8th International Conference on
BIG DATA
& Data Science for Official Statistics
BILBAO 2024

Informing Climate Change and
Sustainable Development Policies
with Integrated Data

BILBAO. SPAIN   10-14 JUNE 2024   #UNBigData2024

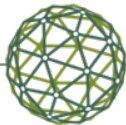# Demonstrations of the use of Mobile Phone Data and of Privacy Enhancing Technologies

8th International Conference on Big Data and Data Science for Official Statistics
June 2024

Session agenda

1. Introduction to the UN PETs Task Team and PET Lab
   - Matjaž Jug, Statistics Netherlands and Chair of the Task Team
   - Dave Buckley, OpenMined and co-chair of the PET Lab

1. Demo: using PySyft for privacy-preserving data science
   - Irina Bejan and Matei Simtinică, OpenMined

1. Experiences using PETs at Instituto Brasileiro de Geografia e Estatística
   - Augusto Fadel, IBGE and Member of the Task Team

What are PETs?

- PETs are a broad collection of technologies that can shift the tradeoff between the privacy and usability of data.

- PETs can enable different researchers and organisations to perform collaborative data science, without ever seeing each other's private data assets.
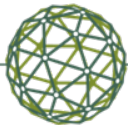
What are PETs?

- Today, high-value data is siloed or locked down, significantly **hampering research and innovation**;

- This is usually due to a combination of privacy, security, IP, cost, or legal reasons;

- When data is shared, things too often go wrong => *data is misused, or data privacy is not properly protected;*

- **PETs**, whilst not a silver bullet, have the potential to overcome this tradeoff between privacy and usability of data
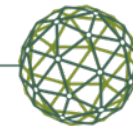
Why PETs for official statistics?

- NSOs have a mandate to serve the public interest by:
  - Protecting citizens' data
  - Maximising the public benefits derived from that data

- Traditional governance and disclosure practices fall short in doing both well

- PETs provide new avenues for enabling microdata to be more broadly accessible for research and innovation, whilst guaranteeing that sensitive information remains confidential
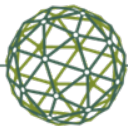
PETs for input and output privacy

Input Privacy

A guarantee that one or more parties can participate in a computation, in such a way that neither party learns anything about the other party's inputs to the computation.

◆ SECURE ENCLAVES

◆ FUNCTIONAL ENCRYPTION

◆ HOMOMORPHIC ENCRYPTION

◆ SECURE MULTI-PARTY COMPUTATION

◆ FEDERATED LEARNING

◆ ZERO-KNOWLEDGE PROOFS

Output Privacy

A guarantee that sensitive information in the data cannot be reverse engineered from the outputs of the computation.

◆ K-ANONYMIZATION

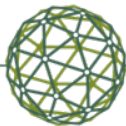◆ DIFFERENTIAL PRIVACY

◆ SYNTHETIC DATA
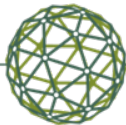
Use cases for NSOs

- Facilitating external researcher access to high-quality datasets where it is currently not possible

- Enhancing and streamlining existing risk management and accreditation processes - researchers can access data faster

- Enabling collaborations between agencies, both domestically and cross-border, via privacy-preserving joins

- Minimising the privacy risk associated with data disclosures, e.g. the use of differential privacy for census data
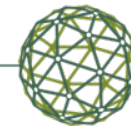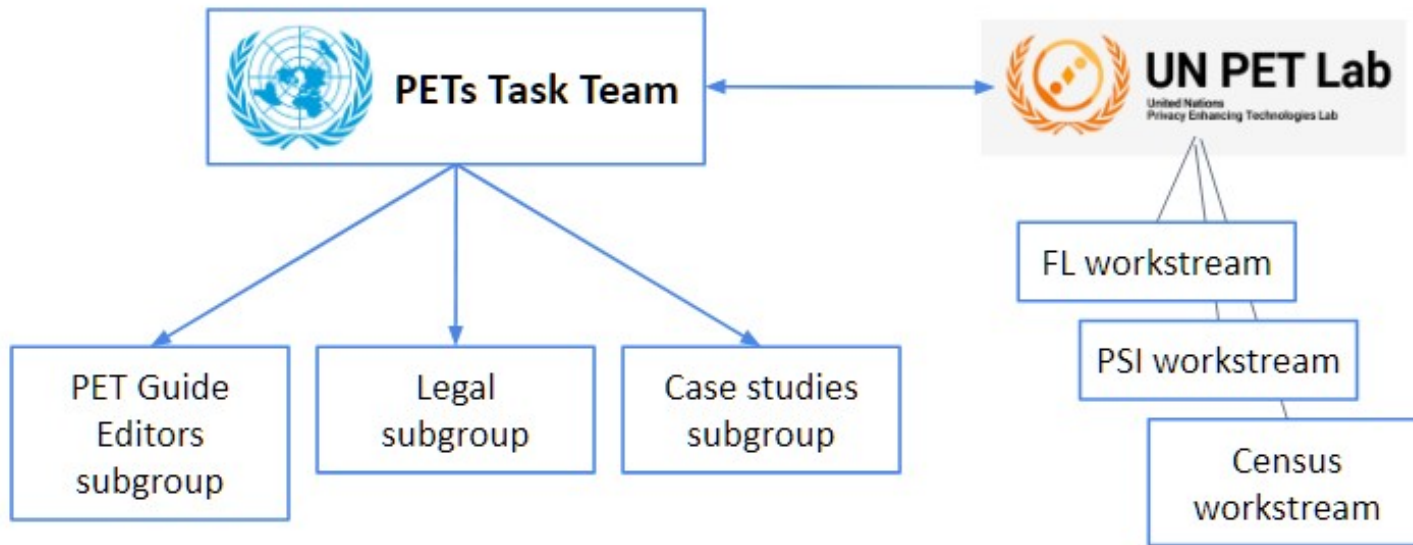
Task Team for Privacy Enhancing Technologies

- Methodology (PET Task Team)
  - Research of applicability of PETs for use cases in Official Statistics.

- Experimentation (PET Lab)
  - Proofs-of-concept and pilot projects focused on the evaluation of PETs for real-world use cases in the official statistics community.

- Outreach & Training (events, webinars, Open Houses)
  - Sharing learnings and insights from the use of PETs with the wider statistical community through training, public events, and educational materials.

- Support Services
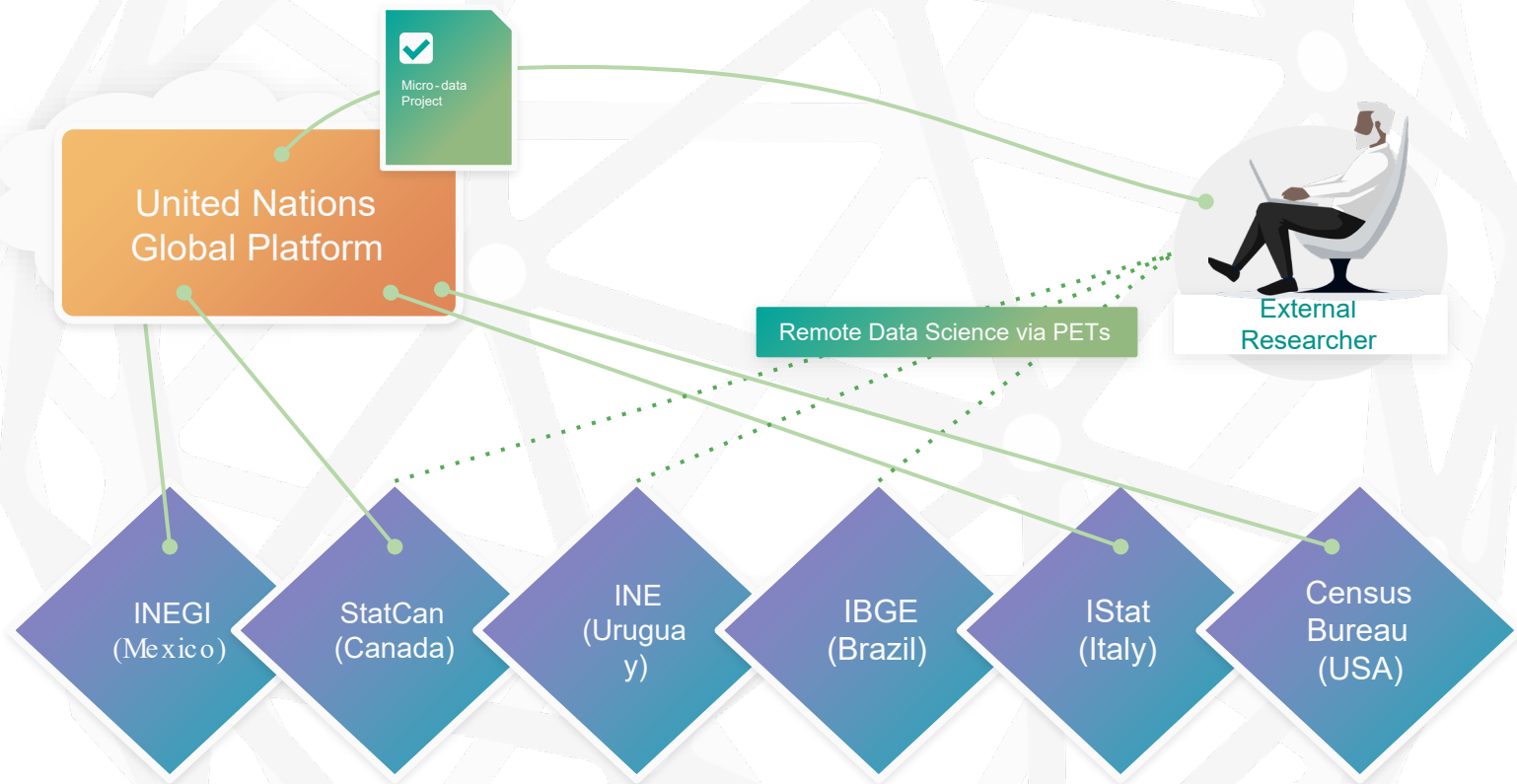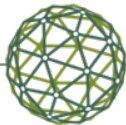 - Advice to organizations utilizing PETs

UN PET Teams

Ongoing PET Lab projects
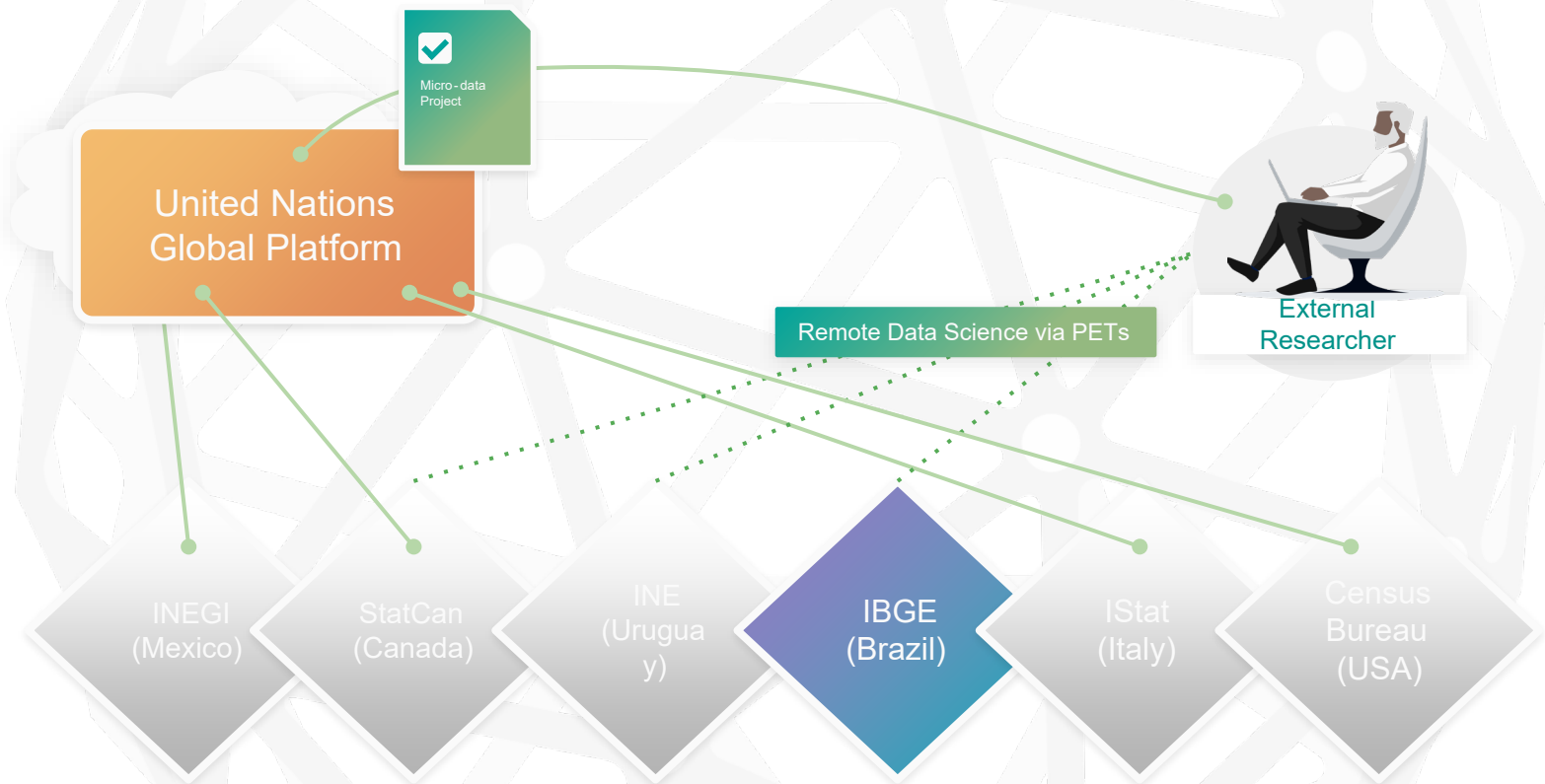

Micro-data Project

United Nations Global Platform

External Researcher

Remote Data Science via PETs

INEGI (Mexico)

StatCan (Canada)

INE (Uruguay)

IBGE (Brazil)

IStat (Italy)

Census Bureau (USA)

BILBAO 2024 · #UNBigData2024

Project 1: using UN Global Platform to enable external researchers to access microdata in Brazil

Micro-data Project

United Nations Global Platform

External Researcher

Remote Data Science via PETs

INEGI (Mexico)

StatCan (Canada)

INE (Uruguay)

IBGE (Brazil)

IStat (Italy)

Census Bureau (USA)
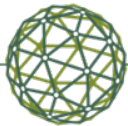
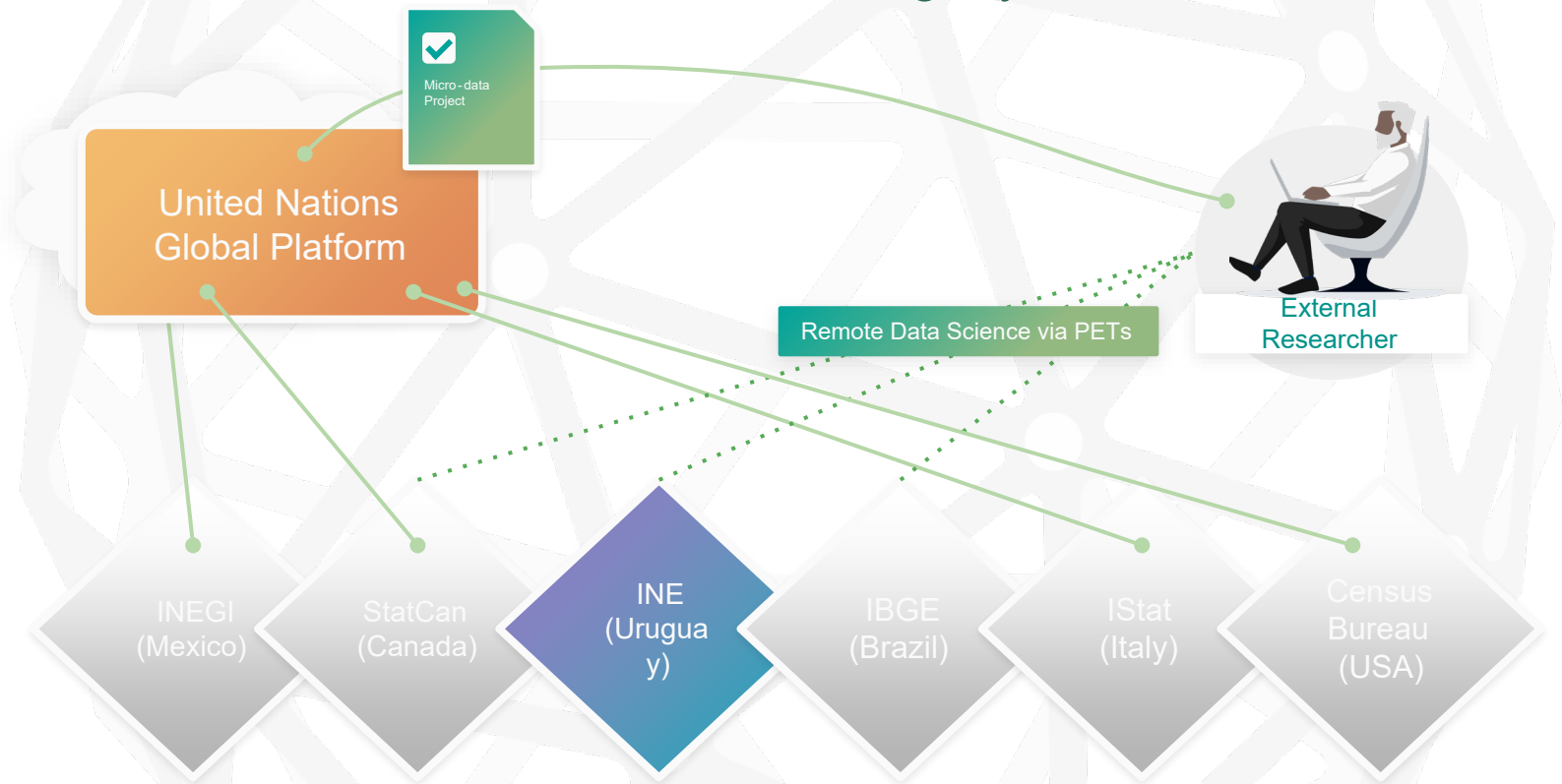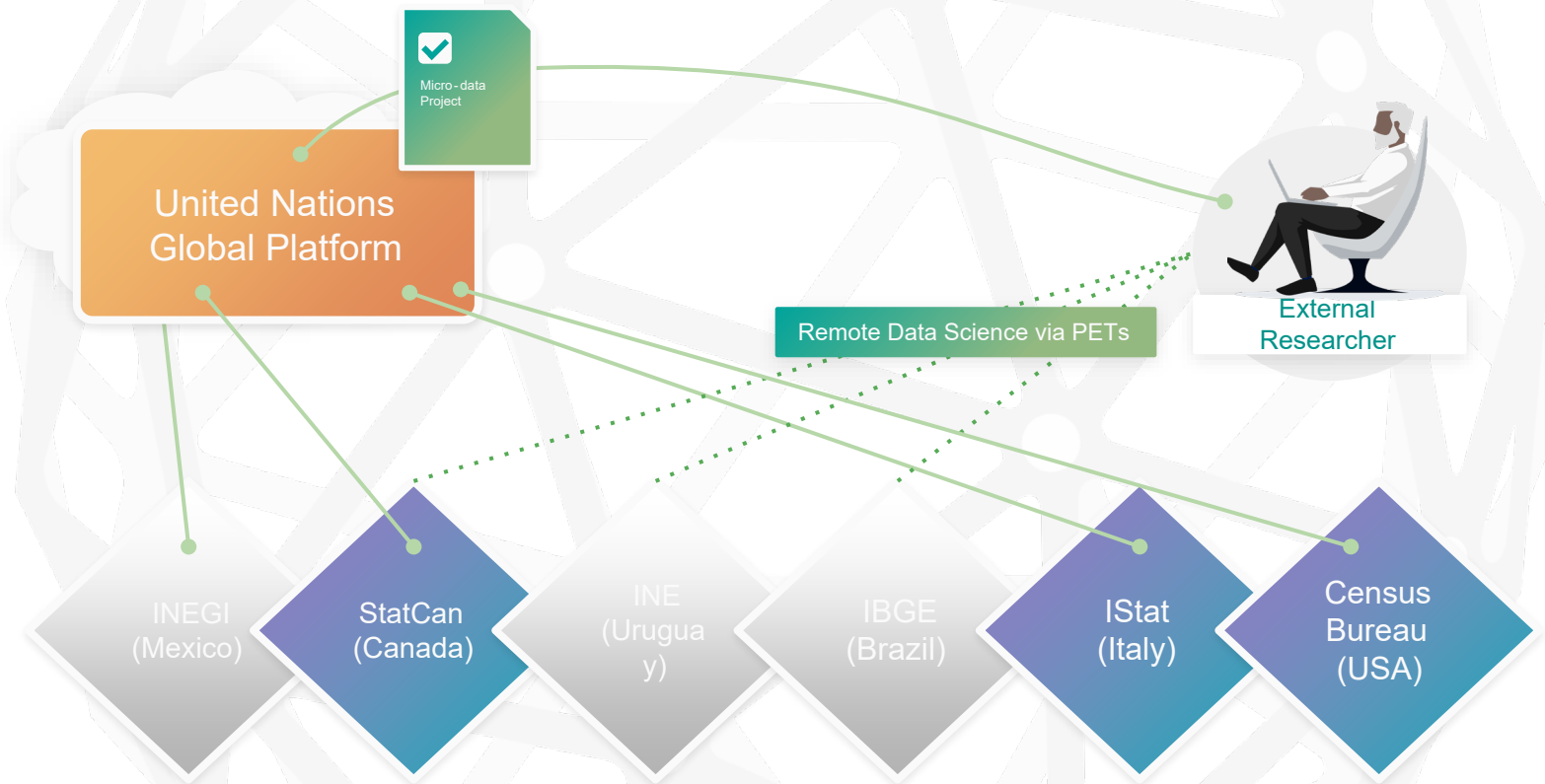# Project 2: using UN Global Platform to enable external researchers to access microdata in Uruguay
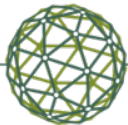
Project 3: demonstrate an international data join between NSOs

Micro-data Project

United Nations Global Platform

External Researcher

Remote Data Science via PETs

INEGI (Mexico)

StatCan (Canada)

INE (Uruguay)

IBGE (Brazil)

IStat (Italy)

Census Bureau (USA)
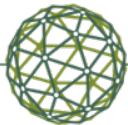
Task team projects: UN PET Guide

- Role of Privacy-Enhancing Technologies in Official Statistics

- Methodologies & Approaches

- 18 Case Studies

- Standards

- Legal and Regulatory Issues

THE PET GUIDE

THE UNITED NATIONS GUIDE ON PRIVACY-ENHANCING TECHNOLOGIES FOR OFFICIAL STATISTICS.
2023

United Nations BigData

PET Case Studies in Official Statistics
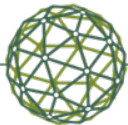
18 case studies

1 Concept
9 Proof of Concept
5 Pilot
3 Production

2 TEE   6 MPC
2 FL   6 HE
3 SD   6 DP

Drawn from NSOs and other organizations in 8 jurisdictions

From 1 to 100s input parties

https://unstats.un.org/bigdata/task-teams/privacy/case-studies/

Case study repository

Created by David Buckley, last modified on Feb 08, 2023

A number of national statistical offices (NSOs) and government agencies are leveraging PETs to enable rich and innovative statistical analysis, whilst protecting the privacy and confidentiality of sensitive information contained within their datasets. This repository provides example case studies of where PETs have been deployed in the real world to facilitate the generation and dissemination of privacy-preserving statistics.

The repository includes the 18 case studies that are provided in Chapter 3 of the 2023 UN Guide on Privacy-Enhancing Technologies for Official Statistics (which can be found here). We intend to update and expand this repository moving forward,

- 1. Boston Women's Workforce Council: Measuring salary disparity using secure multi-party computation
- 2. European Statistical System: Developing Trusted Smart Surveys
- 3. Eurostat: Processing of longitudinal mobile network operator data
- 4. Indonesia Ministry of Tourism: Confidentially sharing datasets between two mobile network operators via a trusted execution environment
- 5. Italian National Institute of Statistics and Bank of Italy: Enriching data analysis using privacy-preserving record linkage
- 6. Office for National Statistics: Trialling the use of synthetic data at the United Kingdom's national statistics institute
- 7. Samsung SDS (Korea): Data aggregation system
- 8. Statistics Canada: Measuring the coverage of a data source using a private set intersection
- 9. Statistics Canada: Training a machine learning model for private text classification using leveled homomorphic encryption
- 10. Statistics Canada: Trialling the use of synthetic data
- 11. Statistics Korea: Developing a privacy-preserving Statistical Data Hub Platform
- 12. Statistics Netherlands: Developing privacy-preserving cardiovascular risk prediction models from distributed clinical and socioeconomic data
- 13. Statistics Netherlands: Measuring effectiveness of an eHealth solution using private set intersection
- 14. Twitter and OpenMined: Enabling Third-party Audits and Research Reproducibility over Unreleased Digital Assets
- 15. United Nations Economic Commission for Europe: Trialling approaches to privacy-preserving federated machine learning
- 16. United Nations PET Lab: International Trade
- 17. United States Census Bureau: Deploying a differentially private Disclosure Avoidance System for the 2020 US Census
- 18. United States Department of Education: Analysing student financial aid data using privacy-preserving record linkage

16. United Nations PET Lab: International Trade

Created by David Buckley, last modified on Feb 09, 2023

| Purpose	Enable multiple national statistical offices (NSOs) to perform reconciliation and joint analysis
Da	
PE	
De co	
Pa re	between parties.
Implementation status	Proof of concept (ongoing)
Resources	What is the UN PET Lab and Why is it Important? The Economists write-up on the PET Lab

The Economist

≡ Menu   Weekly edition   The world in brief   🔍 Search ⌄

Science and technology | Data privacy

The UN is testing technology that processes data confidentially

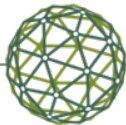How to analyse data without revealing their secrets

OpenMined
Network Node

Trusted Execution Environment
via Secure Enclave
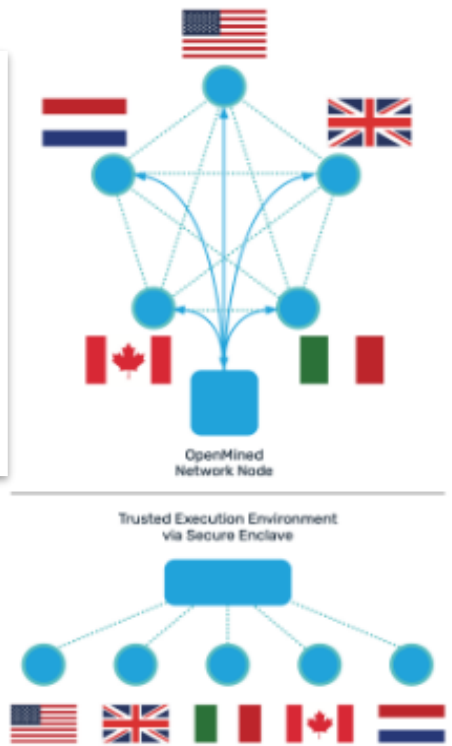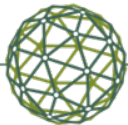
Contributing Case Studies

Interactive Demo of UN PET Case Studies Repository

https://unstats.un.org/wiki/display/UGTTOPPT/Contributing+to+the+repository

Thank you

To join our community, contact: bigdata@un.org

Or reach out to:

- **m.jug@cbs.nl**
- **dave@openmined.org**