# Future of Financial Intelligence Sharing (FFIS)

**Innovation and discussion paper:**

Case studies of the use of privacy preserving analysis to tackle financial crime

RUSI
www.rusi.org

FFIS
Future of Financial Intelligence Sharing

# Future of Financial Intelligence Sharing (FFIS)

## Innovation and discussion paper:
Case studies of the use of privacy preserving analysis to tackle financial crime

June 2020


## Abstract

This Discussion Paper provides an overview of concepts and recent technical developments relevant to privacy preserving analytics and publishes eight case studies of the use of privacy preserving analysis to tackle financial crime, updated to June 2020. These case studies demonstrate how financial institutions are exploring advances in this field of cryptographic technology to enable analysis of data from across multiple participating organisations to inform financial crime risk awareness, without the need for those organisations to share underlying sensitive data. This mapping exercise – highlighting the current state of development of privacy preserving analysis in financial crime prevention use-cases – forms the first component in a wider international FFIS research project into "The Role of Privacy Preserving Data Analytics in the Detection and Prevention of Financial Crime". The case study mapping exercise will be updated throughout the lifetime of the wider research project, up to mid-2021. Overall, this paper is intended to raise awareness about the growth of privacy preserving analytics in the field of anti-money laundering and financial investigations and accelerate discussions about the implications of that growth.



Global strategic partners of the FFIS programme in 2020:

VERAFIN OLIVER WYMAN WESTERN UNION WU

REFINITIV SWIFT INSTITUTE

# About

This Discussion Paper is produced by the Future of Financial Intelligence Sharing (FFIS) programme, as part of our mission to lead independent research into the role of public-private financial information-sharing to detect, prevent and disrupt crime. The FFIS programme is a research partnership between the RUSI Centre for Financial Crime & Security Studies and NJM Research.

Founded in 1831, the Royal United Services Institute (RUSI) is the world's oldest and the UK's leading defence and security think tank. Its mission is to inform, influence and enhance public debate on a safer and more stable world. RUSI is a research-led institute, producing independent, practical and innovative analysis to address today's complex challenges. London | Brussels | Nairobi | Doha | Tokyo | Washington, DC

Every effort has been made to verify the accuracy of the information contained in this report. All information was believed to be correct as of 18 June 2020. Nevertheless, the FFIS programme cannot accept responsibility for the consequences of its use for other purposes or in other contexts. The views and recommendations expressed in this publication are those of the author and do not reflect the views of RUSI or any other institution.

Published on 29 June 2020 by the FFIS programme, Version 1.0
Author: Nick Maxwell

This work is licensed under a Creative Commons Attribution – Non-Commercial – No-Derivatives 4.0 International Licence. For more information, see <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Reference citation:** *Maxwell, N (2020) 'Innovation and discussion paper: Case studies of the use of privacy preserving analysis to tackle financial crime' Future of Financial Intelligence Sharing (FFIS) research programme. Version 1.0 (29 June 2020)*

# Acknowledgements

The FFIS programme would like to thank all those who contribute to the broader FFIS project into "The Role of Privacy Preserving Data Analytics in the Detection and Prevention of Financial Crime", particularly our project sponsors Verafin, Oliver Wyman, Refinitiv, Western Union and the SWIFT Institute. The FFIS team is very grateful for the support of the programme research advisory committee, who contribute in a personal capacity to guide the research process.

To support the mapping exercise of case studies and workshop process to directly contribute to this Discussion Paper, the author is very grateful for contributions from:

Sedicii
Privitar
Inpher
Enveil
Deloitte
Duality Technologies
Exate Technologies
IBM
HSBC
Scotiabank
Société Générale
The Australian Transaction Reports and Analysis Centre (AUSTRAC)
The UK Financial Conduct Authority
The Netherlands Financial Intelligence Unit
The Financial Transactions and Reports Analysis Centre of Canada (FINTRAC)

For more details about the FFIS programme, please visit www.future-fis.com.

# Contents


| | Page |
|---|---|
| Executive Summary | 6 |
| Structure of this Discussion Paper | 7 |
| Objectives and methodology | 8 |
| 1. The need for information sharing to detect financial crime | 9 |
| 2. Understanding privacy preserving analysis | 11 |
| 3. A timeline of the development of privacy preserving analysis | 13 |
| 4. The UK FCA TechSprint | 16 |
| 5. Our approach to mapping case studies | 17 |
| 6. Case studies map and overview table | 19 |
| 7. Case studies in detail | 22 |
| 7.1. Enveil ZeroReveal - Inter-bank secure and private data collaboration - Proof of Concept (2019) | 23 |
| 7.2. Duality SecurePlus - Query for collaborative investigations in Financial Crime and Compliance - Proof of concept (2019-present) | 25 |
| 7.3. Secretarium - 'Project DANIE' Customer reference data validation - Pilot project (2019 - present) | 27 |
| 7.4. Sedicii - Customer address verification - Pilot project (2017-2019) | 28 |
| 7.5. Privitar SecureLink - Privacy-preserving statistical analysis from public and private data owners - Pilot project (2018-2019) | 29 |
| 7.6. Inpher XOR Secret Computing - Cross-border privacy-preserving machine-learning model - Commercially deployed (2017) | 30 |
| 7.7. Deloitte UK / FutureFlow - UK Tri-bank initiative - Proof of concept (2018-2020) | 32 |
| 7.8. AUSTRAC's Fintel Alliance Alerting Project - Project in development for deployment (2019 to present) | 34 |
| 8. Conclusions | 36 |
| 9. Discussion and feedback | 36 |
| Annex A – Summary table of privacy preserving analytical PET techniques | 37 |
| Annex B – Relevant technology companies highlighted in case studies | 39 |
| Endnotes | 40 |

# Executive Summary:

Imagine that an organisation stores a vast amount of data in a safe. The information is secure, but to produce any summary reports of the data or to query that data, the data must be removed from the safe for analysis. At this point of analysis, all of the raw sensitive data is exposed. Now imagine that pre-defined queries and macro analysis could take place on the data without the safe ever being opened. Further still, data in multiple safe environments, across multiple organisations (in multiple countries), could be analysed collectively, by consent of the data owners, without any raw sensitive data leaving their safe environments. Queries that yield no sensitive information can be executed on sensitive data, held within the safe environments, without any sensitive information being disclosed at any point.

This is the promise of privacy preserving analysis, using privacy enhancing technologies (PETs).

In late 2019, the FFIS programme launched a new multi-year international study into "The Role of Privacy Preserving Data Analytics in the Detection and Prevention of Financial Crime". This research project follows three years of research, and activity in more than 20 countries, exploring more traditional information-sharing challenges within the context of anti-money laundering (AML) and financial intelligence objectives.

PETs are not new. The first theories relating to privacy preserving analysis, as we now understand it, were developed in the 1970s. The theories found life in mathematical and cryptographic models in the 2000s, but were typically too expensive in computational cost to be practical. Since 2010, as a result of advances in underlying techniques, PETs have started to be deployed in real-life scenarios – with national intelligence and healthcare as key sectors of early adoption. Since 2019, the development of PETs in relation to AML and financial crime prevention have been spurred by significant and supportive activity by both UK and Australian AML supervisors.

In this Discussion Paper, FFIS is publishing eight case studies of current innovation, pilots and projects of privacy preserving analysis related to AML and financial crime detection use-cases. This document will be updated throughout the lifetime of the wider research project with additional case studies as they come online.

As technical capabilities develop to support information-sharing in a form that may not have been possible previously – or, rather, that were not legally permissible without PETs – policy makers, supervisors, FIUs and private sector stakeholders have an opportunity to reflect on their desired outcomes for financial crime detection, disruption and prevention and the most efficient processes that can support those outcomes. In particular, privacy preserving analytical capabilities should encourage decision-makers to provide clarity about what specific analysis they wish to permit and enable across the financial system – and what contributing information they wish to remain undisclosed within that process, and by whom. It is intended that this paper supports dialogue and feedback related to the implications of, and appropriate framework for, the growth of privacy preserving analytics in the field of AML and financial intelligence.

Throughout 2020 and 2021, supported by feedback on this Discussion Paper, the FFIS programme will be engaging in such issues as:

- desired outcomes in policy frameworks for tackling financial crime and respective information-sharing requirements;
- the legal significance of the use of PETs for those information-sharing requirements;
- wider ethical, regulatory, policy, governance and cultural adoption considerations and potential risks or unintended consequences; and
- practical technical conditions relevant to applying PETs to those use-cases.

We look forward to feedback and engagement in the content of this paper.

## Structure of this paper:

This paper is structured as follows:

The objectives, methodology and contact information for feedback on this Discussion Paper are set out as introductory material.

Section 1. Describes the need, at a high-level, for information sharing to detect financial crimes and typical information-sharing challenges within AML frameworks.

Section 2. Explains privacy preserving analysis, including the contributing Privacy Enhancing Technologies (PETs), and the relevance of this technology to information sharing goals.

Section 3. Provides a timeline of the development of PETs, from the 1970s to 2019.

Section 4. As a key contextual event, this section summarises the UK Financial Conduct Authority 2019 TechSprint, which supported innovation in the use of privacy enhancing technologies to tackle financial crime.

Section 5. Bringing the story up to 2020, this section sets out our approach to mapping relevant case studies of pilots and projects for the use of this technology in AML and financial investigation fields, at this important time in the development of the technology.

Section 6. Presents the case studies in a summarised form, both as a map and an overview table.

Section 7. Presents the case studies in detail, as described by the relevant project owners.

Section 8. Highlights conclusions from this mapping exercise and calls for further reflection, particularly from policy-makers and supervisors, about what an appropriate framework for the development of the technology should be. This paper emphasises the importance of current moment for policy-makers and supervisors to achieve greater clarity about desired outcomes in the AML system and, accordingly, the role that privacy preserving analysis should play in the overall system. This will require consideration about what specific analytical capabilities should be accessible for different AML stakeholders and what information should remain protected and undisclosed in that process.

Section 9. Signposts the opportunity for feedback on this exercise.

Annex A provides a technical summary table of different privacy preserving analytical PET techniques and Annex B provides links to technology companies relevant to the case studies in this paper.

## The objectives of this Discussion Paper:

This discussion paper is the first of a series of papers, contributing to an international FFIS research project into "The Role of Privacy Preserving Data Analytics in the Detection and Prevention of Financial Crime". The broader research project is expected to run up to late 2021 and aims to:

- explore specific privacy enhancing technology use-cases relevant to financial information-sharing partnerships and their role in disrupting serious and organised crime and terrorist financing;
- examine how privacy enhancing technology and privacy preserving analytics may contribute to increased effectiveness of information-sharing within national and international AML regimes;
- identify the implications of such technologies in terms of additional privacy protections and intrusions; and
- provide greater clarity regarding policy, legal, cultural, data protection and data governance adoption considerations relevant to this field.

The purpose of this Discussion Paper is to:

1. provide an overview of concepts and technical developments relevant to privacy preserving analytics;
2. collate up-to-date reference material for case studies of innovations, pilots and projects currently underway that seek to utilise privacy preserving analysis relevant to AML or financial crime objectives;
3. facilitate expert-level dialogue and feedback on the broader implications of recent technical and practical developments in support of the broader FFIS research project.

## Methodology:

This Discussion Paper benefits from a review of the technical and academic literature on relevant Privacy Enhancing Technology, supplemented by key stakeholder interviews. To inform the strategic direction of the broader multi-year FFIS research project, a number of research workshops were convened, covering London, Ottawa, Washington D.C. and The Hague from November 2019 to January 2020. To directly support this Discussion Paper and the associated case study mapping exercise, survey invitations were distributed to a range of technology and banking stakeholders. These stakeholders were identified as active in the field either as a result of their participation in relevant TechSprints or by recommendations from our Research Advisory Committee, relevant public agencies and/or senior figures from a financial intelligence and financial crime compliance community. Case studies submissions have not been independently assessed, but material is presented as provided by authoring project managers and FFIS encourages feedback on the mapping exercise as a whole.

## Feedback on this Discussion Paper:

The case study references will be updated throughout the lifetime of the broader research project, including with additional case studies as they come online and are submitted to the FFIS programme. As a 'living document', feedback and additional contributions relevant to this document are invited from stakeholders in the field, particularly from communities in financial crime intelligence (public and private), relevant technologists, policy-makers, supervisors and data-protection legal experts. The geographic scope of this project is international.

Comments and feedback are welcome on this Discussion Paper and invited to be sent to:

PETproject@future-fis.com

# 1. The need for information sharing to detect financial crime

According to the Financial Action Task Force (FATF), 'effective information-sharing is [a] cornerstone of a well-functioning AML/CTF framework'.[1] Since 2017, the FFIS programme has explored the value of public–private financial information sharing partnerships as innovations in the field of tackling financial crime, including their impact in supporting financial intelligence, risk management and criminal justice outcomes around the world.[2]

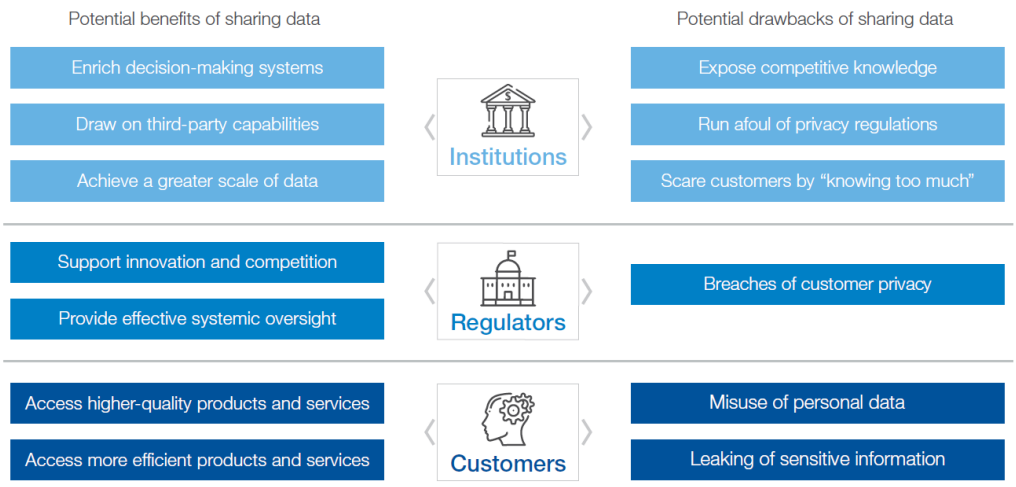
For financial institutions seeking to comply with anti-money laundering regulatory obligations and to understand financial crime risks in their business, information sharing can help to achieve greater insight of potential suspicions of criminality, including through:

- information exchange between public and private entities (including both operational information relevant to law enforcement investigations and strategic intelligence relating to the nature and dynamics of criminal groups or threats);
- co-development of financial crime risk intelligence through a collaborative and iterative process involving a number of institutions (either at the operational or strategic level);
- allow for querying and integration of third-party data and analytical capabilities to complement an institution's (or partnership's) understanding of financial crime risk.

However, there can be a number of limitations on the nature and extent of such information-sharing, including due to:[3]

- legal uncertainty or restrictions related to sharing information, including data privacy legislation and other duties to protect the confidentiality of information; [4]
- technical challenges and other costs in sharing information;
- lack of trust between owners of information and other stakeholders in a control and processing chain of information;
- competitive concerns about sharing information; and
- beyond direct privacy harms, broader public trust and reputational concerns about how private information is collected and used.

Resolving the general tensions between the value and risks of sharing information is a balancing act for individuals and all manner of organisations throughout modern society. More specifically focused on financial services and regulatory compliance, the World Economic Forum published a description[5] of the benefits and risks of sharing data – for institutions, regulators and customers respectively:



Broadly, current limitations and risks of information-sharing in the context of obligations and desired outcomes in the AML/CTF regime create a number of tensions which are, arguably, highly suboptimal.

| Desired outcome in the AML/CTF regime[6] | Typical information-sharing restrictions[7] |
|---|---|
| Regulated entities identify suspicions of crime within their business. | Outside of public–private financial information-sharing partnerships or other bi-lateral exchanges, regulated entities must identify indications of criminality without any information from public agencies that may indicate which entities are under investigation for criminal activity. |
| Criminal networks that seek to conceal money laundering schemes through the use of multiple accounts, spanning multiple financial institutions, are uncovered. | However, outside of a small number of jurisdictions[8], regulated entities are not permitted to share information with counterpart financial institutions relating to their suspicions of financial crime risk. |
| Criminal networks that operate money laundering schemes across borders are identified. | However, financial institutions can often be prohibited from sharing information about their suspicions of financial crime outside of national borders, even within their own financial groups.[9] |
| Regulated entities are prohibited from 'tipping off' parties that they have filed a suspicious activity report on a customer. | Outside of filing a suspicious activity report to a government Financial Intelligence Unit, this prohibition can prevent financial institutions from sharing information about clients who have been the subject of suspicious reports or closed accounts to others who may be at risk of being exposed to the same customer.[10] |
| Regulated entities conduct 'Know Your Customer' checks and identify any risks of mis-represented identifying information about their customers. | However, financial institutions may not have access to authoritative reference information for a customer, independent of documentation provided by the customer. |

As a result, in general, criminals operating professional money laundering schemes[11] can:

- be highly networked;
- adapt rapidly to avoid detection;
- operate internationally with ease;
- conceal their activity across multiple financial institutions; and
- conceal beneficial ownership through layers of legal entities spanning multiple jurisdictions.

In contrast, beyond third-party data that can be procured, private sector regulated entities are required to identify indications of suspicious criminal activity by:

- observing their own business data and customer interactions, in isolation, to spot patterns indicative of any form of criminality.[12]

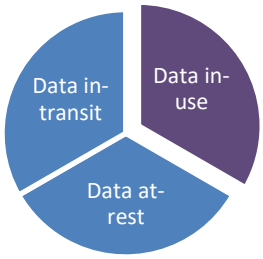
Against this context, 'Privacy Enhancing Technologies' (PETs) and privacy preserving analysis hold some promise to enable institutions to share data in a way that helps to achieve a balance between competing opportunities and obligations, allowing for data-sharing that is compliant with regulatory principles, appropriately protects the privacy of individuals and safeguards the confidentiality of business processes.[13]

# 2. Understanding privacy preserving analysis

Privacy preserving analysis relies on 'privacy enhancing technologies', or PETs, which are specialist cryptographical capabilities. Such privacy-preserving analysis allows for computations to take place on underlying data, without the data owner necessarily divulging that underlying data. The same technology can ensure that the data owner does not have visibility over the search query, with the query and the results remaining encrypted (or not disclosed) and only visible to the requester.

While encryption for data while at-rest (in storage) or data while in-transit (during communication) may be familiar concepts for many, PETs offer capabilities for encryption for data in-use (the ability to process computations without having access to the raw data).

***PETs – The third segment of managing the security of data: allowing data to be used, without exposing raw data to decryption.***




These techniques enable data owners or data stewards to provide analysts (or processors) with an opportunity to undertake limited computations and to provide guarantees that the analyst will not have access to raw data. This provides a number of benefits in that the data owner can cooperate with the processor in situations where:

- the data owner does not have authority to release the underlying data to the processor;
- the data owner does not trust the data processor with the full underlying data; and
- the data owner can have confidence that if the data processor were subject to a data breach, then their underlying data would not be compromised.

In the AML and financial crime setting, PET capabilities therefore have the potential to support information-sharing to enhance (for example):

- public to private and private to public sharing;
- private to private domestic sharing; and
- cross-border sharing (public–public, public–private and/or private–private)

There are various forms of PETs, including the following:

- (Partial, Somewhat or Full) Homomorphic Encryption (HE)
- Secure Multi Party Computation (SMPC)
- Trusted Execution Environments (TEE)
- Zero Knowledge Proofs (ZKP)
- Federated Learning

For more details on these individual PET techniques, please see ANNEX A to this document 'Summary table of privacy preserving analytical PET techniques'. It is important to be aware that there is no consensus definition of the term 'privacy enhancing technology' and different authors may refer to a wide range of technologies as PETs, including personal consumer services such as Virtual Private Networks (VPNs). For the purposes of this study, we are concerned only with PETs that enable computational results to be revealed, without disclosure of relevant input data to requesting parties – i.e. **privacy preserving analysis**.

Even within the small subset of privacy preserving analytical PETs, which are the focus of this paper, there are considerable differences in the nature of the underlying technology. The technologies can be complementary, used in conjunction with one another and, to some extent, are incomparable. However, they support analysis or 'learning' without needing to divulge the data.

A general concept in understanding the use of PETs in practice is the concept of a utility / privacy trade-off. In essence, access to information 'in the clear' (unencrypted) would provide a maximum level of utility, but PETs may limit the potential range of query options available to the requester. The use of PETs has increased the level of privacy protection (for the requested party) and reduced the range of query capabilities (for the requester).[14]

However, for specific use-cases, i.e. those where all data queries are not required and access to all underlying data is not required, PETs may provide a similar or, even, identical (practical) utility compared to analysis 'in the clear'. There may be other costs, such as computational time or data engineering issues, but in these cases, privacy will be protected with no loss of practical utility (in terms of the result).

Therefore, at the appropriate level of technical maturity, PETs could enable secure and controlled data usage, and a range of new data-enabled products or services, in circumstances that would otherwise raise legal, reputational, political, business or competition concerns due to the potential for privacy risks.[15]

# 3. A timeline of the development of privacy preserving analysis

**Since the 1970s**

PETs start to be theorised as mathematical and cryptographical concepts.

**2000 to 2010**

A number of computational and cryptographical breakthroughs take place.[16] In 2004, the first prototypes of Secure Multi Party Computation are produced.[17] In 2008, the first commercial TEE solution based on ARM TrustZone technology is launched.[18] Also in 2008, the first live implementation of SMPC is demonstrated to determine sugar beet market prices in Denmark without revealing individual farmers' economic position.[19] In 2009, the first fully homomorphic encryption (FHE) system was proposed by Craig Gentry, although at this stage requiring enormous computational resources to run.[20]

**2010 to 2020**

The global R&D effort is focused on improving the efficiency of PET techniques.[21]

| | |
|---|---|
| 2010 | The US Defense Advanced Research Projects Agency (DARPA) initiates the *'PROgramming Computation on EncryptEd Data'* programme to develop advances in theory, prototype implementation and application of PETs, specifically HE and MPC. The project resulted in substantial improvements in computation efficiency for these schemes for general applications; the first practical real-world applications of HE[22], design of HE hardware acceleration[23] and a legacy of open-source software for applied HE[24,25]. RAND reports describe the results and early implications of the success on this program.[26] |
| 2011 | Sharemind[27], an SMPC solution developed by Cybernetica, is applied for the first time to analyse Key Performance Indicators (KPIs) for the Estonian Association of Information Technology and Telecommunications, where 17 participating companies provided financial performance metrics to be processed, without disclosing the underlying data which was commercially sensitive. The data was processed within 2 minutes.[28] |
| 2011 | The U.S. Intelligence Advanced Research Projects Activity (IARPA) initiated the Security and Privacy Assurance Research (SPAR) programme to develop (i) prototype implementations of efficient cryptographic protocols for querying a database that keep the query confidential, yet still allow the database owner to determine if the query is authorized and, if so, return only those records that match it; (ii) prototype implementations of efficient cryptographic protocols for subscribing to topics in a stream of documents such that the subscription is kept confidential and only the matching documents are delivered; and (iii) efficient homomorphic encryption techniques to implement queries on encrypted data. A report released in 2015 describes the conclusions of this project.[29] |
| 2015 | From 2015, the EU funded Homomorphic Encryption Applications and Technology (HEAT) project starts to publish relating to advanced cryptographic technologies to process sensitive information in encrypted form and to support a step change in the efficiency and applicability of this technology.[30] |
| 2016 | Australia's national science agency, CSIRO, founded Data61 to support national strategic data-driven projects and Data61 later establishes the 'Confidential Computing' platform to combine distributed machine learning with homomorphic encryption and secure multiparty computing to provide the ability to learn models across multiple datasets without any of the data leaving its secure source. Data61 reports that the models provide for encrypted calculations are identical to the results processed in the clear – i.e. there is no loss of accuracy due to the encryption process.[31] |

| | |
|---|---|
| 2016 | Zcash launches as a blockchain platform for transactions utilising ZKP to shield the transaction path while still being verified under the network's consensus rules[32]. The zk-SNARKS version of ZKP developed by Zcash is later adopted by J.P. Morgan.[33] |
| 2017 | An industry, government and academic consortium produces three white papers[34], on security, Application Programming Interfaces (APIs) and applications of homomorphic encryption; and a draft standard for parameter selection. These developments are reported to enable (somewhat) Homomorphic Encryption to be commercially viable. [35] |
| 2018 | The EU HEAT project publishes its use-case analysis of the potential for automated detection of organized crime (ADOC) in response to Europol's 2013 Organised Crime Threat Assessment (OCTA). The project aims to support the systematic environmental scanning for weak signals, searching, fusing and interpreting data from encrypted databases while enabling data aggregation in the Cloud for authorized users.[36] |
| 2018 | NHS Digital, serving the NHS England as the largest employer in the UK and with responsibility for health data for over 55 million individuals[37], engages Privitar SecureLink to enable safe pooling of data from multiple contributors using partially homomorphic encryption. The system provides for protection of sensitive data attributes within datasets to enable sharing with third parties; ensuring consistency to conduct meaningful pattern analysis on safe data and providing enhanced data traceability to improve governance and to deter misuse. Underlying patient identifying information is not disclosed. Contributing NHS trusts are not disclosed. This example for sharing and collaboration technology uses NHS number, in an encrypted form, to link datasets across NHS Trusts.[38] |
| 2019 | A collaboration between Duality Technologies and Dana Farber Cancer Institute (DFCI) supports secure large-scale genome-wide association studies (GWAS) in the U.S. through homomorphic encryption. The project allows for common genetic variants to be revealed within a large sample, without disclosing submitted genetic data in full or other personal data. Analysts have access to the GWAS results, which identify genetic variants that are relatively more frequent in association to a specific trait or disease, but personal data and the submitted DNA sample by study participants remain encrypted and not visible to the research analysts. |
| March 2019 | TensorFlow (a widely used open-source library for machine learning) published TensorFlow Federated,[39] an open-source framework that allows machine learning to be performed on federated datasets. |
| March 2019 | The UN publishes a handbook focused on privacy-preserving computation for national statistical offices.[40] |
| March 2019: | The Royal Society publishes "Protecting privacy in practice: the current use, development and limits of Privacy Enhancing Technologies in data analysis".[41] The study indicated that PETs are an emerging and potentially disruptive set of technologies, "which, combined with changes in wider policy and business frameworks, could enable significantly greater sharing and use of data in a privacy-preserving, trustworthy manner." The report, among other things, called for more action from government and industry to:<br>• *Accelerate the R&D of PETs and promote the innovation ecosystem.*<br>• *Support organisations to become intelligent users of PETs, to drive the adoption of PETs.*<br>• *Raise awareness and provide advice about how suitably mature PETs can mitigate privacy risks and address regulations such as the General Data Protection Regulation (GDPR). (The Information Commissioner's Office and the National Cyber Security Centre are recommended to support guidance and the development of certification standards for quality assurance.)*<br>• *Maximise on the potential of government as an important early adopter, using PETs and being open about their use so that others can learn from government experience. (Government departments are recommended to consider what existing processing might be performed more safely with PETs and how PETs could unlock new opportunities for data analysis, whilst fully addressing privacy and confidentiality concerns.)* |

| | |
|---|---|
| April 2019 | In the 2019/20 Australian federal budget, AUSTRAC is granted AUD$28.4 million over 4 years to expand the Fintel Alliance, including to develop operational capabilities using privacy preserving techniques.[42] |
| July 2019 | UK Financial Conduct Authority holds a Global Anti-Money Laundering and Financial Crime TechSprint on PETs, focused on how PETs can facilitate the sharing of information about money laundering and financial crime concerns. (more details below).[43] |
| September 2019 | The World Economic Forum publishes the White Paper "The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value".[44] The paper provides a high-level overview of how privacy enhancing techniques work and potential application across the breadth of financial services interests in data.<br><br>The paper concludes:<br><br>*"Due to their nascence, there is uncertainty in some cases on how PETs would be treated under privacy regulations around the world. For example, federated analysis or secure multiparty computation in theory should allow institutions to analyse their data across regions where sharing data across international borders would otherwise be prohibited. However, ensuring that this is explicitly permitted by regulation would be important to preventing any fines or other regulatory risks from materializing, and in many cases the required regulatory certainty does not yet exist. Soliciting this certainty will necessitate an increased understanding of PETs as well as open discussions between the public and private sector on what is a safe approach to using these techniques in the financial sector."*[45] |
| October 2019 | The Confidential Computing Consortium is launched dedicated to accelerating the adoption of confidential computing (TEE), including founding members: Alibaba, Arm, Google Cloud, Huawei, Intel, Microsoft and Red Hat. General members include Baidu, ByteDance, decentriq, Fortanix, Kindite, Oasis Labs, Swisscom, Tencent and VMware.[46] |
| February 2020 | The European Commission publishes a draft Data Strategy which describes privacy preserving technologies as "crucial for the next stages of the data economy" and recommends that the 'Horizon Europe programme' will continue to support these technologies as part of the European strategy for data.[47] |
| June 2020 | The European Data Protection Supervisor publishes a formal opinion that commends the recognition of the Commission of the importance of privacy preserving technologies and:<br><br>*"recalls the potential of privacy enhancing technologies (PETs) as enablers of data sharing which is both privacy-friendly and socially beneficial… In addition, in order to optimise the benefits of the various privacy preserving technologies, the EDPS emphasizes the importance of their standardization and interoperability."*[48] |

# 4. In detail: The UK FCA July 2019 Global Anti-Money Laundering and Financial Crime TechSprint on PETs

In July 2019, the UK Financial Conduct Authority (FCA) hosted a week-long Global Anti-Money Laundering and Financial Crime TechSprint focused on how PETs can facilitate the sharing of information about money laundering and financial crime concerns. Over 140 active participants took part in the TechSprint at the FCA's offices and a satellite event took place in Washington D.C.

The TechSprint focused on developing solutions, using PETs, related to the use case challenges below:

1. How can a network of market participants use PETs and data analytics to interrogate financial transactions stored in databases within institutions to identify credible suspicions without compromising data privacy legislation?
2. How can market participants rapidly and accurately codify typologies of crime, in a way that allows them to be quickly disseminated and implemented by other market participants in their financial crime controls?
3. How can a market participant check that the company or individual they are performing due diligence on hasn't raised flags or concerns within another market participant, and/or verify that the data elements they have for the company or individual match those held by another market participant?
4. How can technology be used to assist in identifying an ultimate beneficiary owner (UBO) across a network of market participants and a national register?

Over 200 senior attendees from across public and private sectors attended 'Demo Day', including representation from 42 international regulators. The team demos are available here: https://www.fca.org.uk/events/techsprints/2019-global-aml-and-financial-crime-techsprint

On 20 September 2019, the FCA hosted a follow up workshop on the role of Privacy Enhancing Technologies (PETs) in preventing money laundering and financial crime. Attended by FFIS, the workshop covered issues of:

- Barriers to adoption, including trust; education; motivation; and infrastructure/processes
- Centralised vs decentralised approach to PETs-enabled data sharing
- Discussing the opportunities and challenges of adoption of PETs for AML at scale; including
    - Social purpose
    - Public-private collaboration
    - Building of trust
    - Federated learning usage
    - Growth of synthetic data
- The role of the regulator
- Supporting further research on use cases
- Encouraging active firm engagement in developing the eco-system

The March 2020 FCA 'Fostering innovation through collaboration: The evolution of the FCA TechSprint Approach' paper[49] compiled lessons from across 7 TechSprints covering a range of subjects.

Moving forward, in 2020, the FCA Innovation team is piloting a 'digital sandbox' to support access to high-quality data assets including synthetic or anonymised data sets to enable testing and validation of technology solutions, including PET solutions to financial crime prevention use-cases.[50]

# 5. Our approach to mapping case studies of the use of privacy preserving analysis to tackle financial crime

The overall context in 2020 for using PETs to tackle financial crime can be characterised as follows:

1. While the maturity of individual PET techniques varies, in recent years multiple scientific research and development institutes refer to the growing maturity of privacy preserving analytics and opportunity for commercial and government testing and deployment.
2. The UK Financial Conduct Authority continues to encourage firms to explore PETs for use in financial crime prevention use-cases.
3. A number of PET proofs of concept, pilot and deployment projects are underway in financial services and in financial crime related use-cases. However, awareness is limited - across a broader public and private sector community of relevant stakeholders involved AML and financial crime prevention - about the nature of the technology, its current use and opportunities for further development.

Against this context, in this paper, the FFIS programme is compiling available case studies of the use of privacy preserving analytics for AML or financial intelligence objectives. This case study compilation is intended to further accelerate awareness in the international practitioner and policy community about the use of the technology and to spur considerations around the appropriate policy and supervisory framework.

In our search for case studies, we have prioritised the following privacy preserving use-cases:

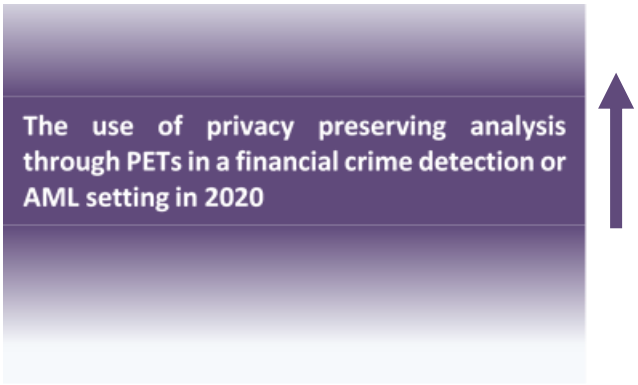
- **Record linking.** Verification/matching of data attributes held by requesting parties against external reference data, without disclosure of the query or disclosure of the reference data, or linking data sets through common data attributes.
- **Network mapping.** Network mapping of connected nodes (e.g. through transactions) across multiple data owners, without data owners' disclosure of underlying data.
- **Prevalence queries.** Macro-insights about the prevalence of certain rule-based queries across a community of data owners, without disclosure of underlying data or an individual data owners prevalence exposure.
- **Regression analysis.** Analysis about the relationship between common data attributes across a community of data owners, without disclosure of underlying data.
- **Federated machine learning.** Analysing patterns across multiple data holdings without any disclosure of underlying data.

A key concept in understanding the growth of technology is Technology Readiness Levels (TRLs). TRLs are a method for understanding the maturity of a technology throughout its research, development and deployment phase progression. Borrowing from NASA's original TRL scale, the EU defines nine TRL levels. The graphic below indicates how the case studies in this paper broadly translate to the EU TRL scheme.

It is intended that this mapping of existing use of privacy preserving analytics yields further clarity on the current state of development (readiness or maturity) of these techniques. While the case studies highlighted in this Discussion Paper have not been subject to independent verification of readiness, we have invited case study authors to describe the current state of deployment (or readiness) that their case study represents and we invite additional feedback and different perspectives on adoption and readiness issues. As stated in the introduction, FFIS intends to update this Discussion Paper case study record with further projects as they come online and are submitted to the FFIS programme.

**The EU TRL scale defining readiness or maturity of technological developments**


**Case studies collated for this paper indicate the following range of potential TRL.**




51

The nature of theoretical use-cases for privacy preserving analysis are limited only by imagination, but practical use-cases in AML financial crime settings will be affected by a large number of factors, including: computational cost; operational cost; technical complexity; efficiency; data availability, quality and inter-operability; legal clarity on the significance of use; clarity on the value proposition of use (compared to sharing 'in the clear'); regulatory support; appropriate data governance; awareness; cultural acceptance; leadership engagement; and broader public and political consent. Technically there also remains an active R&D effort to demonstrate that results are verifiable by requesters, independent of trust in the PET technique, without revealing the underlying data.[52] In collating the following case studies, the paper is intended to advance dialogue and debate with a wide community on these wider adoption considerations.

This version of the Discussion Paper collates:

**8** case studies

Of which:

**5** are executed on real customer data (operational environment)

**3** are executed on synthetic data (a relevant environment)

**3** are proofs of concept

**3** are pilot projects of operational deployment

**1** is commercially deployed in the private sector

**1** is a funded public sector protype in development

# 6. Case studies map

Sedicii - Customer Address Remediation

Verifying bank customer identification and address information against data held by a public utility, in a manner which did not disclose any query of customer data to the public utility.
[Republic of Ireland]

Enveil ZeroReveal - Inter-Bank Secure and Private Data Collaboration

Allowing a financial institution to identify matching customer information in external sets of data, without disclosing the query. [EU]

Inpher XOR Secret Computing - Cross-border privacy-preserving machine-learning model

Enhancing the breadth of training data available to a machine learning algorithm in a cross-border environment where underlying data cannot be shared. [Cross border – The Netherlands, Belgium, Luxembourg - EEA - GDPR]

Secretarium - Project DANIE

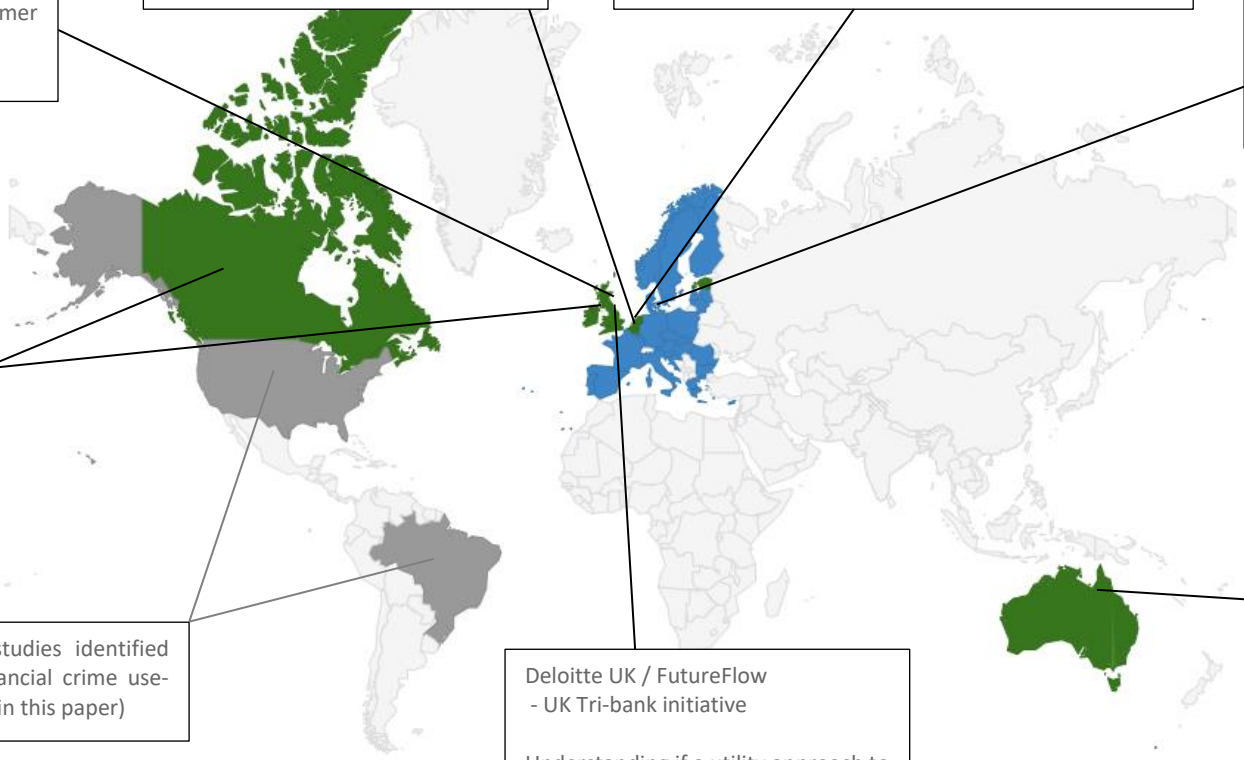
Enabling participant banks to identify ID verification discrepancies between their respective reference data for common clients.
[EU Member State]

Duality SecurePlusTM Query for collaborative investigations in Financial Crime and Compliance

Allowing participants to make queries of data without the query being disclosed to the requested party (data owner). [UK & Canada]

Privitar SecureLink - Privacy-preserving statistical analysis from public and private data owners

Learning statistics about a population from across disparate public and private datasets, without collecting identifiable information.
[Undisclosed jurisdiction]



PET project case studies identified outside of AML/financial crime use-cases (not included in this paper)

Deloitte UK / FutureFlow
- UK Tri-bank initiative

Understanding if a utility approach to Transaction Monitoring (TM) could overcome the asymmetrical knowledge limitations of the currently fragmented approach, whilst respecting appropriate data privacy rules and legislation. [UK]

AUSTRAC's Fintel Alliance Alerting Project (Australia)

Providing the ability to identify whether there are any financial links between two (or more) suspicious accounts across multiple financial institutions and the ability to trace suspicious funds as they move between accounts across financial institutions. [Australia]

# Case study overview table

| Title | **Enveil ZeroReveal Inter-Bank Secure and Private Data Collaboration** | **Duality SecurePlusTM Query for collaborative investigations in Financial Crime and Compliance** | **Project DANIE** | **Sedicii - ESBN utility - Customer Address Remediation** |
|---|---|---|---|---|
| Technology provider | Enveil | Duality Technologies | Secretarium | Sedicii |
| Sector | Banking | Banking | Banking and financial institutions | Banking |
| Category | Anti-money laundering due diligence | Financial crime and compliance investigations | Anti-money laundering due diligence / ID verification | Anti-money laundering due diligence / ID verification |
| Objective | To demonstrate the capability for a financial institution to identify matching customer information in external sets of data, without disclosing the query. | To demonstrate the capability for a financial institution to identify matching customer information in external sets of data, without disclosing the query. | To enable participants to identify discrepancies between their respective reference data for common clients. | To verify customer identification and address information against data held by a public utility, in a manner which did not disclose any query of customer data to the public utility. |
| Information revealed | This engagement demonstrated the capability for the requested party to learn about matching customer information in external sets of data. | This engagement demonstrated the capability for the requested party to learn about matching customer information in external sets of data. | Participants receive individual reports identifying whether there are discrepancies between their client reference data and that held by other institutions in the project. | A query result of a partial discrepancy between ESBN information and queried information from financial institutions is revealed to the requesting party. Additionally, a change in customer attribute data held by ESBN is revealed to parties in cases where a matching customer has previously been identified, on a proactive basis. |
| Privacy preserving qualities | The query by the requesting party is not revealed to the requested party (the data owner). | The query by the requesting party is not revealed to the requested party (the data owner). | No underlying client reference data is shared. | The financial institutions' query, including the customer name and address information, is not disclosed to ESBN. The result of a partial discrepancy in information holdings is not disclosed to ESBN. The attribute data held by ESBN that is in disagreement with the requesting party is not revealed to either party. |
| Participants | A large global, EU-based financial institution (undisclosed) and a third-party data provider | Two pilots, fraud and AML respectively: The fraud pilot involved 5 parties, including 4 UK banks; the AML pilot involved a large Canadian bank. | A core consortium of 3 investment banks and 3 data providers, with several financial institutions involved in evaluating proof of concept. | Irish national electricity company ESB Networks (ESBN) and two banks (undisclosed). |
| Type of project | Proof of concept (2019) | Proof of concept (2019-present) | Pilot project (2019 - present) | Pilot project (2017-2019) |
| Real data | Synthetic data | Synthetic data | Real customer data | Real customer data |
| Jurisdiction | An EU member state | UK & Canada | EU Member State (GDPR) | Republic of Ireland |

| Title | Privacy-preserving statistical analysis from public and private data owners - Privitar SecureLink | Cross-border privacy-preserving machine-learning model.<br>Inpher XOR Secret Computing | UK Tri-bank initiative | AUSTRAC's Fintel Alliance Alerting Project (Australia) |
|---|---|---|---|---|
| **Technology provider** | Privitar | Inpher | Deloitte UK / FutureFlow | AUSTRAC (The Australian Transaction Reports and Analysis Centre) |
| **Sector** | Population data gathered primarily from financial services | Banking | Banking | Banking |
| **Category** | Population statistics to inform public policy | Business intelligence and marketing for a financial product. | Financial intelligence – transactions mapping | Financial intelligence within a public–private partnership framework |
| **Objective** | To learn statistics about a population from across disparate public and private datasets, without collecting identifiable information. | To enhance the breadth of training data available to a machine learning algorithm in a cross-border environment where underlying data cannot be shared. | To understand if a utility approach to Transaction Monitoring (TM) could overcome the asymmetrical knowledge limitations of the currently fragmented approach, whilst respecting appropriate data privacy rules and legislation. | Providing the ability to identify whether there are any financial links between two (or more) suspicious accounts across multiple financial institutions and the ability to trace suspicious funds as they move between accounts across financial institutions. |
| **Information revealed** | The recipient was able to analyse linked data supplied from multiple contributors, with all raw data only presented in a tokenised form. | Additional data points revealed by the Netherlands and Luxembourg subsidiaries were securely shared with the Belgian subsidiary to optimise its predictive machine-learning model without revealing any private inputs. | In effect, transaction flows (dates, amount, and tokenised sender and receiver accounts) were shared anonymised, by participant banks to analysts at the data intermediary. | The algorithm is designed to flag suspicious networks from domestic retail account and transaction data. |
| **Privacy preserving qualities** | The analyst recipient of the data was not able to obtain original, raw contributor data, but was able to link the datasets. | This privacy preserving machine-learning model required no disclosure or sharing by the Netherlands and Luxembourg subsidiary of the relevant underlying data to the Belgian subsidiary. The analyst conducting the computation learned the output of the function without exposing the private inputs. All private inputs remained encrypted throughout the computing protocols without being exposed to the data recipients. | The underlying Personally Identifiable Information (PII) e.g. names and addresses were not disclosed by participant banks. | No additional customer, account and transaction information will be exposed through the results of the algorithm. Where the transactions meet the criminal typology, AUSTRAC will initiate a follow up process that will be undertaken through formal notice to the relevant financial institutions to identify the specific transactions, accounts and customers of interest. |
| **Participants** | A number of public and private institutions. | Subsidiaries in the Netherlands, Belgium, Luxembourg from a single bank | 3 UK banks, intermediated by Deloitte UK | This project is being led by AUSTRAC and is being delivered by and for the Fintel Alliance. |
| **Type of project** | Pilot project (2018-2019) | Commercially deployed (2017) | Proof of concept (2018-2020) | Project in development (2019 to present) |
| **Real data** | Not real customer data | Real customer data | Real customer data | Planned for deployment on real data |
| **Jurisdiction** | An EU member state | Cross border – The Netherlands, Belgium, Luxembourg | UK | Australia |

# 7. Case studies in detail

## 7.1. Inter-Bank Secure and Private Data Collaboration Enveil ZeroReveal - Proof of Concept (2019)

| | |
|---|---|
| **Category of use-case** | Anti-money laundering due diligence |
| **Sector** | Banking |
| **Objective** | To demonstrate, as a proof of concept, the capability for a financial institution to identify matching customer information in external sets of data, without disclosing the query. |
| **Intended outcome** | In this case, to allow for AML queries over external sets of data, without disclosing the query to the requested party (and, therefore, without revealing to the third party any indication of suspicion related to the subject of the query). |
| **Participants** | Three participants: Enveil; A large global, EU-based financial institution (undisclosed); a third-party data provider |
| **Relevant legal Jurisdiction** | An EU member state |

**Information sharing and privacy preservation within this use-case**

| **Information shared or insight achieved** | **Privacy preservation qualities** |
|---|---|
| Enveil designed and developed an approach for performing fuzzy matching of customer profiles over ZeroReveal. This proof of concept demonstrates the capability for the requested party to learn about matching customer information in external sets of data. In terms of auditing the process, the proof of concept engagement also demonstrated that Enveil can process elements of information that can still be made visible for audit, traceability and trust building purposes between a requesting and requested parties. | In this proof of concept, the query by the requesting party is not revealed to the requested party (the data owner).<br><br>1. Enveil's ZeroReveal Client application encrypts the search containing sensitive indicators in the bank's trusted environment.<br>2. The ZeroReveal Server application integrates within the third-party data environment to execute the encrypted operation without ever decrypting anything in the untrusted environment, produces encrypted results, and sends those results back to the source.<br>3. Within the bank's trusted environment, the ZeroReveal Client application decrypts the results and returns them via API to the point of origin. |
| **Nature of data involved in the project** | |
| This proof of concept relied on third-party provided synthetic data, containing a combination of the data variations commonly inspected by financial institutions. A series of different queries were submitted in an iterative progressive way, in order to establish by following a fuzzy matching logic, whether there exists a customer profile match, and what is the level of confidence.<br><br>Encrypted queries were run across datasets comprised of 100k and 1 million customer records. To further demonstrate scalability, a comparable third-party data source containing 5 million, 25 million, and 100 million records respectively was queried. This effort showed how ZeroReveal can scale horizontally to handle larger datasets because of its ability to distribute the load across a cloud architecture for improved linear scalability. | |
| **Privacy Enhancing Technologies (PETs) used** | |
| This project leveraged homomorphic encryption (HE) to allow operations to be performed on ciphertext as if it were plaintext. HE provides the security of encryption while keeping data usable, allowing functions to be performed on the data in its encrypted state. | |

**Outcomes, data quality issues and lessons identified**

- The engagement proved fuzzy matching capabilities and the ability to search external data assets without revealing the contents of the search itself or compromising the security or ownership of the underlying data. This allows financial institutions to improve communications across external parties and gain secure access to new datasets in order to obtain better financial crime insights for new or ongoing investigations.
- The engagement also demonstrated a fully traceable and transparent audit/regulatory control process within Enveil ZeroReveal.
- Fuzzy matching capabilities provide opportunities to resolve or mitigate differences in data standards across participating entities, and therefore support data interoperability between institutions where data quality and consistency are major challenges. Privacy-preserving fuzzy matching queries can also be applied over other types of data (for example, transactions or SARs) using the same approach.
- Data engineering issues can also be minimised by applying Enveil ZeroReveal over enterprise's existing data structure and leverages the authentication, access control, and audit mechanisms that are in place.
- The same capability can also reportedly support matching queries or queries relating to indicators and AML typologies across entities in order to offer additional insight on financial crime.
- The engagement indicated that encrypted queries can be executed at considerable scale and in a timeframe of (single digit) seconds.
- The long-term vision of the operational use case is to support both private-private secure data collaboration between the banking sector as well as private-public secure data sharing between a government entity and the banking sector.

**Technology provider:**

Enveil Inc.

## 7.2. Duality SecurePlusTM Query for collaborative investigations in Financial Crime and Compliance Duality Technologies Inc. - Proof of concept (2019-present)

| | |
|---|---|
| **Category of use-case** | Financial crime and compliance investigations |
| **Sector** | Banking/Financial Services |
| **Objective** | To allow participants to make queries of data without the query being disclosed to the requested party (data owner) in order to accelerate triage and investigations. |
| **Intended outcome** | To demonstrate the privacy preserving capabilities of SecurePlusTM Query for collaborative investigations in both fraud and AML use-cases. |
| **Participants** | Fraud pilot: 5 parties; including 4 UK banks<br>AML pilot: 1 party; a large Canadian bank |
| **Relevant legal Jurisdiction** | UK & Canada |

**Information sharing and privacy preservation within this use-case**

| **Information shared or insight achieved** | **Privacy preservation qualities** |
|---|---|
| This project enabled the requesting parties to query data owned by other parties and receive the results of those queries, without disclosing sensitive query parameters to the data owner. Duality SecurePlus Query can execute complex SQL-like queries, which include multiple features and predicates, and can be performed against any fields in the data set.<br><br>When receiving a query, a data owner has visibility over the query structure, but not the encrypted parameters. This ensures that the data owners know and approve the type of information they might be sharing and can control and approve specific queries execution. Data owners can also see an audit trail, which includes who the inquiring party was, when and how many queries were executed, and the query structure, but not the protected parameters of the queries. | The data owner executes the encrypted query in his/her environment, without being exposed to the query fields nor to the query results.<br><br>Privacy is protected throughout the process – including the privacy of the entity that is the subject of the query, as well as the privacy of the investigation itself (i.e., in order to prevent tip-offs and to protect business confidentiality). In cases with multiple querying parties and data owners, the identities of those sending or responding to a query can also be protected. |
| **Nature of data involved in the project** | |
| Both the fraud and AML pilot were carried out on synthetic data which was generated based on the schema that was agreed on by the collaborating parties. The synthetic data involved in the project was comprised of several tables including account information, telephone numbers, names, transactions, transaction amounts, card information and other fields. Each table contained 1 million records. The data attributes varied for each project, but included numbers (e.g., account numbers, phone numbers), strings (e.g., beneficial owner names), and dates (e.g., transaction dates). | |
| **Privacy Enhancing Technologies (PETs) used** | |
| Fully Homomorphic Encryption (FHE). The underlying homomorphic encryption open-source library, called PALISADE, is used by a large community comprising the public sector (e.g., US Department of Defense), private sector (e.g., Raytheon, CACI), and academia (e.g., MIT, UCSD). PALISADE is also compliant with the security standards for homomorphic encryption as defined by homomorphicencryption.org | |

**Outcomes, data quality issues and lessons identified**

- Data interoperability was facilitated by agreement on a jointly defined data schema and queries. For production deployments, it will be important to ensure that the data controlled by data owners is compliant with the predefined data schema.
- Throughout the course of both pilot initiatives, participants exchanged hundreds of secure, encrypted queries as part of their investigations. The query results through homomorphic encryption were always accurate compared with the same queries that took place on unencrypted data.
- Typical queries on this data set returned results in under 1 minute. Duality's software scales linearly with the size of the data.
- These initiatives succeeded in validating that Homomorphic Encryption, as implemented by PALISADE and used by Duality SecurePlus Query demonstrating readiness for real-world challenges of sensitive data collaborations in financial crime investigation and compliance.

**Technology provider:**

Duality Technologies

## 7.3. Project DANIE - Secretarium - Pilot project (2019 - present)

| | |
|---|---|
| **Category of use-case** | Anti-money laundering due diligence / ID verification |
| **Sector** | Banking and other financial institutions |
| **Objective** | The objective of Project DANIE is to enable participants to identify discrepancies between their respective reference data for common clients. |
| **Intended outcome** | To allow for more effective and efficient detection of potential data inaccuracies in client reference data. |
| **Participants** | A core consortium of 3 investment banks and 3 data providers, with several financial institutions involved in evaluating proof of concept. |
| **Relevant legal jurisdiction** | EU Member State (GDPR) |

**Information sharing and privacy preservation within this use-case**

<table>
<tr><th>Information shared or insight achieved</th><th>Privacy preservation qualities</th></tr>
<tr><td rowspan="6">All participants receive individual & consortium reports identifying whether there are discrepancies between their client reference data and data held by other institutions in the project.<br><br>Through the project, participants learn if other participants share a high probability match for a common client and benchmark indication of potential discrepancies in client reference data.</td><td>Underlying client data or client relationships are not revealed.</td></tr>
<tr><th>Nature of data involved in the project</th></tr>
<tr><td>• Project applied to real data; and<br>• Proof of concept involved 32 data attributes covering approximately 200,000 legal entity identifiers.</td></tr>
<tr><th>Privacy Enhancing Technologies (PETs) used</th></tr>
<tr><td>TEE – Trusted Execution Environment<br>SMPC – Secure Multi-party Computation<br>HE – FHE Homomorphic Encryption<br>ZKP – Zero Knowledge Proof</td></tr>
</table>

**Outcomes, data quality issues and lessons identified**

- Project required data preparation in order to ensure common formats are respected between participants across each data field used by the platform;
- DANIE participants report that experience to date indicates successful discovery of discrepancies in client reference data; and
- This solution is believed by the project managers to be scalable and additional use-cases involving market data and other AML information and in the future DANIE can also be a market place of reference data.

**Technology provider:**

- Secretarium
- Intel SGX (emulator) - TEE

## 7.4. Customer Address Verification - Sedicii - Pilot project (2017-2019)

| | |
|---|---|
| **Category of use-case** | Anti-money laundering due diligence / ID verification |
| **Sector** | Banking |
| **Objective** | In the context of a financial institution's customer onboarding flow, financial institutions required a solution to verify their customer identification and address information against data held by a public utility, in a manner which did not disclose any query of customer data to the public utility. |
| **Intended outcome** | In this use-case, customer address verification query requests are made by banks to the authoritative address records held by ESBN. The Address Remediation use case allows a financial institution to check the accuracy of its meter, name and address records directly against the authoritative source without either party disclosing the underlying data. |
| **Participants** | Irish national electricity company ESB Networks (ESBN) and two financial institutions (undisclosed). |
| **Relevant legal jurisdiction** | Republic of Ireland |

**Information sharing and privacy preservation within this use-case**

| **Information shared or insight achieved** | **Privacy preservation qualities** |
|---|---|
| A query result of a partial discrepancy between ESBN information and queried information from a financial institution is revealed to the requesting party. Additionally, a change in customer attribute data held by ESBN is revealed to parties in cases where a matching customer has previously been identified, on a proactive basis. | The financial institution's query, including the customer name and address information, is not disclosed to ESBN. The result of a partial discrepancy in information holdings is not disclosed to ESBN. The attribute data held by ESBN that is in disagreement with the requesting party is not revealed. |
| **Nature of data involved in the project** | |
| Real customer data holdings queried, covering 2.5M records. Comparison of the following customer data attributes: meter number, name, address, email and phone number. | |
| **Privacy Enhancing Technologies (PETs) used** | |
| Zero-Knowledge Proof (ZKP) to compare meter number, name, address, email and phone number without requiring any of the parties (customer and ESBN) to exchange or disclose them in any way or form. | |

**Outcomes, data quality issues and lessons identified**

- The project achieved an 84% success rate in the address verification process taking place between the financial institutions and the authoritative source (ESBN). We also identified 4 additional normalisation rules that would increase the success rate to 96%. The verification times were measured to be in the millisecond range.
- This use-case demonstrated how ZKP processes can identify data quality or discrepancy issues between participants, by reference to a single authoritative source of information for the relevant attributes.
- To support the use-case, a common data governance model was developed between ESBN and the two financial institutions. A combination of normalisation rules independently run at each party and some rules running inside of a ZKP computation supported interoperability.

**Technology provider:** Sedicii

## 7.5. Privacy-preserving statistical analysis from public and private data owners - Privitar SecureLink - Pilot project with full installation. (2018-2019)

| | |
|---|---|
| **Category of use-case** | Population statistics to inform public policy |
| **Sector** | Population data gathered primarily from financial services |
| **Objective** | To learn statistics about a population from across disparate public and private datasets, without collecting identifiable information. |
| **Intended outcome** | Blind-matching of records using homomorphic encryption. |
| **Participants** | A number of public and private institutions. |
| **Relevant legal jurisdiction** | Undisclosed jurisdiction. |

**Information sharing and privacy preservation within this use-case**

| **Information shared or insight achieved** | **Privacy preservation qualities** |
|---|---|
| The recipient was able to analyse linked data supplied from multiple contributors, with all raw data only presented in a tokenised form. | Privitar SecureLink used a combination of technical and structural controls to allow data contributors to submit encrypted data to the recipient such that:<br><br>1. Only encrypted data left contributor environments.<br>2. Data could not be linked prior to it reaching the recipient. The recipient required a third-party (an intermediary) to convert the encrypted contributor data to a linkable, tokenised dataset.<br>3. The recipient was not able to reverse the processing and obtain original, raw contributor data, but was able to link the datasets. |
| **Nature of data involved in the project** | |
| The data request process involved multiple similar organisations within a sector who were supplying data based on the same schema (as predefined by the authority) and other organisations not in that sector or who were different kinds of organisations, who supplied data based on different, predefined schemas. | |
| **Privacy Enhancing Technologies (PETs) used** | |
| Partially homomorphic encryption. Random and homomorphic properties of the encryption scheme was used and an additional processing step by an intermediary. The intermediary, like the contributors, was not able to see the underlying values or link the datasets. | |

**Outcomes, data quality issues and lessons identified**

- This project demonstrated that the public authority is able to collect and link data about its population for the creation of aggregate statistics to inform public policy. In this process, no directly identifying information is collected. If any party intercepts the data at any point in the process, even if that party were one of the data contributors, they would not be able to decrypt or link the datasets.
- The project required data quality checks to be performed at the data contributor and prior to encryption. The recipient was unable to view or modify the underlying raw values of the linking identifier. When the data arrived at the recipient there would be no way of correcting any errors in the linking identifier and so it was important to ensure checks, such as ensuring there were no spaces at the end of the value, had been performed before the data was submitted otherwise data would not be linkable.

**Technology provider:** Privitar

## 7.6. Cross-border privacy-preserving machine-learning model Inpher XOR Secret Computing - Commercially deployed (2017)

| | |
|---|---|
| **Category of use-case** | Business intelligence and marketing for a financial product. |
| **Sector** | Banking / Consumer Financial Products |
| **Objective** | The objective of this project was to enable a Belgian subsidiary of a European bank to incorporate data from non-Belgian subsidiaries of the same group to inform a machine learning sales-prediction model, without any cross-border disclosure of underlying and contributing data. |
| **Intended outcome** | To enhance the breadth of training data available to a machine learning algorithm in a cross-border environment where underlying data cannot be shared. |
| **Participants** | Subsidiaries in the Netherlands, Belgium, Luxembourg from a single bank. |
| **Relevant legal jurisdiction** | Cross border - Netherlands, Belgium, Luxembourg - EEA - GDPR |

**Information sharing and privacy preservation within this use-case**

| **Information shared or insight achieved** | **Privacy preservation qualities** |
|---|---|
| Additional data points and insights revealed by the Netherlands and Luxembourg subsidiaries were securely shared with the Belgian subsidiary to optimise its predictive machine-learning model without revealing any private inputs. | This privacy preserving machine-learning model required no disclosure or sharing of the relevant underlying data by the Netherlands and Luxembourg subsidiaries to the Belgian subsidiary.<br><br>The analyst conducting the computation only saw the output of the function without revealing the private inputs. All private inputs remained encrypted in-processing throughout the computing protocols without being exposed to the data recipients. |

**Nature of data involved in the project**

- Real customer data.
- This cross-border privacy-preserving project increased the training data for a machine learning model from 24,000 data points in the Belgian bank to 300,000 additional data points from the Netherlands and Luxembourg entities.
- Data inputs related to the predictive analysis of “how likely is a commercial customer to buy a lending product in the next 30 days?”
- Input data included predictive classifiers for creditworthiness such as zip code, cash flow, and annual spending to improve the training model of the regression analysis.

**Privacy Enhancing Technologies (PETs) used**

Privacy-preserving machine-learning model, utilising Secure Multi-Party Computation (MPC) and Fully Homomorphic Encryption (FHE)

**Outcomes, data quality issues and lessons identified**

- Inpher's XOR Secret Computing™ virtual machines (local to each banking subsidiary) securely computed the protocol to classify new data sources and trained the prediction with 300,000 additional data points, producing an optimized model with two additional features for the Belgian subsidiary.
- No significant data interoperability issues were observed as information sharing took place between subsidiaries with similar data structuring.
- MPC and FHE can run logistic and linear regression functions on encrypted data without revealing or transferring any personal information. These simpler models (compared to neural nets, for example) ensure privacy, functionality, and also accountability; with the ability to query back what types of variables were weighed to produce the predictive outcome.
- AI/ML models in the financial sector should be explainable and transparent to ensure fair lending and consumer data protection.
- Building on this project, Inpher reports that it is developing of a privacy-preserving machine learning model for financial institutions to train their fraud models on distributed data from additional participating financial institutions.

**Technology provider:**

Inpher Inc.

## 7.7. UK Tri-bank initiative – Deloitte UK / FutureFlow - Proof of concept (2018-2020)

| | |
|---|---|
| **Category of use-case** | Financial intelligence – transactions mapping |
| **Sector** | Banking |
| **Objective** | To understand if a utility approach to Transaction Monitoring (TM) could overcome the asymmetrical knowledge limitations of the currently fragmented approach, whilst respecting appropriate data privacy rules and legislation. |
| **Intended outcome** | The key project success criteria were to:<br><br>1. Encrypt and extract transaction data from participant institutions in a manner that protects personally identifiable information within the original data;<br>2. Build a network of the encrypted data in order to build a view of the payment behaviour;<br>3. Inject known Financial Crime (FC) typologies into the network and apply a data led approach to identifying unusual behaviour; and<br>4. Identify potential instances of criminality that would not have been identifiable without aggregating data. |
| **Participants** | 3 UK banks, intermediated by Deloitte UK |
| **Relevant legal Jurisdiction** | UK |

**Information sharing and privacy preservation within this use-case**

| **Information shared or insight achieved** | **Privacy preservation qualities** |
|---|---|
| Transaction flows (dates, amount, and tokenised sender and receiver accounts) were shared anonymised, visible to analysts at the data intermediary. | Underlying personally Identifiable Information (PII) e.g. names and addresses were not disclosed by participant banks. |
| **Nature of data involved in the project** | |
| This project involved analysis of real data, comprising 1 year of Small and Medium Sized Enterprise (SME) payments data from the participating financial institutions.<br><br>Data attributes involved in the exercise included account number and IBAN information for both sides of SME transactions at the participant banks. Data was prepared and encrypted by participant banks, using a one-way hash and salt to pseudo-anonymise the data, and then transferred to a centralised protected and independent server (provided by Deloitte as the data analysis intermediary). Analysts mapped transaction networks and identified patterns of potential financial crime risk based on transaction networks.<br><br>Within the dataset provided, 200,000 accounts with 45 million payments, that had a common link within the network were identified and analysed.<br><br>This project required a 12-month process for approvals to extract and data preparation within the participating institutions. The level of linking achieved in the payment network (approximately 40%) indicates that data quality was sufficient for the Proof of Concept. | |
| **Privacy Enhancing Technologies (PETs) used** | |
| One-way hashing encryption, with common encryption key available to participant banks only (not data intermediary.) | |

**Outcomes, data quality issues and lessons identified**

The project reportedly fulfilled the original three objectives, i.e.:

- Encrypt and extract transaction data from participant institutions in a manner that protects personally identifiable information within the original data from disclosure;
- Connect the encrypted data in order to map transaction nodes within the data; and
- Identify payment patterns in the data that reflect typologies of financial crime risk, that could not have been detected with the data from only a single bank.

Approximately 40% of accounts were linked between the three participating financial institutions.

Case study project managers reported the following lessons identified:

- Evaluating data quality issues in pooled encrypted data remained a challenge;
- While some level of detection is possible through analysing transaction data, the information available was insufficient for full AML investigations. Therefore, there is a need in future models to explore the opportunity for networked transaction information to contribute to joint investigation capabilities in a layered approach.

**Technology provider:** FutureFlow
**Data intermediary:** Deloitte UK

# 7.8. AUSTRAC's Fintel Alliance Alerting Project (Australia) - Project in development for deployment (2019 to present)

| | |
|---|---|
| **Category of use-case** | Financial intelligence within a public–private partnership framework |
| **Sector** | Banking |
| **Objective** | The objective of the Alerting Project is to build a platform to identify financial crime crossing the major Australian financial institutions, which can only be identified by connecting the disparate databases held within each organisation. The use of privacy enhancing technologies is a key focus of the project and is being deployed to protect the privacy of data relating to innocent customers (including their personal details, accounts and transactions). |
| **Intended outcome** | The Alerting Project is intended to deliver:<br>1. the ability to identify whether there are any financial links between two (or more) suspicious accounts; and<br>2. the ability to trace suspicious funds as they move between accounts across financial institutions. |
| **Participants** | This project is being led by AUSTRAC and is being delivered by and for the Fintel Alliance, which is a world first private-public partnership set up by AUSTRAC in 2017. There are currently 28 member organisations (as at 30 June 2020). Details of the Fintel Alliance, including the list of members, can be found on the AUSTRAC website here: https://www.austrac.gov.au/about-us/fintel-alliance<br><br>The project is focused on domestic retail banking transactions, and the active participants are therefore the major banking members, AUSTRAC and select government partner agencies. |
| **Relevant legal jurisdiction** | Australia |

**Information sharing and privacy preservation within this use-case**

| **Information shared or insight achieved** | **Privacy preservation qualities** |
|---|---|
| The algorithm is designed to flag suspicious networks from domestic retail account and transaction data, while protecting the privacy of all data. | No additional customer, account and transaction information will be exposed through the results of the algorithm. Where the transactions meet the criminal typology, AUSTRAC will initiate a follow up process that will be undertaken through formal notice to the relevant financial institutions to identify the specific transactions, accounts and customers of interest. |
| **Nature of data involved in the project** | |
| This project focuses on data relating to domestic retail account transactions. Unlike international fund transfer instructions, domestic transactions are not automatically reported to AUSTRAC under Australian legislation and therefore represent an intelligence gap to the agency and our government partners. This is an active, funded project about to exit the "discovery" phase and enter the "alpha" phase. The project will use a federated architecture. Reporting entities will provide the federated platform with access to the agreed dataset via an API. Initially, it is intended that the project will be tested in a simulated federated architecture using a sample of real data, before the platform is implemented in the real environment. The deployed algorithm is expected to cover more than 100m accounts. The number of transactions will depend, in part, on the temporal range covered by queries - which is undetermined at the time of preparing the case study. | |

The specific data fields engaged by this project include:

- Transaction date;
- Transaction time;
- Account BSB and number;
- Transaction counterparty account BSB and number;
- Transaction amount; and
- Transaction description.

Privacy Enhancing Technologies (PETs) used

The Alerting Project will be performing secure multiparty computation (SMPC). Our current algorithm uses two *kinds* of homomorphic encryption (HE):

1. An *additively homomorphic encryption system* - employing the ElGamal encryption scheme with elliptic curves; and
2. A *levelled homomorphic encryption system (i.e. somewhat homomorphic encryption system)* – employing the Brakerski-Fan-Vercauteren (BFV) encryption scheme.

Other encryption systems may also be used as the project progresses.

**Outcomes, data quality issues and lessons identified**

The project is funded by the Australian Government to run from 1 July 2019 to 30 June 2022. Several proof-of-concepts were undertaken by the Fintel Alliance on an unfunded basis prior to 1 July 2019. The project is ongoing and AUSTRAC is currently exiting the "discovery" phase of the funded project.

The project settled on a 'build' rather than 'buy' option. AUSTRAC reports that a survey of the commercial market offering indicated the inability of inspected vendors to offer the transparency in the underlying algorithms and code that is required by AUSTRAC in order to provide provable security.

Data interoperability issues are being assessed and data requirements of the algorithm have been designed to minimise potential data quality and interoperability issues.

A key lesson identified by AUSTRAC during the discovery phase is that there is not a "one size fits all" algorithm for any given use case. Variables in typologies (including the number of accounts under investigation and the number of steps between accounts that AUSTRAC wish to track) impact the algorithm. AUSTRAC expects the algorithm to require tailoring for each specific use case/operation.

**Technology provider:**

AUSTRAC (The Australian Transaction Reports and Analysis Centre)

https://www.austrac.gov.au/

# 8. Conclusions

The case studies in this report provide an indication of the pace of development of privacy preserving analysis in the field of AML and financial investigations.

Given the pace of innovation since 2000, it is reasonable to expect that the next decade will see a significant growth in: the standards framework for the underlying PET technologies, the efficiency and effectiveness of using these techniques in practice, and greater expectations to use privacy preserving analysis where previously data processors may have been comfortable taking liability for analysing raw unprotected data 'in the clear'.

However, significant hurdles remain in the field. At the technical level, data quality and data interoperability remain a key concern – though several of the case studies in this paper point to new techniques to address some data-interoperability issues. Data interoperability challenges in AML processes and inter-institutional sharing are particularly acute for the retail banking sector, which is often characterised as being dominated by legacy IT infrastructure and systems which 'do not talk to each other' even internally. However, this issue may be less relevant for more recent digital payment services and challenger financial institutions. In addition, wider 'open banking' reforms may support inter-operability capabilities which can serve to support AML and privacy preserving analytical process.

As technical capabilities develop to support information-sharing in a form that may not have been possible before, policy makers, supervisors, FIUs and private sector stakeholders have an opportunity to provide greater clarity about specific information-sharing results they wish to support in pursuit of financial crime detection, disruption and prevention and what contributing information they wish to remain undisclosed or protected in that process.

It is intended that this paper supports further dialogue and feedback on the case-studies and broader implications privacy preserving analytics in the field of AML and financial intelligence. Throughout 2020 and 2021, the FFIS programme will be engaging in such issues as:

- Desired outcomes in policy frameworks for tackling financial crime and respective information-sharing requirements;
- The legal significance of the use of PETs for those information-sharing requirements;
- Wider ethical, regulatory, policy, governance and cultural adoption considerations and potential risks or unintended consequences;
- Practical technical conditions of applying PETs to those use-cases.

# 9. Discussion and feedback

This paper is primarily intended to provide a basis for further engagement with policymakers, supervisors and both private sector and public sector leaders involved in financial intelligence outcomes and financial information-sharing partnerships. We hope this paper is a helpful reference document and can support consideration, feedback and the sharing of insight in relation to the issues covered in the wider research project and the forthcoming Discussion Papers envisaged above within the FFIS research project into "The Role of Privacy Preserving Data Analytics in the Detection and Prevention of Financial Crime". This document will be updated throughout the lifetime of the FFIS research project, including with additional case studies as they come online.

Feedback and additional contributions relevant to this document are invited by email:

PETproject@future-fis.com

## Annex A – Summary table of privacy preserving analytical PET techniques[53]

| Technique: (Partial, Somewhat or Full) Homomorphic Encryption (HE). | |
|---|---|
| **PET descriptions and benefits** | **Limitations** |
| Homomorphic encryption is a form of encryption where some operations (like addition, multiplication, or both) can be performed on the ciphertext, and when the result is then decrypted it will have the same result as if the processing had occurred on the plain text. HE allows computations to be run on the encrypted data and then decrypt the result of the computation only. | Traditionally subject to concerns about computational limitations and a lack of widely accepted standards.<br><br>Early Fully Homomorphic Encryption schemes were exceptionally expensive in terms of computational resource requirements. Recent improvements in these techniques allow for some computations to be completed in relatively short order (seconds and minutes), enabling the practical application of homomorphic encryption to protect sensitive data. Likewise, there are initiatives underway (e.g. Homomorphic Encryption Standardization) to define community standards for HE.[54] |
| **Technique**: Secure Multi Party Computation (SMPC) | |
| **PET descriptions and benefits** | **Limitations** |
| SMPC, or multiparty computation (MPC), is a subfield of cryptography concerned with enabling private distributed computations. In particular, it may be used when two or more parties want to carry out analyses on their combined data but, for legal or other reasons, they cannot share data with one another. MPC can also be used to allow private multi-party machine learning: in this case, different parties send encrypted data to each other and they can train a machine learning model on their combined data, without seeing each other's unencrypted data. | Current SMPC systems have relatively high communications costs. SMPC protocols often require a high degree of specificity to the use case, making them hard to generalise. They can also be slower than computing on raw data and are contingent on the availability of the parties involved. However, 'compilers' that abstract the underlying protocols to enable general-purpose computing are reported as under development, supporting data science and machine-learning applications more broadly. |
| **Technique:** Trusted Execution Environments (TEEs) | |
| **PET descriptions and benefits** | **Limitations** |
| TEEs, or secure enclaves, are a secure area within a physical processor where the processing that happens in that area is hidden from the rest of the processor. TEEs could be used to allow a proprietary algorithm to be run by an untrusted party while ensuring the untrusted party cannot see the algorithm. TEEs often perform and scale well with data size.<br><br>The technology is commercially developed with Intel's Software Guard Extensions (SGX)™ providing a leading example of enclave computing in Skylake™ processors and their successors. Virtualization of SGX is an emerging capability. ARM's Trustzone and AMD's Platform Security Processor also offer TEE capability. Multiple cloud providers offer SGX hardware | Use of enclave computation may require the use of specific hardware that includes enclave features. For example, Intel(R) SGX™. Some TEE providers enable virtualization as well, but only virtualization on top of TEE-equipped hardware.<br><br>TEE is considered to be at a relatively high state of technology readiness as a PET. However, much of what an end user expects in terms of usability of a computing product is still very early in development for TEE. A key shortfall at this point in time is the lack of easy to use development environments for TEE, which would enable general programmers to use these capabilities efficiently and configure them correctly. Another current shortfall is that leading TEE's such as Intel SGX require interaction directly |

| | | |
|---|---|---|
| | where one can run these applications when one does not have direct access to such hardware. Microsoft supports Azure Confidential Computing program, IBM Cloud provides machines with SGX support and Alibaba Cloud has SGX machines as well. | with the technology provider in order to properly use these security capabilities.<br><br>TEEs may be vulnerable to certain kinds of side channel attacks. This is where an attacker monitors certain properties of the system, such as the time required to perform an operation, to learn sensitive information. |
| **Technique:** Zero Knowledge Proofs (ZKP) | | |
| | **PET descriptions and benefits** | **Limitations** |
| | ZKPs are a method by which an entity can prove that they know something to another entity, without revealing anything other than that they know that thing. ZKPs can be used for authentication. An entity can prove they know a password that proves their identity, without having to reveal their password. ZKP has applications across a variety of use cases – including payments (Zcash), internet infrastructure (NuCypher), digital identity (Nuggets) and others. and it is expected to be a critical enabler of distributed ledger technologies more broadly. | ZKP has only recently seen real-world operational uses as the methodology continues to mature.<br><br>Scalability can be a technical challenge and, as is common for PETs in 2020, further work is required to develop global community standards for the technology. |
| **Technique**: Federated Learning | | |
| | **PET descriptions and benefits** | **Limitations** |
| | In traditional machine learning data is centralised and brought to the model. In federated learning the data is distributed and the model is sent to the data. What is then centralised is the model updates from all of the federated devices. Federated learning allows a model to be updated without centralising the data the update is based on. As the central party does not see the data, they need to be confident that the data is structured, cleaned, and encoded appropriately, otherwise it can fail or lead to a poorly trained model.<br><br>Federated learning is developed and in use in household mobile applications. In March 2019, TensorFlow (a widely used open-source library for machine learning) published TensorFlow Federated, an open-source framework that allows machine learning to be performed on federated datasets. | Federated learning in isolation is not necessarily privacy preserving, as it can be applied in a manner that there are no meaningful privacy guarantees of the models or of the underlying data.<br><br>It is also important to note that this model does not necessarily produce an equivalent model to the one that would be derived by first combining the training data into a central location; in most cases, a model trained through federated machine learning would be inferior to the one trained on a centralized dataset.<br><br>Again, one of the challenges being faced is the absence of standards, systems and homogeneous languages, which permit distinct actors to interact with services based on this technology. |

## Annex B – Relevant technology companies highlighted in case studies

The Australian Transaction Reports and Analysis Centre (AUSTRAC) / Data61
https://www.austrac.gov.au/
https://data61.csiro.au/

Deloitte UK
https://www2.deloitte.com/uk/en/profiles/pafenton.html

Duality Technologies
https://dualitytech.com/

Enveil
https://www.enveil.com/

FutureFlow
http://www.futureflow.org/

Inpher
https://www.inpher.io/

Privitar
https://www.privitar.com/

Secretarium
https://secretarium.com/

Sedicii
https://sedicii.com/

# Endnotes

[1] FATF (2017) 'Public Consultation on the Draft Guidance for Private Sector Information Sharing', p. 3.
[2] See RUSI 'Expanding the Capability of Financial Information-Sharing Partnerships' (March 2019) and 'the Role of Financial Information-Sharing Partnerships in the Disruption of Crime' (Oct 2017)
[3] World Economic Forum (2019) 'White Paper: The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value' prepared in collaboration with Deloitte.
[4] Note that there are major differences across jurisdictions in how privacy is treated within law. For example, the EU approaches privacy as a fundamental right while the legal tradition in the United States has approached privacy arguably as a commodity subject to market forces – see Peter Blume (1997) Privacy as a Theoretical and Practical Concept in International Review Of Law Computers & Technology, Volume 11, Number 2, Pages 193–202 – However, the development of the California Consumer Privacy Act 2018 provides for a rights-based approach similar to the EU General Data Protection Regulation (GDPR).
[5] World Economic Forum (2019) 'White Paper: The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value' prepared in collaboration with Deloitte.
[6] Summarised from FATF (2012-2019), International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation, FATF, Paris, France
[7] Summarised from Maxwell, N (2019) 'Expanding the Capability of Financial Information-Sharing Partnerships' RUSI
[8] Most notably the USA PATRIOT Act 314(b) makes specific provision to support private-private sharing between regulated entities, prior to a suspicion being formed.
[9] Note in 2017, FATF updated its guidance in FATF (2017), Guidance on private sector information sharing, FATF, Paris to set out expectations of FATF Recommendations 18, 20 and 21 relating to information sharing within financial groups. However, few FATF members have yet instituted changes to legal frameworks to encourage intra-group cross border sharing.
[10] Despite FATF efforts (in FATF (2017), Guidance on private sector information sharing, FATF, Paris) that this is not the intention of the prohibition on tipping off, at the level of national implementation of FATF standards, private–private sharing of financial crime risk information is typically prohibited pre-suspicion.
[11] See FATF (2018), Professional Money Laundering, FATF, Paris, France
[12] Outside of jurisdictions where there exists public–private financial information-sharing partnerships or other bi-lateral exchange and private-private sharing legal gateways.
[13] World Economic Forum (2019) 'White Paper: The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value' prepared in collaboration with Deloitte.
[14] Privitar (2019) Guide to De-identification' Policy and Research Paper, London.
[15] Royal Society (2019) 'Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies' London.
[16] See IBM 2009 research on partially homomorphic encryption https://crypto.stanford.edu/craig/easy-fhe.pdf
[17] Malkhi D et al. 2004 Fairplay – a secure two-party computation system. SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium, 20–20. (see https://www.usenix.org/legacy/event/sec04/tech/malkhi/malkhi.pdf)
[18] See https://web.archive.org/web/20140903041544/http://www.trusted-logic.com/IMG/pdf/TRUSTED_LOGIC_TRUSTED_FOUNDATIONS_OMTP_FINAL.pdf
[19] See https://ercim-news.ercim.eu/en73/special/trading-sugar-beet-quotas-secure-multiparty-computation-in-practice
[20] See https://crypto.stanford.edu/craig
[21] See https://www.theregister.co.uk/2018/03/08/ibm_faster_homomorphic_encryption/
[22] K. Rohloff, D. B. Cousins and D. Sumorok, (2017) "Scalable, Practical VoIP Teleconferencing With End-to-End Homomorphic Encryption," in *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 5, pp. 1031-1041, May 2017, doi: 10.1109/TIFS.2016.2639340.
[23] D. B. Cousins, K. Rohloff and D. Sumorok, (2017) "Designing an FPGA-Accelerated Homomorphic Encryption Co-Processor," in *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 2, pp. 193-206, 1 April-June 2017, doi: 10.1109/TETC.2016.2619669.
[24] See http://homenc.github.io/HElib/
[25] See https://palisade-crypto.org/
[26] See https://www.rand.org/content/dam/rand/pubs/research_reports/RR500/RR567/RAND_RR567.pdf
[27] Based on Archer DW et al. 2018 From Keys to Databases – Real-World Applications of Secure Multi-Party Computation (see https://eprint.iacr.org/2018/450 )

[28] Royal Society (2019) 'Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies' London.
[29] See https://www.intelligence.gov/index.php/ic-on-the-record-database/results/685-iarpa-releases-research-report-on-security-and-privacy-assurance
[30] See https://heat-project.eu/index.html
[31] See https://data61.csiro.au/en/Our-Research/Our-Work/Safety-and-Security/Privacy-Preservation/Confidential-computing
[32] See https://101blockchains.com/zero-knowledge-proof/#7
[33] See https://cointelegraph.com/explained/zero-knowledge-proofs-explained
[34] See homomorphicencryption.org
[35] Albrecht MR et al. 2015 On the concrete hardness of Learning with Errors. J. Mathematical Cryptology 9. (see https://eprint.iacr.org/2015/046.pdf
[36] See https://heat-project.eu/deliverables.html
[37] See https://www.nuffieldtrust.org.uk/resource/the-nhs-workforce-in-numbers
[38] Privitar (2019) Guide to De-identification' Policy and Research Paper, London.
[39] See https://medium.com/tensorflow/introducing-tensorflow-federated-a4147aa20041
[40] UN (2019) 'Handbook on Privacy-Preserving Computation Techniques' Big Data UN Global Working Group.
[41] Royal Society (2019) 'Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies' London.
[42] https://www.lexology.com/library/detail.aspx?g=a2c715b1-0a7a-4c6c-95a7-d1c9537f0c95
[43] https://www.fca.org.uk/events/techsprints/2019-global-aml-and-financial-crime-techsprint
[44] World Economic Forum (2019) 'White Paper: The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value' prepared in collaboration with Deloitte.
[45] World Economic Forum (2019) 'White Paper: The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value' prepared in collaboration with Deloitte.
[46] See https://confidentialcomputing.io/wp-content/uploads/sites/85/2019/12/CCC_Overview.pdf
[47] European Commission (2020) 'European strategy for data' consultation - Brussels, 19.2.2020 COM(2020)
[48] European Data Protection Supervisor (2020) Opinion 3/2020 on the European strategy for data (16 June 2020)
[49] See https://www.fca.org.uk/publication/research/fostering-innovation-through-collaboration-evolution-techsprint-approach.pdf
[50] See pilot description for more details: https://www.fca.org.uk/firms/innovation/digital-sandbox
[51] See https://www.twi-global.com/technical-knowledge/faqs/technology-readiness-levels
[52] World Economic Forum (2019) 'White Paper: The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value' prepared in collaboration with Deloitte.
[53] Compiled from World Economic Forum (2019) 'White Paper: The Next Generation of Data-Sharing in Financial Services: Using Privacy Enhancing Techniques to Unlock New Value' prepared in collaboration with Deloitte; Privitar (2019) 'Guide to De-identification'; Royal Society (2019) 'Protecting privacy in practice: The current use, development and limits of Privacy Enhancing Technologies' London; and UN (2019) 'Handbook on Privacy-Preserving Computation Techniques' Big Data UN Global Working Group.
[54] See also Homomorphic Encryption Computing Techniques with Overhead Reduction (HECTOR) iarpa program - https://www.iarpa.gov/index.php/research-programs/hector