



TARTU ÜLIKOOL

Privaatsustehnoloogiad

Õigusloomejuristide koolitus 2019



Euroopa Liit
Euroopa Sotsiaalfond



Eesti
tuleviku heaks

Kristjan Krips, Riivo Talviste

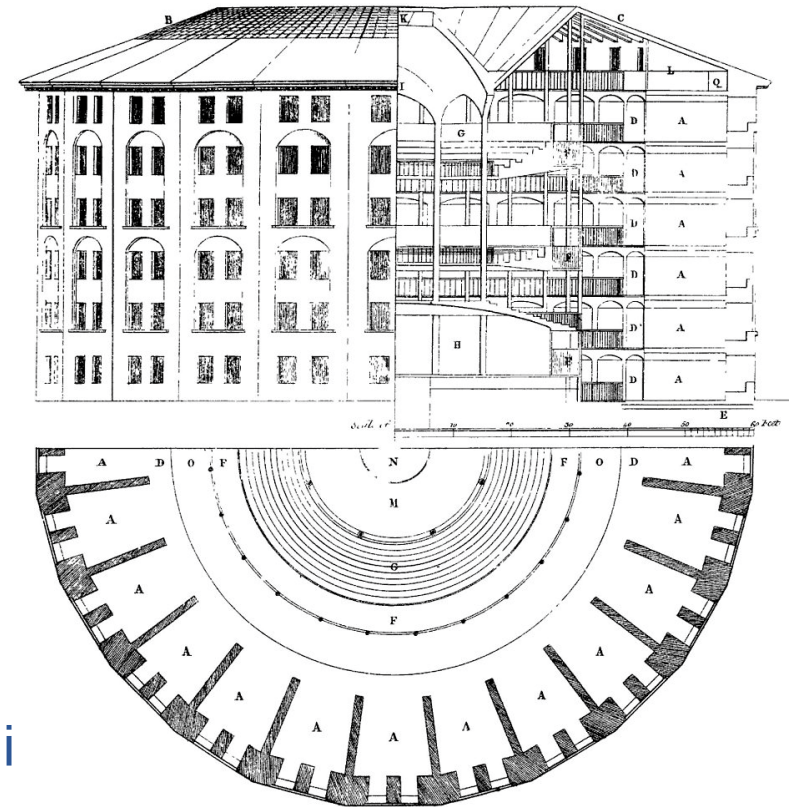
06.06.2019

Privaatsus ja terminoloogia

- Mis on privaatsus?
 - Täpsem tähendus sõltub keskkonnast
 - Üldiselt tähendab privaatsus võimalust piirata enda kohta käivat informatsiooni kolmandate osapoolte eest
- Terminoloogia
 - privaatsus vs anonüümsus
 - andmed vs info
 - konfidentsiaalsus
 - krüptograafia ja krüpteerimine

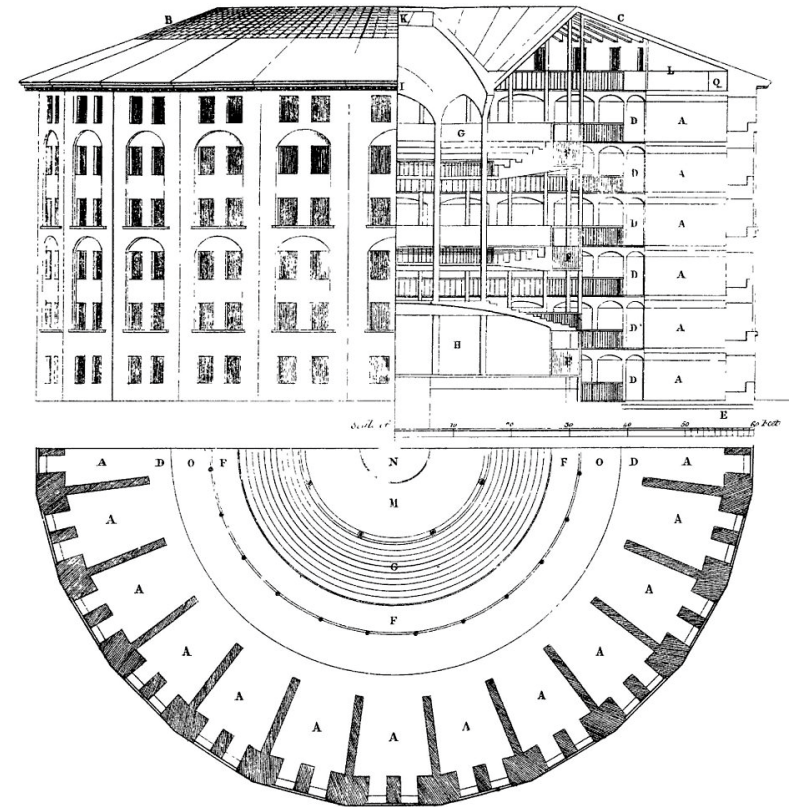
Kas on vaja privaatsust?

- “Mul pole midagi varjata”
 - vabadus vs kontroll
- Ühiskond
 - Sõnavabadus ja demokraatia vs autoritaarne / ebademokraatlik valitsemine
 - Korruptsiooni ja ebaseadusliku tegevuse paljastamine
- Ettevõtted
 - GDPR trahvid võivad tekitada motivatsiooni



Kuidas tagada privaatsust?

- Üksikisiku tasand
 - Saab õpetada privaatsust kaitsma
 - Ei skaleeru
 - Koolitamine on kallis
- Teenuste tasand
 - Rakendada privaatsustehnoloogiaid
 - Skaleerub
 - Tehnoloogia kasutuselevõtt võib olla kallis



Privaatsustehnoloogiad

Privacy Enhancing Technologies (PET)

- (Krüptograafiliselt tõestatud) Tehnilised meetmed andmesubjekti privaatsuse tagamiseks
- Proaktiivsed, preventatiivsed
 - Vastukaaluks lepingust tulenevatele reaktiivsetele meetmetele (nt sanktsioonid)
- Praktikas kasutatakse koos lepingutega
 - Muudavad lepingud kergekaalulisemaks
- Eeldavad eelnevat tehnilist seadistust

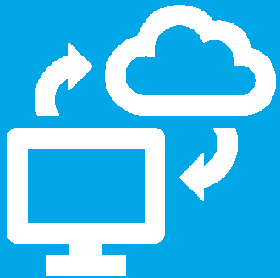
Kasutusvaldkonnad

- Andmete transport
 - Et võõrad ei saaks sõnumeid (sh metaandmeid) lugeda või muuta
- Tuvastamine
 - Et ei avalikustataks rohkem (delikaatseid) isikuandmeid kui vajalik
- Andmete töötlemine
 - Vältimaks ebavajalike andmete lekkimist nende töötlejale





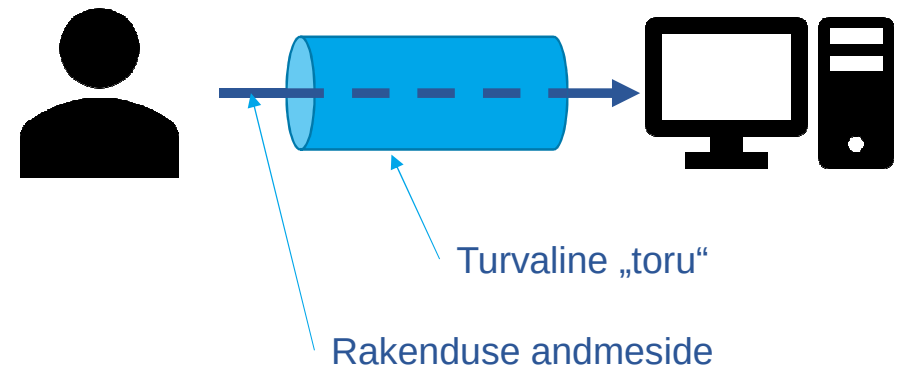
Privaatsustehnoloogiad andmete transpordil



Transport Layer Security (TLS)

- Pakub:
 - Serveri autentimist
 - Kliendi autentimist (valikuline)
 - Andmeside krüpteerimist
 - Andmeside terviklust
 - Tulevikusalastus (*Perfect Forward Secrecy*, valikuline)
- Lekib:
 - Suhtlevad osapooled
 - Side aeg ja maht

- Turvaline andmevahetus üle mitteusaldatud võrgu



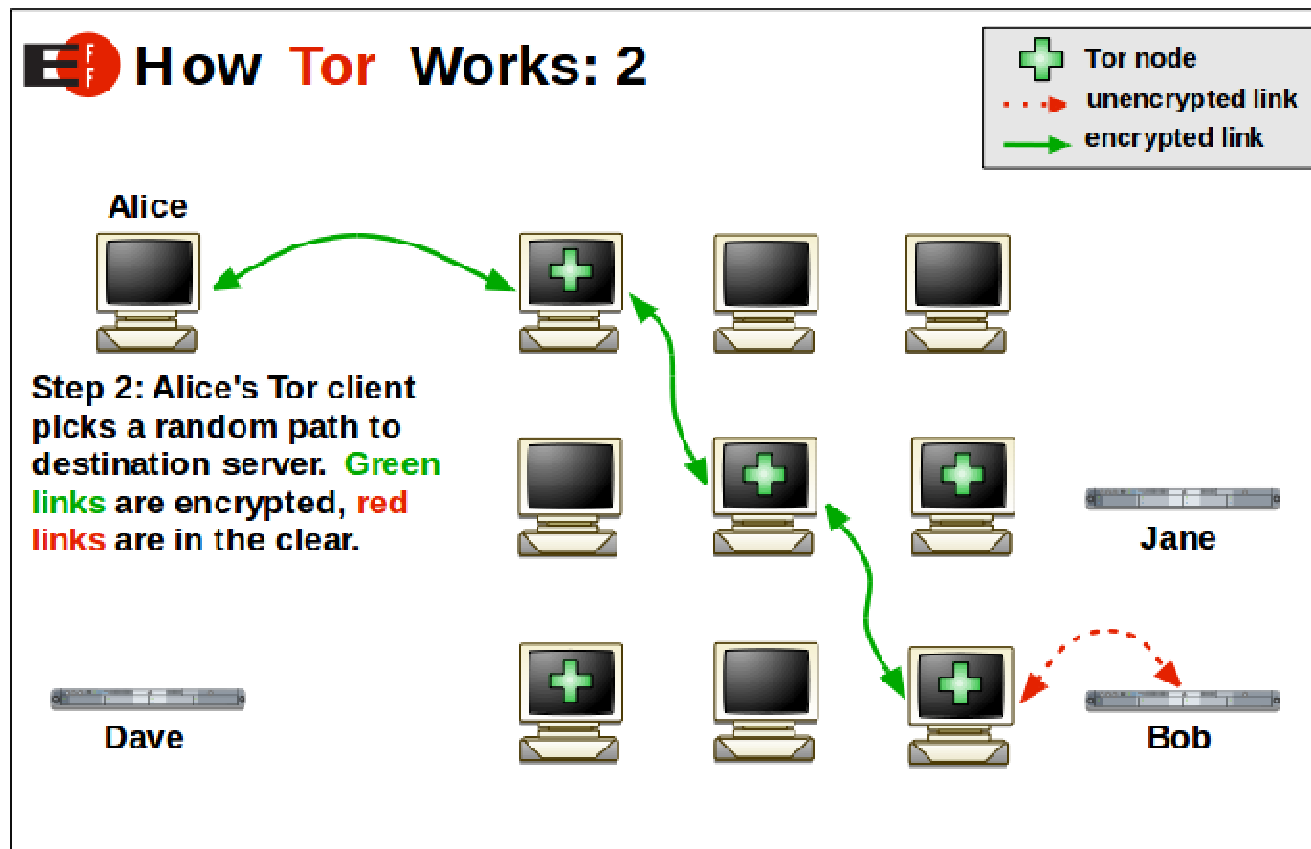
Transport Layer Security (TLS)

- Kasutatakse väga paljudes rakendustes:
 - Turvaline veebiliiklus (HTTPS)
 - Turvaline meilivahetus (IMAPS, SMTPS)
 - X-tee
 - Muud rakendusespetsiifilised protokollid
- Võiks kasutada alati, kui toimub andmevahetus üle interneti
 - Tänapäeval saab sertifikaate tasuta: *Let's Encrypt!*
- Tähelepanu tuleb pöörata ka õigele konfiguratsioonile

Anonüümne suhtlemine - Tor

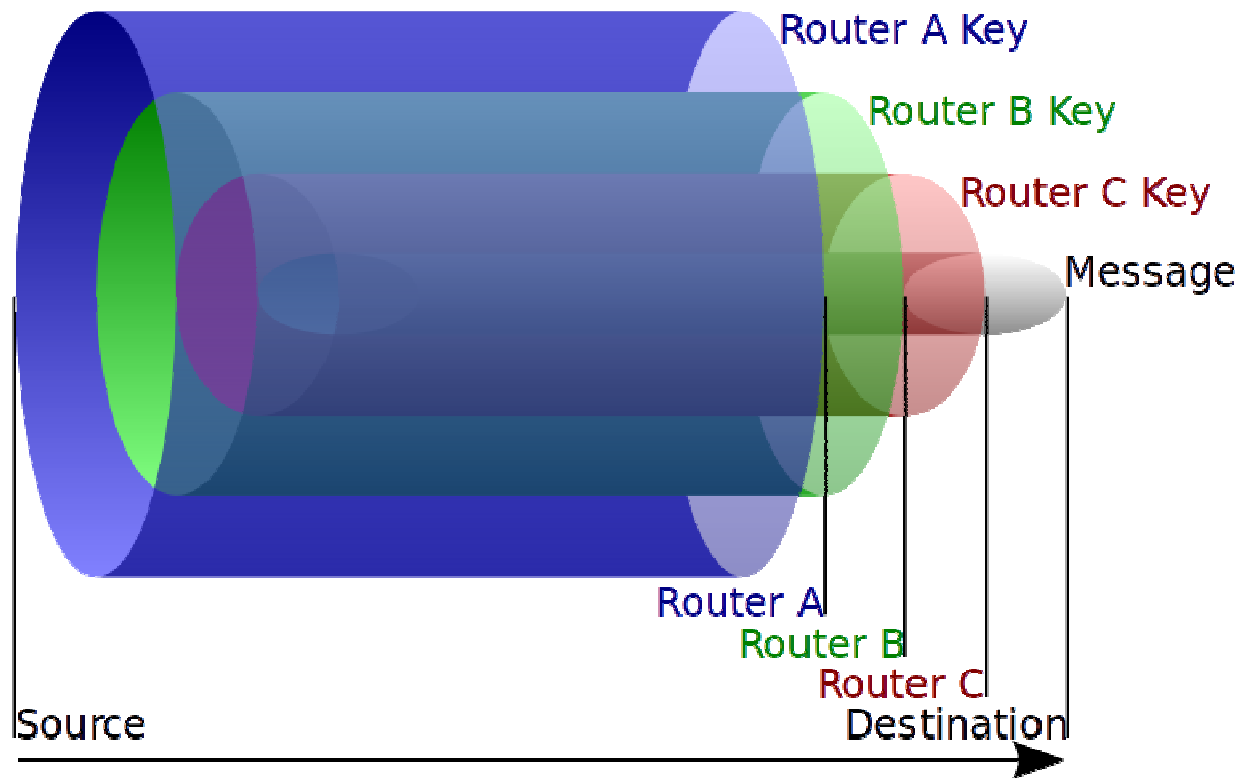
- Tor on süsteem, mis võimaldab internetis anonüümselt pakette ruutida ehk anonüümselt interneti kasutada
 - Peidab kasutaja asukohta (IP)
 - Pakub ka peidetud teenuseid, mille asukoht pole teada
- Peamised omadused:
 - Kasutab nn. sibulmarsruutimise protokoll
 - Põhineb avaliku võtme krüptograafial (nagu ka TLS)
 - Rahastatud osaliselt USA valituse poolt (nt. DARPA)
 - Kättesaadav kõige lihtsamalt Tor Browser-i kaudu, aga võimalik integreerida ka enda lahendustega
 - Kasutatakse sõnavabaduse ja privaatsuse kaitsmiseks
 - Kasutatakse ka kriminaalsete tegevuste jaoks

Anonüümne suhtlemine - Tor



- Iga osapool ahelas näeb ainult oma vahetuid naabreid
- Päring avalikku internetti tehakse viimase Tor osapoole (*exit node*) nimel

Tor - sibulmarsruutimine



Pilt: English Wikipedia user HANtwister (https://commons.wikimedia.org/wiki/File:Onion_diagram.svg), „Onion diagram“, CC-BY-SA-3.0

Otspunktkrüpteeritud suhtlus

End-to-end Encryption (E2EE)

- Side on krüpteeritud ühest otspunktist teise otspunktini
 - Kahe otspunkti vahel ei ole serverit, kes haldaks võtmeid ja saaks ühendust vajadusel pealt kuulata
- Laiatarberakendustes on kõige levinum Signali protokoll
 - Signal (originaalrakendus, mis kasutab Signali protokoll)
 - WhatsApp
 - Facebook Messenger (E2EE ei ole vaikimisi kasutusel)
 - Skype (E2EE ei ole vaikimisi kasutusel)



Privaatsustehnoloogiad tuvastamisel



Privaatne vastavuskontroll

- Nullteadmustõestuse (*zero-knowledge proof*) abil on võimalik kontrollida reeglitele vastavust ilma selleks sisendandmeid jagamata
 - Nt kuulumine mingisse nimekirja
 - Nt täisealisuse kontroll ilma vanust avaldamata
- Vajab subjektilt (või tema identiteedipakkujalt) ja verifitseerijalt eelnevalt tehniliste parameetrite kokkuleppimist

Privaatne grupikuuluvus

- Grupisignatuurid (Chaum, van Heyst - EuroCrypt 1991)
 - Vaja on nn. grupi haldajat, kes määrab ära liikmed ja jagab võtmed
 - Anonüümsus sõltub grupi haldajast ja seda saab eemaldada
 - Näited: TPM, VSC (vehicle safety communication)
- Ringisignatuurid (Rivest, Shamir, Tauman - ASIACRYPT 2001)
 - Allkirjastada saab ringi nimel suvaline ringi kuuluv liige
 - Ei saa tuvastada, kes liikmetest allkirjastas
 - Liikmete ringi saab määrata / valida allkirjastamise hetkel
 - Näited: krüptoraha, e-hääletamine, info anonüümne lekitamine - klassikaline näide on nn. “high ranking White House official”.



Privaatsustehnoloogiad andmete töötlemisel

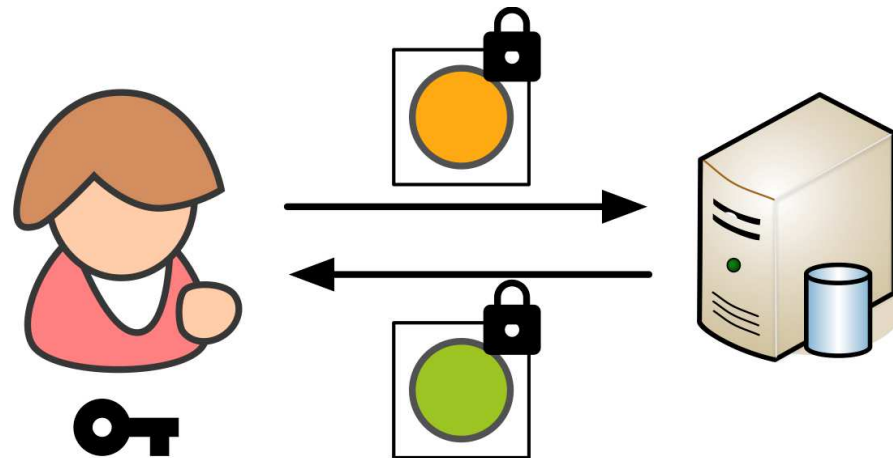


Anonüümimine ja pseudonüümimine

- Kasutatakse andmekogude jagamisel või avalikustamisel, et vältida liigsete (delikaatsete) isikuandmete lekkimist
- Lihtne ja arusaadav lähenemine
- Anonüümimine muudab andmekogude linkimise võimaluks, pseudonüümimise puhul tuleb seda teha andmekogude vahel koordineeritult (st nõuab kokkuleppeid)
 - Räsimine ei ole hea lahendus!
- Anonüümimisel analüüsitulemuste täpsus langeb
- Kaudsete tunnuste, mustrite või väliste andmekogude kaudu võib olla võimalik andmesubjekte isikustada

Homomorfne krüptograafia

- Krüpteerimise homomorfsus – krüptogrammidega tehteid tehes sooritatakse tehe ka krüpteeritud andmetel
 - Seejuures ei ole vaja andmeid lahti krüpteerida, st pole vaja võtit
- Täishomomorfne krüptograafia (*fully homomorphic encryption*, FHE) on pilvearvutuse Püha Graal



Homomorfne krüptograafia

- Üldine täishomomorfne krüptograafia on veel ebapraktiline, kuna praegustes skeemides tuleb „krüptoteksti krüpteeritult krüpteerida“
- Teatud arvutusülesannete jaoks on tehnoloogia aga juba praktiline
 - Nt akadeemilised teostused rämpspostifiltrist ja krüpteeritud VoIP kõnedest

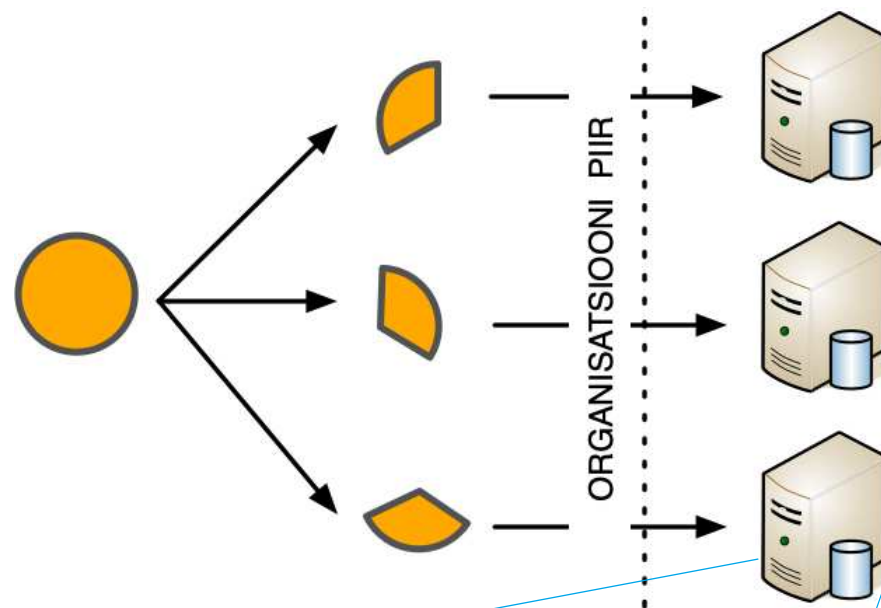
Turvaline ühisarvutus

Secure Multi-party Computation (MPC)

- Mitu **sõltumatut** osapoolt arvutavad üheskoos kõigi sisendandmetel mingit funktsiooni väärtust, kusjuures ükski osapool ei saa selle käigus teada teise sisendväärtusi ning arvutuse tulemus on nähtav vaid autoriseeritud osapool(t)ele

Turvaline ühisarvutus

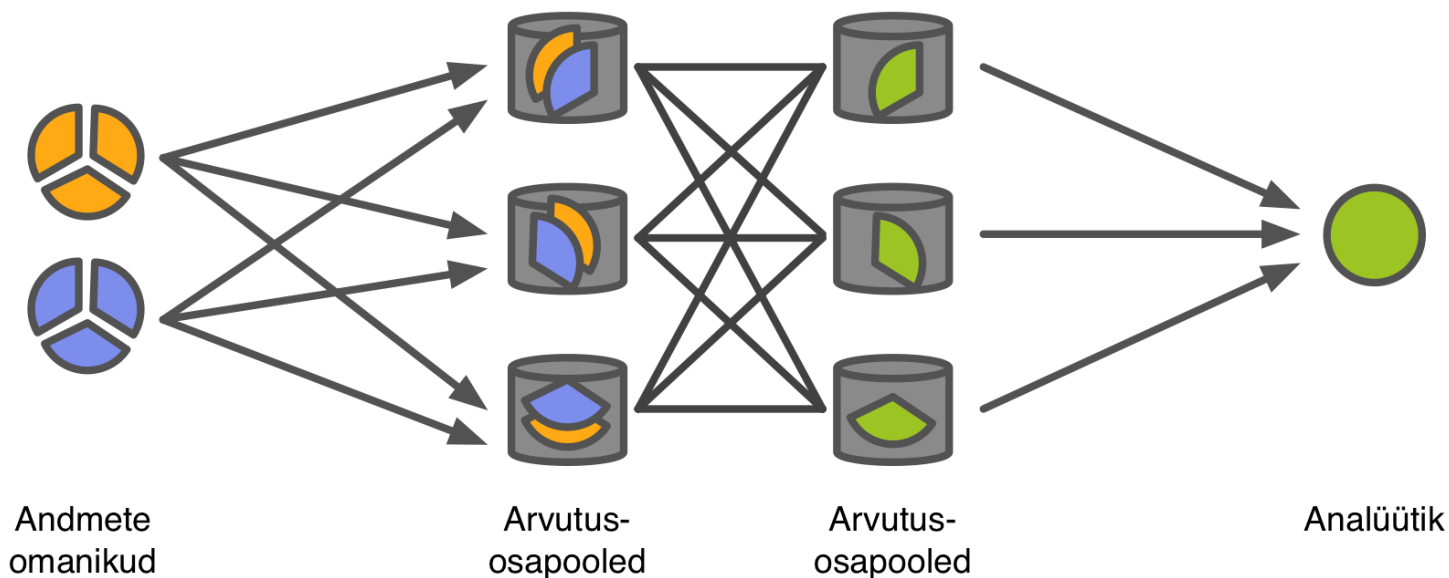
- Enne arvutamist *ühissalastavad* andmeomanikud ja jagavad need arvutajate vahel
 - Ühissalastamise käigus muudetakse salajased andmed juhuslikeks tükkideks
- Üks arvutaja näeb lokaalselt andmemudelit, aga mitte väärtuseid



#	Eesnimi	Perenimi	Sugu	Vanus
1	?	?	?	?
2	?	?	?	?
3	?	?	?	?
4	?	?	?	?

Turvaline ühisarvutus

- Arvutuse ajal ühissalastatud andmeid ei rekonstrueerita
 - Arvutajad väärtusi teada ei saa
- Arvutuse tulemus on ka (enamasti) ühissalastatud ja nähtav vaid autoriseeritud osapoolle



Turvaline ühisarvutus

- Eriti sobiv andmehulkade linkimiseks ja statistiliseks analüüsiks, kus andmed tulevad erinevatelt osapooltelt
 - Algandmeid pole tarvis eelagregeerida ega anonüümida, tänu millele on analüüsitulemused täpsed
- On edukalt rakendatud praktikas ning leidub nii akadeemilisi kui ka kommertstooteid.

Turvalised käivituskeskkonnad

Trusted Execution Environment (TEE)

- Arvuti protsessori (CPU) sees olevad turvalised alamprotsessorid (enklaavid), mille sisse teised tavaprotsessid ei näe
 - Mälu ja kettale kirjutatavad andmed on igal ajal krüpteeritud, võti on enklaavi sees
 - Võrgusuhtlus käib läbi vahendava tavaprotsessi, kuid see võrgusõnumeid ei näe
- Tooted: Intel Software Guard Extensions (SGX), ARM TrustZone, iPhone Secure Element
- Juurutustes tuleb mingil määral protsessoritootjat usaldada
- Peamiseks ohuks on kõrvalkanaliründed

Diferentsiaalprivaatsus

- Arvutustulemusele lisatakse täpselt niipalju juhuslikku müra, et ei oleks võimalik tuvastada ühegi andmesubjekti kuulumist või mittekuulumist antud andmekogusse
 - Müra suurus ei sõltu andmekogust vaid analüüsist (arvutatavast funktsioonist)
- Parameeter ϵ määrab ära maksimaalse avalikustatavate tulemuste arvu, rohkem arvutades hakkab avalikustatav väärtus statistiliselt koonduma „õige“ väärtuse ümber.

Diferentsiaalprivaatsus

- Diferentsiaalprivaatsust (*differential privacy*) ei rakendata mitte andmekogule vaid sellel tehtud arvutustulemustele
 - Kasutatav koos eelpool mainitud tehnoloogiatega: homomorfne krüptograafia, turvaline ühisarvutus ja turvalise käivituskeskkonnad

