# Smarter decisions with no privacy breaches

**Dan Bogdanov & the Sharemind team**

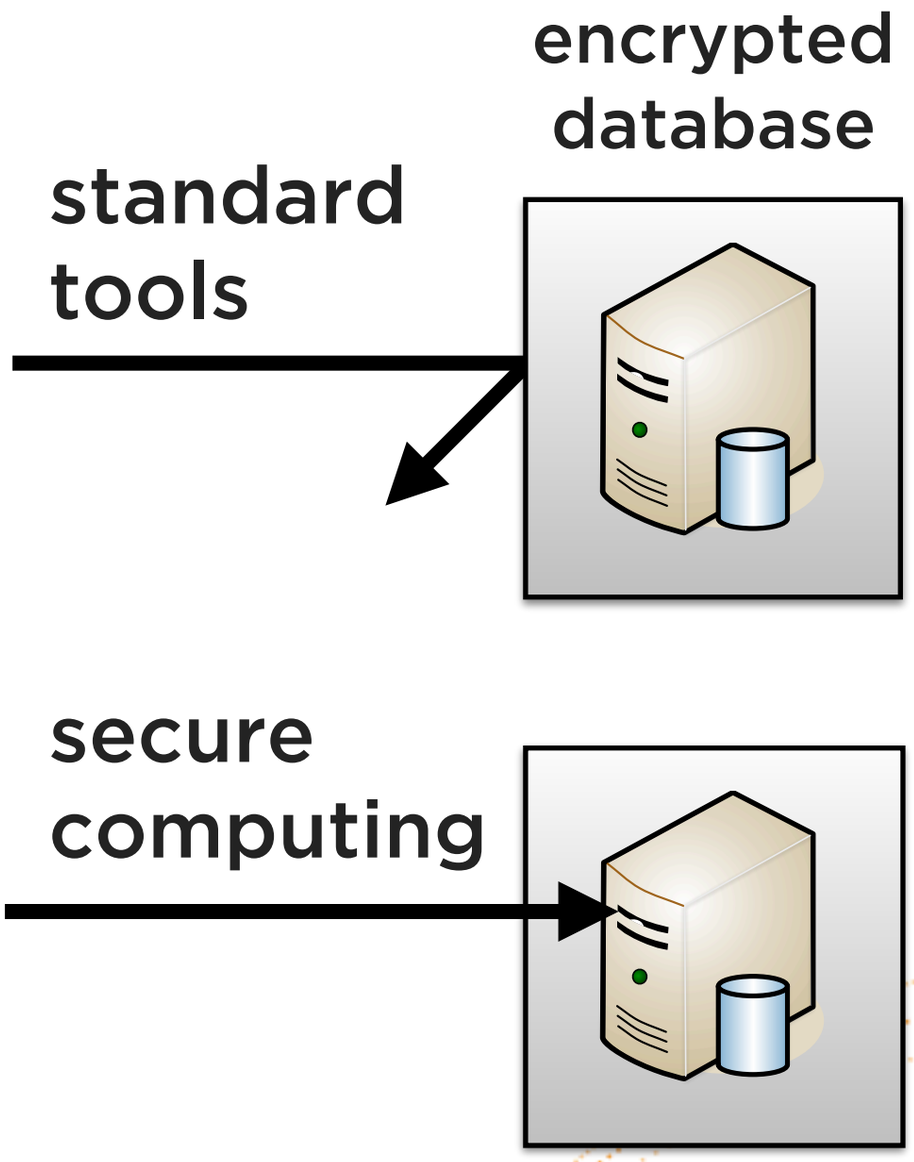
dan@cyber.ee
http://sharemind.cyber.ee/

sharemind

# Secure computing

When a standard database encrypts data, it must be decrypted before analysis

Secure computing systems can analyse data without removing the encryption.




sharemind

# Flavours of “practical”

*Paper-practical*
“The proposed technique solves a practical problem and performs rather nicely in the lab.”

*Dan Bogdanov. Sharemind: programmable secure computations with **practical** applications. PhD thesis. 2013.*

*Real-world practical*
“We had a customer, figured out the legal aspects, solved the deployment problems and made it work in a less-than-optimal environment.”

sharemind

# Massive lack of IT talent?

**Software developer shortage transcends international boundaries**

**Shortage brings demand for overseas engineers**

**The Myth of America's Tech-Talent Shortage**

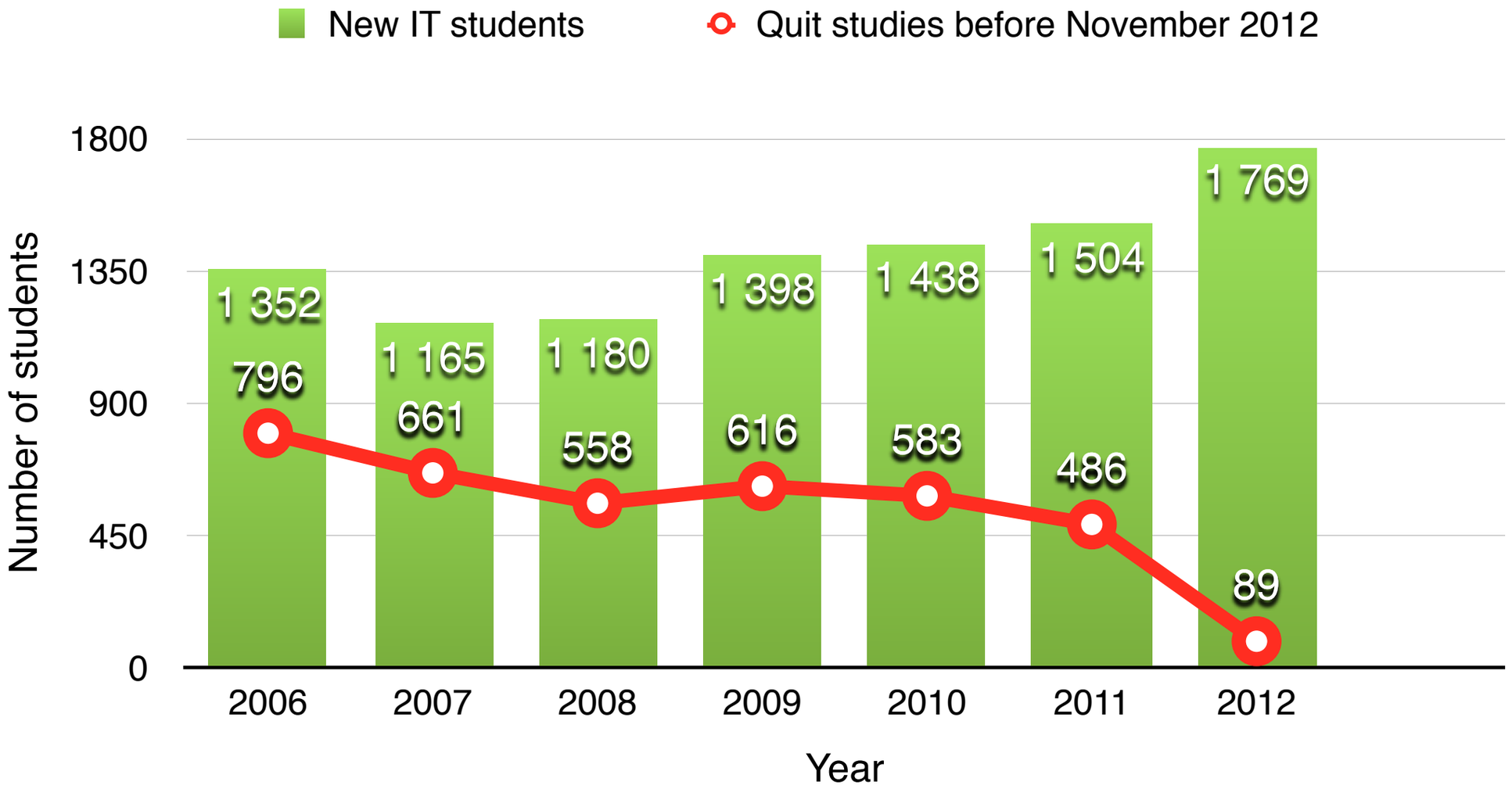
Computer science graduates struggle to find work despite IT skills shortage

The fact that up to 900 000 jobs in the ICT sector remain unfilled because of a skills gap gives the clearest indication possible of what needs to be done," says Manuel Kohnstamm, Liberty Global's senior vice president and chief policy officer.

http://careers.ieee.org/article/European_Job_Outlook_0414.php

sharemind

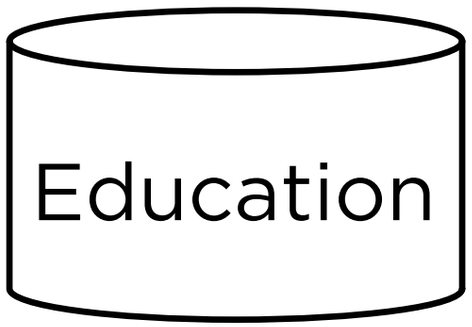
# Training has a failure rate




*By 2012, a total of 43% of students enrolled in in the four largest IT higher learning institutions in Estonia during 2006-2012 had quit their studies.*
*Source: Estonian Ministry of Education and Research, CentAR.*

sharemind

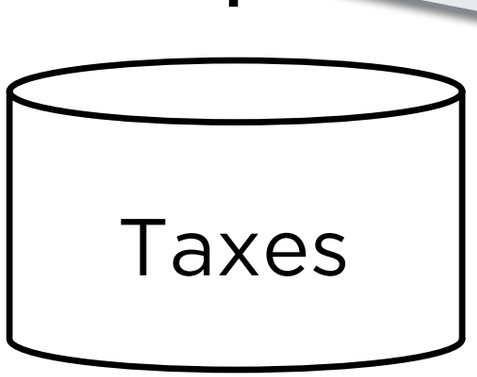
# Government knows, why




The ministry knows if you've been studying.

+

However, this operation breaches
1. the Estonian Personal Data Protection Act,
2. the Estonian Taxation Act and (possibly)
3. the EU Data Protection regulations.




The tax board knows if you've been working.

=


Does working cause school failure?

sharemind

# The current workaround

1. The analyst describes demographic groups.
2. The Ministry of Education provides person codes for each group to the Tax Board.
3. The Tax Board merges education records with income tax records, ensuring that no group has less than three people (smaller groups are removed).

This directly causes **54%** of Master's students and **78%** of PhD students to be left out of the study.

*Source: Experiment carried out by CentAR and Cybernetica in 2014.*

sharemind

# Our achievement

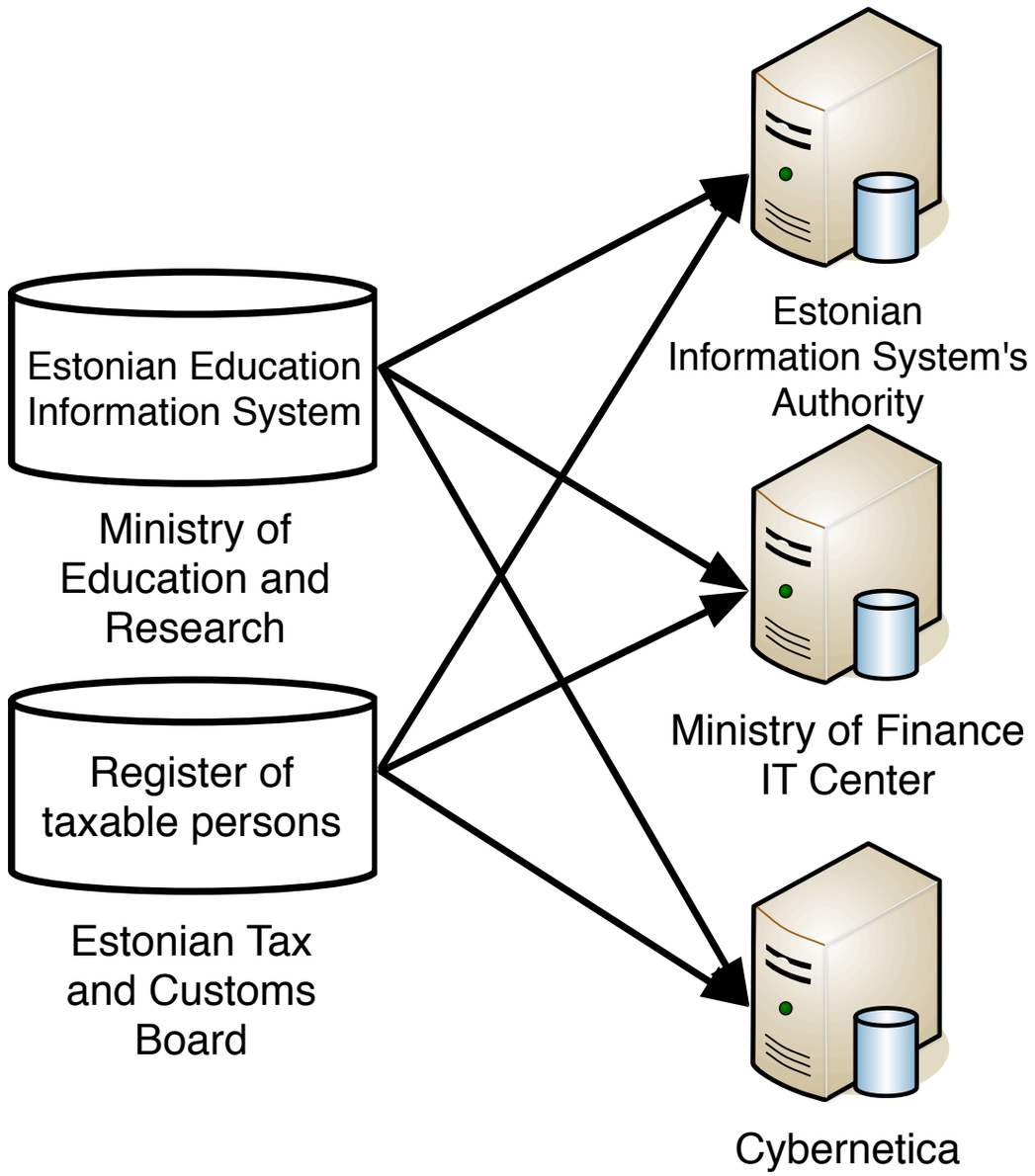
We built a privacy-preserving system to securely collect tax and education records, link them and perform the necessary statistical analysis.

The solution is based on the Sharemind secure multi-party computation platform and provides cryptographic protection during data processing.

**It runs on real tax and education records.**

sharemind

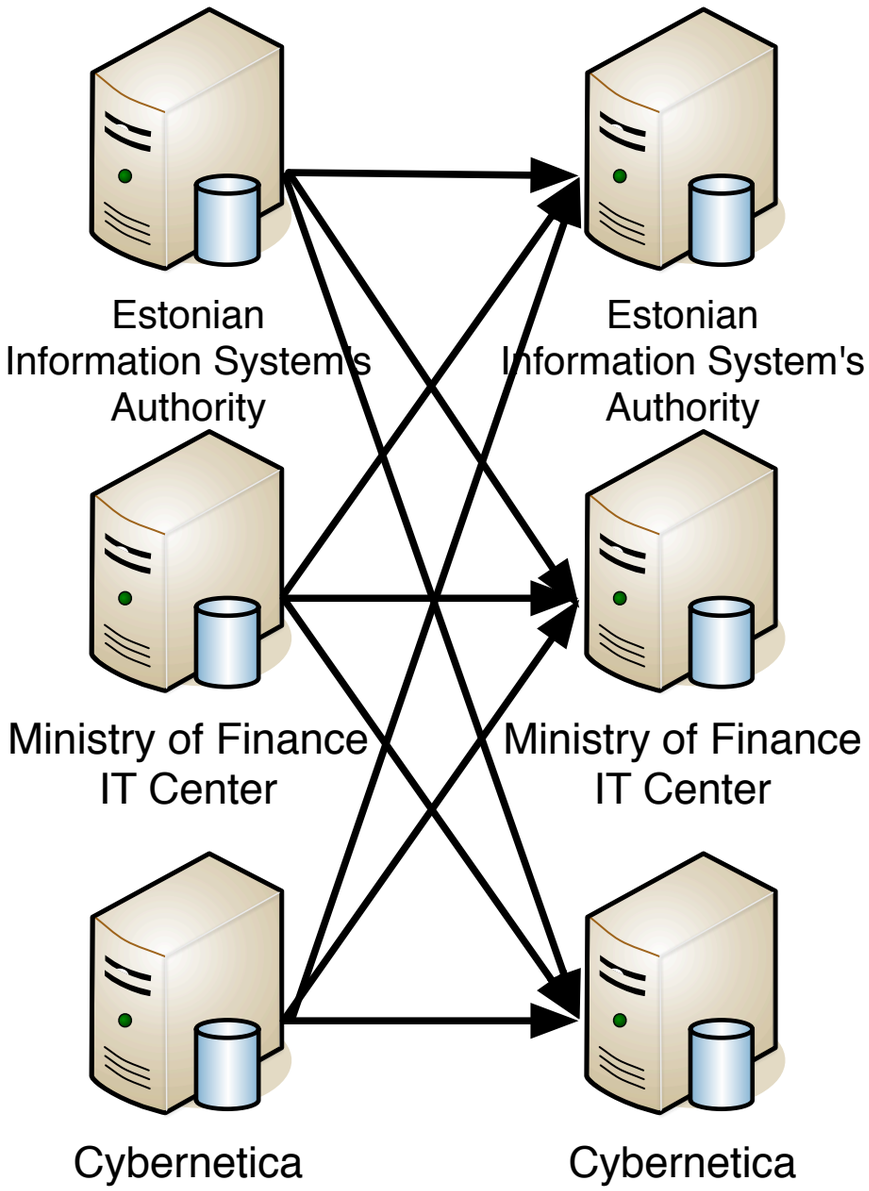
# Step 1: Import the data




- Sharemind importer tool loads CSV files.
- Each value is secret-shared at the source
- Private data never leaves the organisation.
- **800 000** study records.
- **20 million** tax records.
- **Largest secure MPC app ever.**

sharemind

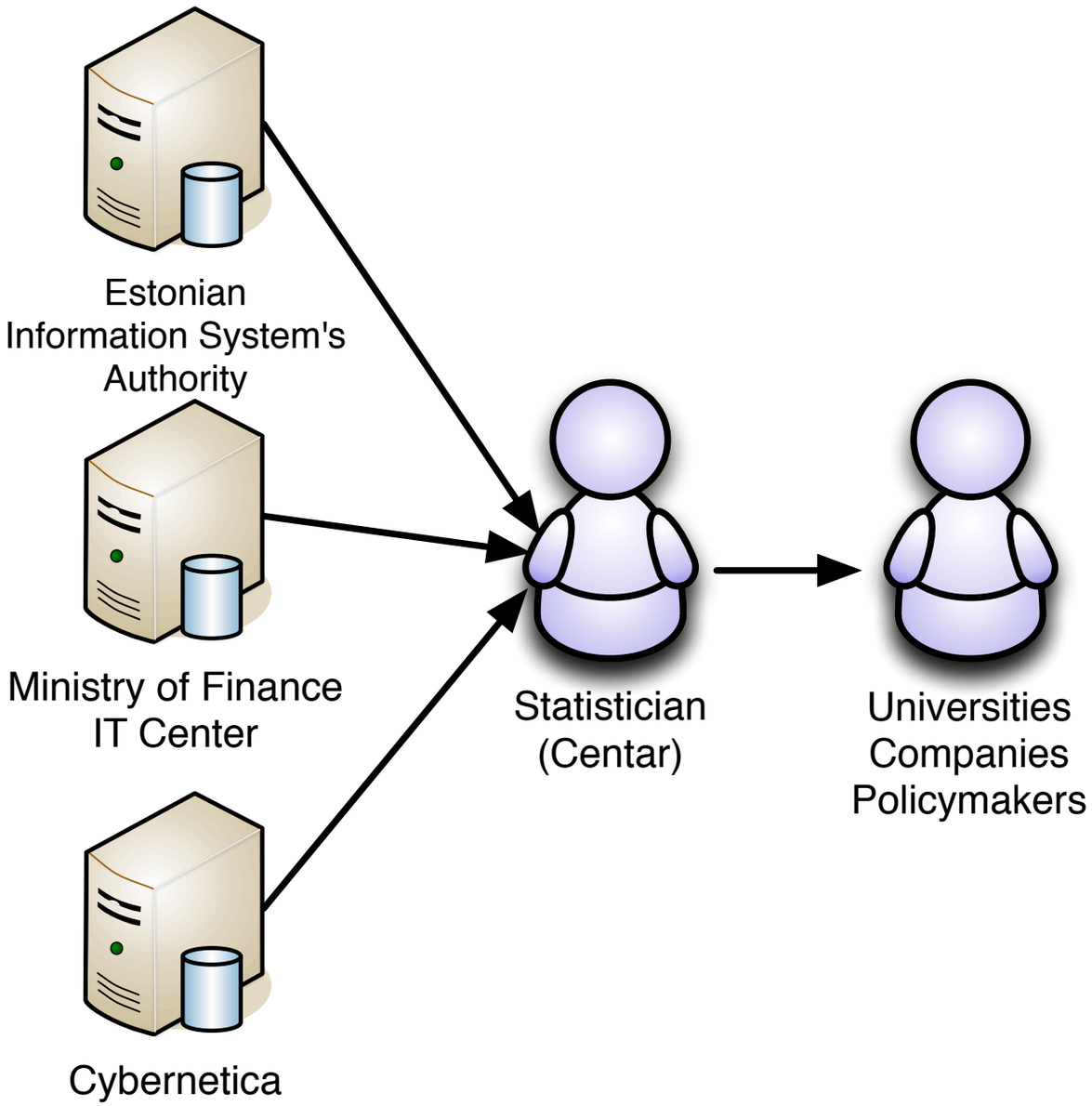
# Step 2: Run the analysis




- We implement data processing algorithms in a C-like language that uses privacy types.
- The Sharemind virtual machine automatically uses secure operations on private data.
- **Sharemind processes secret-shared inputs without reconstructing.**

sharemind

# Step 3: Publish the results




- The analyst uses an R-like analysis tool to perform queries.
- Sharemind hosts ensure that only queries in the study plan are actually executed.
- **The analyst cannot post arbitrary queries that all hosts do not agree to.**

sharemind

# Data protection? Check.

**Problem**: the Ministry of Education and Tax Board can't just share Personally Identifiable Information.

**What we did**: we described the private data flow and the use of encryption to the national DPA.

**January 2014**: The DPA responded that we don't need to apply for any special permissions,  as we are not processing personal information.

sharemind

# Tax secrecy? Covered.

**Problem**: the Taxation Act is an extra restriction.

**What we did**: we set up an installation of the Sharemind tools and reviewed it (and the source code) jointly with the Tax Board people.

**January 2015**: the internal oversight people in the Tax Board agreed to upload actual income tax records into the Sharemind-based analytics system.

sharemind

# Secure multi-party contracting

**Problem**: even with legal hurdles removed, parties asked for agreements to formalise roles and responsibilities.

**What we did**: we drafted agreements between Sharemind hosts and data owners, following the security model.

**Next few weeks**: the Tax Board, Ministry of Education, Information Systems Authority, Ministry of Finance IT Center and Cybernetica will sign the world's first secure multi-party data analysis agreement.

sharemind

# Take-home messages

1. Secure multi-party computation is mature enough to be used for statistically analysing personal data.
2. We are setting a precedent that this is a legal thing to do. Our end users agree on this.
3. We will still need agreements between entities, but the responsibilities are reduced, as technology enforces privacy guarantees.

sharemind

sharemind

https://sharemind.cyber.ee/

sharemind@cyber.ee