

Preserving privacy in audio data for health research and diagnostics

The opportunity

Audio data containing verbal content and nonverbal vocalisations (coughing, breathing, speech pauses) can be used to train machine learning models for predicting disease[144]. Alzheimer's Disease (AD) is a type of dementia that affects an individual's memory, motor skills and cognition. AD researchers seek non-invasive techniques for screening and detecting AD. Speech and audio data are growing areas for research and diagnosis of AD and other diseases[145, 146, 147]. AD may affect the content of speech – such as the range of a person's vocabulary – or the cadence of speech – such as increased hesitation due to difficulty finding words. Other neurological conditions such as Parkinson's disease may alter speech characteristics including pitch, cadence and articulation. Vocal biomarkers are thus a promising avenue for AD and other research, particularly when coupled with AI[148, 149].

The challenge

Vocal data is vulnerable because there are large open datasets of identifiable audio publicly available (eg on YouTube)[150], making reidentification relatively straightforward. Beyond the content of verbal data, the very presence of an individual within a dataset reveals sensitive information.

As biometric data, vocal data is personal data. It is not considered anonymous under UK and EU GDPR. Additionally, audio data may also be transcribed, doubling the data used (vocal and textual data)[151]. In this case, data minimisation would mitigate risk of information leakage, for example, only retaining transcripts or audio.

144 Wroge TJ, Özkanca Y, Demiroglu C, Si D, Atkins D C, Ghomi RH. 2018 Parkinson's disease diagnosis using machine learning and voice. See https://www.ieeespmb.org/2018/papers/l01_01.pdf (accessed 23 April 2022).

145 König A, Satt A, Sorin A, Hoory R, Toledo-Ronen O, Derreumaux A, Manera V, Verhey F, Aalten P, Robert PH, David R. Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimers Dement.* 1, 112—124. (https://doi.org/10.1016/j.dadm.2014.11.012)

146 Haulcy R, Glass J. 2021 Classifying Alzheimer's Disease Using Audio and Text-Based Representations of Speech. Front. Psychol. Sec. Human-Media Interaction. 11 (https://doi.org/10.3389/fpsyg.2020.624137)

147 University College London (Meet the C-PLACID Audio-Recording Research Team). See https://www.ucl.ac.uk/drc/c-placid-study/audio-recording-c-placid/meet-c-placid-audio-recording-research-team (accessed 1 September 2022).

148 Fagherazzi G, Fischer A, Ismael M, Despotovic V. 2021 Voice for health: The use of vocal biomarkers from research to clinical practice. *Digit Biomark.* 5, 78—88. (https://doi.org/10.1159/000515346)

149 Arora A, Baghai-Ravary L, Tsanas A. 2019 Developing a large scle population screening tool for the assessment of Parkinson's disease using telephone-quality voice. J Acoust Soc Am. 145 5 2871. (https://doi.org/10.1121/1.5100272)

150 Examples include: Mozilla Labs (Common Voice). See https://labs.mozilla.org/projects/common-voice/ (accessed 15 August 2022); Google Audio Set: Gemmeke JF *et al*. 2017 Audio Set: An ontology and human-labaled dataset for audio events. Proc. IEEE ICASSP 2017 New Orleans. See https://research.google/pubs/pub45857/, https://research.google.com/audioset/dataset/index.html (accessed 2 June 2022), and open data sets such as Oxford University's VoxCeleb: Oxford University (VoxCeleb). See https://www.robots.ox.ac.uk/~vgg/data/voxceleb/ (accessed 14 May 2022).

151 Haulcy R, Glass J. 2021 Classifying Alzheimer's Disease Using Audio and Text-Based Representations of Speech. Front. Psychol. Sec. Human-Media Interaction. 11 (https://doi.org/10.3389/fpsyg.2020.624137)

Privacy preservation in biometric audio data

When using biometric audio data, PETs should be layered with audio-specific approaches to anonymisation. For example, voice transformation techniques may be used to alter a patient's voice quality[152]. Transcription of audio data can be automated using AI-based applications (eg Google Cloud's Speech API), then scanned using a machine learning algorithm that tags identifiers such as names, dates, ages, or geographical location. By highlighting identifiable elements, identifiers can be swiftly redacted.

Audio data collection techniques may include phone or web-based recording[153]; these can entail potential for eavesdropping. Voice Over IP (VOIP) can include end-to-end homomorphic encryption, ensuring that no other parties listen during data collection[154]. It is also possible to encrypt voice data for cloud storage[155], or to split voice data into random fragments, which are each processed separately.

Privacy-preserving synthetic data (PPSD) may be generated from audio recordings prior to sharing or querying[156]. However, this is an emerging application of PPSD[157, 158]. New synthetic datasets may need to be created specific to various research queries[159, 160], which could become costly.

Conclusions

Voice recognition technology is becoming ever more sophisticated, such that speaker identification is now feasible even under noisy conditions. These methods may be applied even where masking techniques such as transformation have been used[161]. Without greater sharing of audio data there is a risk that audio-trained models become biased according to language-, accent-, age-, and culture-specific biomarkers. This could be countered through open and crowd-sourced initiatives[162], which could be rolled out most safely with PETs.

152 Jin Q, Toth AR, Schultz T, Black AW. 2009 Voice convergin: Speaker de-identification by voice transformation. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing. (https://doi.org/10.1109/icassp.2009.4960482)

153 Fagherazzi G, Fischer A, Ismael M, Despotovic V. 2021 Voice for health: The use of vocal biomarkers from research to clinical practice. *Digit Biomark*. 5, 78—88. (https://doi.org/10.1159/000515346)

154 In this case, VOIP signals from multiple parties are mixed at a central server, improving the scalability of the solution and protecting the data held on the central server, were the server to be compromised. See: Rohloff K, Cousins D B, Sumorok D. 2017 Scalable, Practical VoIP Teleconferencing with End-to-End Homomorphic Encryption. *IEEE T Inf Foren* Sec. 12, 1031—1041. (https://doi.org/10.1109/tifs.2016.2639340)

155 Shi C, Wang H, Hu Y, Qian Q, Zhao H. 2019 A speech homomorphic encryption scheme with less data expansion in cloud computing. KSII T *Internet Inf*. 13, 2588—2609. (https://doi.org/10.3837/tiis.2019.05.020)

156 Fazel A *et al*. 2021. SynthASR: Unlocking Synthetic Data for Speech Recognition. (https://doi.org/10.48550/arXiv.2106.07803)

157 Tomashenko N *et al*. 2020 Introducing the VoicePrivacy initiative. See https://doi.org/10.48550/arXiv.2005.01387 (accessed 30 March 2022).

158 Shevchyk A, Hu R, Thandiackal K, Heizmann M, Brunschwiler T. 2022 Privacy preserving synthetic respiratory sounds for class incremental learning. *Smart Health*. 23. (https://doi.org/10.1016/j.smhl.2021.100232)

159 Fazel A *et al*. 2021 SynthASR: Unlocking Synthetic Data for Speech Recognition. See https://doi.org/10.48550/arXiv.2106.07803 (accessed 10 October 2022).

160 Rossenbach N, Zeyer A, Schlüter R, Ney H. 2020 Generating synthetic audio data for attention-based speech recognition systems. See https://doi.org/10.48550/arXiv.1912.09257 (accessed 10 October 2022).

161 Chung J S, Nagrani A, Zisserman A. 2018 VoxCeleb2: Deep speaker recognition. See https://doi.org/10.48550/arXiv.1806.05622 (accessed 2 September 2022).

162 Fagherazzi G, Fischer A, Ismael M, Despotovic V. 2021 Voice for health: The use of vocal biomarkers from research to clinical practice. *Digit Biomark*. 5, 78—88. (https://doi.org/10.1159/000515346)

USE CASE 2

PETs and the internet of things: enabling digital twins for net zero

The opportunity

The UK has committed to reach net zero carbon emissions by the year 2050 as part of a wider effort to mitigate climate change. Data-driven digital technologies are poised to play a key role in meeting these targets[163]. Digitalising energy systems will be an important step in decarbonising sectors such as energy, heat, and transport, as well as supporting a greener, circular economy.

Digital twins are an emerging area of focus in climate technologies. A digital twin is a relevant, virtual counterpart of a physical object (such as a wind turbine or electric motor) or process (such as patterns of economic transactions). When integrated with other models and physical-virtual systems through sensors, digital twins can function as decision-support tools.

While small-scale digital twins are in use, large-scale digital twins are at a relatively early stage of development, where security and privacy are emerging concerns[164]. Establishing best practice and privacy solutions will be key to the acceptability of digital twins, as well as ensuring interoperability and other technical requirements are met. A digital twin of the UK's energy system would help balance real-time energy 'smart' grids. This will be important alongside wider uptake of decentralised and intermittent sources of renewable energy.

A digital twin is a relevant, virtual counterpart of a physical object (such as a wind turbine or electric motor) or process (such as patterns of economic transactions).

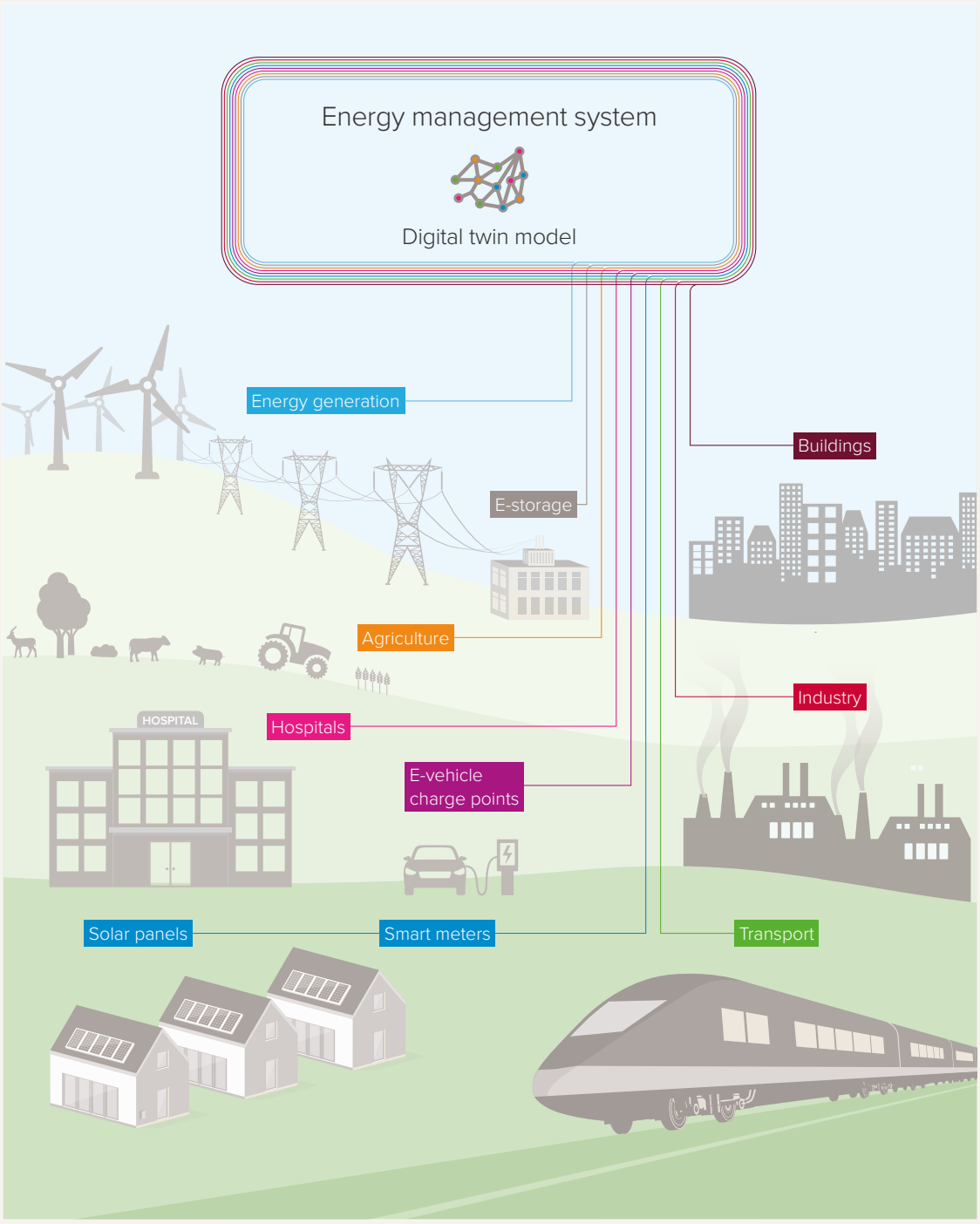
163 The Royal Society. 2020 Digital technology and the planet: Harnessing computing to achieve net zero. See https://royalsociety.org/-/media/policy/projects/digital-technology-and-the-planet/digital-technology-and-the-planet-report.pdf (accessed 20 September 2022).

164 Dietz M, Putz B, Pernul G. 2019 A Distributed Ledger approach to Digital Twin secure data sharing. See https://core.ac.uk/download/pdf/237410573.pdf (accessed 27 September 2022).

FIGURE 3

A digital twin of the UK energy system

Data is needed from a range of sources to develop, evaluate, and 'fuel' a digital twin of the UK energy system. Emerging privacy and security concerns must be addressed to allow the safe flow of data between digital twin models and real-world assets.



The challenges

Data is needed to develop, evaluate, and 'fuel' digital twins. Presuming energy data as open[165] will help unlock research and innovation potential (such as through digital twinning). At the same time, emerging privacy and security concerns must be addressed.

In this case, data sharing issues concern several stakeholder groups:

- Individuals: UK energy consumers' metering data was once only read monthly, but readings can now be taken at a more granular level (typically half-hourly), meaning energy usage patterns can be used to track household activities[166];
- Industry: Energy sector actors may be disincentivised to share data that is commercially sensitive (eg algorithmically derived pricing models);
- Government: Data pertaining to the built environment could expose vital infrastructure or utilities to attack, leading to national security concerns;
- Regulators: Perception of data misuse could lead to loss of public trust, compromising efforts to use data for net zero (for example, leading to low uptake of smart meters).

A flow of data must enable communication between digital twins and real-world assets[167]. Data infrastructure must be able to link physical assets, accounting for different data types, components, technical standards, and analytical capabilities – a lightweight 'digital spine'[168]. Some steps have already been taken, as with the creation of the Information Management Framework within the National Digital Twin Programme[169, 170].

165 Catapult Energy Systems. 2019 A strategy for a Modern Digitalised Energy System: Energy Data Taskforce report. See https://esc-production-2021.s3.eu-west-2.amazonaws.com/2021/07/Catapult-Energy-Data-Taskforce-Report-A4-v4AW-Digital.pdf (accessed 27 September 2022).

166 In one smart heating example, analysts demonstrated the ability to uncover users' sleeping patterns, location within a home, even whether a user was sitting or standing. While this level of detail goes beyond what is possible with typical smart metering, it is one example where perceived potential invasiveness of smart fixtures in the home may prevent uptake of this technology: Morgner P, Müller C, Ring M, Eskofier BM. 2017 Privacy Implications of Room Climate Data. *Lecture Notes in Computer Science vol 10493*. See https://doi.org/10.1007/978-3-319-66399-9_18 (accessed 27 September 2022).

167 Dietz M, Putz B, Pernul G. 2019 A Distributed Ledger approach to Digital Twin secure data sharing. See https://core.ac.uk/download/pdf/237410573.pdf (accessed 27 September 2022).

168 Catapult Energy Systems (Energy Digitalisation Taskforce publishes recommendations for a digitalised Net Zero energy system). See https://es.catapult.org.uk/news/energy-digitalisation-taskforce-publishes-recommendations-for-a-digitalised-net-zero-energy-system/ (accessed 22 September 2022).

169 University of Cambridge (Centre for Digital Built Britain). See https://www.cdbb.cam.ac.uk/subject/information-management-framework-imf (accessed 20 September 2022).

170 More specifically, one of CReDo's aims is to trial the MFI to evaluate the framework's capacity to operate at a national level.

Privacy solutions should be implemented at several critical points in the coupled digital twin-asset ecosystem. This use case focusses on energy consumption, where private data may disclose:

- What appliances are used and when[171];
- What behaviour patterns might be revealed by consumers' energy usage – particularly occupancy patterns[172];
- What information might be inferred about the building / utilities and other features, leading to security risks in national energy systems assets[173];
- How energy companies' processing algorithms might give away proprietary knowledge and commercially sensitive behavioural insights;
- What billing or other pseudonymised records might reveal private information about consumers[174], including consumer responsiveness to changes in price;
- Appliance and usage patterns that might be used in unsolicited targeted marketing, for example, ads or messages prompting consumers to have their boiler serviced.

While these inferences could be made using contemporary smart meter data, future versions may take readings at shorter intervals, allowing for detection of which appliances are used, or which TV channels are watched (through discernible electromagnetic interference signatures)[175].

Individual privacy solutions: Smart meter data privacy

Smart meter data is personal data[176]. Privacy concerns around smart meter data have gained attention with the roll-out of devices in Europe and the UK[177], However, smart meter data holds substantial value for renewable energy integration: there is no other way of measuring energy consumption in real time, or so close to consumer end-use.

171 Molina-Markham A, Shenoy P, Fu K, Cecchet E, Irwin D. 2010 Private memoirs of a smart meter. *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*. See https://doi.org/10.1145/1878431.1878446 (accessed 2 September 2022).

172 Lisovich MA, Mulligan DK, Wicker SB. 2010 Inferring Personal Information from Demand-Response Systems. *IEEE Secur Priv*. 8, 11—20. (http://dx.doi.org/10.1109/MSP.2010.40)

173 Beckel C, Sadamori L, Staake T, Santini S. 2014 Revealing household characteristics from smart meter data. Energy. 78 397—410. (http://dx.doi.org/10.1016/j.energy.2014.10.025)

174 Jawurek M, Johns M, Rieck K. 2011 Smart metering de-pseudonymization. ACSAC 2011 Proceedings of the 27th Annual Computer Security Applications Conference. See http://doi.acm.org/10.1145/2076732.2076764 (accessed 20 March 2022).

175 Enev M, Gupta S, Kohno T. 2011 Televisions, video privacy, and powerline electromagnetic interference. See http://doi.acm.org/10.1145/2046707.2046770 (accessed 2 September 2022).

176 The UK government's Smart Metering Implementation Programme (2018) outlined the smart metering Data Access and Privacy Framework, which aimed to 'safeguard consumers' privacy, whilst enabling proportionate access to energy consumption data'. HM Government. 2018 Smart metering implementation programme: Review of data access and privacy framework. See https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/758281/Smart_Metering_Implementation_Programme_Review_of_the_Data_Access_and_Privacy_Framework.pdf (accessed 22 September 2022).

177 For example: Pöhls HC, Staudemeyer RC. 2015 Privacy enhancing techniques in Smart City applications. See https://cordis.europa.eu/docs/projects/cnect/4/609094/080/deliverables/001-RERUMdeliverableD32Ares20153669911.pdf (accessed 26 September 2022).

A number of solutions have been proposed in handling smart meter data. Non-cryptographic methods may meet the resource constraints of smart meters most effectively; for example, differential privacy could be used to add 'noise' to datasets. Other approaches include spatial aggregation, where smart meters are geographically clustered (such as in a block of houses), allowing for load balancing without collecting household-level information[178].

BOX 7

Secure multi-party computation for smart meter data privacy

The Netherlands implemented smart metering in 2006, along with mandates for data sharing at 15-minute intervals. This was subsequently found to violate Article 8 of the European Convention on Human Rights (respect for private and family life)[179]. As a result, new legislation was passed allowing Dutch customers to opt out entirely or retain smart meter administrative and shutdown capabilities.

More recently, the privacy officer of the DSA has approved the use of smart meter data in cohorts of six neighbouring households. However, this requires averaging six numbers without an analyst seeing those six numbers.

Secure multi-party computation (SMPC) is being piloted in the Netherlands through a public-private partnership with Roseman Labs. SMPC is used to total and average the energy use of six neighbouring houses 'in the blind'. This provides mid-level network views of power consumption for the first time. This solution is currently in trial phase using hardware, which is retrofitted onto smart meters. In the future, the SMPC software could be run as part of software built into smart meters, with data encrypted locally before being sent to the secured server.

178 UN Conference of European Statisticians. 2019 Protecting Consumer Privacy in Smart Metering by Randomized Response. See https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2019/mtg1/SDC2019_S4_URV_Protecting_Consumer_Privacy_AD.pdf (accessed 24 August 2022).

179 Cuijpers C, and Koops B-J. 2013 Smart Metering and Privacy in Europe: Lessons from the Dutch Case. In: Gutwirth S, Leenes R, de Hert P, Poullet Y. (eds) *European Data Protection: Coming of Age*. Berlin: Springer, Dordrecht.

Government, regulators and national security harms

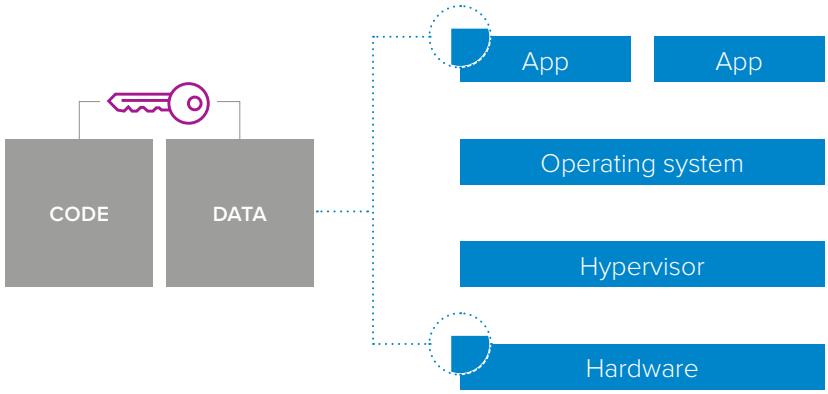
Combined summary statistics of energy data sets will be key to maximising the benefits of an energy digital twin. Privacy-preserving synthetic data (PPSD) could be used to share relevant properties of rich microdata – in essence, how the datasets relate to one another – collected through smart systems. Simpler, differentially private summary statistics could be shared (where the privacy-utility trade-off would be more transparent). This would enable decision-making by government and regulators without releasing full datasets. However, the utility and privacy trade-offs of PPSD must be better understood and will be highly case-dependent[180].

Data coming from physical assets may be used to control the grid and national power distributions. TEEs – potentially coupled with homomorphic encryption – could safeguard collaborative cloud computing from attacks, protecting security to critical national infrastructure[181]. Homomorphic encryption can be highly compute-intensive and would require significant development to be used at a large scale.

FIGURE 4

Trusted Execution Environment (TEE)

TEEs are secure areas inside a processor, which are isolated from the rest of the system. The code contained in the TEE cannot be read by the operating system, nor the hypervisor (a process that separates a computer's operating system and applications from the underlying physical hardware).



180 Jordon J *et al*. 2022 Synthetic data: What, why and how? See https://arxiv.org/pdf/2205.03257.pdf (accessed 2 September 2022).

181 Archer *et al*. 2017 Applications of homomorphic encryption. See https://www.researchgate.net/publication/320976976_APPLICATIONS_OF_HOMOMORPHIC_ENCRYPTION/link/5a051f4ca6fdcceda0303e3f/download (accessed 23 April 2022).

Commercially sensitive data solutions for digital twins

Energy providers could use insights from smart meter data to provide new service models (eg heating as a service). In addition to SMPC, federated learning could allow users' data to stay localised while training models are used by energy providers. For example, a machine learning model could be sent to individual smart home systems and 'learn' locally about certain energy consumption patterns in order to predict demand[182].

Conclusions

Digital twins hold significant potential in enabling the net zero transition. A privacy-enhanced digital twin using PETs should be bolstered with basic security measures, including the physical restriction of access to critical infrastructure, servers and computers (eg using hardware keys). For PETs to be embedded into the realisation of an energy digital twin, data protection regulation and related guidance should consider what mandates or advice would be effective and ethical in promoting the uptake of smart meters.

Ofgem and other regulatory bodies should ensure that data usage reflects consumer interests. In a digital twin, this could entail allowing users to audit and challenge their smart meters' outputs, for example[183]. Where algorithms are trained on real-time data, every effort must be made to ensure sections of the population are not over- or under-represented, as this could reproduce systemic biases and promote inaccuracies. A consumer consent dashboard, such as the one proposed by the Energy Digitalisation Taskforce[184] in the UK, may provide a greater sense of control and encourage consumer trust.

182 Fuller A, Fan Z, Day C, Barlow C. 2020. Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access*. 8, 108952—108971. (https://doi.org/10.1109/access.2020.2998358)

183 The Royal Society. 2020 Digital technology and the planet: Harnessing computing to achieve net zero. See https://royalsociety.org/-/media/policy/projects/digital-technology-and-the-planet/digital-technology-and-the-planet-report.pdf (accessed 20 September 2022).

184 HM Government 2022. Energy Digitalisation Taskforce report: joint response by BEIS, Ofgem and Innovate UK. See https://www.gov.uk/government/publications/digitalising-our-energy-system-for-net-zero-strategy-and-action-plan/energy-digitalisation-taskforce-report-joint-response-by-beis-ofgem-and-innovate-uk (accessed 24 August 2022).

USE CASE 3

Social media data: PETs for researcher access and transparency

The opportunity

Over 4 billion people use social media – including networking platforms, online games, wellbeing applications, and budgeting tools – to upload and share media and messages, log activities and access entertainment[185]. The extent to which people interact with, and generate content on, these platforms has made social media services an increasingly valuable source of data for research. The Royal Society has recommended that social media platforms establish ways to allow independent researchers to access data in a secure and privacy compliant manner[186], particularly for audit and to encourage the accountability of platforms.

Data generated through social media includes content posted and shared (such as text posts or photos) as well as metadata (such as demographic information, location, time of upload, and behavioural patterns – for example, how often a user opens a fitness app or inferred relationships depicted in a user's photos)[187]. User data is often volunteered by users, such as an uploaded profile photo, or self-described location. Most metadata is logged automatically, such as the geotag on an image, or the timestamp on a message.

© Unsplash / Gian Cescon.

185 Statista (Number of internet and social media users worldwide as of July 2022). See https://www.statista.com/statistics/617136/digital-population-worldwide/ (accessed 18 August 2022).

186 The Royal Society. 2022 The online information environment: Understanding how the internet shapes people's engagement with scientific information. See https://royalsociety.org/-/media/policy/projects/online-information-environment/the-online-information-environment.pdf?la=en-GB&hash=691F34A269075C0001A0E647C503DB8F (accessed 30 March 2022).

187 See Lomborg S, Bechmann A. 2014 Using APIs for Data Collection on Social Media. The Information Society 30 4 256—265. (https://doi.org/10.1080/01972243.2014.915276)

BOX 8

Social media data for research and decision making

An increasing number of studies use social media platforms and mobile applications as rich data sources[188]. Interdisciplinary research using social media data includes:

- Disaster management and emergency response[189];
- Social patterns of influence and the dynamics of social movements;
- Information cascades (how information propagates in social media sites, understanding the spread and impact of misinformation)[190];
- Event monitoring by location to enhance physical safety and security;
- Vulnerability management[191], identifying and communicating with communities most at risk of natural disaster, climate emergencies or disease outbreak;
- Studying online harms (including bullying and harassment, toxicity, radicalisation);
- Political research and opinion forecasting[192].

However informative, the use of social media data can be resource intensive and invasive. Accessing and curating social media data is hindered by technical capabilities and public distrust.

Researcher access: APIs and PETs

Researchers typically use an API (Application Programming Interface) to access social media data logs (or data streams), which can be analysed for patterns. An API is a backend interface that connects social media services and their data to third parties. Making APIs available to researchers can be part of a social media company's business model. Some large companies like Facebook and Twitter provide free, if restricted, access to datasets of public-facing data. Private user data is released through APIs only to approved researchers, who may submit queries for specific datasets[193] or use the Twitter 1% sampled stream, which delivers a random selection of roughly 1% of public Tweets in real-time[194].

188 For example: Giglietto F, Rossi L, Bennato D. 2012 The Open Laboratory: Limits and Possibilities of Using Facebook, Twitter, and YouTube as a Research Data Source. *Journal of Technology in Human Services*. 30, 145–159. (https://doi.org/10.1080/15228835.2012.743797)

189 Teodorescu H-N. 2015 Using analytics and social media for monitoring and mitigation of social disasters. *Procedia Engineer.* 107 325—334. (https://doi.org/10.1016/j.proeng.2015.06.088)

190 Harvard Kennedy School Misinformation Review (Tackling misinformation: What researchers could do with social media data). See https://misinforeview.hks.harvard.edu/article/tackling-misinformation-what-researchers-could-do-with-social-media-data/ (accessed 20 November 2021).

191 Gundecha P, Barbier G, Huan L. 2011 Exploiting Vulnerability to Secure User Privacy on a Social Networking Site., *Proceedings of the 17th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*, KDD, 2011. 511–519.

192 Sobkowic P; Kaschesky M; Bouchard G. 2012 Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web. *Gov Inform Q*. 29, 470–479. (https://doi.org/10.1016/j.giq.2012.06.005)

193 For example, Meta's Graph API. Meta for Developers (Graph API Overview). See https://developers.facebook.com/docs/graph-api/overview/ (accessed 17 July 2022).

194 Twitter Developer Platform (Volume streams). See https://developer.twitter.com/en/docs/twitter-api/tweets/volume-streams/introduction (accessed 27 September 2022).

BOX 9

Public health is an emerging and growing area for research using social media data

For example:

- Economists are using mobile game scores and geolocation data from Lumocity, a brain-training game, to see whether local air pollution spikes correlate with declines in cognitive function and productivity. This research could establish exposure to particulate matter as a mechanism for inequality in the workforce[195];
- Following the hashtag #cheatmeal on Instagram, kinesiologists and psychologists analysed tagged images to characterise an emerging dietary trend, which they linked to binge eating[196];
- In April 2020, Facebook's Data for Good programme released new visualisations and datasets including Movement Range Maps, Co-Location Maps and symptom surveys to enable researchers, international agencies, non-profits and public sector institutions track and combat COVID-19[197]. The usage of this data influenced international public policy responses and helped researchers identify economic, health and social impacts in communities[198]. Researchers may now access recent survey datasets on the future of business[199] and gender equality at home[200].

Greater visibility is needed around the full lifecycle of social media data for researchers to fully utilise social media data[201]. Transparency in social media data, including how it is used by platforms, would also promote the rights of data subjects to exercise informed consent around how their data is used.

195 La Nauze A, Severnini ER. 2021 Air pollution and adult cognition: Evidence from brain training. See https://www.nber.org/system/files/working_papers/w28785/w28785.pdf (accessed 30 April 2022).

196 Pila E, Mond JM, Griffiths S, Mitchison D, Murray SB. 2017 A thematic content analysis of #cheatmeals images on social media: Characterizing an emerging dietary trend. *Int J Eat Disord*. (https://doi.org/10.1002/eat.22671)

197 Meta (Data for Good: New Tools to Help Health Researchers Track and Combat COVID-19). See https://about.fb.com/news/2020/04/data-for-good/ (accessed 27 September 2022).

198 Office for National Statistics Data Science Campus (Using Facebook data to understand changing mobility patterns). See https://datasciencecampus.ons.gov.uk/using-facebook-data-to-understand-changing-mobility-patterns/ (accessed 24 August 2022).

199 Humanitarian Data Exchange (Future of Business Survey—Aggregated Data). See https://data.humdata.org/dataset/future-of-business-survey-aggregated-data (accessed 21 February 2022).

200 Humanitarian Data Exchange (Survey on Gender Equality At Home). See https://data.humdata.org/dataset/survey-on-gender-equality-at-home (accessed 21 February 2022).

201 COVID-19 Mobility Data Network (Facebook Data for Good Mobility Dashboard). See https://visualization.covid19mobility.org/?date=2021-09-24&dates=2021-06-24_2021-09-24&region=WORLD (accessed 27 September 2022).

The challenge

Social media data entails a variety of personal data, including a user's age, gender, political orientation[202] and moods[203]. Images can expose location[204], residence[205] and relationship status between individuals in a photo. These privacy issues are related to contextual downstream harms, for example, inferring sexual orientation in countries where homosexuality is a criminal offence. Mobility data can provide detailed history of whereabouts, leading to novel inferences (eg cultural background)[206].

Under the UK GDPR, an identifiable person includes someone who can be identified indirectly. In this sense, social media metadata is personal data. Using metadata, even pseudonymised datasets can be reidentified – for example, by comparing the structure of social networks[207] to uncover a third party's approximate whereabouts[208]. Inferring an individual's identity or location through metadata without consent – for example, with targeted advertising – violates the UK Data Protection Act 2018.

There are technical challenges around collecting and using social media data in a privacy-preserving way. One challenge is deletion: data shared on social media platforms as unrestricted (available to anyone without logging onto the platform) may be collected for research purposes without violation of terms of use. However, data subjects have the right to request their data be excluded from studies, regardless of how it was shared or accessed. For example, researchers using a Twitter stream must also verify whether Tweets used in analysis have been deleted. This can be particularly difficult in longitudinal studies. Social media users posting anonymously or using pseudonyms may not be matched across platforms, making cross-platform studies at user level difficult or impossible.

202 Rao D, Yarowsky D, Shreevats A, Gupta M. 2010 Classifying latent user attributes in twitter. See https://www.cs.jhu.edu/~delip/smuc.pdf (accessed 30 March 2022).

203 Tang J, Zhang Y, Sun J, Rao J, Yu W, Chen Y, and Fong A C M. 2012 Quantitative Study of Individual Emotional States in Social Networks. *IEEE T Affect Comput*. 3, 132–144.

204 Hays J, Efros A. 2008 Im2gps: estimating geographic information from a single image. *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) 2008*. http://graphics.cs.cmu.edu/projects/im2gps/im2gps.pdf (accessed 27 September 2022).

205 Jahanbakhsh K, King V, Shoja GC 2012. They Know Where You Live! See https://arxiv.org/abs/1202.3504 (accessed 10 October 2022).

206 Silva TH, de Melo POSV, Almeida JM, Musolesi M, Loureiro AA F 2014. You are What you Eat (and Drink): Identifying Cultural Boundaries by Analyzing Food & Drink Habits in Foursquare. See https://arxiv.org/abs/1404.1009 (accessed 27 September 2022).

207 Narayanan A, Shmatikov V 2009. De-anonymizing social networks. See https://www.cs.utexas.edu/~shmat/shmat_oak09.pdf (accessed 15 August 2022).

208 Li R, Wang S, Deng H, Wang R, Chang K C C. 2012 Towards social user profiling: Unified and discriminative influence model for inferring home locations. KDD 2012: P*roceedings of the 18th ACM SIGKDD International conference on Knowledge Discovery and Data Mining*. 1023–1031. (https://doi.org/10.1145/2339530.2339692)

BOX 10

Facebook data: Researcher access and Cambridge Analytica

Researchers can request access to non-public Facebook data by creating a Facebook app using Facebook's Open Graph API. Apps created by researchers often take the form of games, which can then be installed by Facebook users who agree to the app's terms and conditions. These terms list which types of data the app will collect from your Facebook activity and share with the app developers.

Cambridge Analytica used this method in creating the *thisisyourlife* app. The app included in its terms and conditions access to the data of app users *as well as* their friends' data. While just over 300,000 consenting users installed the *thisisyourlife* app, data from 87 million profiles was collected[209, 210].

The Cambridge Analytica controversy highlights how amalgamated data used for research can be experienced as an invasion of collective and individual privacy. The result was a tightening of access to APIs and reformed policies, particularly at large social media companies[211]. This case also demonstrates how so-called privacy mechanisms, such as APIs that restrict access to approved researchers, can be applied in ways that do not preserve privacy. The API performed to its technical specifications, but the use case violated the intent of data subjects. Facebook and Google have since experimented with homomorphic encryption, federated learning and differential privacy to enable advertising and market research[212]. In these ways, PETs are being used to support business-as-usual, enhancing user profiling and targeted advertising.

Preserving privacy in social media data use

Privacy by design in social media data should address two primary concerns. The first is poor information scoping, where access to user's private information may expose more than is required (eg sharing a user's entire calendar rather than one calendar event). The second is the tracking of individuals, for example, through user 'fingerprinting' or cookies, or by logging the user's unique metadata (eg screen resolution, plugins installed, list of fonts and time zone).

APIs can be designed with consideration of user interface and data minimisation approaches. API users could mediate access themselves, for example, through prompts that contextualise the data request. The request could be embedded in the flow of the data subject's intended action (not diverting their attention). Data minimisation can be used to expose minimal information by limiting queries to specifics.

209 Rosenberg M, Dance GJX. 2018 You Are the Product': Targeted by Cambridge Analytica on Facebook. New York Times. 8 April 2018. See https://www.nytimes.com/2018/04/08/us/facebook-users-data-harvested-cambridge-analytica.html (accessed 14 May 2022).

210 Lawmakers publish evidence that Cambridge Analytica work helped Brexit group. *Reuters*. 16 April 2018. See https://www.reuters.com/article/us-facebook-cambridge-analytica-britain/lawmakers-publish-evidence-that-cambridge-analytica-work-helped-brexit-group-idUSKBN1HN2H5 (accessed 2 March 2022).

211 Kelly H. 2018 California just passed the nation's toughest data privacy law. *CNN*. 29 June 2018. See https://money.cnn.com/2018/06/28/technology/california-consumer-privacy-act/index.html (accessed 16 March 2022).

212 Ion M *et al*. 2017 Private Intersection-Sum Protocol with Applications to Attributing Aggregate Ad Conversions. See https://eprint.iacr.org/2017/738.pdf (accessed 25 March 2022).

Differential privacy can be used to safeguard datasets for release to researchers by obscuring information pertaining to specific users in a dataset . In social media datasets, this could mean sharing regional or other cohort-based data to prevent reidentification of individuals. There are limitations around combining data (such as layering spatial data using maps) from multiple sources, alongside the addition of noise. This is one area for further research[213].

Facebook's Data for Good programme[214], launched in 2017, has used differential privacy to provide access to researchers studying crucial topics, including disease transmission, humanitarian responses to natural disasters and extreme weather events. Where public datasets are considered sensitive in aggregation, noise is added to prevent reidentification using a Differential Privacy Framework[215, 216, 217]. Facebook's Data for Good programme has received criticism for its execution; researchers have been denied access to the programme, or provided with inaccurate data, invalidating months of research[218].

PETs may also be used to share social media data between researchers, or to enable open access social media databases without compromising privacy. For example, centralised data stores could be built and queried. This could include specific attributes, keywords, locations or other demographics in a centralised model. Homomorphic encryption or other cryptographic tools may be applied to social network data, allowing researchers to query to the data holders without requesting data. The data holder could then run the query and release differentially private results. Synthetic data may also be used to release versions of datasets.

213 With regard to mobility data, for example, 'as various providers stack up different sources of data in a collaborative project such as the Network, it often erodes corrections made for differential privacy noise in a single dataset.' Open Data Institute COVID-19 Mobility Data Network's Use of Facebook Data for Good Mobility Data. See http://theodi.org/wp-content/uploads/2021/04/5-COVID-19-Mobility-Data-Networks-Use-of-Facebook-Data_v2.pdf (accessed 7 October 2022).

214 Facebook (Data for Good). See https://dataforgood.fb.com/ (accessed 18 August 2022).

215 Facebook Research (Privacy protected data for independent research on social media data). See https://research.fb.com/blog/2020/02/new-privacy-protected-facebook-data-for-independent-research-on-social-medias-impact-on-democracy/ (accessed 2 September 2022).

216 Jin KX, McGorman L. Data for Good: New tools to help health researchers track and combat COVID-19. Facebook News. 6 April 2020. See https://about.fb.com/news/2020/04/data-for-good/ (accessed 15 March 2022).

217 Facebook Research (Protecting privacy in Facebook mobility data during the Covid-19 response). See https://research.fb.com/blog/2020/06/protecting-privacy-in-facebook-mobility-data-during-the-covid-19-response/ (accessed 23 September 2022).

218 Moon M. Facebook has been giving misinformation researchers incomplete data. Engadget. See https://www.engadget.com/facebook-misinformation-researchers-incomplete-data-050143486.html (accessed 30 August 2022).

BOX 11

PETs for transparency: Twitter and OpenMined partnership for algorithmic accountability

In January 2022, Twitter's ML Ethics, Transparency, and Accountability (META) team announced a partnership with OpenMined to explore the use of PETs for public accountability over social media data. OpenMined is an open-source non-profit organisation that aims to build and promote the use of PETs through educating data owners and making privacy-preserving technologies more accessible to private and public organisations.

The Twitter-OpenMined partnership proposes the use of PETs as a tool for accountability. Currently, one barrier to algorithmic accountability is external researchers and third parties lack access to proprietary algorithms and the data they use, rendering it difficult to conduct independent investigations and audits. PETs in this instance may allow companies to share internal algorithms and datasets for algorithmic audits and replicating research, while avoiding concerns around privacy, security or intellectual property.

The first project will involve developing a method of replicating internal research findings on algorithmic amplification of political content on Twitter by using a synthetic dataset. Long-term, Twitter suggests they will share their actual internal data through PETs to enable external researchers to conduct their own investigations on currently non-public data.

Conclusions

In this use case, PETs are used as tools for privacy and confidentiality, as well as accountability and transparency through external audit. While social media data is not usually sold, social media business models depend on personal data – and derived insights – collected and analysed through opaque processes. A privacy-enhanced strategy for enhancing access to data and increasing transparency will improve user trust and mitigate legal or reputational risks for social media platforms. Furthermore, the amount of compute power required to analyse large social media datasets may motivate platforms to use networked PETs to provide analysis as a service[219].

As the types and scale of personal data shared on social media continues to expand, novel privacy concerns will emerge. For example, the linking of consumer genomics products with social media platforms is increasingly popular on sites like Ancestry.com, or opensource genetics databases such as GEDmatch or Promethease. While open DNA databases have prompted some users to consider the risks associated with making their genome public[220], the implications of linking an individual's DNA to social media metadata (such as location, behavioural patterns or social networks) are less understood.

219 The Royal Society. 2022 The online information environment: Understanding how the internet shapes people's engagement with scientific information. See https://royalsociety.org/-/media/policy/projects/online-information-environment/the-online-information-environment.pdf?la=en-GB&hash=691F34A269075C0001A0E647C503DB8F (accessed 30 March 2022).

220 Mittos A, Malin B, De Cristofaro E. 2018 Systematizing genome privacy research: A privacy-enhancing technologies perspective. See https://arxiv.org/abs/1712.02193 (accessed 23 March 2022).

USE CASE 4

Synthetic data for population-scale insights

The opportunity

A vast amount of national-scale data and microdata is held in various public records controlled by different institutions. This data enables greater understanding of population-level behaviour, forecasting and 'nowcasting' important metrics (such as GDP or disease prevalence) and monitoring regional development across the UK.

The UK's Office for National Statistics (ONS) is the UK's largest independent producer of national statistics and serves as the national statistical institute. As the body responsible for collecting and sharing official statistics relevant to the UK economy and population, the ONS stores and controls a wealth of high-value data and microdata[221] and substantial national datasets, including census data[222].

There is significant appetite across the UK public sector to use national data to drive innovation and growth, to support better policy and decision-making and to use AI to improve service efficiencies. In 2017, an Office for Statistics Regulation investigation found that the UK's statistical system's capacity to link data and provide insights to users was lacking, causing a significant loss of value to society[223].

The ONS is currently exploring how PETs might help reverse this trend by supporting anonymisation at population scale.

In 2017, an Office for Statistics Regulation investigation found that the UK's statistical system's capacity to link data and provide insights to users was lacking, causing a significant loss of value to society.

221 The Digital Economy Act 2017 provides a gateway for the ONS to access the data of all public authorities and Crown bodies in support of the production of National Statistics and other official statistics, including the census. It also entails powers to mandate data from some UK businesses. In some (limited) circumstances, ONS-held data may also be shared with devolved administrations for statistical purposes. HM Government (Digital Economy Act 2017). See https://www.legislation.gov.uk/ukpga/2017/30/contents/enacted (accessed 13 May 2022).

222 HM Government (Census Act 1920). See https://www.legislation.gov.uk/ukpga/Geo5/10-11/41/contents (accessed 23 April 2022).

223 The Office for Statistics Regulation (Joining up data for better statistics). See https://osr.statisticsauthority.gov.uk/publication/joining-up-data/ (accessed 30 March 2022).

Synthetic data is data that is modelled to represent the statistical properties of original data.

The challenge: Anonymisation in big data

Data controllers are often unable to share datasets without compromising legal or ethical requirements to protect confidentiality. The growing availability of population-scale data, linked datasets, access to powerful analytical techniques and compute power means that the risk of 'hacking' or 'reverse-engineering' anonymised datasets is growing[224].

Privacy-preserving synthetic data

Synthetic data is data that is modelled to represent the statistical properties of original data. New data values are created which, taken as a whole, reproduce the statistical properties of the 'real' dataset without including any original datapoints. Users of synthetic datasets optimised for privacy may be virtually unable to identify any information pertaining to original datapoints. For this reason, synthetic data has significant privacy-preserving potential.

Privacy-preserving synthetic data (PPSD) is synthetic data generated from real-world data to a degree of privacy that is deemed acceptable for a given application[225]. PPSD may be used to enable broader access to high-value datasets to drive exploration and innovation. It may also reduce the time for development of new data products by allowing early access to 'good enough' models, and to develop models and build pipelines while access to 'real' data is negotiated. This could also unlock sensitive datasets, for example, by synthesising microdata currently held in the Secure Research Service to provide access to a wider range of users[226].

High-value synthetic population-level datasets

The Data Science Campus at the ONS is working in partnership with the Alan Turing Institute to explore the role of PPSD in using national datasets for public benefit[227]. This does not include the use of PPSD for decision-making, but rather to supply provisional datasets to researchers who wish to test systems or develop methods in non-secure environments. It may also be used to educate, promoting the use of ONS data sources[228]. The programme is exploring the generation of PPSD with an aim to develop a robust framework for assessing privacy-utility trade-offs.

224 For example Rocher L, Hendrickx JM, de Montjoye Y-A. 2019 Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun* 10 3069. (https://doi.org/10.1038/s41467-019-10933-3)

225 Gartner Research 2022. Top strategic technology trends for 2022: Privacy-Enhancing Computation. See https://www.gartner.co.uk/en/information-technology/insights/top-technology-trends#:~:text=Trend%203%3A%20Privacy%2Denhancing%20Computation,well%20as%20growing%20consumer%20concerns (accessed 23 September 2022).

226 Government Statistical Service (Examples of data linking within the government statistical service). See https://gss.civilservice.gov.uk/examples-of-data-linking-within-the-government-statistical-service/ (accessed 23 September 2022).

227 Office for National Statistics (Office for National Statistics and the Alan Turing Institute join forces to produce better and faster estimates of changes to our economy). See https://www.ons.gov.uk/methodology/methodologicalpublications/generalmethodology/onsworkingpaperseries/onsmethodologyworkingpaperseriesnumber16syntheticdatapilot (accessed 23 September 2022).

228 Office for National Statistics (Synthetic data pilot working paper). See https://www.ons.gov.uk/methodology/methodologicalpublications/generalmethodology/onsworkingpaperseries/onsmethodologyworkingpaperseriesnumber16syntheticdatapilot (accessed 23 September 2022).

Synthetic data can also be used to improve the quality of data. This is achieved through data augmentation and other techniques[229] that address incompleteness in datasets, particularly where populations are small or less represented. There are potential issues with skew or bias in these cases, which must be addressed.

Although synthetic data techniques may be applied to virtually any data, ranging from imagery to textual, three high-value datasets illustrate the potential for this technology:

- **A high-quality synthetic version of the Census-Health-Mortality dataset** (the 'health asset') would allow the ONS to share realistic data quickly with many research partners, speeding up research and innovation by allowing a wide variety of users to rapidly develop models and hypotheses, and build pipelines which can then be applied to the real data for decision-making;
- **Synthetic versions of telecoms mobility data** would enable ONS and cross-government partners to fully assess the opportunities for this data, before going to procurement. This would provide better value for money and would improve official mobility-based statistics such as those relating to COVID-19 analysis
- **Synthesis of administrative data** would allow for the off-line exploration of synthetic data allowing for a single, well defined data extract request being made to the data owners. If this is not practical, a full tested and robust data pipeline could be developed to process and analyse the sensitive data in situ.

There are several prerequisites to implementing PPSD. The first is a consistent and comprehensive way to evaluate synthetic datasets. The ONS is addressing this issue through a framework, which will be in the form of a Python library. The framework will assess the performance of synthetic datasets in terms of both utility and privacy.

Second is the investigation and assessment of synthetic data generation methods. This means exploring off-the-shelf methods such as Synthpop[230], as well as more sophisticated machine and deep learning methods such as Generative Adversarial Networks and Evolutionary Optimisation. This requires a deal of technical expertise in implementation, as well as deep knowledge of the context, risk factors (adversaries and threat models) and potential for downstream harms.

A synthetic dataset with all the utility of the original dataset cannot offer privacy. For this reason, high-dimensional datasets (which contain many variables) may not be suitable for PSSD generation. Rather, an external researcher or client might request a custom synthesised dataset pertaining to a specific question (calling on a limited number of attributes or variables). In this way, greater utility may be offered without higher risk of privacy loss[231].

PPSD may also be layered with other PETs to enhance its privacy preserving potential. For example, synthetic data can be generated with differential privacy guarantees, offering greater assurance of privacy. However, further erosion of utility must be considered when adding noise to a synthetic dataset[232].

229 For example, missing value imputation and removing class imbalances.

230 Synthpop (Homepage). See https://cran.r-project.org/web/packages/synthpop/vignettes/synthpop.pdf (accessed 23 September 2022).

231 Jordon J *et al.* 2022 Synthetic data: What, why and how? See https://arxiv.org/pdf/2205.03257.pdf (accessed 2 September 2022).

232 Jordon J, Yoon J, van der Schaar M. 2019 PATE-GAN: Generating synthetic data with differential privacy guarantees. See https://openreview.net/pdf?id=S1zk9iRqF7 (accessed 26 September 2022).

BOX 12

Privacy-preserving synthetic data framework for population-scale patient data

The Clinical Practice Research Datalink (CPRD)[233] is the Medicine and Healthcare products Regulatory Agency (MHRA's) real world data research service created to support retrospective and prospective public health and clinical studies. CPRD is jointly sponsored by the MHRA and the National Institute for Health Research (NIHR) as part of the Department of Health and Social Care.

CPRD collects anonymised patient data from a network of GP practices across the UK. Since 2018, CPRD has working on the development of synthetic datasets based on GP patient data to maximise the benefit of this valuable data, while balancing privacy concerns and preventing downstream harm to data subjects. These synthetic datasets can be used as sample datasets, enabling third parties to develop, validate and test analytic tools. They can also be used for training purposes, and for improving algorithms and machine learning workflows.

CPRD has now made two high-fidelity synthetic datasets available[234]: a cardiovascular disease synthetic dataset and a COVID-19 symptoms and risk factors synthetic dataset. Both synthetic datasets are generated from anonymised real primary care patient data extracted from the CPRD Aurum database[235] and are available to researchers for a nominal administrative fee.

The MHRA was motivated to explore synthetic data generation methods to support regulatory requirements for external validation of machine learning (ML) and AI algorithms. Anonymised health datasets have high utility, but still carry residual privacy risks which limit their wider access[236]; a fully synthetic approach can substantially mitigate these risks[237]. In some cases, synthetic data may even improve the utility of anonymised data – its potential to be clinically meaningful. This is because anonymised data may entail gaps, which can lead to biased inferences. Synthetic data can be used in these cases to supplement real data by filling the gaps or boosting underrepresented subgroups in the dataset[238].

233 Clinical Practice Research Datalink (Homepage). See https://cprd.com/ (accessed 17 September 2022).

234 Clinical Practice Research Datalink (Synthetic data CPRD cardiovascular disease synthetic dataset). See https://cprd.com/synthetic-data#CPRD%20cardiovascular%20disease%20synthetic%20dataset (accessed 23 September 2022).

235 CPRD Aurum contains routinely collected data from practices using EMIS Web® electronic patient record system software. Clinical Practive Research Datalink (Primary care data public health research). See https://cprd.com/primary-care-data-public-health-research (accessed 23 September 2022).

236 Sweeney L. 2000 Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paper 3.

237 Park Y, Ghosh J. 2014 PeGS: perturbed Gibbs samplers that generate privacy-compliant synthetic data. *Trans Data Privacy*. 7, 253—282.

238 Wu, L., He, H., Zaïane, O. R. 2013 Utility of privacy preservation for health data publishing. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*. 510—511.

CPRD uses the Synthetic Data Generation and Eva luation Framework[239] to guide synthetic data generation. It consists of a set of procedures, including a ground truth selection process as input, a synthetic data generation procedure, and an evaluation process.

The Synthetic Data Generation Framework has been proven to produce effective synthetic alternatives to 'real' health data. This is particularly beneficial when 1) access to the ground truth data is restricted; 2) when the sample size is not large enough, or representative of a population; 3) when lacking machine learning, AI training or testing datasets. There are limitations and challenges to consider during synthetic data generation outside the framework, including data missingness and the complex interactions between variables. The Synthetic Data Generation Framework used by CPRD is flexible enough to allow for generation of different types of synthetic datasets, while at the same time enabling researchers to demonstrate that they have balanced data utility with patient privacy needs.

Access to the synthetic datasets requires a data sharing agreement with the applicant's organisation for access (this is in line with advice received from the ICO Innovation Hub)[240].

Conclusions

Synthetic data can be useful for expediting data projects and enabling partnerships. For example, organisations can test whether a partnership is worthwhile and start building models while waiting for access (such as through data sharing agreements or other means). Whether or not synthetic data will provide a stand-in for useful and sufficiently private data for analytical use cases remains an open question.

The generation of synthetic datasets, even 'good enough' synthetic versions, is challenging. As yet, there are no standards related to privacy in PPSD generation, though emerging synthetic data standards may include privacy metrics[241]. Further research is required to quantify the privacy-utility trade-offs[242]. To these ends, the ONS plans to test with data owners and the wider data community as part of their synthetic data project.

239 Wang Z, Myles P, Tucker A. 2021 Generating and evaluating cross-sectional synthetic electronic healthcare data: Preserving data utility and patient privacy. *Comput Intell*. 37, 819—851.

240 The Synthetic Data Generation and Evaluation Framework, owned by the MHRA, was developed through a grant from the Regulators' Pioneer Fund launched by BEIS and managed by Innovate UK. Further development of the COVID-19 synthetic data and refinement of synthetic data generation methods was funded by NHSX.

241 Institute of Electrical and Electronics Engineers (Synthetic data standards). See https://standards.ieee.org/industry-connections/synthetic-data/ (accessed 18 August 2022).

242 One recent publication finds the privacy gain is highly variable, and utility loss unpredictable, when used in high-dimensional datasets: Stadler T, Oprisanu B, Troncoso C *et al*. 2021 Synthetic Data—Anonymisation Groundhog Day. See https://arxiv.org/abs/2011.07018 (accessed 27 September 2022).

USE CASE 5.1

Pets in the public sector: Collective intelligence, crime prevention and online voting

Collaborative analysis for collective intelligence

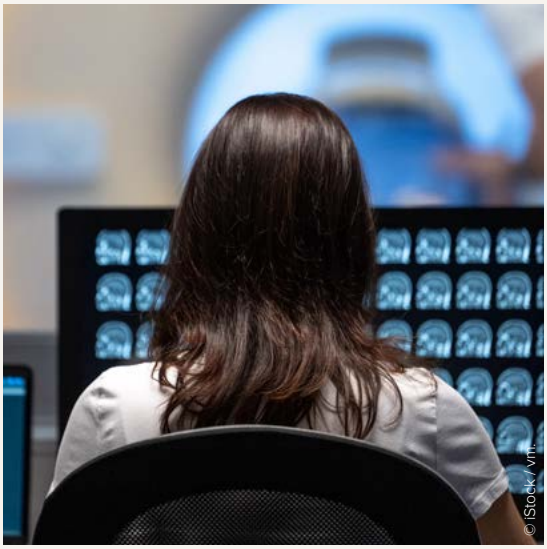
The opportunity

A wealth of data is collected and stored across government departments and non-public bodies in the UK and abroad. This data potentially holds insights that could save substantial money, make government services more efficient and effective, drive the transition to net zero by 2050 (see page 67), guide life-saving choices during a pandemic or understand the effect of regional policies.

Much of the data required to tackle social challenges is sensitive. Particularly where politically sensitive data is used, there are inherent security risks. While there are some special accessions to using health data during emergencies, intra-departmental collaboration must adhere to privacy legislation, including data protection. As such, the risk of collaboration between departments may be deemed larger than potential benefits.

Collaborative analysis with SMPC

Secure multi-party computation (SMPC) allows multiple parties to jointly compute a function using inputs from all parties, while keeping those inputs private. In this way, SMPC is a tool for securely generating insights using data held by different departments or organisations. For example, in a health study, patient data may be input from different hospitals, or even combined with other datasets – such as social demographic data – without researchers ever seeing or accessing the data directly.



SMPC has been demonstrated using large-scale studies on government data since 2015[243]. The performance of SMPC relates to the analysis, or functions, to be computed. Summations (adding numbers together) are faster than more complex computations[244]. This is a rapidly advancing technology with the potential for use in long-term data governance; this is because SMPC depends on access control by all parties involved, meaning analysis can only be performed if all parties agree. SMPC protocols ensure input privacy (no information can be obtained or inferred by any party aside from their own input and the output). As such, SMPC may provide a generic, standardised – and potentially certifiable – method for computation on encrypted data[245].

243 Bogdanov D, Kamm L, Kubo B, Rebane R, Sokk V. 2015 Students and taxes: a privacy-preserving social study using secure computation. See https://eprint.iacr.org/2015/1159.pdf (accessed 25 September 2022).

244 UN PET Lab Handbook. See https://unstats.un.org/bigdata/task-teams/privacy/UN%20Handbook%20for%20Privacy-Preserving%20Techniques.pdf (accessed 17 July 2022).

245 Archer DW *et al*. 2018 From keys to databases: real-world applications of secure multi-party computation. See https://eprint.iacr.org/2018/450 (accessed 10 October 2022).

There are a number of potential uses of SMPC in the public sector and in cross-sector partnerships; a few examples include:

- Combining cyber intelligence housed in various government departments and ministries to identify cyber threats and incidents (such as the Dutch Ministry of Justice and National Cyber Security Centre partnership, Secure Net)[246];
- Combining data from different social domains to ensure government funds and interventions are well targeted; for example, combining detailed social statistics with healthcare costs to see where government actions on preventions should be targeted and;
- Establishing a decentralised register for businesses, law enforcement and banks to log fraud incidents (including details on company name, individuals and account numbers). Parties would only be able to test whether a name or account number has been registered in a previous fraud, such as during due diligence checks. Parties could also run analysis to identify trends in modus operandi, allowing them to take preventative measures.

BOX 13

Improving data quality and accuracy: Collaborative analysis for compliance

Part of Société Générale's London-based Greenhouse incubator programme, Secretarium is an 'integrity and confidentiality platform' that uses PETs to help financial institutions meet EU reporting requirements[247]. The EU Markets in Financial Instruments Directive II (MiFID II) requires financial organisations to report trade data to regulators. This entails using reference data of varying quality, making the task difficult and potentially ineffective. Data quality can be improved if multiple firms compare their client reference data to identify inaccuracies. However, financial institutions are not inclined to share data with competitors, as this would include client lists and sensitive personal data.

Secretarium uses a distributed, confidential computing platform to enable multiple institutions to compare data in a 'blind' fashion. It uses a distributed confidential computing format to benchmark reference data quality. A group of secured computers contain the organisations' reference data in an encrypted form, and the computers process the data without providing access to any individuals or organisations, even Secretarium itself.

246 Hazebroek E, Jonkers K, Segers T. 2021 Secure net NCSC partnership for rapid and safe data sharing. See https://emagazine.one-conference.nl/2021/secure-net-ncscs-partnership-for-rapid-and-safe-information-sharing/ (accessed 23 September 2022).

247 Secretarium (Homepage) See https://secretarium.com/ (accessed 27 September 2022).

SMPC can drive public sector efficiency by allowing for safe and rapid collective intelligence. While technical challenges such as performance and compute power were once primary challenges to implementation, this is no longer the case. One of the biggest challenges around SMPC is the understanding of legal implications, for instance the impact of EU and UK GDPR requirements. Other challenges include alignment of data structures and formats (interoperability), reliability and auditability, data availability and complexity of ongoing management of SMPC[248].

Only registered parties can contribute to SMPC analyses. Registered parties should not have intent to input information that is invalid (eg reporting false information).

SMPC applications may be purchased as a software package, which enables different parties to collaborate on sensitive data through analysis 'in the blind'. While open frameworks require deep knowledge of SMPC, suppliers are trialling software that will be usable by data scientists with no previous experience with SMPC.

248 The Financial Action Taskforce. 2021 Stocktake on data pooling, collaborative analytics and data protection. See https://www.fatf-gafi.org/media/fatf/documents/Stocktake-Datapooling-Collaborative-Analytics.pdf (accessed 22 September 2022).

BOX 13

Public-private partnerships for PETs in the Dutch public sector

Roseman Labs in the Netherlands is encouraging the uptake of PETs in the Dutch public sector through creative, low-risk collaborations that demonstrate the value of SPMC.

First, they identify use cases for SMPC relevant to a given public body. In some instances, a use case idea is generated through scoping conversations between public sector stakeholders. The use case idea is then formulated into a pilot project, or proof-of-concept, which can be carried out within the low-risk public procurement threshold (for example, conducting a six-month trial). Once the economic and social value of the SMPC solution becomes clear, the public sector organisation may begin an informed RFI process with an aim to scale-up the solution long-term.

This has resulted in successful applications, including:

- Increasing digital resilience with the Dutch National Cyber Security Centre (NCSC). The NCSC collects cybersecurity intelligence from organisations across the Netherlands, which report risks such as hacking or ransomware incidents. Organisations are not motivated to publish data on security breaches, which could compromise their reputation and marketability. An SMPC system now allows the NCSC to collect intelligence on cyber security risks from over tens of organisations (scaling to 15.000 over time) in the Netherlands in a private fashion: each organisation inputs data on cyber attacks and breeches in a fully anonymous and confidential way on a weekly basis. The NCSC does not see provenance information but is able to identify trends and take action accordingly;
- Reducing money laundering. Where multiple banks are able to generate graphs using transaction data, these graphs can be compared using SMPC for patterns that suggest money laundering (namely, money going in circles, or 'smurfing', where many small transactions that are ultimately deposited with one entity). This cross-bank identification of patterns is far more reliable than each individual bank looking at their own data – often generating a very large number of false positives. With this cross-bank approach, the likelihood of spotting true positives increases and Banks and LEAs can then focus resources. This allows law enforcement to set priorities. This should open up private partnerships between banks, for example, where thousands of employees are dedicated to identifying potential money laundering incidents (compared to just hundreds at the national public sector level)[249].

Roseman Labs has technical and in-house legal expertise (complemented with external privacy experts), meaning they are able to prescribe a data-use solution that meets current data protection requirements, helping clients to complete the Data Protection Impact (DPI) process together. This added value bolsters their work with public sector clients.

249 Roseman Labs (Secure data collaboration in financial services). See https://rosemanlabs.com/blog/financial_services.html (accessed 10 October 2022).

USE CASE 5.2

Online safety: Harmful content detection on encrypted platforms

The challenge

In April 2019, the UK Government published the Online Harms White Paper[250]; this paper identified the need to address the negative consequences that arise from individuals being online, both for social cohesion and for democratic society. The paper set out a programme of action to tackle content or activity that harms individual users, particularly children, either by undermining national security, or by destabilising shared rights and responsibilities. Many of the measures suggested in the white paper require social media platforms to take action[251]; Social media companies are required to identify and prevent the sharing of harmful and illegal content for a number of legal reasons[252]. Other motivations for regulating harmful content on private platforms include fear of new legislation and potential public relations backlash influencing their user base[253, 254].

One of the most serious forms of online offending is child sexual exploitation and abuse (CSEA). Since CSEA material can be shared and disseminated through social media platforms, the Online Harms White Paper identified social media companies as responsible for protecting their users from harm.

Detection of harmful and illegal content online

Many social media companies use live moderation, employed teams of human moderators, or automated detection systems to help combat harmful and illegal content on their platforms. Automated detection systems range from simple approaches such as matching images (using hashes) to the use of deep learning models trained on material that may be illegal. This poses a particular challenge to an automated approach to detecting harmful content[255].

Efforts to detect CSEA material can be severely restricted by end-to-end encryption: a secure method of transferring information (including messages or images). Due to the privacy afforded, encrypted messaging platforms can mask the sharing of illegal content.

250 HM Government. Online Harms White Paper. See https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973939/Online_Harms_White_Paper_V2.pdf (accessed 23 January 2022).

251 HM Government (Online Harms White Paper: consultation outcome). See https://www.gov.uk/government/consultations/online-harms-white-paper (accessed 15 March 2022).

252 Current legislation is complex, and includes the Malicious Communications Act 1988, the Communications Act 2003, the Public Order Act 1986, and the Investigatory Powers Act 2016.

253 Internet Watch Foundation (Our MOU, the law and assessing content). See https://www.iwf.org.uk/what-we-do/how-we-assess-and-remove-content/laws-and-assessment-levels (accessed 28 July 2022).

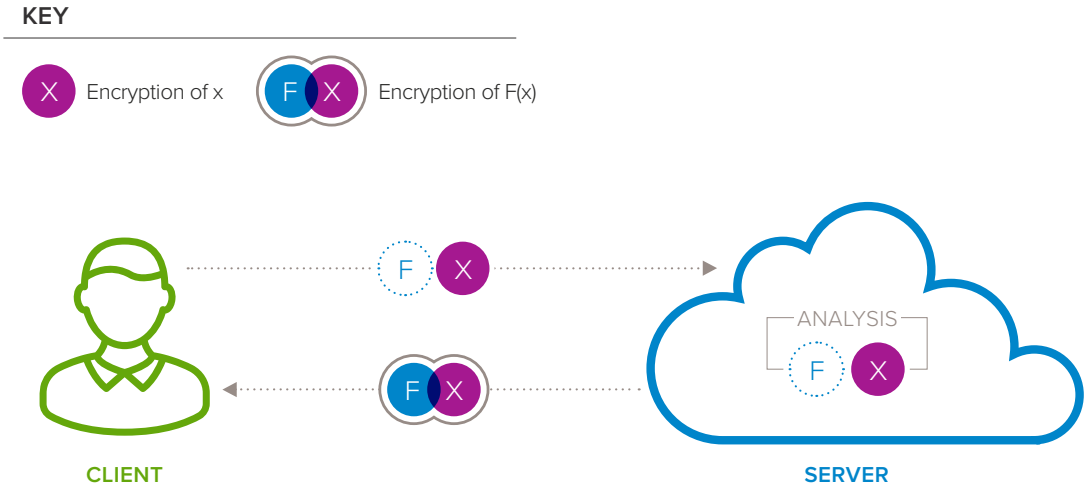
254 House of Commons Library. 2022 Regulating online harms (research briefing). See https://researchbriefings.files.parliament.uk/documents/CBP-8743/CBP-8743.pdf (accessed 27 September 2022).

255 Gillespie T. 2020 Content moderation, AI and the question of scale. Big Data & Society. 7. (https://doi.org/10.1177/2053951720943234)

FIGURE 7

Homomorphic encryption depicted in the context of a client-server model.

The client sends encrypted data to a server, where a specific analysis is performed on the encrypted data, without decrypting that data. The encrypted result is then sent to the client, who can decrypt it to obtain the result of the analysis they wished to outsource.



While the UK Government has considered the banning of end-to-end encryption in efforts to stymie CSEA material sharing[256], end-to-end encryption offers critical benefits to private citizens and must be preserved and promoted[257]. Recent technical advances may provide a solution to detecting harmful content without ending end-to-end encryption or the privacy of individual users.

Homomorphic encryption (HE) has been demonstrated as a PET that allows for the analysis of encrypted data, and which could be used as a tool for identifying CSEA material on encrypted platforms. Apple's planned roll-out of a very similar programme received criticism from privacy rights groups. The Apple case illustrates how PETs may be applied in ways perceived to violate, rather than preserve, privacy. This use case is intended to provide an explanation, rather than an endorsement, of how PETs could be used to detect illegal material on encrypted platforms.

256 HM Government (International statement: End-to-end encryption and public safety). See https://www.gov.uk/government/publications/international-statement-end-to-end-encryption-and-public-safety (accessed 20 September 2022).

257 The Royal Society. 2016 Progress and research in cybersecurity: Supporting a resilient and trustworthy system for the UK. See https://royalsociety.org/-/media/policy/projects/cybersecurity-research/cybersecurity-research-report.pdf (accessed 27 September 2022).

Underpinning technology: Image matching through hashing

Image hashing is a simple way to detect a specific image being shared, or to identify an image contained in high volumes of data. Hashing algorithms produce an output called a message digest, which is a unique text 'fingerprint' derived from an image (or other input). Message digests are short and easy to compare, yet unique to individual images. Message digests cannot be reversed from their text form back into an image. As such, these 'fingerprints' are well suited for detecting matching images held digitally without revealing the images themselves. First, an example image is hashed to a message digest, then this is compared to the digests of candidate files. If a match is found, this means the image with an identical digest is the same as the example image. Such lists of illicit or illegal hashes are known as 'matching databases'.

Typical hashing does not account for similarity between images. If one bit in an image is changed, the hash will change completely. This property helps ensure a low false detection rate; however, this means that small alterations – like subtle changes in colour, rotations, skewing, or mirroring – could enable the image to evade typical hashing.

Alternative hashing techniques may address these issues. One alternative is Locality Sensitive Hashing (LSH), which accounts for visually similar images (it intentionally hashes similar inputs to close or identical outputs). LSH and similar alternatives are useful where small variations in the input image are expected. However, transformations of the image, such as mirroring, could result in completely distinct raw data and would not be matched. The digital definition of 'similar' is not necessarily comparable to human perceptions of similarity.

PETs and image matching

PETs can help to ensure that image matching is done securely. A system used to detect CSEA material may be held within the image library of a mobile device. A verified third party (such as a law enforcement agency) would retain a matching database with hashes of images that have been categorised as CSEA material. The system could check whether the hashes of the images stored on the mobile device match any of the known illegal material in the matching database without sending the matching database out to the mobile device or revealing the user's photos.

Private Set Intersection (PSI) allows two parties who independently hold data elements to find the intersection of their data – that is, the elements held in common between two parties. In this system, PSI could be used to allow a third party to detect any image hashes which match their matching database without sharing the hashes. In this way, the third party learns only about any images which match their own hash list, but nothing about any images which do not match. Security is preserved for all non-matching elements. In not sharing the hash lists, the risks of bad actors being able to circumvent the detection using this knowledge is eliminated.

The ability to securely detect CSEA on mobile devices requires the combination of several techniques. Perceptual hashing can help to match images with a matching database of illicit material, even images with small perceptual changes. Combining this with private set intersection preserves the security of the matching database, whilst providing privacy for individuals. Together, these technologies could be used to develop a robust CSEA detection system that does not compromise end-to-end encryption.

Risks and challenges

One challenge is the potential for 'scope creep' – the adding of additional functionality above and beyond the detection of CSEA material. This may include, for example, state actors using the technology to counter digital piracy, digital rights management, or for national security and surveillance purposes. Platforms may face reputational risk and loss of users if systems were perceived as disproportionate surveillance tools.

While perceptual hashing algorithms allow for modified images to be matched against the database, they are also more likely to flag false positives. Innocent images may appear close enough to illicit images to be flagged by the machine learning system. This could lead to innocent people being identified as possessing CSEA material, entailing negative impact for the individual. The performance of the perceptual hashing system would need to be closely tested and monitored to measure the false positive rates. Ultimately, a human moderator should always verify whether flagged material is illegal or harmful; a user should never be charged based on the automated system detection alone.

Legal challenges and public trust must also be addressed. Law enforcement agencies would need to be clear on the legal basis for running such systems, which may constitute a passing on of their legal duties to third parties. This has implications for public trust, particularly where on-device screening is used.

The UK government aims to minimise the existence of spaces online where illegal material can be securely shared. Likewise, social media companies are motivated to ensure users are not breaching their terms of use, even in encrypted spaces. A PETs-enabled system for identifying illegal material is an alternative to privacy rollbacks such as the outright banning of end-to-end encryption.

BOX 14

Apple Tech child safety features

In August 2021 Apple announced new child safety features to be implemented on its US devices. Three planned changes aimed to mitigate child sexual abuse. One change related to iCloud Photos, which would scan images to find CSEA. While cloud service companies such as Google, Microsoft and Dropbox already scan material for CSEA, Apple planned to conduct scans on personal iPhone devices using a technology called NeuralHash.

NeuralHash scans images without revealing them to moderators. It translates the image into a unique number (a hash) based on its features. Before uploading to iCloud Photos, the hash is compared on-device against a database of known CSEA hashes provided by child safety organisations. Any matches prompt the creation of a cryptographic safety voucher. If a user reaches a threshold of safety vouchers, they are decrypted and shared with Apple moderators for review[258].

The proposals were welcomed by many, including child safety organisations[259]. However, the image hashing feature faced criticism from privacy advocates, cryptographers and other tech companies, who viewed Apple's proposals as introducing a backdoor on their devices.

Critics argued this could make the system vulnerable to state censorship of political dissent or LGBTQ+ content, or flagging of innocent images, causing unnecessary distress. Further criticism targeted the efficacy of the system: researchers reverse-engineered the hashing algorithm and were able to create images that were falsely flagged by the system[260].

In September 2021, Apple announced it was pausing implementation of CSAM scanning to collect feedback and make improvements. In April 2022, Apple announced its intention to introduce the parental control safety feature on the Messages app on iPhones in the UK[261].

It is unclear how an image hashing program would operate under UK and EU data protection law. On-device screening would likely entail explicit consent and user opt-in (rather than opt-out)[262]. User images are not necessarily personal data under the GDPR; they must depict identifiable living people, or be linked to a living person, to constitute personal data. However, neural hashes may constitute personal data. These emerging legal questions, as well as general public scepticism, suggest that an on-device detection system may face barriers in the UK or EU contexts.

258 Apple. CSAM detection: technical summary. See https://www.apple.com/child-safety/pdf/CSAM_Detection_Technical_Summary.pdf (accessed 20 March 2022).

259 O'Neill PH. 2021 Apple defends its new anti-child-abuse tech against privacy concerns. *MIT Technology Review*. 6 August 2021. See https://www.technologyreview.com/2021/08/06/1030852/apple-child-abuse-scanning-surveillance/ (accessed 22 March 2022).

260 Brandom R. 2021 Apple says collision in child-abuse hashing system is not a concern. *The Verge*. 18 August 2021. See https://www.theverge.com/2021/8/18/22630439/apple-csam-neuralhash-collision-vulnerability-flaw-cryptography (accessed 10 October 2022).

261 Hern A. 2022 Apple to roll out child safety feature that scans messages for nudity to UK iPhones. *The Guardian*. 20 April 2022. See https://www.theguardian.com/technology/2022/apr/20/apple-says-new-child-safety-feature-to-be-rolled-out-for-uk-iphones (accessed 23 April 2022).

262 Cobbe J. 2021 Data protection, ePrivacy, and the prospects for Apple's on-device CSAM Detection system in Europe. SocArXiv Papers. See 10.31235/osf.io/rhw8c (accessed 10 October 2022).

USE CASE 5.3

Privacy and verifiability in online voting and electronic public consultation

The opportunity

Remote online voting offers to bring the ballot to the voter, allowing convenience, flexibility and greater access to the democratic process[263].

Cryptography plays a critical role in providing electronic voting and counting[264]. Online voting has been used in elections in Estonia since Cybernetica's IT Lab built the first online voting solution in 2005[265]. Following the launch of multi-channel voting (in which votes can be cast using mail, traditional written ballots or online), voter participation has risen in Estonia, with 47% of voters voting online in 2021. Online voting has reportedly reduced public spending on elections[266].

The challenge

Democratic elections depend on security and auditability for fair and accurate collection and counting of votes. In online voting these analogue threats become digital, requiring solutions that can ensure votes are kept accurate, secret, anonymous and auditable simultaneously.

Many approaches can be used for electronic and internet voting, some of which include PETs. For example, homomorphic encryption can be used in electronic voting to achieve security; election results can be tallied without decrypting the votes[267]. This works well in small-scale elections. However, compute power entails high costs when scaled up.

It may be that hybrid schemes, which use cryptographic tools layered with other approaches (such as blockchain) may be the most robust. One such solution has recently been prototyped by the Smartmatic-Cybernetica Centre of Excellence for Internet Voting (SCCEIV)[268].

263 WebRoots Democracy. 2020 The Cratos Principles. See https://webrootsdemocracy.files.wordpress.com/2020/04/the-cratos-principles-webroots-democracy-v2.pdf (accessed 20 August 2022).

264 National Democratic Institute (The important uses of cryptography in electronic voting and counting). See https://www.ndi.org/e-voting-guide/examples/cryptography-in-e-voting (accessed 2 September 2022).

265 Smartmatic (Estonia: the world's longest standing, most advanced voting solution). See https://www.smartmatic.com/case-studies/article/estonia-the-worlds-longest-standing-most-advanced-internet-voting-solution/ (accessed 10 October 2022).

266 Krimmer R, Duenas-Cid D, Krivonosova I, Vinkel P, Koitmae A. 2018 How much does an e-vote cost? Cost comparison per vote in multichannel elections in Estonia. *Lecture Notes in Computer Science* (Conference paper) 117—131. (https://doi.org/10.1007/978-3-030-00419-4_8)

267 National Democratic Institute (The important uses of cryptography in electronic voting and counting). See https://www.ndi.org/e-voting-guide/examples/cryptography-in-e-voting (accessed 2 September 2022).

268 Smartmatic (Smartmatic—Cybernetica awarded European Commission blockchain research project). See https://www.smartmatic.com/media/article/smartmatic-cybernetica-awarded-european-commission-blockchain-research-project/ (accessed 10 October 2022).

In 2014, Cybernetica partnered with Smartmatic to develop TIVI, an online voting solution that aims to guarantee end-to-end integrity in remote voting. TIVI has been used in Estonia, Norway, Chile and parts of the US. With this technology, a voter verifies their identity using a digital or mobile ID. i-Voting is optional. A voter can cast multiple i-Votes, with only the final vote counted; a paper vote always takes precedence over an i-Vote.

PETs and distributed ledgers: Tiviledge

SCCEIV more recently developed Tiviledge, a prototype for privacy-preserving, auditable i-voting. It can be used with the TIVI platform and includes PETs. It focuses on making election data available for independent audits while meeting the condition of a secret ballot[269].

Tiviledge uses zero knowledge proofs and secure multi-party computation to verify votes and summate totals. It writes the results on an immutable, auditable distributed ledger. A distributed ledger is a shared database, which can serve as a public record. It may only be added to: any tampering attempt is made obvious because it is synchronised and distributed across multiple hosts. This guarantees integrity of the record. Today, elections rely heavily on a central organisation, and trusting the integrity of an election means trusting a single entity. A distributed ledger means election results are verifiable to external auditors.

Tiviledge is currently a research prototype for experimental and developmental use; it is not open source. Several key areas must be addressed prior to any legally binding use of the technology for voting. First, compatibility with legal requirements must be considered within a given jurisdiction, particularly where there are complex voting protocols (eg beyond standard 'winner takes all' models). Second, a more robust system of verification will be key to avoid fraud or breach of voter privacy. Third, the protection of the election privacy key should be considered (for example, using hardware security) to prevent an attacker from gaining access to information.

Tiviledge is one prototype developed by the PRIViLEDGE project[270], funded by Horizon Europe (see page 28).

269 Archer DW *et al.* 2018 From keys to databases: Real-world applications of secure multi-party computation. *Comput J.* 61, 1749—1771. (https://doi.org/10.1093/comjnl/bxy090)

270 PRIViLEDGE Project (Homepage). See https://priviledge-project.eu/ (accessed 30 March 2022).

USE CASE 6

PETs and the mosaic effect: Sharing humanitarian data in emergencies and fragile contexts

The opportunity

In the last ten years there has been a substantial rise in the volume and variety of data produced during and for humanitarian responses and development programmes. Humanitarian data may contain telecommunications, messaging and other ICT data, information from mobile money or cash transfer applications, banking or smart cards, as well as social media data[271]. It can include contextual data (such as damage assessment or geospatial data), information about people affected by a crisis (including their needs) or information related to response efforts (such as transportation infrastructure, food prices, or the availability of education facilities)[272]. Emergency or crisis-related data may include traditional humanitarian data, as well as user generated data, such as social media posts or locations entered through GPS tracking apps.

At a larger scale, and over time, humanitarian and crises data can inform understandings of patterns – such as environmental catastrophes, or cycles of social conflict – assisting in anticipatory action and the reduction of negative impacts. The use of long-term crisis insights during the COVID-19 pandemic has led to wider reflections on data governance in the early outbreak of COVID-19 and the role PETs might have played[273, 274].

271 Privacy International. 2018 The humanitarian data problem: 'doing no harm' in the digital era. See https://privacyinternational.org/sites/default/files/2018-12/The%20Humanitarian%20Metadata%20Problem%20-%20Doing%20No%20Harm%20in%20the%20Digital%20Era.pdf (accessed 10 October 2022).

272 OCHA Centre for Humanitarian Data 2021. Data Responsibility Guidelines. See https://data.humdata.org/dataset/2048a947-5714-4220-905b-e662cbcd14c8/resource/60050608-0095-4c11-86cd-0a1fc5c29fd9/download/ocha-data-responsibility-guidelines_2021.pdf (accessed 10 October 2022).

273 El Emam K. 2020 Viewpoint: Implementing privacy-enhancing technologies in the time of a pandemic. *Journal of Data Protection & Privacy*. 3, 344—352.

274 Shainski R, Dixon W. 2020 How privacy enhancing technologies can help COVID-19 tracing efforts. *World Economic Forum Agenda*. 22 May 2020. See https://www.weforum.org/agenda/2020/05/how-privacy-enhancing-technologies-can-help-covid-19-tracing-efforts/ (accessed 10 October 2022).

The challenge

Today, big data plays a fundamental role in responses to humanitarian crises and other emergency scenarios. At the same time, new technologies – such as biometrics, mobile banking and drones – simultaneously provide new avenues for security and privacy risks.

Humanitarian datasets contain information about some of the world's most at-risk people, including refugees and internally displaced people fleeing their homes due to persecution, conflict, and disaster[275]. The risk of reidentification of an individual, or the disclosure of a personal attribute or characteristic, often entails magnified harms in these fragile contexts: as the Resolution on Privacy and International Humanitarian Action outlines, 'data that would normally not be considered as sensitive under data protection laws may be very sensitive in humanitarian emergencies' context'[276]. Responsible data stewardship must ensure the safety of these groups by understanding potential harms and working toward prevention.

As datasets accumulate, so too does the likelihood of content overlap between them. This is especially true in concentrated settings, such as refugee camps. The more information that is common across multiple datasets, the higher the disclosure risk posed by the 'mosaic effect': the potential for individuals or groups to be re-identified through using datasets in combination, even though each dataset has been made individually safe.

The UN Global Pulse recommended in 2015 that risks in humanitarian data use be assessed according to level of data security and availability of PETs; they also recommended using PETs in conjunction with other methods (such as anonymisation) to employ privacy by design principles from the outset. However, few examples of PETs in humanitarian data exist, possibly because they are highly technical and so can be expensive to deploy.

The mosaic effect risk

The mosaic effect risk is described as the potential for 'disparate items of information, though individually of limited or no utility to their possessor, [to] take on added significance when combined with other items of information'[277]. The mosaic effect suggests that even de-identified or pseudonymous data can be reidentified if other datasets or complementary information are combined, revealing significant new information. This could disclose, for example, the identity and location of people from minoritised groups. While such information could be used to inform effective humanitarian responses, it could also be used to do harm.

275 Inter-Agency Standing Committee 2021. Operational Guidance on Data Responsibility in Humanitarian Action. See https://interagencystandingcommittee.org/system/files/2021-02/IASC%20Operational%20Guidance%20on%20Data%20Responsibility%20in%20Humanitarian%20Action-%20February%202021.pdf (accessed 10 October 2022).

276 Global Privacy Assembly. 2015 37th International Conference of Data Protection and Privacy Commissioners. See http://globalprivacyassembly.org/wp-content/uploads/2015/02/Resolution-on-Privacy-and-International-Humanitarian-Action.pdf (accessed 10 October 2022).

277 Pozen DE. 2005 The mosaic theory, national security, and the freedom of information act. Yale L J. 115.

BOX 15

Collective intelligence in disaster management and emergency response

Emergency and disaster management are significant areas of research that use social media data. Using near-real time data, researchers identify communities impacted, the geographical spread of an event, and gain understanding of public behaviour during a disaster[278]. In these cases, social media data represents the collective intelligence of users on the ground who act as 'sensors', relaying real-time observations from the field.

Some examples include:

- A team of international researchers used Twitter data and location metadata following the 2012 Hurricane Sandy evacuation in Florida, USA. They found a correlation between per-capita hurricane-related Twitter activity and per-capita economic hurricane damage, suggesting disaster-related social media could be used for rapid damage assessments[279];
- A machine learning algorithm and semantic analysis were used to classify tweets to detect earthquakes in Japan. The team at The University of Tokyo was able to detect earthquakes registering magnitude 3+ with high probability (93% of those detected by the Japan Meteorological Agency) simply by monitoring tweets, delivering notifications much faster than national broadcasted announcements[280] and;
- Researchers at the University of Georgia analysed images of lost-and-found tornado debris shared on social media following a 2011 outbreak in the south-eastern US. Using Geographic Information System mapping and trajectory modelling techniques, this was the most comprehensive study to date on debris trajectory from a tornado outbreak[281].

278 Chae J, Thom D, Jang Y, Kim S, Ertl T, Ebert DS. 2014 Public behavior response analysis in disaster events utilizing visual analytics of microblog data. *Comput. Graph*. 38, 51–60. (https://doi.org/10.1016/j.cag.2013.10.008)

279 Kryvasheyeu Y *et al*. Rapid assessment of disaster damage using social media activity. *Sci Adv*. 2. (https://doi.org/10.1126/sciadv.1500779)

280 Sakaki T, Okazaki M, Matsuo Y. Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Trans. Knowl. Data Eng*., 25, 919–931. (https://doi.org/10.1109/tkde.2012.29)

281 Knox AJ *et al*. 2013 Tornado debris characteristics and trajectories during the 27 April 2011 super outbreak as determined using social media data. *Bulletin of the American Meteorological Society*. 94, 1371—1380.

The mosaic effect risk is related to the increased use of metadata, or *data about other data*, in humanitarian contexts. This could be the time and location of a message sent, rather than the content of the message itself. Communications with people affected by crises can include social media or SMS messaging, sharing information-as-aid, mobile cash transfer programmes, and monitoring and evaluation systems (such as those used to detect fraud), all of which entail rich and potentially compromising metadata[282]. Privacy International has therefore recommended that humanitarian organisations practice *do no harm* principles by understanding how the data and metadata they store and use may be employed for purposes beyond aid – such as for profiling, surveillance or political repression. This highlights the need for mitigation tools to be developed (such as PETs) and the importance of data minimisation.

The Centre for Humanitarian Data is focused on increasing the use and impact of data in the humanitarian sector and are interested in the potential for PETs to address the mosaic effect and enhance collaboration[283]. They make the following recommendations:

- **Technical actions** Humanitarian organisations should invest in further strengthening metadata standards and interoperability, enabling monitoring of related datasets to counter mosaic effect risks;
- **Procedural actions** A data asset registry and data ecosystem mapping assessment should be completed as per the recommendations included in the IASC Operational Guidance on Data Responsibility in Humanitarian Action (2021);
- **Governance actions** Sector-wide fora should be used to ensure that datasets are not shared on different platforms at different levels of aggregation, and determine consistent standards for approaches such as anonymisation and;
- **Legal actions** Humanitarian organisations can also improve licensing for datasets by adding clauses that prohibit joining datasets or analysing data with the purpose of re-identification or attribute disclosure. While this will not prevent intentional misuse, it will help explain what type of use goes against the use allowed by the sharing organisation.

282 Privacy International. Humanitarian Metadata Problem Doing No Harm in the Digital Era. See https://privacyinternational.org/sites/default/files/2018-12/The%20Humanitarian%20Metadata%20Problem%20-%20Doing%20No%20Harm%20in%20the%20Digital%20Era.pdf (accessed 28 September 2022).

283 Weller S. 2022 Minimizing privacy risks in humanitarian data. *Privitar* blog. 9 March 2022. See https://www.privitar.com/blog/fragility-forum-minimizing-privacy-risks-in-humanitarian-data/ (accessed 10 October 2022).

The role of PETs in countering the risk of mosaic effect

PETs could help safeguard personal data while still allowing researchers to utilise it in humanitarian efforts. Differential privacy could be used to add 'noise', to make any one true datapoint more difficult to trace to a real individual. The resulting 'noisy' dataset can then be shared between organisations more safely. The noise can be adjusted for extra privacy (and reduced utility), allowing data controllers to make contextual privacy-utility trade-offs.

Federated learning could be used on geospatial datasets, such as people's locations, without sharing the data used to train the model. This would entail training a model, or a predictive algorithm, on a local geospatial dataset. The model would then be shared for training on remote datasets at other organisations, which are never revealed to the model owner. External organisations holding relevant data might include telecoms, other humanitarian organisations, or social media sites, all of which may not have established data partnerships or sharing agreements.

The model would return to the owner with improved ability to predict peoples' movements or locations. This type of model would be incredibly valuable for humanitarian organisations making decisions about where to direct resources during crises. There are already examples of federated learning being used in medical research (see Use case 1.1, page 57). Homomorphic encryption has been used to perform large-scale studies on cross-border health data (see Use case 1.1, page 57), including multiple institutions in collaboration and crowdsourced materials (such as genomics)[284].

PETs could help safeguard personal data while still allowing researchers to utilise it in humanitarian efforts.

284 Blatt M, Gusev A, Polyakov Y, Goldwasser S. 2020 Secure large-scale genome-wide association studies using homomorphic encryption. *P Natl Acad Sci USA*. 117, 11608—11613. (https://doi.org/10.1073/pnas.1918257117)

BOX 16

Tackling human trafficking through cooperation between law enforcement and NGOs.

A partnership was established between Dutch law enforcement and NGOs working against human trafficking. The law enforcement agency (LEA) wanted to shadow potential trafficking victims from a long list of identified individuals.

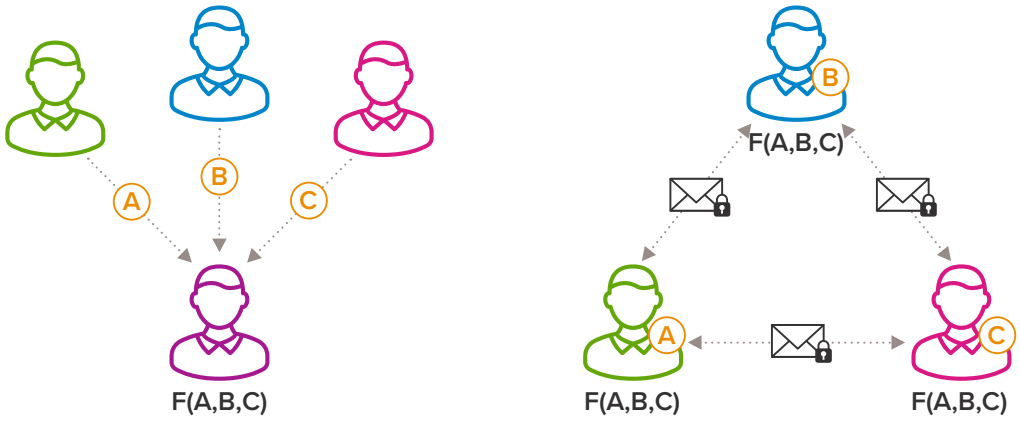
However, local human trafficking NGOs also held informant lists, with potential overlap between their lists and that of the LEA. The NGOs were concerned that their informants would feel confidentiality had been breached by the NGO if they were shadowed by the LEA. The long list from law enforcement was compared to the short list from the NGO 'in the blind' using SMPC[285].

The result was a random list of 20 people who were candidates for LEA shadowing. A future SMPC application may include tracking the movements of potential trafficking victims across agencies and NGOs without sharing their names, to identify trends and potential trafficking routes. This approach could also shed light on the extent of human trafficking crimes more widely – an issue otherwise impossible to measure.

FIGURE 8

Private multi-party machine learning with MPC

Using MPC, different parties send encrypted messages to each other, and obtain the model F(A,B,C) they wanted to compute without revealing their own private input, and without the need for a trusted central authority.



Central trusted authority **Secure multi-party machine learning**

285 Pinsent Masons. Data sharing coalition helps flag victims of human trafficking. See https://www.pinsentmasons.com/out-law/news/data-sharing-coalition-helps-flag-victims-of-human-trafficking (accessed 7 July 2022).

Conclusions

There are no known cases of the mosaic effect causing harm in humanitarian, crises or development scenarios. At the same time, there is a cost to not sharing or linking data in such cases, particularly where lives may be saved. Humanitarian organisations seek to consolidate and strengthen approaches to reduce risk through data responsibility practices and increasing cross-organisational work (including with NSOs [Use Case 4]).

Some humanitarian organisations reserve the right to nondisclosure[286]; however, in fast-developing situations data flows become difficult to control. Anticipatory tools, such as Privacy Impact Assessments or Data Protection Impact Assessments, could be used to better mitigate downstream harms stemming from the mosaic effect. Further, crisis situations can change the calculus of harm vs potential benefit[287].

PETs may expand the bounds of possibility in providing strong privacy and high utility from a dataset. However, a broader approach should also consider the ethics of humanitarian projects, training and vetting of trusted researchers, robust data sharing agreements and other legal controls, as well as security access controls and locked down physical hardware. Historically, organisations have used approaches such as vetting an internal team of trusted researchers, drafting bilateral data sharing agreements, or other legal tools. Even with additional technical safeguards, these non-technical solutions remain important.

286 Council of the European Union. 2012 Applicability of the General Data Protection Regulation to the activities of the International Committee of the Red Cross. See http://data.consilium.europa.eu/doc/document/ST-7355-2015-INIT/en/pdf (accessed 26 June 2022).

287 Veale M, Binns R, Edwards L. 2018 Algorithms that remember: model inversion attacks and data protection law. *Philos T R Soc A.* 376. (https://doi.org/10.1098/rsta.2018.0083)

Conclusions

Left
Shanghai, China. © iStock / AerialPerspective Works.

Conclusions

PETs may be used in any scenario where data benefits those with exclusive access, or where open access could cause harm. This could include, for example, data pertaining to natural resources (to prevent over-exploitation).

This report sets out to refresh perspectives on PETs following the society's 2019 report *Protecting privacy in practice*. In doing so, it considers the role of PETs beyond data protection and highlights the secondary effects of PETs in motivating partnerships and enabling collaboration across sectors and international borders. The risk of personal data use is considered in terms of privacy attack (what is technically possible) as well as the severity of potential downstream harms of compromised data (which is contextual).

Several questions remain beyond the scope of this report and suggest areas for further research. First, very little is known about the potential market value of PETs as discreet technologies, or their true significance in data use in collaborative scenarios. It is therefore difficult to estimate what value would be unlocked with widespread uptake of PETs, whether in economic terms or in social benefit. The market value of PETs may also depend on trends in use cases, whether PETs are employed as security tools or for increased collaborative learning and analysis.

Second, this report has not explored the full range of potential follow-on effects of PETs adoption. These include potential harms which may stem from greater monitoring and surveillance on the part of governments and private sector actors, leading to enhanced profiling and resulting in increased distrust of public services and loss of privacy in online spaces (such as through highly targeted advertisement). In some cases, PETs are already being used to facilitate business-as-usual in online advertising[288], easing companies' access to, and use of, customer data to the usual ends.

Given their multipurpose nature, networked PETs that allow for collaborative analysis might be viewed as an upgrade to traditional systems of information sharing, such as the internet, rather than new privacy compliance tools. For this reason, in the future, PETs may be used for any sufficiently valuable data, not just sensitive category data (such as personal or commercially advantageous data). Rather, PETs may be used in any scenario where data benefits those with exclusive access, or where open access could cause harm. This could include, for example, data pertaining to natural resources (to prevent over-exploitation).

Finally, more work is required to integrate PETs into wider data governance systems. The tendency for PETs to be developed as discreet technologies has led users to approach PETs as a set of tools, each with unique problem-solving capabilities. In the future, PETs may operate more like complementary pieces of machinery which, when combined with other technological, legal and physical mechanisms, will amount to automated data governance systems. These systems could help to enact an organisation's data policy and facilitate responsible information flows at unprecedented scales. This next level of PETs abstraction will require collaboration between PETs developers and leading organisations to develop and test use cases.

PETs can play an important role in a privacy by design approach to data governance when considered carefully, informed by appropriate guidance and assurances. Given the rapid development of these technologies, it is a critical time to consider how PETs will be used and governed for the promotion of human flourishing.

288 See Box 10, page 78 on Cambridge Analytica.

Appendices

APPENDIX 1:

Definitions

Differential privacy: security definition which means that, when a statistic is released, it should not give much more information about a particular individual than if that individual had not been included in the dataset. See also privacy budget.

Distributed Ledger Technology (DLT): an open, distributed database that can record transactions between several parties efficiently and in a verifiable and permanent way. DLTs are not considered PETs, though they can be used (as some PETs) to promote transparency by documenting data provenance.

Epsilon (Ɛ): see privacy budget.

Fully homomorphic encryption (FHE): a type of encryption scheme which allows for any polynomial function to be computed on encrypted data, which means both additions and multiplications.

Homomorphic encryption (HE): a property that some encryption schemes have, so that it is possible to compute on encrypted data without deciphering it.

Metadata: data that describes or provides information about other data, such as time and location of a message (rather than the content of the message).

Mosaic effect: the potential for individuals of groups to be re-identified through using datasets in combination, even though each dataset has been made individually safe.

Noise: noise refers to a random alteration of data/values in a dataset so that the true data points (such as personal identifiers) are not as easy to identify.

Privacy budget (also differential privacy budget, or epsilon): a quantitative measure of the change in confidence of an individual having a given attribute.

Privacy-preserving synthetic data (PPSD): synthetic data generated from real-world data to a degree of privacy that is deemed acceptable for a given application.

Private Set Intersection (PSI): secure multiparty computation protocol where two parties compare datasets without revealing them in an unencrypted form. At the conclusion of the computation, each party knows which items they have in common with the other. There are some scalable open-source implementations of PSI available.

Secure multi-party computation (SMPC or MPC): a subfield of cryptography concerned with enabling private distributed computations. MPC protocols allow computation or analysis on combined data without the different parties revealing their own private inputs to the computation.

Somewhat Homomorphic Encryption (SHE): a type of encryption scheme which supports a limited number of computations (both additions and multiplications) on encrypted data.

Synthetic data: data that is modelled to represent the statistical properties of original data; new data values are created which, taken as a whole, reproduce the statistical properties of the 'real' dataset.

Trusted Execution Environment (TEE): secure area of a processor that allows code and data to be isolated and protected from the rest of the system such that it cannot be accessed or modified even by the operating system or admin users. Trusted execution environments are also known as secure enclaves.

APPENDIX 2:

Acknowledgements

Working Group members

The members of the Working Group involved in this report are listed below. Members acted in an individual and not a representative capacity and declared any potential conflicts of interest. Members contributed to the project on the basis of their own expertise and good judgement.

Chair
Professor Alison Noble FRS FREng OBE, Technikos Professor of Biomedical Engineering and Department of Engineering Science, University of Oxford

Members
Professor Jon Crowcroft FRS FREng, Marconi Professor of Communications Systems in the Computer Lab, University of Cambridge; Alan Turing Institute
George Balston, Co-Director, Defence and Security, Alan Turing Institute
Professor Sir Anthony Finkelstein CBE FREng, President, City University London
Guy Cohen, Independent
Dr Benjamin R Curtis, Senior Researcher, Zama
Professor Emiliano de Cristofaro, Professor of Security and Privacy Enhancing Technologies, University College London
Dr Marion Oswald, Associate Professor in Law, University of Northumbria
Professor Carsten Maple, Professor of Cyber Systems Engineering, University of Warwick Cyber Security Centre
Dr Suzanne Weller, Head of Research, Privitar

Royal Society staff

Royal Society secretariat
Dr June Brawner, Senior Policy Adviser and Project Lead
Areeq Chowdhury, Head of Policy, Data
Dr Natasha McCarthy, Head of Policy, Data (until February 2022)
Dr Franck Fourniol, Senior Policy Adviser (until July 2021)

Royal Society staff who contributed to the development of the project
Dr Rupert Lewis, Chief Science Policy Officer
Dr Mahi Hardalupas, Project Coordinator (until July 2022)
Helena Gellersen, Patricia Jimenez, Louise Parkes, UKRI work placement (various periods)

Reviewers

This report has been reviewed by expert readers and by an independent Panel of experts, before being approved by Officers of the Royal Society. The Review Panel members were not asked to endorse the conclusions or recommendations of the report, but to act as independent referees of its technical content and presentation. Panel members acted in a personal and not a representative capacity. The Royal Society gratefully acknowledges the contribution of the reviewers.

Reviewers
Dr Clifford Cocks CB FRS
Andrew Trask, Founder and Leader, OpenMined
Alex van Someren FREng, Chief Scientific Adviser for National Security, UK Government

Event participants

The Royal Society would like to thank all those who contributed to the development of this project, in particular through participation in the following events.

PETs Contact Group Session One: *Evidence and advice needs* (21 April 2021)
15 participants from UK government, regulators and civil society.

PETs Contact Group Session Two: *Use cases and outputs development* (18 October 2021)
13 participants from UK government, regulators and civil society.

Contributors of use cases and standards

The use cases and standards chapters received domain-specific input from a range of experts in research, industry and civil society. Domain experts were not asked to endorse the conclusions or recommendations of the report, but to act as independent referees of the use cases and standards chapters, their technical content and presentation. Contributors acted in a personal and not a representative capacity. The Royal Society gratefully acknowledges their contributions.

Domain experts who consulted on use cases and standards
Gerry Reilly, Health Data Research UK
Greg A Johnston, Energy Systems Catapult
Alex Howard, Octopus Energy Centre for Net Zero (until September 2021)
Dr Louisa Nolan, Office for National Statistics (until May 2022)
Dr Sergey M Plis, Dr Vince D Calhoun and Eric Verner, COINSTAC
Sahar Danesh, British Standards Institution
Annemarie Büttner, Independent expert

Commissioned evidence-gathering and reviews

- Hattusia 2022. The current state of assurance in establishing trust in PETs. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/
- Jordon J *et al*. 2022 Synthetic data: What, why and how? See https://arxiv.org/pdf/2205.03257.pdf (accessed 2 September 2022).
- London Economics and the Open Data Institute. 2022 Privacy Enhancing Technologies: Market readiness, enabling and limiting factors. The Royal Society. See https://royalsociety.org/topics-policy/projects/privacy-enhancing-technologies/ This project was partly funded by a grant from the Centre for Data Ethics and Innovation.

Market research London Economics / Open Data Institute PETs market research

The Royal Society worked with London Economics and the Open Data Institute on exploratory research into the state of PETs adoption, barriers and incentives within key UK public sector data institutions. A sample of seven public sector organisations were interviewed by invitation, chosen to represent a cross-section of criteria including function relevant to data (eg storage, processing) and type of data used.

Public sector organisations profiled for PETs market readiness research
Competition and Markets Authority
DataLoch
Department for Transport
Government Digital Service
Greater London Authority
National Archives
Office for National Statistics

The Royal Society is a self-governing Fellowship of many of the world's most distinguished scientists drawn from all areas of science, engineering, and medicine. The Society's fundamental purpose, as it has been since its foundation in 1660, is to recognise, promote, and support excellence in science and to encourage the development and use of science for the benefit of humanity.

The Society's strategic priorities emphasise its commitment to the highest quality science, to curiosity-driven research, and to the development and use of science for the benefit of society. These priorities are:

- The Fellowship, Foreign Membership and beyond
- Influencing
- Research system and culture
- Science and society
- Corporate and governance

For further information
The Royal Society
6 – 9 Carlton House Terrace
London SW1Y 5AG

T +44 20 7451 2500
W royalsociety.org

Registered Charity No 207043

ISBN: 978-1-78252-627-8
Issued: January 2023 DES7924