Confidential sharing of datasets of two mobile network operators:
A case study for tourism statistics
from the perspective of a technology provider

Baldur Kubo, baldur.kubo@cyber.ee
Agenda

- Cybernetica
- What is privacy
- Stakeholders
- Problem
- Technology selection
- Stakeholder roles
- What was created
- Key documents
Cybernetica is a **knowledge-intensive SME** based in Estonia

- started as an applied research unit of the Institute of Cybernetics of the Academy of Sciences of Estonia in 1960
- established as a private limited company in 1997

- 190+ employees (10% PhD)
- Clients in 35+ countries
- Inhouse R&D dept.
- AA Credit Rating
- ISO certified
Extensive information security and privacy expertise

- Inventors and engineers of e-governance solutions since 2001
- Pioneers in privacy-enhancing technologies (PETs) since 2007

- Developed and maintains the first and only national online voting solution.
- Developed and maintains Estonia’s government data exchange platform (X-Road). Distributes a product version of the software (UXP).
- Carried out a pilot for Estonia’s eID smart card service. Developed next generation eID (SplitKey).
- Developed a clinical decision support tool for GlaxoSmithKline.
- Designed an encrypted genome data storage and querying mechanism for Sophia Genetics.
- Information security & privacy risk analysis of the COVID 19 tracing app in Estonia; privacy-preserving statistics based on mobile location data for Eurostat.
What is Privacy?

● Broad
  – the right to be let alone, or freedom from interference or intrusion

● In information context
  – the right to have some control over how your personal information is collected and used

  – Source: https://iapp.org/about/what-is-privacy/
Privacy is use case specific

Any data processing use case is defined by:
1. purpose (why?)
2. data (what?)
3. processors (who?)

Data processing usually relies on a legal basis, be it for privacy/data protection or confidentiality/data ownership reasons.
Stakeholders

- Ministry of Tourism
- Statistics Indonesia (BPS)
- Mobile network operator1 (MNO1)
- Mobile network operator2 (MNO2)
- Mobile Subscribers
- Positium
- Cybernetica AS
- Intel
Problem

- Mobile positioning data characterizes quantities and movements of tourists, information needed by the NSI and MoT. (see use case)
- Tourists are using mobile phones by roaming through local mobile network operators (MNOs).
- Cross-roaming - a person might use two or more different MNOs, resulting in overcounting.
- Cross-roaming can be analyzed when unique subscriber information (IMSI) is compared across several MNOs.
- Input data is both privacy sensitive and business confidential.
Data

Mobile positioning data from largest operator in 29 areas
## Choice Between Two Security Technologies

<table>
<thead>
<tr>
<th></th>
<th>Sharemind MPC</th>
<th>Sharemind HI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technology</strong></td>
<td>Multi-Party Computation based on Homomorphic Secret Sharing</td>
<td>Trusted Execution Environments using Intel® SGX</td>
</tr>
<tr>
<td><strong>Deployment</strong></td>
<td>Distributed application server</td>
<td>Single node application server</td>
</tr>
<tr>
<td><strong>Cloud support</strong></td>
<td>Any cloud provider</td>
<td>Any cloud provider supporting Intel® SGX technology</td>
</tr>
<tr>
<td><strong>Applications</strong></td>
<td>Healthcare, finance, government, statistics, ML/AI and more</td>
<td></td>
</tr>
</tbody>
</table>
Sharemind Technology

- Back-end solution for built-in cryptography in data analytics
  - data minimization pushed to the maximum

- Technologically enforced data governance policies
  - new level of data transparency

- State of the art endorsed by data protection supervisors
  - anonymization tool, appropriate protection measure
Data holders
Unlocking the value of data
New kind of control

Service providers
Delivering the value of data
New kinds of services

Data users
Realizing the value of data
New insights

✓ License
✓ Maintenance
✓ Privacy engineering services
Choise Between Two Security Technologies

- Both TEE Sharemind HI or Sharemind MPC were suitable
  - From scalability requirements of the specific project
  - Deployment model point of view
- Positium’s development plans of the future version of Positium Data Mediator (PDM) determined the technology choice
  - Solution with Sharemind HI as the next implemented module of PDM
  - Cross-check of cross-roaming on the cross-border
Stakeholders roles I/II

- Ministry of Tourism (MoT)
  - Coordinator; Output Consumer; Solution host
- Mobile network operator 1 & 2
  - Data Provider; Enforcer
- Mobile Subscribers
  - Data subject
Stakeholders roles II/II

- **positium**
  - Domain expert, Methodology provider, Tester, Auditor, Enforcer

- **Cybernetica AS**
  - Security technology provider (Sharemind HI); Attestation service proxy
  - Designer of the solution; Developer of the solution

- **Intel**
  - Security technology provider Intel SGX®; Attestation Service provider
What did we do? I/II

- Developed the technical solution using Sharemind HI confidential computing platform which uses the Intel® Software Guard Extensions (Intel® SGX) technology
  - to analyse mobile positioning data (gathered by two MNO-s) for
  - calculating weights to correct counts of tourists by region for national tourism statistics
  - protecting privacy of subscribers and confidential business information of MNO-s. (see longer post)
What did we do? II/II

• Developed the methodology
• Tested the technical solution
• Deployed it onsite in MoT
• Organized MNO-s to come together and encrypt data for the solution
• Ran the secure calculations
• Produced report for BPS and MoT
Key documentation

- Secure IMSI list intersection with Sharemind® HI – problem, security goals, stakeholders and their roles, solution and technology intro, business process
- Technical documentation – key management practice, deployment key setup, installation guides (client, server), Sharemind® HI security technology overview
- Positium’s report to MoT
DEMONSTRATING

Privacy Engineering can become a core competence of Statistical Offices

Privacy-Enhancing Technologies are mature and usable
Future opportunities

- Selecting the MNO with whom to analyze tourism statistics
- Pilots of privacy-preserving location data analytics a la Eurostat
  - [https://ec.europa.eu/eurostat/cros/content/eurostat-cybernetica-project_en](https://ec.europa.eu/eurostat/cros/content/eurostat-cybernetica-project_en)