



Centrālā statistikas
pārvalde

Statistiskā konfidencialitāte un statistikas (datu) pieejamība

Uldis Ainārs

Informācijas, izdevniecības un poligrāfijas departamenta direktors



Centrālā statistikas
pārvalde

Lai runātu vienotā valodā

Dati?

Statistika?

Statistiskā konfidencialitāte?



Centrālā statistikas
pārvalde

Saturs

1. Kas ir statistiskā konfidencialitāte

- Kāpēc
- Ko nosaka tiesību normas
- Kā izpaužas

2. Konfidenciālu datu izmantošana

- Kāpēc nepieciešams
- Galvenie principi
- Svarīgākie projekti



Centrālā statistikas
pārvalde

Statistiskās konfidencialitāte nozīme

Abpusēja uzticēšanās
(atbildētības līmenis, atbilžu patiesums)



Nostiprināts tiesību normās
(izmantošana vienīgi statistiskiem
mērķiem)



Statistikas kvalitāte



Centrālā statistikas
pārvalde

Svarīgākās tiesību normas

ANO, Oficiālās statistikas pamatprincipi

- *2013/21. Fundamental Principles of Official Statistics.*
- *6. princips – individuālie dati var tikt izmantoti tikai statistiskiem mērķiem*

Regula 223/2009 Par Eiropas statistiku

- *20. pants – Statistikas konfidencialitāte*

Statistikas likums

- *Statistikas iestādes pienākums nodrošināt konfidencialitāti (7.p.)*
- *Datu apstrāde un statistiskā konfidencialitāte (17.pants.)*
- *Datu pieejamības vispārējie noteikumi (25.pants)*
- *Datu izmantošana pētniecībā (26. pants)*



Centrālā statistikas
pārvalde

Datu pieejamības vispārējie noteikumi, 25.pants

- (1)(2) Datu apmaiņa statistikas nodrošināšanas vajadzībām
- (3) Rakstveida atļauja konfidenciālo datu publicēšanai
- (4) Datu izmantošana pētnieciskiem mērķiem
- (5) Izņēmumi attiecībā uz privātpersonu datiem (rekvizīti, nodarbināto skaits, izstrādājumu veidi, emisijas vidē)
- (6) Izņēmumi attiecībā uz valsts institūcijām (viss kas neskar privātpersonu vai ierobežotas pieejamības informāciju)



Centrālā statistikas
pārvalde

Datu aizsardzība

Lai arī Likums pasaka, ka ir jānodrošina konfidencialitāte, izpaušanas kontrole ir statistikas iestādes ziņā.

- Tiešā identificēšana – identificēša pēc tiešajiem identifikatoriem (vārds, uzvārds, adrese, nosaukums, personas kods)
- **Netiešā identificēšana – identificēšana izmantojot raksturojošu informāciju (vienīgais uzņēmums nozarē)**

Jo detalizētāki dati tiek publicēti, jo sarežģītāk ir aizsargāt no netiešās identificēšanas



Centrālā statistikas
pārvalde

Izpaušanas kontroles metodes

- **Pirmsatabulārās (tiek veiktas izmaiņas ierakstu līmenī) CSP izmanto tikai anonimizētu datu nodošanā.**
- **Tabulu pārveidošana (dažādi grupējumi un apvienošana).**
- **Pēctabulārās metodes:**
 - **Šūnu vērtību noklusēšana,**
 - **Vērtību maiņa (noapaļošana, troksnis).**
- **Organizatoriski pasākumi:**
 - **Līgums, attālinātā piekļuve.**
- **Ģeotelpiskās informācijas aizsardzības metodes**

Jebkuras aizsardzības metodes izmantošana rada datu/statistikas analītiskā potenciāla zudumu.

Statistiķa uzdevums ir nodrošināt minimālu datu zudumu, t.i. būtiski nemazināt statistikas/datu izmantošanas iespējas.



Centrālā statistikas
pārvalde

Izpaušanas kontroles metodes tabulārajā statistikā

Farmācijas nozare, dominances kritēriji

RU18. APGROZĪJUMA PĀRMAIŅAS NOZARĒS PA MĒNEŠIEM

	2016M01	2016M02	2016M03	2016M04
% pret iepriekšējo periodu, sezonāli izlīdzināti dati				
(21) farmaceitisko pamatvielu un farmaceitisko preparātu ražošana				
PAVISAM
Vietējais tirgus
Eksports
(31) mēbeļu ražošana				
PAVISAM	-4.4	6.4	2.9	-2.0
Vietējais tirgus	-16.3	13.5	6.7	-6.2
Eksports	13.6	-1.6	-2.1	3.9



Centrālā statistikas
pārvalde

Statistiskās konfidencialitātes nodrošināšanas principi

1. Jāsaprot vai nepieciešama
«identificēšanas aizsardzība».
 - a. viens tūrists no Kenijas ir pavadījis piecas
naktis viesnīcā Kolkas novadā?
 - b. 90% tūristu Kolkas novadā ir no Kenijas?

Jāaizsargā no informācijas iegūšanas par
respondentu, to var noteikt vienīgi veicot
«ielaušanās» [penetration] testus.

2. Jāpielieto izpaušanas kontroles metodes
3. Rezultātam jāveic «ielaušanās» testi.



Centrālā statistikas
pārvalde

Believe me, it is not fun having your disclosure control decisions examine in the High Court. This happened to us, **when the court examined our decision not to publish some small numbers in abortion statistics.** We were asked "why have you **not suppressed the number 9**, but you have **suppressed the number 4**?" We answered that we followed policy and that this is what the methodology concluded. The court then asked "but what **evidence do you have that your policy and methods were correct and therefore that a 9 really is an acceptable risk and a 4 really is an unacceptable risk?**" ... and at that point, we knew we had lost the case. We had no evidence.



Centrālā statistikas
pārvalde

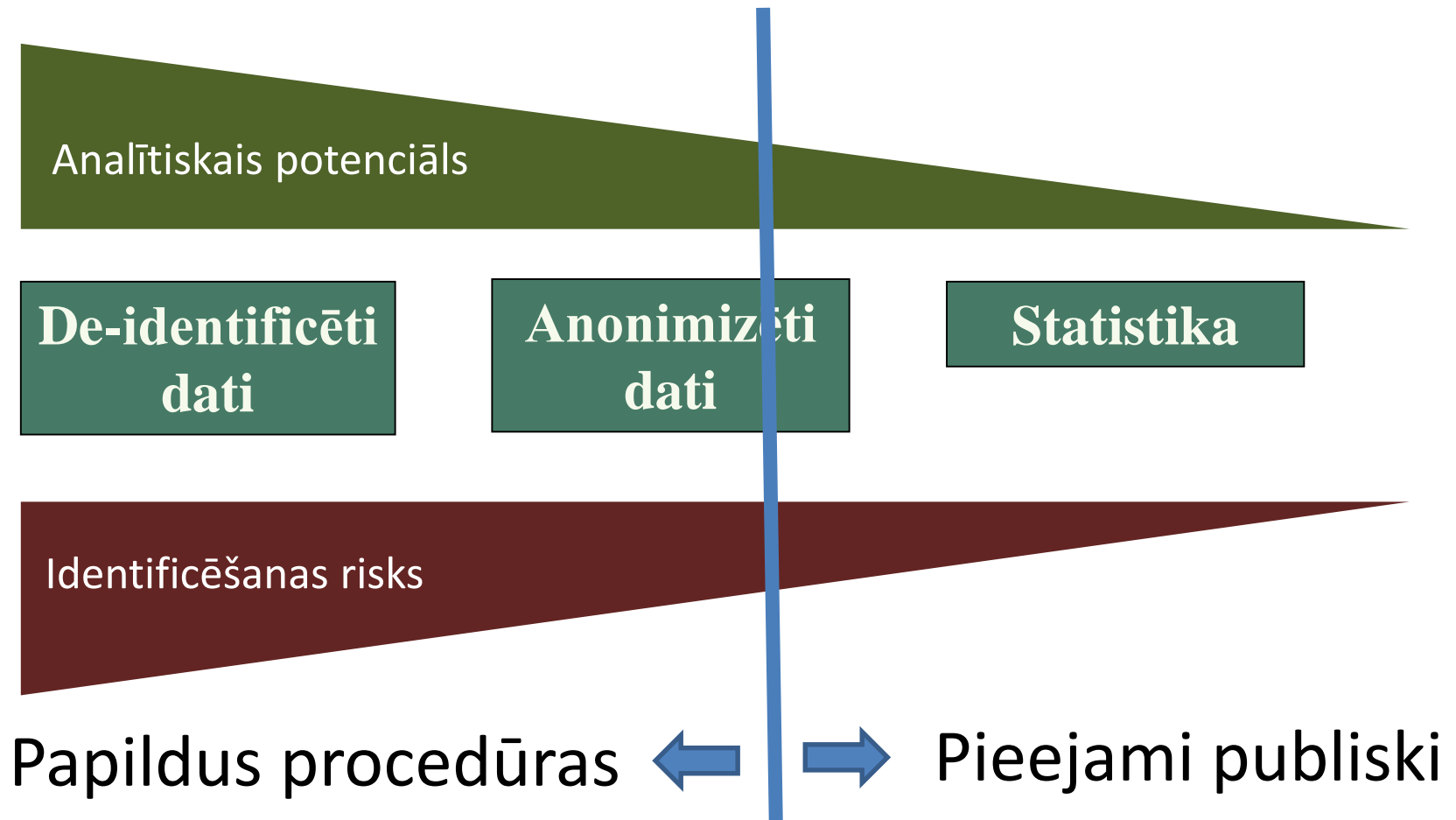
Izpaūšanas kontroles metodes publiskās anonimizētās datnēs (datiem nevis statistikai)

PY100N	VECUMA PENSIJAS (neto) (anonimizēts (0.5% no lielākajām vecuma pensijām ir pielīdzinātas vecuma pensiju 99.5-tai procentīlei))	1 - 999999.99	Ienākumi
		0	Nav ienākumu
reg	LATVIJAS STATISTISKIE REĢIONI	1	Rīga
		2	Pierīga
		3	Vidzeme
		4	Kurzeme
		5	Zemgale
		6	Latgale
laupil	TERITORIJA	1	Pilsētas
		2	Lauki
vec_gr	VECUMA GRUPA	1	0 – 15 gadi
		2	16 – 25 gadi
		3	26 – 35 gadi
		4	36 – 45 gadi
		5	46 – 55 gadi
		6	56 – 65 gadi
		7	66 – 75 gadi
		8	76 un vairāk gadi



Centrālā statistikas
pārvalde

Konfidenciālu datu izmantošana. Kāpēc nepieciešams





Centrālā statistikas
pārvalde

Galvenie principi

Sabiedrības ieguvumi

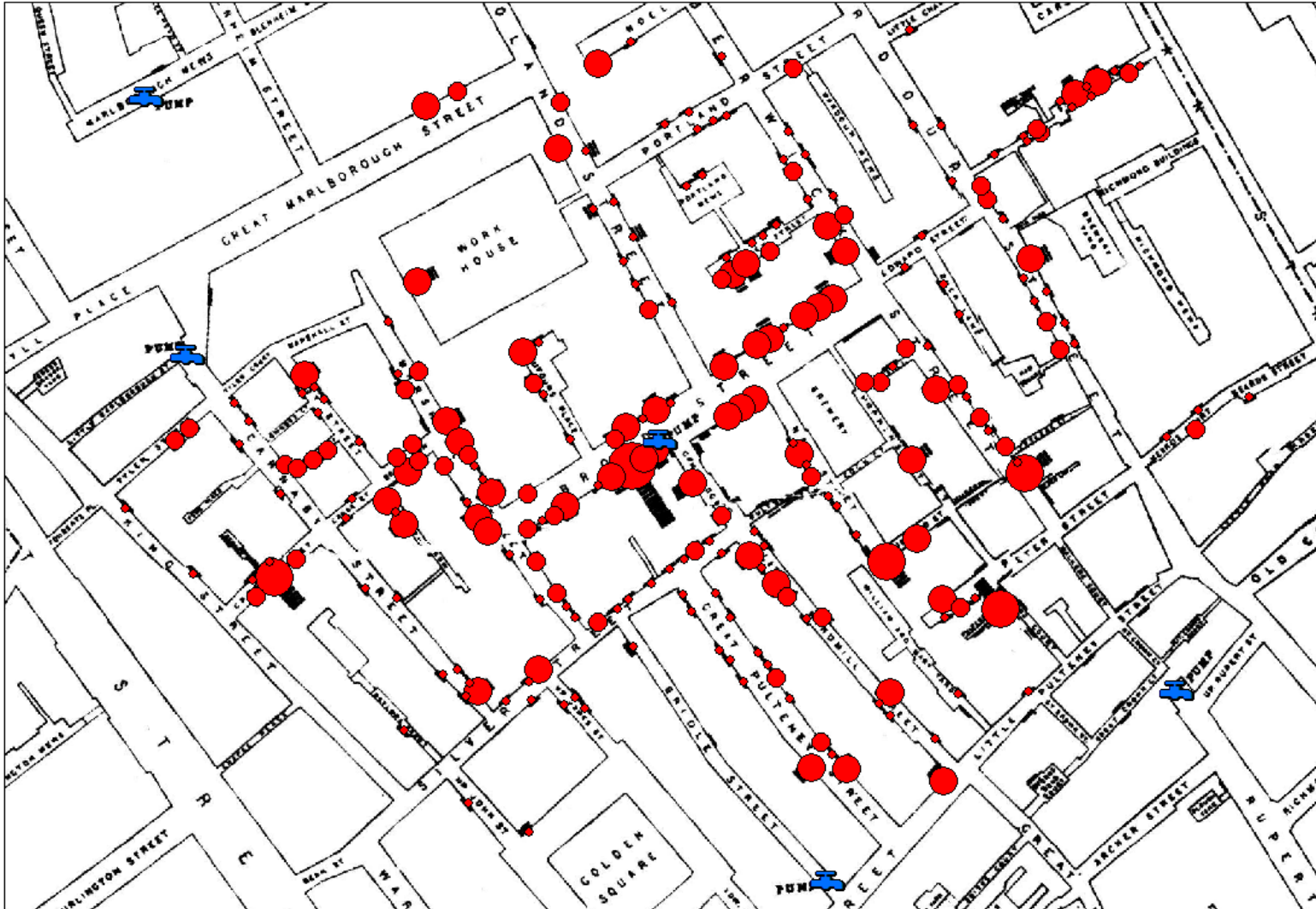


Privātuma/konfidenc.
apdraudēj.



Centrālā statistikas
pārvalde

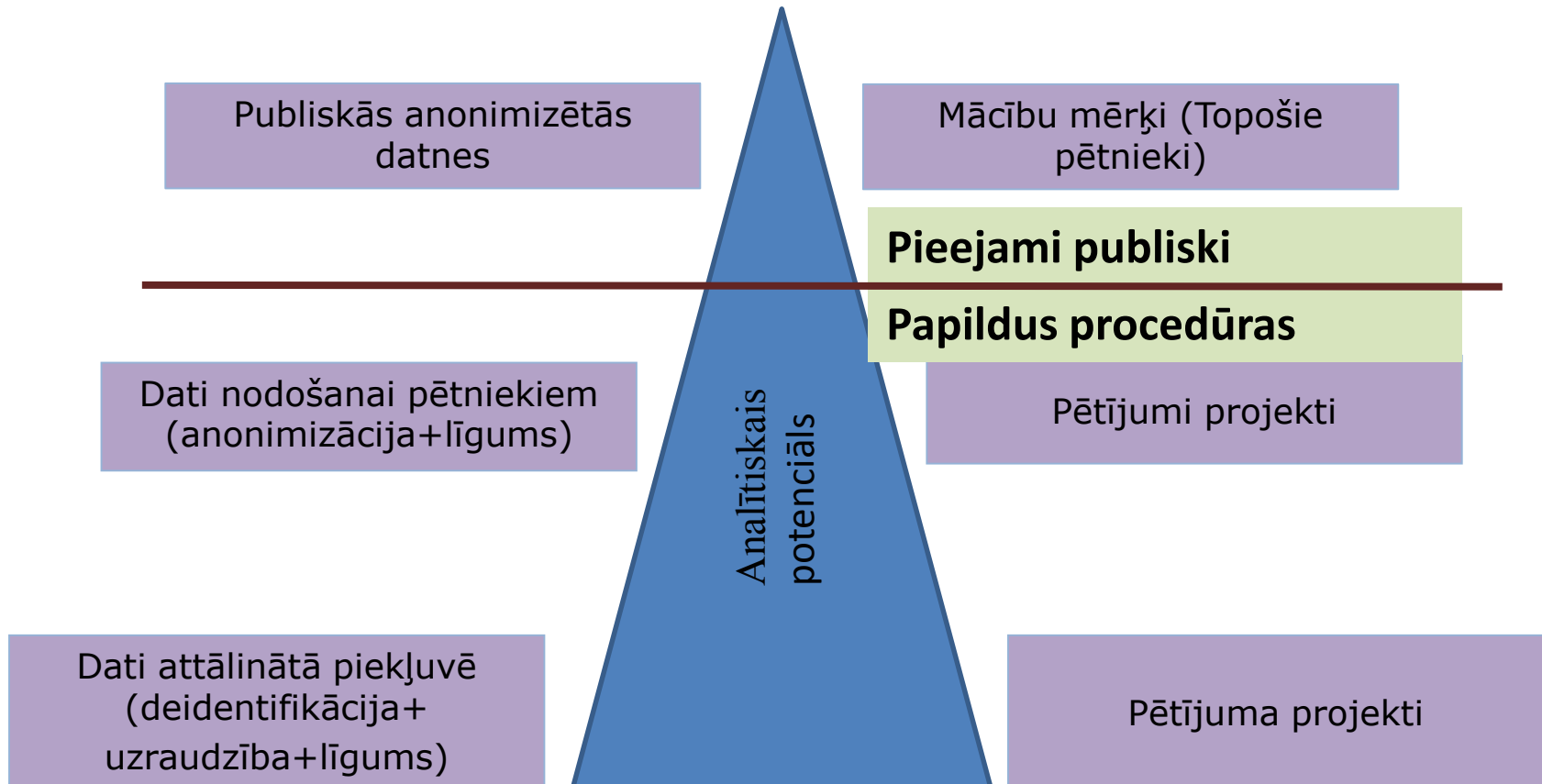
Holeras izplatības analīze. Londona 1854. gads.





Centrālā statistikas
pārvalde

CSP prakse datu nodošanā pētniecības vajadzībām





Centrālā statistikas
pārvalde

Nozīmīgākie pētījumu, kuros tiek izmantoti CSP sagatavotie dati

Latvijas banka

- Mērķis: Latvijas uzņēmumu konkurētspējas pētījums,
- [Uzņēmumu finanses, ārējā tirdzniecība darbinieku atalgojums.](#)
- VID dati par personu atalgojumu

Pasaules Banka (*pasūtītājs* *Veselības ministrija*)

- Mērķis: Veselības aprūpes sistēmas pilnveidošana,
- TS2011 dati, apvienoti ar NVD datiem.

Pasaules Banka (*pasūtītājs* *Finanšu ministrija*)

- Mērķis: Latvijas nodokļu sistēmas izvērtējums,
- EU SILC, LFS, HBS + VID ienākumi.

Latvijas universitāte (Ģeogrāfi)

- Mērķis: Mobilitātes un migrācijas pētījumi,
- [TS2000, TS2011 \(ģeoreferencēti dati\).](#)

BICEPS

- Mērķis: Valsts pētījumu programmas ietvaros, dažādi pētījumi
- [Uzņēmumu finanses, ārējā tirdzniecība darbinieku atalgojums.](#)
- VID dati par personu atalgojumu



Centrālā statistikas
pārvalde

Uldis.Ainars@csb.gov.lv

67366920