

# Põhiandmed ja andmete kvaliteet

Sisukord

## 1. Printsiibid

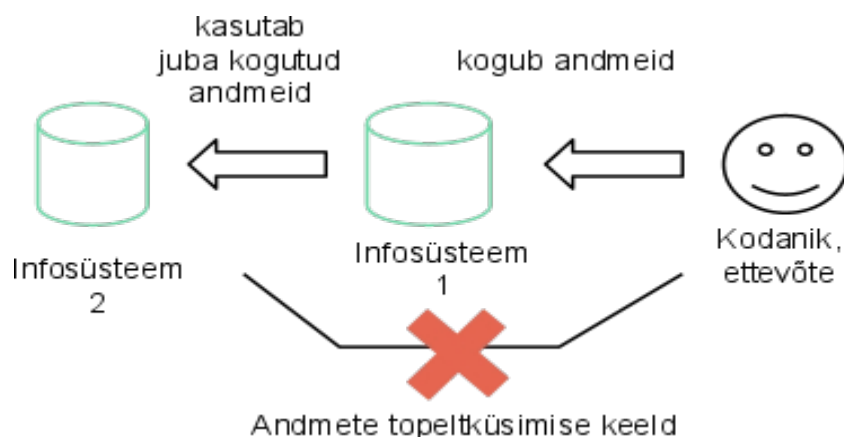
1. Andmete dubleeriva küsimise keeld
  2. Andmete võtmine võimalikult tekkimise (või riigi infosüsteemi hõlvamise) kohalt
  3. Andmete võtmine kõige kvaliteetsemast allikast.
2. Eesmärgipärasuse printsiip ehk mis on "samad andmed"
3. Põhiandmete mõiste AvTSis
4. Põhiandmete määramine
5. Kaks viisi põhiandmete määratlemiseks: tekkepõhine ja kasutuspõhine

Põhiandmed on avaliku teabe seaduses (AvTS) sätestatud oluline õigusmõiste. Sellega reguleeritakse riigi infosüsteemide arhitektuuri kujundamise ja koosvõime kindlustamise mitmeid tähtsaid aspekte. Praktikas tekitab põhiandmete määratlemine andmekogude ja infosüsteemide omanikele tihti raskusi. Käesoleva kirjutise eesmärgiks on näidata, et põhiandmete mõistes on koos tegelikult mitu keerukat momenti ja põhiandmete mõiste õigeaks kohaldamiseks on otstarbekas käsitleda kogu asja läbi andmete kvaliteedi prisma.

## 1. Printsiibid

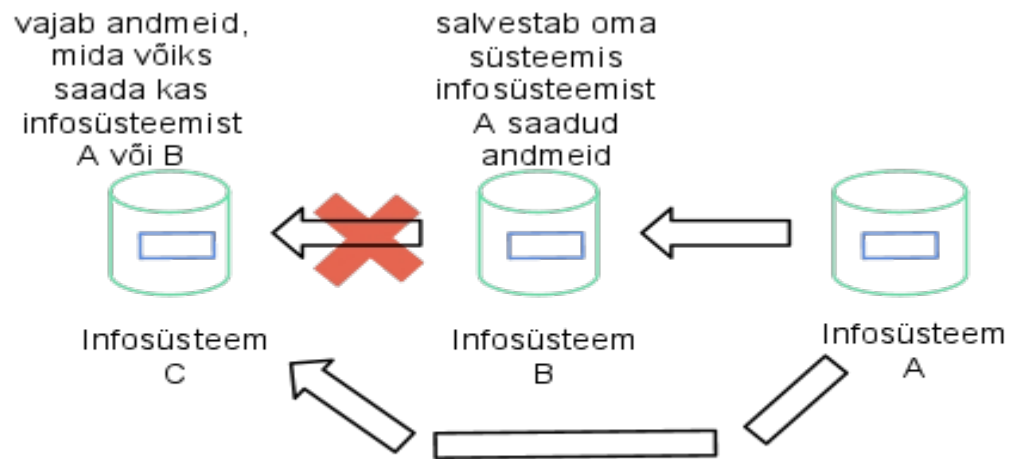
Mida põhiandmete mõistega tahetakse saavutada? Kolm printsiipi:

1.1 Andmete dubleeriva küsimise keeld. Kui andmeid on kodaniku või ettevõtte käest juba küsitud, siis ei tohiks neid enam uuesti küsida, vaid tuleks kasutada riigi infosüsteemi juba hõivatud andmeid.



## 1.2. Andmete võtmine võimalikult tekkimise (või riigi infosüsteemi hõlvamise)

kohalt. Põhjendub sellega, et tekkekohal on andmete autentsuse säilumine kõige tõenäolisem. Andmete liikumisega tekkekohalt teistesse süsteemidesse võib kaasas käia andmete muutumine. See põhimõte ei ole absoluutne. Laiem eesmärk on andmekvaliteedi tagamine. Andmekvaliteedi oluline mõõde on ka käideldavus - andmetele ligipääsetavus. Kui andmed tekivad hajutatult, paljudes kohtades, siis ei tarvitse nende kasutamisel alati minna tekkekohale tagasi olla efektiivne.

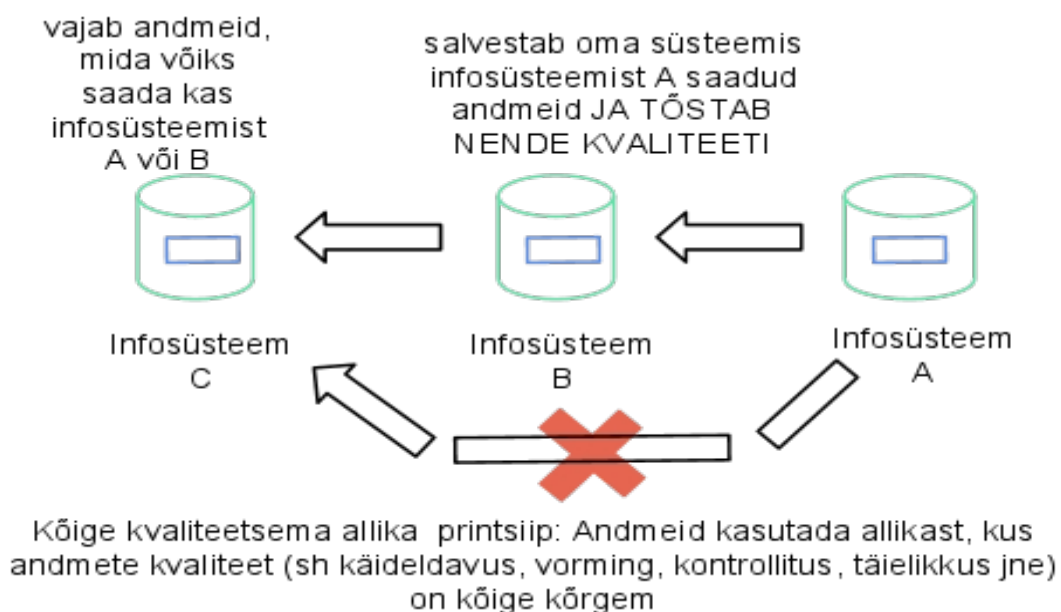


- Algallika kasutamise printsiip: Andmeid kasutada võimalikult nende algallikast, sest tõenäoliselt on seal andmed kõige autentsemad
- 1.

Näiteks terviseandmed tekivad inimese teenindamisel erinevates meditsiinasutustes. Nende andmete kättesaadavaks tegemiseks kogutakse terviseandmed kesksesse e-tervise süsteemi.

**1.3. Andmete võtmine kõige kvaliteetsemast allikast.** See printsiip on eelmise laiendus. (Kuid see printsiip ei ole kehtivas õiguses välja toodud!) Kui on loodud keskne, kõrgkäideldav süsteem, kuhu erinevates kohtades tekkivad üheliigilised andmed kokku kogutakse, siis on andmeid lihtsam võtta kesksest süsteemist. Andmete kättesaadavaks tegemine ongi kesksete süsteemide üks peamisi eesmärgi.

Näiteks: E-toimiku süsteem koondab väär- ja kuritegude menetlemisel erinevates asutuste tekkinud informatsiooni ja teeb selle tsentraliseeritult kättesaadavaks (neile, kes seda teavet vajavad ja on õigustatud seda teavet saama, nt kohtud). Rahvastikuregister koondab ja teeb kättesaadavaks teavet, mis tegelikult ei teki keskselt, vaid kohalikes omavalitsustes perekonnaseisutoimingute tegemise käigus. Tervise infosüsteem (e-tervis) koondab ja teeb digitaalselt kättesaadavaks patsiendi tervikliku haigusloo.



## 2. Eesmärgipärasuse printsiip ehk mis on “samad andmed”?

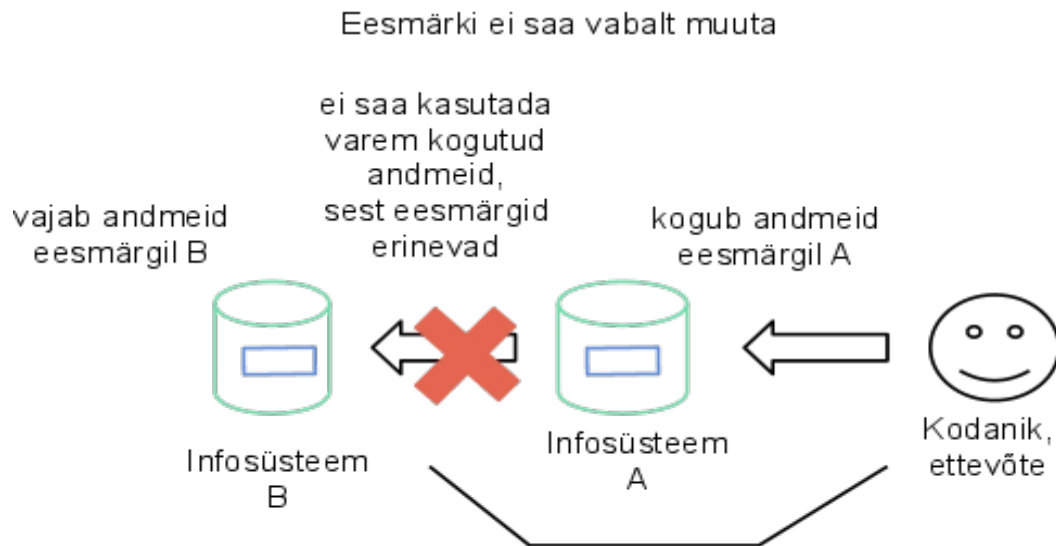
Andmeid ei saa vaadata lahus nende kogumise ja kasutamise eesmärkidest. Eesmärgipärasus on isikuandmete kaitse üks tähtsamaid põhimõtteid. Andmeid, mis on kogutud ühe eesmärgi täitmiseks, ei saa lihtsalt niisama kasutada teistel eesmärkidel. Näiteks riikliku statistika tegemiseks kogutud andmeid ei tohi kasutada uurimiseks kas kodanikud või ettevõtted on teinud midagi seadusevastast.

Isikuandmete kaitse seaduses ja teistes seadustes on päris üksikasjalikult määratletud milliseid isikuandmeid millistel eesmärkidel saab koguda ja kasutada. Eesmärgipärasuse põhimõte võib

tähendada seda, et samu andmeid on vaja korduvalt koguda. Järelikult mõistet 'samad andmed' tuleb vaadata andmete kasutamise eesmärkide kontekstis.

Eesmärgipärasuse põhimõte piirub isikuandmetega - mida seaduses määratletakse kui isikut identifitseerivaid või isikuga vahetult seotud andmeid. Kuid avalikus sektoris on suur osa andmetest isikuandmed. Seetõttu on eesmärgipärasuse mõistel tähtis roll. Kuidas eesmärgipärasus on seotud põhiandmetega?

- Ühed ja samad andmed võivad tegelikult olla erinevad andmed - kui nende kogumise ja kasutamise eesmärgid on erinevad.
- Andmete olemine põhiandmeteks ei tähenda automaatselt, et kõik võivad neid kasutada. Kui eesmärgid on erinevad, siis peab välja selgitama kuidas kasutamine on õiguslikult reguleeritud.



Ühel eesmärgil kogutud andmete kasutamine TEISE eesmärgi täitmiseks võib olla takistatud

Andmete unikaalsust, millest räägib seaduses antud põhiandmete definitsioon, tuleb samuti nagu andmete samasust, käsitleda andmete kogumise ja kasutamise eesmärkide lõikes.

“Samad andmed” võivad ka lihtsa teise kohta liigutamisega omandada uue tähenduse. Andmete tähendus sõltub andmete kontekstist. Rangelt võttes on “samad andmed” on tihti üldse nonsens: andmete teise kohta tõstmine või nendest koopia tegemine ei ole üldjuhul tähenduseta operatsioon; andmete omandavad uue või lisatähenduse või tähendusvarjundi. Mõnedel juhtudel ei ole lisatähendused olulised ja need võib kõrvale jätta; teistel juhtudel ongi lisatähendused peamised. Näiteks inimese isikukoodi hoitakse kümnetes andmekogudes ja infosüsteemides. Isikukood on kõikjal kus seda kasutatakse sama. Kuid sama siiski erinevates kontekstides. Maksukohustuslaste registris väljendab isikukood mitte seda, et sellisel isikul on selline isikukood, vaid seda, et isik on maksumaksjat. Hariduse infosüsteemis väljendab isikukood seda, et isik on õpilane jne.

## Mis on 'samad andmed'?

Andmed VÕIVAD jääda samaks... ... kui tõsta või kopeerida neid teise süsteemi ... kui siduda neid teiste andmetega ... paigutades andmed teise konteksti ... andes andmetele teise tähenduse ... kui muuta andmete kasutuseesmärki	Andmed EI TARVITSE jääda samaks... ... kui tõsta või kopeerida neid teise süsteemi ... kui siduda neid teiste andmetega ... paigutades andmed teise konteksti ... andes andmetele teise tähenduse ... kui muuta andmete kasutuseesmärki
---	--

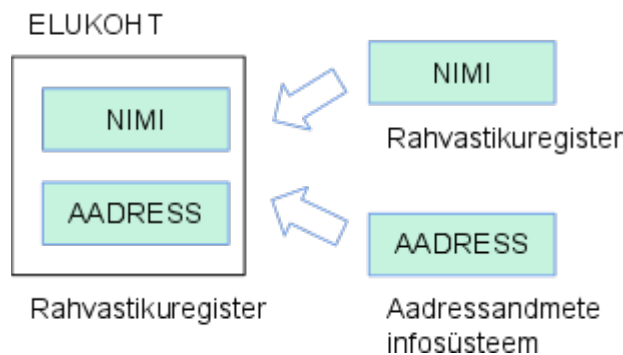


Infosüsteem  
A



Infosüsteem  
B

Andmete tähendus tuleneb nende seostamisest. Andmed mudeldavad reaalse maailma keerukust, moodutades keerukaid struktuure. Nendes keerukates struktuurides, seostes andmete vahel peituvadki sageli kõige olulisemad tähendused. Põhiandmete mõiste peab arvestama seda keerukust. Andmeid atomaarsetena (homogeense massina) käsitades ei suuda me põhiandmete mõistet õigesti tööle panna. Näiteks inimese NIMI ja elamu AADRESS omavad tähendust eraldiseisvate andmetena. Neid hoitakse põhiandmetena eraldi infosüsteemides (Eestis - Rahvastikuregistris ja Aadressandmete infosüsteemis). Inimese ELUKOHT on vaadeldav seosena inimese ja elamu vahel. ELUKOHT on andmeobjektide NIMI ja AADRESS agregaat, kuid ei taandu nende lihtsale summale. Mitte igasugune nime ja aadressi kombinatsioon ei moodusta elukohta. Määrav on tähendus, mis nime ja aadressi summale omistatakse.



NIMI on põhiandmeteks Rahvastikuregistris. AADRESS on põhiandmeteks Aadressandmete infosüsteemis. ELUKOHT on põhiandmeteks samuti Rahvastikuregistris. Kuid ei ole välistatud, et mõnes teises infosüsteemis (nt politsei, kohtu või riikliku statistika infosüsteemis) salvestatakse inimese TEGELIKKU ELUKOHTA (mis võib ju Rahvastikuregistris näidatud ametlikust elukohast erineda). Tegelik elukoht koosneb samuti nagu ametlik elukohtki nime ja aadressi seosest, kuid selle seose tähendus siiski erineb.

### 3. Põhiandmete mõiste AvTSis

Põhiandmed on õiguslikult määratletud järgmiselt:

§ 43<sup>6</sup>. *Põhiandmed ja andmete tähendus*

(1) *Põhiandmed on riigi infosüsteemi kuuluvasse andmekogusse kogutavad andmekogu unikaalsed andmed, mis tekivad andmekogu haldaja avalike ülesannete täitmise käigus.*

(2) *Andmete töötlemisel, mida kogub põhiandmetena teine riigi infosüsteemi kuuluv andmekogu, tuleb aluseks võtta vastava teise andmekogu põhiandmed.*

*(3) Andmete põhiandmeteks olek määratakse kindlaks käesoleva seaduse § 43<sup>3</sup> lõike 3 kohaselt kooskõlastatud tehnilise dokumentatsiooni alusel riigi infosüsteemi haldussüsteemis. Põhiandmete kindlaksmääramisel lähtutakse andmekogu asutamise eesmärgist. (AvTS)*

Põhiandmeteks määratletud andmete kasutamise kohustus on sisse kirjutatud ka Rahvastikuregistri seadusse<sup>1</sup>.

#### 4. Põhiandmete määramine

Siin on mitu momenti, mis praeguse seaduse sõnastusest ei tule kõige selgemal kujul välja. Ühelt poolt on põhiandmed defineeritud - andmed on põhiandmeteks seal, kus nad tekivad. Samas ütleb seadus, et alles RIHA menetluses määratakse kus andmed loetakse põhiandmeteks. Need kaks momenti on ülalnimetatud printsiipide väljendused. Ühelt poolt tahame andmeid kasutada võimalikult nende tekkekoha lähedalt. Teiselt poolt on sageli siiski ökonoomsem andmete kogumine kesksesse varamusse ja sealt kasutamine.

#### 5. Kaks viisi põhiandmete määratlemiseks: tekkepõhine ja kasutuspõhine

Kodakondsuse andmed tekivad mitmes kohas. Samas vist enamusel juhtudel, kus kodakondsuse andmeid kasutatakse, ei ole üldse tähtis millisel teel kodakondsus tekkis. Kodakondsuse andmete kokkukoondamine ühte andmekogusse on ilmselt mõistlik. Tundub ikkagi, et on kaks viisi põhiandmete määratlemiseks: tekkepõhine ja kasutuspõhine. Tekkepõhine – andmed on põhiandmed seal, kus nad tekivad. Kasutuspõhine – andmed on põhiandmed seal, kust neid kasutamiseks tuleb võtta. Tekkepõhine ja kasutuspõhine ideaaljuhul langevad kokku, kuid ei tarvitse!

---

<sup>1</sup>“Avaliku ülesande täitmisel peab lähtuma rahvastikuregistri põhiandmetest. Teistesse andmekogudesse rahvastikuregistri põhiandmete esmane kogumine on keelatud.” - RRS (jõustub 01.01.2011) § 6 lg 4.