

# Andmete kvaliteet ja selle olulisus uuringutel Sälli näitel

Meri Liis Treimann

**Andmeteadlane**

1. Miks on andmete kvaliteet oluline?
2. Millised on kvaliteetsed andmed?
3. Milliseid takistusi tekitavad madala kvaliteediga andmed?

# Miks on andmete kvaliteet oluline?

$$Q_{\text{mudel}} \leq Q_{\text{andmed}}$$



Millised on  
kvaliteetsed  
andmed?



# Millised on kvaliteetsed andmed?

## 1. Võrreldavad

Olemus	Pikkus	Laius	Kõrgus	Paksus	Sügavus
raamat	20 cm	15 cm	2 cm	200 lk	
raamat	20 cm	15 cm		2 cm	
raamat		15 cm	20 cm	2 cm	

...

laud		160 cm	70 cm		80 cm
riiul					
vaas	5 cm	10 cm	30 cm	1 cm	

# Millised on kvaliteetsed andmed?

## 1. Võrreldavad

3 erinevat kahjustust



Olemus	Kahjustused
raamat	niiskuslaigud
raamat	niiskuskahjustus
raamat	niiskuse <u>l</u> kahjustus

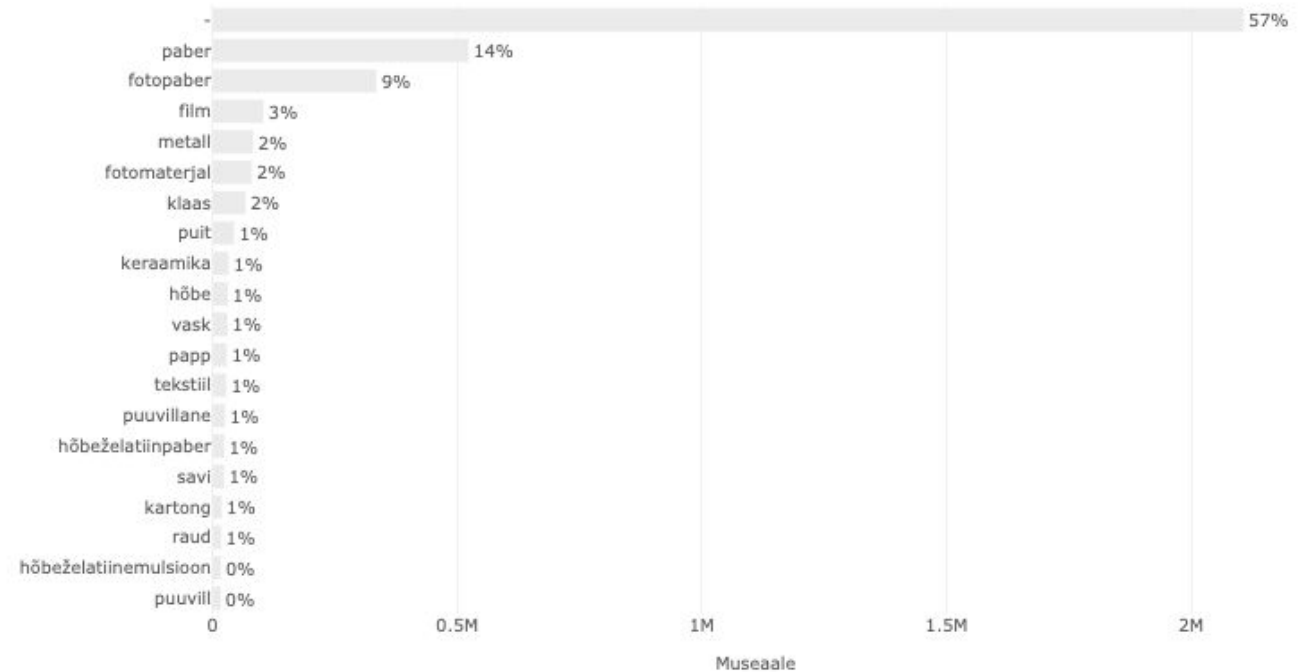
# Millised on kvaliteetsed andmed?

1. Võrreldavad

**2. Täielikud**

## Populaarseimad materjalid

Top 20 materjali



# Millised on kvaliteetsed andmed?

1. Võrreldavad

**2. Täielikud**

Olemus	Materjal	Tehnika	Dateering
dokument	paber	trükkimine, käsitsi kirjutamine	
			1925
mööbel			
foto	fotopaber	fotograafia	1954

**~20%  
kõik täidetud**



# Millised on kvaliteetsed andmed?

1. Võrreldavad

**2. Täielikud**

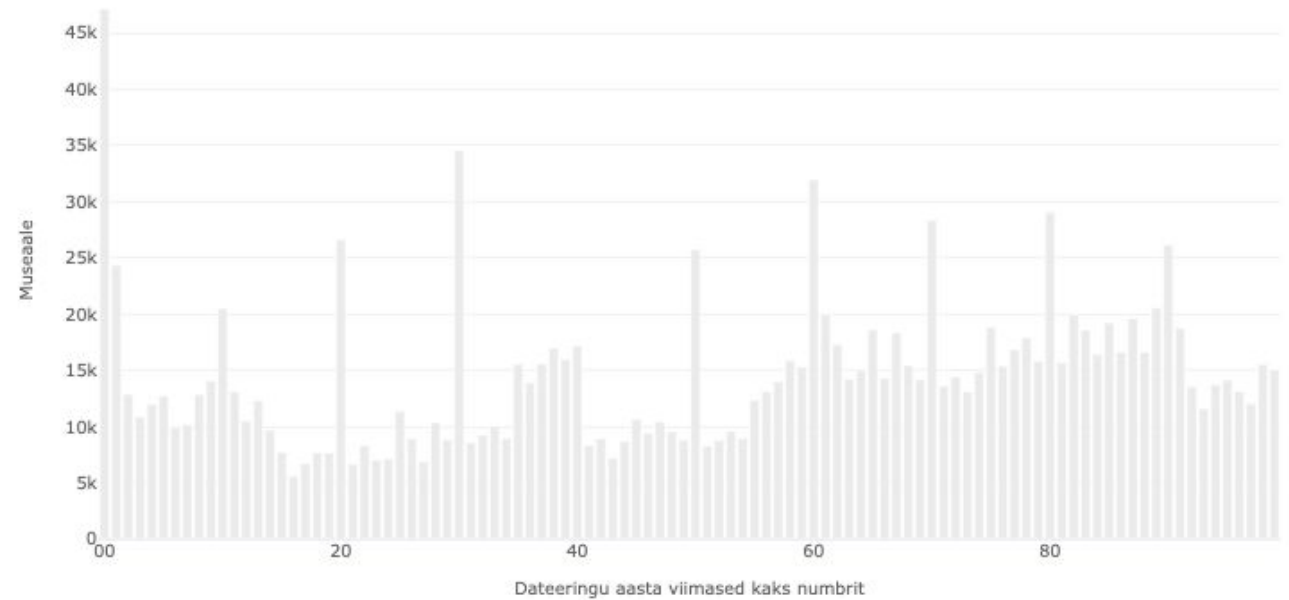
Näitus	Algus	Lõpp	Tagastatud
...	11.01.2012	11.01.2016	

**~12%**

**lõppenud näitustest tagastamise aeg puudu**

# Millised on kvaliteetsed andmed?

1. Võrreldavad
2. Täielikud
- 3. Täpsed**



# Millised on kvaliteetsed andmed?

1. Võrreldavad
2. Täielikud
3. Täpsed
- 4. Õiged**

<b>Dateeringu nimetus</b>	<b>Algus</b>	<b>Lõpp</b>
valmistamine	08.10.2209	
pildistamine	29.03.3000	
pildistamine	01.01.9418	
valmistamine	01.01.9500	01.01.1150
valmistamine	01.01.1995	01.01.1956

# Millised on kvaliteetsed andmed?

1. Võrreldavad
2. Täielikud
3. Täpsed
- 4. Õiged**

Olemus	Diameeter
maakera, ...	6400 km

# Millised on kvaliteetsed andmed?

1. Võrreldavad
2. Täielikud
3. Täpsed
4. Õiged
- 5. Ajakohased**

Olemus	Seisund	Aeg
paber	väga halb	04.10.2001
nahk, paber, vask	väga halb	30.01.2002
paber	väga halb	30.01.2002

**~50%**  
**museaali seisund vanem kui 5 aastat**

**~10%**  
**museaali seisund vanem kui 10 aastat**

# Millised on kvaliteetsed andmed?

1. Võrreldavad
2. Täielikud
3. Täpsed
4. Õiged
5. Ajakohased
- 6. Selged**

## kujutis/teema vs eseme tegelik olemus (säilivuse vaatenurgast)



Nimetus	Olemus	Materjal
Aruküla mõisa aiatöölised	kasvuhoone, hobune foto ←	fotopaber, tint
hobused saanidega	regi/saan, hobune foto ←	fotopaber
saan	regi/saan	puit
saan	regi/saan	

...

vankriratas	hobune, vanker/käru, ratas	puit
jalgratta number	jalgratas, silt/märk	metall

# Millised on kvaliteetsed andmed?

1. Võrreldavad
2. Täielikud
3. Täpsed
4. Õiged
5. Ajakohased
- 6. Selged**

mis sajandist?



Nimetus	Dateeringu nimetus	Dateering
Helga Amanda Kilk karjaaia peal istumas	pildistamine	01.01.20
Pitsat. Aadli. Kirjade pitseerimiseks XIX saj. Klaas, vask, puu.	valmistamine	01.01.19

Milliseid takistusi  
tekitavad madala  
kvaliteediga  
andmed?





# Milliseid takistusi tekitavad madala kvaliteediga andmed?

## 1. Andmete puhastamisel

- **Andmed ei ole võrreldavad: vii andmed ühele kujule**, nt
  - teisenda samale ühikule
  - ühenda samatähenduslikud väärtused

# Milliseid takistusi tekitavad madala kvaliteediga andmed?

## 1. Andmete puhastamisel

- Andmed ei ole võrreldavad: vii andmed ühele kujule, nt
  - teisenda samale ühikule
  - ühenda samatähenduslikud väärtused
- **Andmed ei ole täielikud: täida lüngad**
  - palju võimalusi
  - võib tekitada valesid andmeid

# Milliseid takistusi tekitavad madala kvaliteediga andmed?

## 1. Andmete puhastamisel

- Andmed ei ole võrreldavad: vii andmed ühele kujule, nt
  - teisenda samale ühikule
  - ühenda samatähenduslikud väärtused
- Andmed ei ole täielikud: täida lüngad
  - palju võimalusi
  - võib tekitada valesid andmeid
- **Andmed ei ole täpsed: ?**

# Milliseid takistusi tekitavad madala kvaliteediga andmed?

## 1. Andmete puhastamisel

- Andmed ei ole võrreldavad: vii andmed ühele kujule, nt
  - teisenda samale ühikule
  - ühenda samatähenduslikud väärtused
- Andmed ei ole täielikud: täida lüngad
  - palju võimalusi
  - võib tekitada valesid andmeid
- Andmed ei ole täpsed: ?
- **Andmed ei ole õiged: eemalda valed andmed**
  - nt andmed tulevikust
  - ei pruugi olla tuvastatav

# Milliseid takistusi tekitavad madala kvaliteediga andmed?

## 1. Andmete puhastamisel

- Andmed ei ole võrreldavad: vii andmed ühele kujule, nt
  - teisenda samale ühikule
  - ühenda samatähenduslikud väärtused
- Andmed ei ole täielikud: täida lüngad
  - palju võimalusi
  - võib tekitada valesid andmeid
- Andmed ei ole täpsed: ?
- Andmed ei ole õiged: eemalda valed andmed
  - nt andmed tulevikust
  - ei pruugi olla tuvastatav
- **Andmed ei ole ajakohased: ?**

# Milliseid takistusi tekitavad madala kvaliteediga andmed?

## 1. Andmete puhastamisel

- Andmed ei ole võrreldavad: vii andmed ühele kujule, nt
  - teisenda samale ühikule
  - ühenda samatähenduslikud väärtused
- Andmed ei ole täielikud: täida lüngad
  - palju võimalusi
  - võib tekitada valesid andmeid
- Andmed ei ole täpsed: ?
- Andmed ei ole õiged: eemalda valed andmed
  - nt andmed tulevikust
  - ei pruugi olla tuvastatav
- Andmed ei ole ajakohased: ?
- **Andmed ei ole selged: täpsusta andmeid**
  - mis reeglite alusel?

# Milliseid takistusi tekitavad madala kvaliteediga andmed?

1. Andmete puhastamisel

**2. Andmete ettevalmistamisel**

Materjal
savi
paber
puit, metall
-

→

savi	paber	puit	metall
1	0	0	0
0	1	0	0
0	0	1	1
0	0	0	0

# Milliseid takistusi tekitavad madala kvaliteediga andmed?

1. Andmete puhastamisel

**2. Andmete ettevalmistamisel**

Dateering
1939
1800
-

→

Dateering
1939
1800
0 <b>X</b>
-1 <b>X</b>



# Milliseid takistusi tekitavad madala kvaliteediga andmed?

1. Andmete puhastamisel
2. Andmete ettevalmistamisel
- 3. Mudelite arendamisel**

Väli	Saagis	
	Täitmata	Täidetud
Olemus	64%	75%
Materjal	68%	75%

# Küsimused

# Aitäh!

Meri Liis Treimann

**Andmeteadlane**

meriliis.treimann@stacc.ee