



RIIGI INFOSÜSTEEMI AMET



Euroopa Liit  
Euroopa  
Regionaalarengu Fond



Eesti  
tuleviku heaks

---

# *Andmekvaliteedi uuring*

## Lõpparuanne

*Riigi Infosüsteemi  
Amet*

*26. august 2016*





**Taimar Peterkop**  
Riigi Infosüsteemi Amet  
Pärnu mnt 139a  
15169 Tallinn

26. august 2016

Austatud Taimar Peterkop

Oleme läbi viinud uuringu „Andmekvaliteedi tagamise juhend“. Töö teostamisel lähtusime kahepoolselt allkirjastatud lepingust nr 4.2-3/16-0089-001 kuupäevaga 26. veebruar 2016.

Käesoleva dokumendi näol on tegemist lõpparuandega. Juhul kui järeldused lõppversioonis erinevad meie varasemate diskussioonide käigus väljaöeldust või vahearuandes kirjeldatust, tuleb eelistada neid, mis esitatakse lõpparuandes.

Meie ülesannete hulka ei kuulunud lähteandmete õigsuse kontrollimine, mistõttu AS PricewaterhouseCoopers Advisors (PwC) ei vastuta algandmete õigsuse eest ega ka tulemuste eest juhul, kui need põhinevad puudulikel või ebaõigetel algandmetel. Meie töö oli piiratud lepingus sätestatud tegevustega.

Lisaks juhime Teie tähelepanu, et AS PricewaterhouseCoopers Advisors ei võta vastutust kolmandate osapoolte ees, kellele käesolev dokument on avaldatud või mõnel muul moel kättesaadavaks saanud.

Lugupidamisega

*/allkirjastatud digitaalselt/*  
Teet Tender  
AS PricewaterhouseCoopers Advisors

# Sisukord

---

<b>Uuringu autorid</b> .....	<b>4</b>
<b>Uuringu juhtrühm</b> .....	<b>4</b>
<b>Mõisted ja lühendid</b> .....	<b>5</b>
Mõisted.....	5
Lühendid .....	6
<b>Kokkuvõte</b> .....	<b>7</b>
<b>Uuringu lähteülesanne ja tulemid</b> .....	<b>9</b>
<b>Andmekvaliteedi tagamise uuring</b> .....	<b>13</b>
Andmekvaliteedi tunnuste uuring .....	13
Tunnuste allikad.....	13
Tunnuste komplekti valik .....	14
Andmekvaliteedi nõuete analüüs .....	15
Andmekvaliteedi juhtimise uuring .....	31
Rakendamise meetodika .....	33
Andmekvaliteedi juhtimise raamistik .....	33
Kokkuvõte .....	34
<b>Andmekvaliteedi uuring andmekogudes</b> .....	<b>35</b>
<b>Andmekvaliteedi tagamise juhend riigile</b> .....	<b>36</b>
<b>Lisad</b> .....	<b>37</b>
Lisa 1: Tunnuste allikate loetelu .....	37
Lisa 2: Tunnuste sünteesi tabel .....	38
Lisa 3: Hinnang tegurite kaetuse osas aspektidega .....	50
Lisa 4: Andmekvaliteedi küpsusmudel .....	51

---

# Uuringu autorid

Janar Linros

Mihkel Lauk

Viljar Kähari

Jaak Tepandi

# Uuringu juhtrühm

*Tabel 1: Andmekvaliteedi uuringu juhtrühma liikmed*

<b>Juhtrühma liige</b>	<b>Asutus/organisatsioon/ettevõte</b>
Priit Rospel	Riigi Infosüsteemi Amet
Irina Ivahnenko	Riigi Infosüsteemi Amet
Urve Pastarus	Riigi Infosüsteemi Amet
Helen Simisker	Riigi Infosüsteemi Amet
Mihkel Lauk	AS PricewaterhouseCoopers Advisors
Janar Linros	AS PricewaterhouseCoopers Advisors
Jaak Tepandi	Tepinfo OÜ
Viljar Kähari	Advokaadibüroo PricewaterhouseCoopers Legal OÜ

Täname uuringu autoreid, juhtrühma ja uuringus osalejaid!

Käesolev projekt „Andmekvaliteedi tagamise juhend“ on Riigi Infosüsteemi Ameti (edaspidi RIA) projekt, mis on toetatud Euroopa Liidu struktuurifondide toetuskeemist „Avalike teenuste koosvõime loomine“.

# Mõisted ja lühendid

## Mõisted

Mõiste	Selgitus
Andmeatribuudid	on andmeobjekti eeldefineeritud omadused, kus igal omadusel on konkreetne väärtus.
Andmekogu omanik	on riigi infosüsteemi haldussüsteemis registreeritud andmekogu vastutav töötleja.
Andmekvaliteedi juhtimise raamistik	on andmekvaliteedi juhtimiseks mõeldud töövahendite ja nende rakendamise meetodikatest koosnev kogum.
Andmekogu andmekvaliteedi keskne teadmusbass	on infosüsteem või andmekogu, mis sisaldab konkreetse andmekogu osas andmekvaliteedi juhtimise praktikate kirjeldusi (protseduurid, juhised, ametijuhendid, jms. igapäevast tööd korraldavad dokumendid) ja mille ülesandeks on selle teadmuse säilitamine andmekvaliteedi protsesside jätkusuutlikkuse tagamiseks.
Andmekvaliteedi mõõtmine	on andmete kvaliteedi kontrollimine.
Andmekvaliteedi nõue	on tingimus või kriteerium, millele andmed peavad vastama, et saaks öelda, et need andmed on kvaliteetsed.
Andmekvaliteedi nõuete määratlemine	on andmekogule rakenduvate andmekvaliteedi nõuete väljaselgitamine ja dokumenteerimine.
Andmekvaliteedi poliitika	on andmekvaliteedi juhtimise aluseks olev nõuete kogum, mis määratleb ühtlasi ka andmekvaliteedi tegevuste põhimõtted.
Andmekvaliteedi profiil	on andmekvaliteedi nõuetel põhinev andmekontrollide kogum, mille abil selgub andmete kvaliteet.
Andmekvaliteedi protsessid	on andmekvaliteeti mõõtva või tagava iseloomuga protsessid, mille käigus toimub reaalne andmekvaliteedi kontrollimine ja parendamine.
Andmekvaliteedi sihtportfell	on andmekvaliteediga seotud infosüsteemide arendusvajaduste nimekiri, mida planeeritakse ellu viia.
Andmekvaliteedi tsoneerimine	on ühe andmekogu andmekvaliteedi juhtimisega tegeleva organisatsiooni määratlemine.
Andmekvaliteedi tsoon	on ühe andmekogu andmekvaliteedi juhtimisega tegeleva organisatsiooni mõjuala.
Andmekvaliteedi tunnus	on andmekvaliteedi nõuete klassifikatsioon, mis iseloomustab nende nõuete poolt käsitletavat andmekvaliteedi vaatepunkti/aspekti.
Andmekvaliteedi võtmeisikud	on organisatsioonis andmekvaliteedi kompetentse ja vastutusi omavad isikud.
Andmekvaliteet	on andmete vastavus andmekvaliteedi nõuetele (ja headele tavadele).
Andmete omanik	on osapool, kellel on õigused ja võimalused toota või muuta andmekogude põhiandmeid. Andmete omanikeks võivad olla nii füüsilised kui ka juriidilised isikud (sh. riigi- ja kohaliku omavalitsuse üksused).
IT-grundschutzhandbuch	on Saksamaa infoturbeameti poolt välja töötatud IT etalonturbe käsiraamat, millele baseerub Eestis väljatöötatud ISKE.
Keskne tõeallikas	on riigi põhiandmeid haldav andmekogu.
Kriitilised andmed	on prioriteetsed andmed, mille kvaliteedile mittevastavuse tõttu pole võimalik

	oluliste äriprotsesside toimimine.
Semantiline õigsus	tähendab, et andmed on kasutatavad tähenduslikus ja interpreteeritavas mõttes antud kasutuskontekstis, nt mingi avaliku teenuse osutamise eesmärgil.
Süntaktiline õigsus	on andmete tehniline korrektsus, st et andmetes puuduvad kirjavead ja samad andmed on andmekogus läbivalt ühesugused.

## Lühendid

Lühend	Selgitus
IKS	Isikuandmete kaitse seadus
AvTS	Avaliku teabe seadus
ISKE	Infosüsteemide kolmeastmeline etalonturbe süsteem
OPDCA	Kvaliteedi 5-etapiline parendamistsükkel ( <i>observe-plan-do-check-act</i> )
PDCA	<i>Plan-Do-Check-Act</i> kvaliteedi juhtimise raamistik
RIHA	Riigi infosüsteemi haldussüsteem

# Kokkuvõte

Eesti riigi infosüsteem põhineb paljudel erinevatel andmekogudel. Andmekogudes olevate andmete kvaliteet mõjutab otseselt riigi toimimist. Mida parem on andmekogudes olevate andmete kvaliteet, st. mida väiksema viivitusega ja täpsemalt vastavad andmed tegelikkusele, seda paremini suunatud, mugavamaid ja lihtsamalt kättesaadavaid teenuseid on võimalik riigi elanikkonnale pakkuda ning seda täpsemaid otsuseid on võimalik vastu võtta riigi juhtimisel.

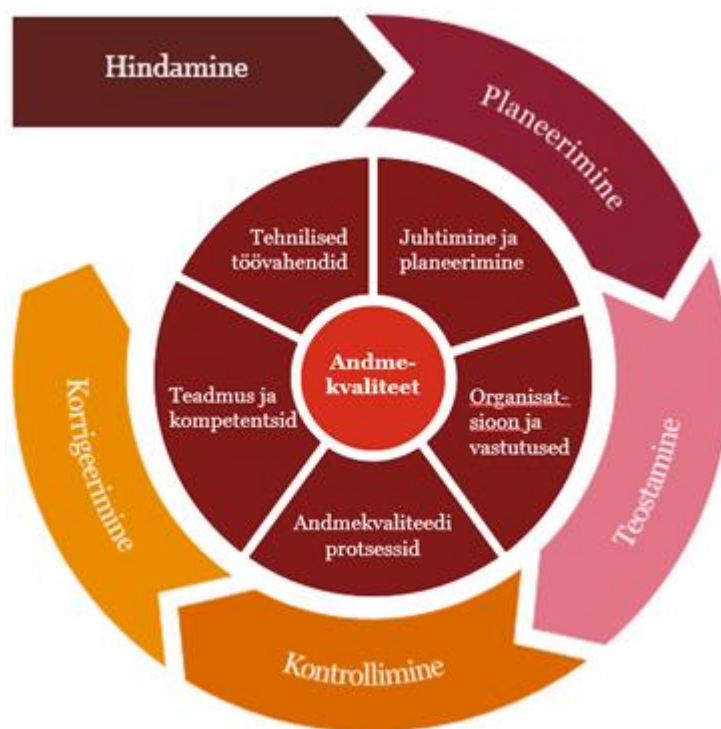
Riigi Infosüsteemi Ameti (edaspidi RIA) tellimusel viidi läbi andmekvaliteedi uuring, mille tulemusena töötati välja andmekvaliteedi tagamise juhend andmekogu omanikele. Lisaks töötati välja juhend RIA-le, mis annab soovitusi juhtida andmekvaliteeti riigi tasandil. Käesolevas dokumendis on kirjeldatud uuringu läbiviimise meetodikat ja tulemusi, mille põhjal nimetatud juhendid loodi.

Aruande esimeses osas „Uuringu lähteülesanne ja tulemid“ antakse ülevaade sellest, kuidas töö käigus adresseeriti tellija ootuseid tuues välja lähteülesandes toodud nõuete vastavused teostatud töodes.

Aruande teises jaotises „Andmekvaliteedi tagamise uuring“ keskendutakse sellele, kuidas töötati välja andmekvaliteedi tagamise juhendi jaoks andmekvaliteedi juhtimise raamistik ja selle sisu komponendid. Uuringu raames väljatöötatud andmekvaliteedi juhtimise raamistik koosneb kolmest komponendist:

1. Andmekvaliteedi küpsusmudel, mis on tööriist andmekvaliteedi juhtimisega seotud kategooriate mõõtmiseks ja parendamiseks. Väljatöötatud mudel vaatleb viit juhtimise kategooriat viies küpsusastmes. Mudeli osas on kirjeldatud rakendamise meetodika;
2. Andmekvaliteedi tunnuste komplekt, mis on vahend andmete kvaliteedi kontrollimiseks erinevatest kvaliteediga seotud aspektidest. Komplekt koosneb 9-st tunnusest, mille osas tuuakse välja andmekvaliteedi nõuded, tagamise meetmed ja kontrollküsimused;
3. Andmekvaliteedi juhtimise raamistik, mis on OPDCA kvaliteedi juhtimise mudelil põhinev ja andmekvaliteedi tagamiseks iteratiivselt rakendatav tegevuste kogum.

Raamistik kombineerib omavahel loetletud komponendid ühtseks kvaliteedi juhtimise süsteemiks.



Joonis 1: Andmekvaliteedi juhtimise raamistik

Andmekogu omanikele mõeldud andmekvaliteedi tagamise juhendi valmimise järel teostati andmekvaliteedi uuring andmekogudes, mille eesmärkideks olid testida juhendi rakendatavust ja täiendada juhendit praktilise rakendamise tulemusel saadava tagasiside abil. Aruande kolmandas jaotises „Andmekvaliteedi uuring andmekogudes“ kirjeldatakse kokkuvõtva ülevaatenähtena juhendi piloteerimist kolmel andmekogul: Rahvastikuregister, Aadressiandmete süsteem ja Riigi infosüsteemi haldussüsteem. Põhjalikum ülevaade juhendi piloteerimisest on kirjeldatud eraldi aruandena.

---

Uuringu raames töötati välja andmekvaliteedi tagamise juhend riigile, mis sisaldab metoodikaid riigi infosüsteemi kui terviku andmekvaliteedi mõõtmiseks ja tagamiseks. Juhend on kirjeldatud eraldi dokumendina. Lühike kokkuvõtte juhendis sisalduvast leiab käesoleva aruande peatükist „Andmekvaliteedi tagamise juhend riigile“.

Aruande lisadena tuuakse välja:

1. Allikate loetelu, mille andmed olid aluseks tunnuste komplektile;
2. Tunnuste sünteesi tabel, mis oli aluseks valikule tunnuste komplekti;
3. Hinnang andmekvaliteedi tegurite kaetuse osas andmekvaliteedi aspektidega;
4. Andmekvaliteedi küpsusmudel.



# Uuringu lähteülesanne ja tulemid

Käesolev uuring on läbi viidud Riigi Infosüsteemi Ameti tellimusel projekti „Andmekvaliteedi tagamise juhend“ raames. Projekti lähteülesanne seab tingimused projekti töödele ja oodatavatele tulemitele. Järgnevas tabelis on kirjeldatud teostatud tööde vastavus uuringu lähteülesandele.

Tabel 2: Teostatud tööde vastavus uuringu lähteülesandele

Lähteülesande tingimus töödele	Vastavus teostatud töödes
<b>Viia läbi andmekvaliteedi uuring</b>	
Uuring peab sisaldama mujal maailmas kasutatavaid andmete kvaliteedi hindamise meetodikaid. Tehtud valikuid tuleb põhjendada.	Uuringus on kasutatud andmekvaliteedi tunnuste komplekti ja küpsusmudeli väidete väljatöötamiseks mujal maailmas kasutatavaid andmekvaliteedi raamistikke.
<b>Koostada andmekogu omanike juhend, mis sisaldab järgnevaid juhiseid:</b>	
andmekogu hetkel kehtiva andmekvaliteedi taseme välja selgitamise ja kirjeldamise meetodika;	Andmekvaliteedi hetke taseme väljaselgitamist on võimalik teostada juhendi abil, mis sisaldab meetodika kirjeldusi.
andmekogu toimimiseks vajaliku andmekvaliteedi taseme määramise ja kirjeldamise meetodika	Andmekvaliteedi toimimiseks vajaliku andmekvaliteedi taseme (andmekvaliteedi sihttaseme) määramise meetodika on kirjeldatud juhendis.
andmekogu andmekvaliteedi hetkel kehtivalt andmekvaliteedi tasemelt kirjeldatud vajalikule andmekvaliteedi tasemele viimise tegevuste planeerimise ja planeeritud tegevuste läbi viimise meetodika;	Planeerimise tegevuste läbiviimise meetodika kirjeldatud juhendis „Planeerimine“ peatükis.
andmekogu andmekvaliteedi seire, järelevalve ja auditeerimise meetodika;	Andmekvaliteedi seire, järelevalve ja auditeerimine toimub juhendis kirjeldatud küpsusmudeli ja selle kasutamise meetodika abil.
organisatsioonilised jms soovitusel andmekogu omanikele;	Andmekogu omanikele suunatud juhised kirjeldatu juhendis ja andmekogude uuringu aruandes.
juhend andmekvaliteedi tagamiseks riigi tasemel (rakendusjuhend meetodika rakendamiseks riigis, seadusandluse ja ajakava planeerimisega seotud soovitusel jms).	Uuringu tulemusena töötati välja juhend andmekvaliteedi tagamiseks riigi tasemel, mis sisaldab rakendamisega seotud soovitusi.
Meetodikad tuleb luua universaalsena selliselt, et neid on võimalik rakendada kõigile riigis olemasolevate ja tulevikus loodavate andmekogude andmekvaliteedi optimeerimiseks ja auditeerimiseks.	Meetodikad on loodud universaalselt ja andmekogude üleselt ja on rakendatavad kõikidel riigi infosüsteemi andmekogudel.
Meetodika loomisel tuleb võtta arvesse ISKE erinevate tervikluse (T) ja konfidentsiaalsuse (S) nõuete rakendamist erinevatele andmekogudele ja sellest tulenevat mõju andmekvaliteedi juhtimise vajadustele ja võimalustele ning kehtivatele ja kehtestatavatele nõuetele ning piirangutele.	Meetodika loomisel on arvesse võetud ISKE-st tulenevaid nõudeid andmekvaliteedi tunnuste osas nõuete selgitamisel..
Tööde elluviimisel tuleb lähtuda eesmärgist luua KASUTATAV meetodika riigi andmekogude andmekvaliteedi tõstmiseks. See tähendab seda, et meetodika väljatöötamise üheks väga oluliseks aspektiks on see, kuidas meetodikat tulevikus rakendada hakatakse ja kui arusaadav on meetodika selle kasutajatele.	Väljatöötatud juhendi ja selles sisalduvate meetodikate rakendatavust testiti edukalt uuringu käigus kolmel riigi infosüsteemi andmekogul. Seega saab öelda, et juhend on rakendatav.
Juhend tuleb luua sellisena, et selle järgi on võimalik süstemaatiliselt hinnata andmekogude andmekvaliteedi hetkeseisu, planeerida vajalikku andmekvaliteedi taset, võrrelda hetkeseisu planeeritud andmekvaliteedi tasemega ja planeerida tegevusi ühelt tasemelt teisele üleminekuks.	Juhendis kirjeldatud andmekvaliteedi juhtimise raamistik võimaldab hinnata olemasolevat ja planeerida andmekvaliteedi taset, ning planeerida tegevusi ühelt tasemelt teisele üleminekuks.
Juhendi koostamiseks läbiviidav uuring ja loodavad juhised peab katma mudelis toodud kõiki tegureid/valdkondi.	Andmekvaliteedi uuringu käigus veenduti, et juhendis kirjeldatud küpsusmudeli erinevad kategooriad, katavad kõik lähteülesandes kirjeldatud andmekvaliteedi aspektid, mis katavad omakorda kõik lähteülesandes kirjeldatud valdkonnad/tegurid.

<p>Andmekvaliteedi hindamise mudeli välja töötamiseks tuleb täitjal vaadelda erinevaid andmekvaliteedi hindamise mudeleid ja meetodikaid, valida välja sobivaim ja täiustada seda teiste (valikust välja jäänud) meetodikate alusel selliselt, et see sobiks andmekogude andmekvaliteedi hindamiseks. Seejuures tuleb arvesse võtta seda, et meetodika peab katma ka andmekogude vahelise semantilise andmekvaliteedi hindamise.</p>	<p>Andmekvaliteedi hindamise meetodika väljatöötamisel kasutati algallikatena mitmete mujal maailmas rakendust leidnud kvaliteediraamistike alusel. Meetodika on kohandatud andmekogude andmekvaliteedi hindamiseks. Arvestatud on ka andmekogude vahelise semantilise andmekvaliteedi hindamist.</p>
<p>Uuringus peab andmekvaliteeti hindama minimaalselt järgmiste tunnuste alusel: ajakohasus, täpsus, ühilduvus, veakäsitus ja tagasiside, liiasused andmestruktuurides, liiasused andmekäsitusprotsessides, andmete juurdepääs ja turvalisuse tagamine, once only printsiip</p>	<p>Uuringus hinnati kõiki lähteülesandes kirjeldatud nõudeid, kui võimalike liikmeid välja töötatavas andmekvaliteedi tunnuste kompleksis. Lõplik tunnuste kompleks moodustati uuringu juhtrühmas kokkulepitud kriteeriumite alusel.</p>
<p>Loodav meetodika peab olema üles ehitatud selliselt, et selle järgi saab hinnata ja planeerida andmekogu andmekvaliteeti andmekogu kui terviku kohta ja ka andmekogu erinevate, hindaja meelevaldsuse alusel määratud osade kaupa eraldi.</p>	<p>Loodavat meetodika on võimalik rakendada ka andmekogu erinevate osade kaupa. Seda toetavad juhendis kirjeldatud andmete ulatuse määramise ja andmekvaliteedi organisatsiooni tsooneerimise tegevused.</p>
<p>Loodav meetodika peab olema koostatud vähemalt kahetasemelisena (astmelisena), kus esimene tase annab reeglid üldisemaks ja teine osa täpsemaks andmete kvaliteedi juhtimiseks. Esimene tase peab olema loodud viisil, mis võimaldab seda kasutada eraldiseisva tervikliku meetodikana. Esimene tase peab katma kõik olulisemad andmekogu kvaliteedi taseme välja selgitamise ja juhtimise aspektid ning määratlema iga andmekogu juures kohustuslikult läbi viidavate meetmete kogumi. Kvaliteeditasemete eraldamisel tuleb lähtuda ISKE tervikluse (T) ja konfidentsiaalsuse (S) nõuetest tingitud gradatsioonist ja selle mõjust andmete kvaliteedi tasemete tõlgendamisele. Meetodika teine tase peab laiendama (täpsustama) esimese taseme meetodikat selliselt, et lisanduksid kõik esimesel tasemes mitte sisaldunud meetmed. On lubatud meetodika jagamine ka rohkemateks tasemeteks. Olenemata tasemete arvust koondab meetodika esimene tase endas alati kõik kohustuslikult rakendatavad meetmed.</p>	<p>Juhendis kirjeldatud meetodika esimese tasemena vaadeldakse andmekvaliteedi juhtimisprotsesside parendamist, mille osas antakse üldisemad nõuded ja teise tasemena vaadeldakse andmete kvaliteedi juhtimist tunnuste komplekti ja sellega seotud andmekvaliteedi nõuete ja tagamise meetmete abil.</p>
<p>Andmekvaliteedi hindamise ja parendamise võimalusi tuleb analüüsida vähemalt järgmiste aspektide kaupa: andmekvaliteedi juhtimine, organisatsiooni võimekus, organisatsiooni struktuur, vahendid ja personal. Kirjeldatud aspekte on lubatud grupeerida ja ka võtta lahti detailsemateks aspektideks. Kui uuringu käigus tuvastatakse toodud loetelu laiendavaid aspekte, tuleb need lisada sellele loetelule</p>	<p>Andmekvaliteedi aspektid on jaotatud küpsusmudelisse 5-e kategooriasse, mis sündisid uuringu käigus aspektide ümbergrupeerimise tulemusena.</p>
<p>Uuritud aspektidele tuleb teostada riskianalüüs. Uuritud aspektid tuleb järjestada andmekvaliteedile mõju kahanemise järjekorras (riskianalüüsi alusel).</p>	<p>Kõikide lähteülesandes kirjeldatud aspekte analüüsiti uuringu käigus. Hinnanguliselt kõige mõjukamate aspektide osas viidi vastavasisulisel väidet andmekvaliteedi küpsusmudelisse.</p>
<p>Iga aspekti jaoks tuleb kirjeldada hindamiskriteeriumite skaala (näiteks 5-tasemelisel jaotusskaalal: "ei ole ja ei ole veel plaanitud", "plaanitud, aga sisuliselt veel puudub", "rakendatud nõrgalt", "rakendatud rahuldavalt", "kõik on korras). Täitja poolt välja töötatud hindamisskaala peab olema põhjendatud. Igal hindamistasemel peab olema olema loend meetmetest, mis määravad selle taseme. Igal järgmisel tasemel lisanduvad uued meetmed, mis eristavad seda eelmisest tasemest. Maksimaalsel tasemel peavad olema rakendatud kõik meetmed.</p>	<p>Küpsusmudelisse on jagatud andmekvaliteedi aspektide 5 kategooriat 5-ks tasemeks. Iga kategooria igal tasemel on kirjeldatud andmekvaliteedi väited, mis aitavad kontrollida, kas vastav tase selles kategoorias on saavutatud või mitte.</p>
<p>Välja töötatud meetodika peab võimaldama hinnata andmekogu hetke andmekvaliteedi taset ja kirjeldada</p>	<p>Küpsusmudeli rakendamise meetodika võimaldab hinnata olemasolevat andmekvaliteedi küpsustaset ja seada</p>

	soovitavat andmekvaliteedi taset. Nende kahe taseme võrdluse tulemusena peab olema võimalik välja töötada ühelt tasemelt teisele üle mineku meetmed.	soovitavat sihttaset. Küpsusmudeli rakendamine enne ja pärast parendustegevusi võimaldab võrrelda muudatusi tasemetes.
	Uuringu tulemusena tuleb uuringus osalenud andmekogudele anda tagasisidet nende hetke andmekvaliteedi, selle parendamise võimaluste ja meetodite kohta. Sisuliselt tähendab see iga uuringus osalenud andmekogu andmekvaliteedi hindamist, kasutades uuringu tulemusena loodud meetodikat.	Uuringu lõppedes anti tagasisidet uuringus osalenud andmekogudele. Samuti on uuringu tulemused esitatud aruandes „Andmekvaliteedi uuring andmekogudes“.
<b>Viia läbi uuring andmekogudes</b>		
	konkreetsete uuringus käsitletud andmekogudele tehtud soovitused peavad olema lisaks üldistatud sellisele tasemele, et need on kasutatavad ka teistes, nii hetkel olemasolevates, kui tulevikus loodavates andmekogudes.	Uuringu käigus andmekogudele tehtud soovitused põhinesid andmekvaliteedi küpsusmudelil. Samad soovitused on rakendatavad samadel tasemetel olevatele kolmandatele andmekogudele.
	Uuringu käigus tuleb välja selgitada, kas (igas) uuringus osalevas andmekogus on tegeletud andmekvaliteedi hindamisega, milliseid tegevusi ning millistest kaalutlustest lähtudes on andmekvaliteedi parandamiseks tehtud ja milliseid tulemusi on saavutatud.	Andmekvaliteedi uuringu käigus andmekogudes kirjeldati andmekogudes kasutusele võetud andmekvaliteedi tagamise meetmed. (vt. „Andmekvaliteedi uuring andmekogudes“) aruannet.
<b>Koostada juhend riigile</b>		
	Riigi juhtimisega seotud tegureid tuleb analüüsida kui kogu meetodikale suurt mõju avaldavaid tegureid, mille mõjud avalduvad mudeli kõigis teistes osades.	Riigi juhtimisega seotud tegureid on arvestatud andmekvaliteedi nõuete selgitamise ja planeerimise kontekstis, kus riigi poolt üle andmekogude kehtestatud regulatsioonid mõjutavad andmekvaliteeti enim.
	Kui tööde käigus selgub, et olemasolev seadusandlus on piiranguks andmekvaliteedi tõstmisele, tuleb töö tulemustes esitada soovitud seadusandluse muutmiseks.	Soovitused seadusandluse muutmiseks on antud andmekvaliteedi juhendis riigile.

Järgnevas tabelis on kirjeldatud uuringu tulemite vastavus oodatud tulemitele.

*Tabel 3: Uuringu tulemite vastavus oodatud tulemustele*

<b>Tulem</b>	<b>Oodatud tulemi sisu</b>	<b>Tulemi vastavus</b>
	<b>Uuringu tulemuste ülevaade, mis sisaldab:</b>	Käesolev dokument
	uuringu meetodit	Uuringu meetod kirjeldatud läbivalt.
	kasutatud materjalide loendit. Uuringus kasutatavad materjalid peavad olema viidatud selliselt, et tellijal on võimalus nendega tutvuda. Vajadusel tuleb materjal lisada uuringule	Kasutatud materjalid on viidatud dokumendis läbivalt ja dokumendi lisades.
	uuringu tulemusi, mis sisaldavad põhjalikult kommenteeritud loendit andmekvaliteeti mõjutavatest aspektidest. Dokumendis peab olema põhjendatult hinnatud erinevate aspektide osatähtsust (kaalu) andmekvaliteedi tagamisel ja grupeeritud ning järjestatud need tähtsuse järjekorras (tähtsamad eespool). Iga aspekti kohtapeab olema piisava täpsusega kirjeldatud seda aspekti maksimaalselt katvad andmekvaliteedi tagamise meetmed	Uuringu tulemused on kirjeldatud Andmekvaliteedi tagamise uuring peatükis.
	<b>Andmekvaliteedi uuring igale uuringus osanud andmekogule, mis sisaldab:</b>	„Andmekvaliteedi uuring andmekogudes“ aruanne
	andmekogu andmekvaliteedi hetke taseme hinnangut	Hinnangud kirjeldatud uuringu tulemite peatükis.
	andmekogule vajaliku andmekvaliteedi taseme kirjeldust	Andmekogude sihttasemed seatud uuringu tulemite peatükis.
	meetmeid vajaliku taseme saavutamiseks	Andmekvaliteedi parendusplaan RIHA jaoks toodud dokumendi lisas
	<b>Andmekvaliteedi tagamise juhend andmekogude omanikele, mis sisaldab:</b>	„Andmekvaliteedi tagamise juhend andmekogude omanikule“ juhend
	Andmekogu kvaliteedi hindamise meetodika juhendit, mis võimaldab hinnata ja põhjendatult kirjeldada olemasolevat andmete kvaliteeti, kirjeldada andmekogu andmekvaliteedi soovitava/vajaliku taseme ja planeerida meetmed ühelt tasemelt teisele üleminekuks. Meetodika peab sisaldama kvaliteeditasemete gradatsiooni (sh auditeerimist)	Hindamise ja parendamise meetodikad kirjeldatud juhendis.

	Andmete RIHA-s kirjeldamise mustreid, s.o kuidas ja milliste omaduste alusel tuleb andmed kirjeldada RIHA-s, et see toetaks andmete kvaliteedi juhtimist. RIHA-s rakendatavad mustrid ja protseduurid andmekvaliteedi tagamiseks	Soovitused RIHA-s kirjeldamise jaoks antud „Korrigeerimine“ peatükis
	Näidismaterjali näidisandmekogu andmekvaliteedi taseme hindamisest, soovitava/vajaliku taseme kirjeldamisest ja üleminekutööde planeerimisest. Dokument peab olema esitatud õppematerjali vormis metoodilise materjalina koos teadmisekontrolli testidega ja/või kontrollküsimustega	Juhend on täiendatud andmekogudes läbiviidud uuringute näidete varal. Tunnuste nõuete väljaselgitamise ja tagamise meetmete osas on kirjeldatud igale tunnusele kontrollküsimused.
	Andmekvaliteedi tõstmise organisatsioonilised soovitused (organisatsiooniline struktuur, vajalikud protsessid)	Küpsusmudel pakub vajaliku raamistiku andmekvaliteedi parendamiseks ka organisatsioonilistest aspektidest.
<b>Andmekvaliteedi tagamise juhend Riigi Infosüsteemi Ametile ja riigile üldisemalt, mis sisaldab:</b>		
	Riigis (keskselt) andmekogude kvaliteedi juhtimise seire ja järelevalve meetrika ning selle meetrika kasutamise metoodika, st milliste näitajate alusel ja kuidas tuleb riigis andmekvaliteeti kontrollida	Seire ja järelevalve meetrika ja metoodika on kirjeldatud riigi juhendis vastavanimelises peatükis.
	Rakendusjuhis loodud andmekvaliteedi taseme tõstmise metoodika rakendamiseks riigis (juhised RIA-le)	Rakendusettepanekud andmekvaliteedi taseme tõstmiseks on toodud riigi juhendis eraldi peatükina.
	Soovitused seadusandluse muutmiseks	Soovitused seadusandluse muutmiseks on kirjeldatud riigi juhendisse eraldi peatükina.
	Realistlik ajakava loodud metoodika rakendamiseks riigis	Juhendi rakendamise ajakava osas on toodud vastavasisulised soovitused iga rakendusettepaneku juures.
<b>Seminar andmekvaliteedi juhendi tutvustamiseks</b>		Juhendit tutvustav seminar viidi läbi andmekogude omanikele 18.-ndal augustil.

---

# *Andmekvaliteedi tagamise uuring*

Andmekvaliteedi tagamise uuringu eesmärk oli välja töötada andmekvaliteedi tagamise juhend andmekogu omanikele. Uuringu ülesandeks oli välja töötada meetodikad andmekvaliteedi mõõtmiseks ja tagamiseks ning luua struktureeritud juhend nende tegevuste rakendamiseks.

Uuringu tulemina sündis andmekvaliteedi tagamise juhendi esmane versioon, mille rakendatavust testiti andmekvaliteedi uuringus andmekogudes (vt Andmekvaliteedi uuring andmekogudes).

## *Andmekvaliteedi tunnuste uuring*

Andmekvaliteedi tunnuste uuringu eesmärgiks oli määratleda universaalne andmekvaliteedi tunnustekomplekt, mida on võimalik rakendada kvaliteedi mõõtmiseks ja tagamiseks igas riigi infosüsteemi andmekogus.

Tunnustekomplekti koostamiseks viidi läbi internetipõhine uuring leidmaks erinevaid tunnuseid käsitlevad andmekvaliteedi raamistikke (tunnuste allikaid). Allikatest kogutud tunnuste informatsiooni baasil sünteesiti tunnuste nimetusi ja definitsioone ning valiti välja tunnused, mis on aluseks andmekvaliteedi mõõtmiseks ja tagamiseks.

Komplektis sisalduvate tunnuste osas viidi läbi andmekvaliteedi nõuete uuring ja analüüs, et tuvastada ja analüüsida kehtivates õigusaktides ja regulatsioonides iga tunnuse andmekvaliteedi nõuded.

Andmekvaliteedi tunnuste uuringu tulemusena valiti välja andmekvaliteedi tunnuste komplekt, tuvastati andmekogude ülesed andmekvaliteedi nõuded, kirjeldati soovituslikud tagamise meetmed ja loodi kontrollküsimused, mis aitavad andmekogu omanikke nõuete väljaselgitamisel ja tagamise meetmete väljatöötamisel.

Järgnevalt on kirjeldatud andmekvaliteedi tunnuste uuringu tegevusi ja tulemusi detailsemalt.

## *Tunnuste allikad*

Tunnuste allikate tuvastamiseks viidi läbi internetipõhine uuring, et leida avalikust inforuumist kirjeldusi hüpoteetiliste või praktilist rakendust leidnud andmekvaliteedi raamistike osas, mis sisaldaksid defineeritud tunnuseid andmekvaliteedi mõõtmiseks.

Uuringu tulemusena vaadeldi ca. 20 leitud andmekvaliteedi raamistiku kirjeldusi, milles sisaldasid andmekvaliteedi standardid, andmekvaliteedi kasutuspraktikad rahvusvahelistes pankades ja ettevõtetes, andmekvaliteedi juhtimispraktikad eri riikide riigiasutustes, valdkonna ekspertide avaldatud artiklid ja blogipostitused üldtunnustatud teaduslikes ja informaatika-alastes veebi-väljaannetes.

Vaadeldud raamistikest käsitlesid andmekvaliteedi tunnuseid 12 andmekvaliteedi raamistikku (vt Lisa 1: Tunnuste allikate loetelu), milles sisalduvad tunnused ja nende definitsioonid võeti aluseks tunnuste komplekti valikul.

## Tunnuste komplekti valik

Tunnuste komplekti valiku ülesandeks oli valida vaadeldavates allikates esinevate kvaliteedi tunnuste seast selline komplekt tunnuseid, mis vastaks uuringu juhtrühmaga kokku lepitud kriteeriumitele. Valiku kriteeriumid olid järgnevad:

5. tunnuste komplekt peab katma andmekvaliteedi valdkonna;
6. tunnused peavad kompleksis olema unikaalsed (ei dubleeri üksteist);
7. komplekti kuuluvad tunnused on olulised;
8. komplekti kuuluvad tunnused on mõõdetavad;
9. komplekti kuuluvad tunnused peavad andmekogu omaniku poolt olema mõjutatavad.

Vaadeldud 12-s allikas käsitleti kokku 108-t andmekvaliteedi tunnust. Tunnuste ja nende definitsioonide osas viidi läbi tunnuste nimede süntaktiline ja nende definitsioonide semantiline analüüs, mille tulemusena selgus, et 108-liikmelises tunnuste kompleksis leidis:

1. süntaktiliselt sama nimega tunnuseid, millede definitsioonid olid semantiliselt sarnased või samad;
2. süntaktiliselt sama nimega tunnuseid, millede definitsioonid olid semantiliselt erinevad;
3. süntaktiliselt eri nimega tunnuseid, millede definitsioonid olid semantiliselt sarnased või samad;
4. süntaktiliselt eri nimega tunnuseid, millede definitsioonid olid semantiliselt erinevad (unikaalsed tunnused teistega võrreldes).

Tulenevalt tunnuste komplekti valikute kriteeriumitest, ei tohtinud tunnuste kompleksis esineda tunnuste osas dubleerimist. Seega tuli olemasolev tunnuste nimekiri sünteesida, et tekitada olukord, kus vaadeldavas tunnuste nimekirja jääks alles vaid süntaktiliselt unikaalse nime ja semantiliselt unikaalse definitsiooniga tunnused. Selleks teostati nende tunnuste osas, kus sellist unikaalsust polnud tagatud, konsolideerimine ja koondati sellised tunnused ühise koondnimetaja alla (vt Lisa 2: Tunnuste sünteesi tabel).

Konsolideeringu tulemusena jäi nimekirja alles 53 unikaalset tunnust. Tunnused olid nimekirjas läbiseegi eesti- ja inglisekeelsed. Pärast inglisekeelsete tunnuste tõlkimist eesti keelde, analüüsiti uuesti läbi tunnuste nimed ja nende algsed definitsioonid semantilisest aspektist, et veenduda nende unikaalsuses ka tõlgete järel. Selgus, et nimekirjas sisaldas palju selliseid tunnuseid, mille osas tuli läbi viia täiendav konsolideerimine (nt. 2 tunnust oli inglisekeelsena süntaktiliselt unikaalsed, kuid eestikeelsena tekkis nimekirja 2 süntaktiliselt sama tunnust: accuracy->täpsus, precision->täpsus). Seega tuli unikaalsuse tagamiseks läbi viia täiendav süntees, mille raames osad tunnused konsolideeriti täiendavalt või käsitleti unikaalsetena.

Sünteeside tulemusena jõuti 28-st eestikeelsete nimetustega unikaalsetest tunnustest koosneva komplektini, mis on kirjeldatud tabelis Tabel 4: Tõlgitud tunnuste tabel. Tunnuste olulisuse kriteeriumi täitmiseks vaadeldake tunnuste esinemist vaadeldud 12-s raamistikus. Tunnuste esinemiste arv allikates indikeerib kuivõrd sagedasti on raamistikes seda tunnust andmekvaliteedi mõõtmisel või tagamisel käsitletud. Esinemiste arvu allikates saab tõlgendada ka kui tunnuse olulisuse näitajat.

Tabel 4: Tõlgitud tunnuste tabel

Nr	Eestikeelne nimetus	Esinemine allikas												Esinemiste arv allikates kokku
		1	2	3	4	5	6	7	8	9	10	11	12	
1	Täpsus	x	x	x	x	x	x		x	x	x	x	x	11
2	Täielikkus	x	x		x	x	x	x	x	x		x	x	10
3	Metoodiline järjepidevus	x	x		x	x	x		x	x			x	8
4	Kättesaadavus	x	x	x		x	x	x					x	7
5	Usaldusväarsus	x	x			x		x			x	x	x	7
6	Ajakohasus	x	x		x		x	x	x				x	7
7	Arusaadavus	x		x	x	x		x	x					6
8	Korrektus				x	x	x	x	x		x			6
9	Õigeaegsus			x		x				x	x	x		5

10	Reeglipärasus	x	x							x			x	4
11	Detailsus	x					x				x	x		4
12	Asjakohasus		x	x				x					x	4
13	Konfidentsiaalsus	x						x					x	3
14	Töödeldavus		x						x					3
15	Unikaalsus						x		x	x				3
16	Objektiivsus								x				x	2
17	Ülekantavus	x				x								2
18	Vormisobivus					x								1
19	Koherentsus				x									1
20	Võrreldavus				x									1
21	Sisutihesus								x					1
22	Vormiline järjepidevus								x					1
23	Tõhusus	x												1
24	Osaline sobivus						x							1
25	Vormipaindlikkus					x								1
26	Olemikordumatus										x			1
27	Taastatavus	x												1
28	Jälgitavus	x												1

Tõlgitud tunnuste tabeli abil tehti uuringu juhtrühmale ettepanek lõplikuks tunnustekomplektiks. Uuringu juhtrühmas hinnati tabeli abil täiendavalt tunnuste mõõdetavust, mõjutatavust ja valdkonna kaetust. Hindamise tulemusena jõuti 9-liikmelise tunnustekomplektini, mida otsustati võtta aluseks riigi infosüsteemi andmekogudes andmekvaliteedi hindamisel. Valitud tunnustekomplekt on kirjeldatud tabelis Tabel 5: Andmekvaliteedi tunnuste komplekt.

Tabel 5: Andmekvaliteedi tunnuste komplekt

Tunnus	Definitsioon
<b>Õigsus</b>	Andmete õigsus iseloomustab seda, kui võrd andmeatribuudid esitavad spetsiifilises kasutuskontekstis korrektselt mõiste või sündmuse tõeseid väärtusi.
<b>Täielikkus</b>	Andmete täielikkus iseloomustab seda, kui võrd on väärtused olemas kõikide olemite kõigil andmeatribuutidel, mille puhul väärtuse olemasolu nõutakse.
<b>Kooskõla</b>	Andmete kooskõla iseloomustab seda, millisel määral andmed on vabad vasturääkivustest ja kooskõlalised teiste andmetega.
<b>Usaldusväarsus</b>	Andmete usaldusväarsus iseloomustab seda, kui võrd andmetel on atribuudid, mida kasutajad spetsiifilises kasutuskontekstis võivad pidada tõesteks ja usutatavaks.
<b>Ajakohasus</b>	Andmete ajakohasus iseloomustab seda, kui võrd ajaline viide andmeid tekitanud sündmusest vastab vajadustele ja nõuetele.
<b>Konfidentsiaalsus</b>	Andmete konfidentsiaalsus on andmete kättesaadavus ainult selleks volitatud tarbijaile (isikutele või tehnilistele süsteemidele) ning kättesaamatus kõigile ülejäänutele.
<b>Ühekordsus</b>	Andmete ühekordsus iseloomustab seda, kui võrd on andmeesitajatelt andmete kogumisel järgitud riigis kehtestatud ühekordse kogumise põhimõtteid.
<b>Mitteliiasus</b>	Mitteliiasus iseloomustab seda, kui võrd andmekogu kavandamisel ja arenduses on välditud liiasuste tekkimist andmekogu andmestruktuurides ning kui võrd välditakse liiasust andmetöötlusprotsessides.
<b>Reeglipärasus</b>	Andmete reeglipärasus iseloomustab seda, mil määral andmetel on atribuudid, mis vastavad spetsiifilises kasutuskontekstis andmekvaliteeti puudutavatele standarditele, kokkulepetele või kehtivatele õigusaktidele ning samalaadsetele normidele.

## Andmekvaliteedi nõuete analüüs

Andmekvaliteedi nõuete analüüsi eesmärgiks oli välja selgitada need andmekvaliteedi nõuded, mille tagatuse korral on võimalik väita, et andmekogu andmed on kvaliteetsed. Analüüsi ülesandeks oli tuvastada need regulatsioonid, mis seavad iga tunnuse osas need üldised nõuded, mis rakenduvad igale riigi infosüsteemi andmekogule.

Analüüsi käigus arvestati nõuete väljaselgitamisel järgnevate õigusaktide ja regulatsioonidega:

1. Eesti Vabariigi üldised seadused, vabariigi valitsuse määrused ja muud dokumendid:
  - a. Avaliku teabe seadus (AvTS)
  - b. Isikuandmete kaitse seadus (IKS)
  - c. Riigisaladuse ja salastatud välisteabe seadus
  - d. Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsioon
  - e. Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsiooni järelevalveasutusi ja andmete liikumist üle piiri käsitlev lisaprotokoll
  - f. Riigisaladuse ja salastatud välisteabe kaitse kord
  - g. Riigi Infosüsteemi Ameti põhimäärus
2. Andmekogusid puudutavad seadused ja määrused:
  - a. [Infosüsteemide turvameetmete süsteem](#). Vabariigi Valitsuse määrus.
  - b. [Infosüsteemide kolmeastmeline etalonturbe süsteem \(ISKE\)](#).
3. Muud riigi juhtimisega seotud materjalid:
  - a. Avaandmete roheline raamat; ([https://opendata.riik.ee/sites/default/files/manuals/avaliku-teabe-masinloetava-avalikustamise-roheline-raamat-20141125\\_0.odt](https://opendata.riik.ee/sites/default/files/manuals/avaliku-teabe-masinloetava-avalikustamise-roheline-raamat-20141125_0.odt)) (26.02.2016)
  - b. Avaandmete loomise ja avaldamise juhend. Riigi Infosüsteemi Amet. Versioon 1.0, Tallinn 2015 ([https://www.ria.ee/public/Arhitektuur/Avaandmete\\_loomise\\_juhend\\_v1.0.pdf](https://www.ria.ee/public/Arhitektuur/Avaandmete_loomise_juhend_v1.0.pdf)) (26.02.2016)
  - c. Riigi infosüsteemi koosvõime raamistik (<https://www.mkm.ee/et/riigi-infosusteemi-koosvoime-raamistik>) (26.02.2016)
  - d. Riigi infosüsteemide semantilise koosvõime raamistik (<https://www.mkm.ee/et/riigi-infosusteemi-koosvoime-raamistik>) (26.02.2016)
  - e. Veebide koosvõime raamistik (<https://www.mkm.ee/et/riigi-infosusteemi-koosvoime-raamistik>) (26.02.2016)
  - f. RIHA andmete ajakohasuse seire. Riigi infosüsteemi amet ([https://www.ria.ee/public/RIHA/RIHA\\_andmete\\_ajakohasuse\\_seire.pdf](https://www.ria.ee/public/RIHA/RIHA_andmete_ajakohasuse_seire.pdf)) (26.02.2016)
  - g. Uurimisprojekti „Andmeaitade (teiseste andmekogude) loomise põhimõtete väljatöötamine“ lõpparuanne, 2013 ([https://www.riigikantselei.ee/valitsus/valitsus/et/riigikantselei/strateegia/poliitika-analuusid-ja-uuringud/tarkade-otsuste-fondi-uuringute-kokkuvotted/Andmeaidad\\_1%C3%B5pparuanne.pdf](https://www.riigikantselei.ee/valitsus/valitsus/et/riigikantselei/strateegia/poliitika-analuusid-ja-uuringud/tarkade-otsuste-fondi-uuringute-kokkuvotted/Andmeaidad_1%C3%B5pparuanne.pdf)) (26.02.2016)
  - h. andmekooslusi kehtestavad määrused (nt. Pagulase reisidokumendi vormi, tehnilise kirjelduse ja reisidokumendi kantavate andmete loetelu kehtestamine )
  - i. andmete esitamise ja andmetega toimigute tegemise korrad /määrused. (nt. Pensionistaaži andmete kogumise ja riiklikku pensionikindlustuse registrisse kandmise kord)

Lisaks ülaltoodud üldistele õigusaktidele tuleb iga andmekogu puhul arvestada spetsiifilisi õigusakte (nt andmekogu põhimäärus või valdkonda puudutavad õigusaktid), mis sätestavad konkreetseid kvaliteedinõudeid antud andmekogu jaoks. (Näiteks Rahvastikuregistri seadus, Ruumiandmete seadus, Vabariigi Valitsuse määrus "Aadressiandmete süsteem", RIHA põhimäärus).

Nõuete analüüsi tulemusena sündis iga tunnuse osas:

- loetelu vastava tunnusega seotud andmekogude ülestest andmekvaliteedi nõuetest, millega iga andmekogu omanik peab arvestama;
- loetelu vastava tunnusega seotud soovituslikest andmekvaliteedi tagamise meetmetest;
- andmekvaliteedi nõuete kontrollküsimused, mis aitavad andmekogu omanikke andmekvaliteedi nõuete tuvastamisel ja tagamise meetmete väljatöötamisel.



## Andmekvaliteedi nõuded

Järgnevat esitatakse välja toodud iga tunnuse osas analüüsitud andmekogude ülesed andmekvaliteedi nõuded.

### Õigsus

AvTS § 9 lg 8 kehtestab teabevaldaja kohustuse mitte anda teadvalt eksitavat, tegelikkusele mittevastavat või ebaõiget teavet ning kontrollida kahtluse korral väljastatava teabe õigsust ja vastavust tegelikkusele.

IKS § 6 lõige 5 sätestab andmete kvaliteedi põhimõtte isikuandmete töötlemisel – isikuandmed peavad olema ajakohased, täielikud ning vajalikud seatud andmetöötluse eesmärgi saavutamiseks.

IKS § 24 sätestab, et isikuandmete töötleja on isikuandmete töötlemisel kohustatud: eesmärkide saavutamiseks mittevajalikud isikuandmed viivitamata kustutama või sulgema, kui seadus ei näe ette teisiti; tagama, et isikuandmed on õiged ja, kui see on eesmärkide saavutamiseks vajalik, viimases seisus; mittetäielikud ja ebaõiged isikuandmed sulgema ning nende täiendamiseks ja parandamiseks võtma viivitamata kasutusele vajalikud abinõud; ebaõiged andmed säilitama märkusega nende kasutamise aja kohta koos õigete andmetega; isikuandmed, mille õigsus on vaidlustatud, sulgema kuni andmete õigsuse kindlakstegemiseni või õigete andmete väljaselgitamiseni; isikuandmete parandamise korral viivitamata teavitama sellest kolmandaid isikuid, kellelt isikuandmed saadi või kellele isikuandmeid edastati, kui see on tehniliselt võimalik ega too kaasa ebaoproportsionaalselt suuri kulutusi.

Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsiooni artikkel 5 sätestab, et automatiseeritud töödeldavad isikuandmed peavad olema õiged ja vajadusel täiendatavad.

IKS seletuskiri lisab, et ühest küljest väljendab see põhimõte, et kui juba isikuandmeid töödeldakse, siis on soovitatav seda teha korrektselt. Põhjendamatu oleks isikuandmete töötlemisega sekkumine isiku eraellu, kui kas kogutud andmete ebapiisavuse või nende ebatäpsuse või mitteajakohasuse tõttu jääks saavutamata see oluline töötlemise eesmärk, mis sellist isiku privaatsfääri sekkumist õigustas. Teisest küljest piirab see põhimõte aga isikuandmete kogumist suuremas ulatuses, kui on vältimatult vajalik andmetöötluse eesmärgi saavutamiseks. Andmete kvaliteedi põhimõtte sisaldab ka kohustust töödelda andmesubjekti tuvastamist võimaldavaid isikuandmeid üksnes juhul, kui see on vajalik seoses andmete kogumise või hilisema töötlemise eesmärkidega. Erandi saab siiski teha teaduse ja statistika vajadusi silmas pidades.

Vabariigi Valitsuse määruse "[Infosüsteemide turvameetmete süsteem](#)" § 6 sätestab, et nõutav turvatase määratakse vastavalt infoturbe eesmärkidele tervikluse, konfidentsiaalsuse ja käideldavuse parameetrite kaudu. Andmete terviklus on andmete õigsuse, täielikkuse ja ajakohasuse tagatus ning päritolu autentsus ja volitamata muutuste puudumine.

Seega tulevad õigsuse nõuded kaudselt sisse andmete turvaklassi määramisel - tuleb arvestada seadustest ja lepingutest tulenevaid nõudeid, põhitegevuse protsessidest tulenevaid nõudeid ning tagajärgede kaalukust, sealhulgas õigsuse osas. Eelkõige ongi ISKE ning selle aluseks oleva *IT-Grundschutzhandbuch* kataloogides andmete tervikluse all silmas peetud andmete õigsuse ja täielikkuse tagatust ning päritolu autentsust ja volitamata muutuste puudumist.

### Täielikkus

IKS § 6 lõige 5 sätestab andmete kvaliteedi põhimõtte isikuandmete töötlemisel – isikuandmed peavad olema ajakohased, täielikud ning vajalikud seatud andmetöötluse eesmärgi saavutamiseks.

IKS § 24 sätestab muuhulgas, et isikuandmete töötleja on isikuandmete töötlemisel kohustatud mittetäielikud ja ebaõiged isikuandmed sulgema ning nende täiendamiseks ja parandamiseks võtma viivitamata kasutusele vajalikud abinõud ning isikuandmete parandamise korral viivitamata teavitama sellest kolmandaid isikuid,

kellelt isikuandmed saadi või kellele isikuandmeid edastati, kui see on tehniliselt võimalik ega too kaasa ebaproportsionaalselt suuri kulutusi.

Vabariigi Valitsuse määruse "Infosüsteemide turvameetmete süsteem" § 6 sätestab, et nõutav turvatase määratakse vastavalt infoturbe eesmärkidele tervikluse, konfidentsiaalsuse ja käideldavuse parameetrite kaudu. Andmete terviklus on andmete õigsuse, täielikkuse ja ajakohasuse tagatus ning päritolu autentsus ja volitamata muutuste puudumine.

Seega tulevad täielikkuse nõuded kaudselt sisse andmete turvaklassi määramisel - tuleb arvestada seadustest ja lepingutest tulenevaid nõudeid, põhitegevuse protsessidest tulenevaid nõudeid ning tagajärgede kaalukust, sealhulgas täielikkuse osas. Eelkõige ongi ISKE ning selle aluseks oleva *IT-Grundschutzhandbuch* kataloogides andmete tervikluse all silmas peetud andmete õigsuse ja täielikkuse tagatust ning päritolu autentsust ja volitamata muutuste puudumist.

## Kooskõla

Vabariigi Valitsuse määruse "Riigi infosüsteemi haldussüsteem" §-s 5 sätestatakse ühtsuse põhimõtte riigi infosüsteemi haldamisel – riigi infosüsteemi kuuluvad andmekogud peavad olema omavahel ühildatavad ning olema võimelised koostoitumiseks, andmevahetuseks ja andmete kontrolliks.

Vabariigi Valitsuse määruse "Riigi infosüsteemi haldussüsteem" §-s 5 sätestatakse, et andmekogu asutatakse ja selle muutmine toimub, lähtudes riigi IT koosvõime raamistikest. Vastavalt IT koosvõime raamistikule on semantiline koosvõime organisatsioonide võime mõista vahetatud informatsiooni (andmete) tähendust ühtemoodi. Koosvõime raamistiku juurde kuulavas dokumentatsioonis kirjeldatakse infosüsteemi semantikavarad ja andmete semantilise rikastamise reeglid, sealhulgas:

- Infosüsteemid peavad toetama mitmekeelseid ja rahvusvahelisi semantikavarasid.
- Semantilises ja tehnilises mõõtmes reguleeritakse koosvõimet mitmesuguste kokkulepete, standardite või soovitusetega.
- Iga infosüsteemi koosvõime arhitektuuri dokument peab sisaldama semantilist koosvõimet käsitlevat osa.
- Osapooltel on vabadus valida, milliseid töövahendeid ja standardeid nad semantikavarade loomisel ja infovarade semantilisel rikastamisel kasutavad. Semantikavarad ja annoteeritud infovarad peab avaldama RIHAs semantika juhise nõuete kohaselt.

Dokumentides "Riigi infosüsteemi koosvõime raamistik", "Riigi infosüsteemide semantilise koosvõime raamistik" ja "Veebide koosvõime raamistik" on põhjalikult esitatud riigi infosüsteemi koosvõime, sealhulgas semantilise koosvõime põhimõtted.

Vastavalt AvTS § 3<sup>1</sup> lg 4 võimaldab, kui see on võimalik ja asjakohane, teabevaldaja juurdepääsu avaandmetele failivormingus, mis on struktureeritud selliselt, et tarkvararakendused suudavad spetsiifilisi andmeid, sealhulgas üksikuid faktiväiteid, ja nende sisemist struktuuri kergesti tuvastada, ära tunda ja välja lugeda (edaspidi masinloetav kuju), ning platvormist sõltumatus vormingus, mis tehakse üldsusele kättesaadavaks dokumendi taaskasutamise piiranguteta (edaspidi avatud vorming). Kui avaandmete digitaalsele kujule, masinloetavale kujule või avatud vormingusse viimine ei ole võimalik või see nõuab ebaproportsionaalselt suuri pingutusi, võimaldab teabevaldaja avaandmetele juurdepääsu nende algkujul või mis tahes muus vormis.

Eesti avaliku teabe masinloetava avalikustamise roheline raamat toob ära avatud riigi töörühma poolt defineeritud nõuded, mille täidetuse korral võib andmehulga andmeid lugeda avaandmeteks. Üks nõuetest on masinloetavus - andmed on mõistetava struktuuriga ja automaatselt töödeldavad. Ühilduvusega võib seostada ka avatud standardite kasutamise nõuet - andmed on esitatud avatud formaadis, mis ei ole ühegi ettevõtte ega isiku ainuomand. Tuuakse ära avaandmete kirjeldamiseks soovitatavad avatud standardid. Formaaside ja kodeeringute kasutajasõbralikkuse hindamisel lähtutakse sir Timothy Berners-Lee viie täрни süsteemi põhimõtetest.

Avaandmete loomise ja avaldamise juhend toob ära ühilduvusega seotud nõuded, millest tuleb lähtuda avaandmete planeerimisel, projekteerimisel, loomisel ja avaldamisel, sealhulgas esitusviisi ja mõistetavuse. Pakutakse andmekogumite esitlusvormingu, vorminduse ja andmete semantilise mõistetavuse põhimõtted.

## Usaldusväärsus

Avaliku teabe seaduse § 435 sätestab, et andmekogu põhimääruses sätestatakse andmekogu pidamise kord, sealhulgas andmekogu vastutav töötaja (haldaja), andmekogusse kogutavate andmete koosseis, andmeandjad ja vajaduse korral muud andmekogu pidamisega seotud korralduslikud küsimused. Andmeandjaks on riigi- või kohaliku omavalitsuse asutused või muud avalik-õiguslikud või eraõiguslikud isikud, kui neil on seadusega või selle alusel antud õigusaktiga sätestatud kohustus andmekogusse andmeid esitada või kui nad teevad seda vabatahtlikult. Nende andmeandjate usaldusväärsus võib erineda.

Avaliku teabe seaduse § 436 sätestab, et andmete töötlemisel, mida kogub põhianndmetena teine riigi infosüsteemi kuuluv andmekogu, tuleb aluseks võtta vastava teise andmekogu põhianndmed. Kuna andmete usaldusväärsus erinevates andmekogudes võib erineda, on soovitatav analüüsida, kas teise andmekogu andmete usaldusväärsus on piisav.

IKS § 25 lõige 3 sätestab, et isikuandmete töötaja on isikuandmete töötlemisel kohustatud tagama, et tagantjärele oleks võimalik kindlaks teha, millal, kelle poolt ja milliseid isikuandmeid salvestati, muudeti või kustutati või millal, kelle poolt ja millistele isikuandmetele andmetöötlussüsteemis juurdepääs saadi. IKS § 25 lõige 5 sätestab, et isikuandmete töötaja on isikuandmete töötlemisel kohustatud tagama andmete olemasolu isikuandmete edastamise kohta: millal, kellele ja millised isikuandmed edastati, samuti selliste andmete muutusteta säilimise.

Vabariigi Valitsuse määruse "Infosüsteemide turvameetmete süsteem" § 6 sätestab, et nõutav turvatase määratakse vastavalt infoturbe eesmärkidele tervikluse, konfidentsiaalsuse ja käideldavuse parameetrite kaudu. Andmete terviklus on andmete õigsuse, täielikkuse ja ajakohasuse tagatus ning päritolu autentsus ja volitamata muutuste puudumine.

Seega tulevad usaldusvääruse nõuded (päritolu autentsus) kaudselt sisse andmete turvaklassi määramisel - tuleb arvestada seadustest ja lepingutest tulenevaid nõudeid, põhitegevuse protsessidest tulenevaid nõudeid ning tagajärgede kaalukust ka usaldusvääruse, sealhulgas autentsuse osas.

## Ajakohasus

IKS § 6 lõige 5 sätestab andmete kvaliteedi põhimõtte isikuandmete töötlemisel – isikuandmed peavad olema ajakohased, täielikud ning vajalikud seotud andmetöötluse eesmärgi saavutamiseks.

IKS § 24 sätestab isikuandmete töötlemise nõuded, mille järgi isikuandmete töötaja on isikuandmete töötlemisel kohustatud sh tagama, et isikuandmed on õiged ja, kui see on eesmärkide saavutamiseks vajalik, viimases seisus.

AvTS § 12 "Nõuded dokumendiregistrile" lg 1 sätestab nõuded dokumentide registreerimisele, nt dokumendiregistris registreeritakse asutusele saabunud dokumendid ja asutusest väljasaadetud dokumendid hiljemalt saabumise või väljasaatmise päevale järgneval tööpäeval.

Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsiooni artikkel 5 sätestab, et automatiseeritud töödeldavad isikuandmed peavad olema säilitatud vormis, mis lubab andmesubjekti teha kindlaks mitte kauemaks, kui perioodiks, mis on vajalik andmete kogumise eesmärkide saavutamiseks.

Vabariigi Valitsuse määruse "Infosüsteemide turvameetmete süsteem" § 6 sätestab, et nõutav turvatase määratakse vastavalt infoturbe eesmärkidele tervikluse, konfidentsiaalsuse ja käideldavuse parameetrite kaudu. Andmete terviklus on andmete õigsuse, täielikkuse ja ajakohasuse tagatus ning päritolu autentsus ja volitamata muutuste puudumine.

Seega tulevad ajakohasuse nõuded kaudselt sisse andmete turvaklassi määramisel - tuleb arvestada seadustest ja lepingutest tulenevaid nõudeid, põhitegevuse protsessidest tulenevaid nõudeid ning tagajärgede kaalukust, sealhulgas ajakohasuse osas. Samas on ISKE ning selle aluseks oleva IT-Grundschutzhandbuch kataloogides andmete tervikluse all silmas peetud pigem andmete õigsuse ja täielikkuse tagatust ning päritolu autentsust ja volitamata muutuste puudumist - andmete ajakohasust käsitlevad meetmed puuduvad.

## Reeglipärasus

Vastavalt Eesti Vabariigi põhiseaduse §-le 3 teostatakse riigivõimu põhiseaduse ja sellega kooskõlas olevate seaduste alusel, kusjuures rahvusvahelise õiguse üldtunnustatud põhimõtted ja normid on Eesti õigussüsteemi lahutamatu osa ning täitmiseks kohustuslikud on üksnes avaldatud seadused.

Seega lähtub andmete reeglipärasus eelkõige järgmistest riigi infosüsteemi reguleerivatest õigusaktidest ning nende alusel välja antud dokumentidest.

1. Eesti Vabariigi üldised seadused, vabariigi valitsuse määrused ja muud dokumendid:
  - a. Avaliku teabe seadus (AvTS)
  - b. Isikuandmete kaitse seadus (IKS)
  - c. Riigisaladuse ja salastatud välisteabe seadus
  - d. Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsioon
  - e. Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsiooni järelevalveasutusi ja andmete liikumist üle piiri käsitlev lisaprotokoll
  - f. Riigisaladuse ja salastatud välisteabe kaitse kord
  - g. Riigi Infosüsteemi Ameti põhimäärus
2. Andmekogusid puudutavad seadused ja määrused:
  - a. [Infosüsteemide turvameetmete süsteem](#). Vabariigi Valitsuse määrus.
  - b. [Infosüsteemide kolmeastmeline etalonturbe süsteem \(ISKE\)](#).
3. Muud riigi juhtimisega seotud materjalid:
  - a. Avaandmete roheline raamat; ([https://opendata.riik.ee/sites/default/files/manuals/avaliku-teabe-masinloetava-avalikustamise-roheline-raamat-20141125\\_0.odt](https://opendata.riik.ee/sites/default/files/manuals/avaliku-teabe-masinloetava-avalikustamise-roheline-raamat-20141125_0.odt)) (26.02.2016)
  - b. Avaandmete loomise ja avaldamise juhend. Riigi Infosüsteemi Amet. Versioon 1.0, Tallinn 2015 ([https://www.ria.ee/public/Arhitektuur/Avaandmete\\_loomise\\_juhend\\_v1.0.pdf](https://www.ria.ee/public/Arhitektuur/Avaandmete_loomise_juhend_v1.0.pdf)) (26.02.2016)
  - c. Riigi infosüsteemi koosvoime raamistik (<https://www.mkm.ee/et/riigi-infosusteemi-koosvoime-raamistik>) (26.02.2016)
  - d. Riigi infosüsteemide semantilise koosvoime raamistik (<https://www.mkm.ee/et/riigi-infosusteemi-koosvoime-raamistik>) (26.02.2016)
  - e. Veebide koosvoime raamistik (<https://www.mkm.ee/et/riigi-infosusteemi-koosvoime-raamistik>) (26.02.2016)
  - f. RIHA andmete ajakohasuse seire. Riigi infosüsteemi amet ([https://www.ria.ee/public/RIHA/RIHA\\_andmete\\_ajakohasuse\\_seire.pdf](https://www.ria.ee/public/RIHA/RIHA_andmete_ajakohasuse_seire.pdf)) (26.02.2016)
  - g. Uurimisprojekti „Andmeaitade (teiseste andmekogude) loomise põhimõtete väljatöötamine“ lõpparuanne, 2013 ([https://www.riigikantselei.ee/valitsus/valitsus/et/riigikantselei/strateegia/poliitika-analuusid-ja-uuringud/tarkade-otsuste-fondi-uuringute-kokkuvotted/Andmeaidad\\_1%C3%B5pparuanne.pdf](https://www.riigikantselei.ee/valitsus/valitsus/et/riigikantselei/strateegia/poliitika-analuusid-ja-uuringud/tarkade-otsuste-fondi-uuringute-kokkuvotted/Andmeaidad_1%C3%B5pparuanne.pdf)) (26.02.2016)
  - h. andmekooslusi kehtestavad määrused (nt. Pagulase reisidokumendi vormi, tehnilise kirjelduse ja reisidokumendi kantavate andmete loetelu kehtestamine )
  - i. andmete esitamise ja andmetega toimigute tegemise korrad /määrused. (nt. Pensionistaaži andmete kogumise ja riiklikku pensionikindlustuse registrisse kandmise kord)

## Konfidentsiaalsus

Üldjuhul on riigi ja kohaliku omavalitsuse andmekogude konfidentsiaalsuse nõuded välja selgitatud seoses ISKE rakendamisega.

Avaliku teabe seadus, isikuandmete kaitse seadus ja mitmed muud õigusaktid kehtestavad olulisi kitsendusi andmetele ligipääsuks. Üldised põhimõtted kehtestavad AvTS § 3<sup>1</sup> lg 3: "Teabe üldiseks kasutamiseks andmisel peab olema tagatud isiku eraelu puutumatus, autoriõiguste kaitse, riigi julgeoleku kaitse, ärisaladuse ja muu juurdepääsupiiranguga teabe kaitse. Enne teabe üldiseks kasutamiseks andmist peab teabevaldaja hindama teabe üldisele kasutamisele piirangute kehtestamise vajadust" ja AvTS § 4 lg 3: "Teabele juurdepääsu võimaldamisel peab olema tagatud isiku eraelu puutumatus ja autoriõiguste kaitse". AvTS peatükk 5 kehtestab nõuded piiratud juurdepääsuga teabe töötlemiseks.

IKS § 25 kehtestab isikuandmete organisatsioonilised, füüsilised ja infotehnilised turvameetmed muuhulgas ka andmetele juurdepääsu ja turvalisuse tagamise - andmete käideldavuse ja konfidentsiaalsuse - osas.

Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsiooni artikkel 7 sätestab, et automatiseeritud andmekogudes säilitatavate isikuandmete tahtmatu või tahtliku hävitamise, kaotsimineku, aga ka omavolilise juurdepääsu, muutmise või levitamise eest kaitsmiseks võetakse kasutusele kohased turvameetmed.

Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsiooni järelevalveasutusi ja andmete liikumist üle piiri käsitlev lisaprotokoll Artikkel 1 sätestab, et protokolliosalised näevad ette ühe või mitu järelevalveasutust, kelle ülesanne on tagada tema riigisisese õiguse kohaselt võetud meetmete järgimine, millega rakendatakse konventsiooni II ja III peatükis ning protokollis sätestatud põhimõtteid. Vastavalt Artiklile 2 sätestatakse isikuandmete piiriülene edastamine andmete vastuvõtjale, kes ei kuulu konventsiooniosalise jurisdiktsiooni alla, ainult sel juhul, kui see riik või organisatsioon tagab kavandatava andmeedastuse küllaldase kaitse.

Vabariigi Valitsuse määruse "Infosüsteemide turvameetmete süsteem" § 6 sätestab, et nõutav turvatase määratakse vastavalt infoturbe eesmärkidele tervikluse, konfidentsiaalsuse ja käideldavuse parameetrite kaudu. Sama määruse § 91 sätestab turvameetmete süsteemi rakendamise auditeerimise nõuded riigi infosüsteemi kuuluvate riigi andmekogude pidamisel, § 92 - turvameetmete süsteemi rakendamise auditeerimine kohaliku omavalitsuse riigi infosüsteemi kuuluvate andmekogude pidamisel. Sama määruse § 11 esitab turvameetmete süsteemi rakendamise auditeerimise tähtajad riigi infosüsteemi kuuluvate riigi andmekogude pidamisel.

## Ühekordsus

Avaliku teabe seaduse § 433 lg 2 järgi on keelatud asutada ühtede ja samade andmete kogumiseks eraldi andmekogusid.

Avaliku teabe seaduse § 436 lg 2 järgi tuleb andmete töötlemisel, mida kogub põhiandmetena teine riigi infosüsteemi kuuluv andmekogu, aluseks võtta vastava teise andmekogu põhiandmed.

Vabariigi Valitsuse määruse "Riigi infosüsteemi haldussüsteem" §-s 5 sätestatakse põhiandmete kasutamise põhimõtte riigi infosüsteemi haldamisel - andmed kogutakse võimalikult autentsest allikast ning koondatakse andmekogusse, mis on vastavate andmete suhtes ühtseks allikaks kõikidele riigi infosüsteemi kuuluvatele andmekogudele nende seadusekohaste ülesannete täitmisel.

Eesti avaliku teabe masinloetava avalikustamise rohelises raamatus (Majandus- ja Kommunikatsiooniministeerium, 2014) on välja toodud avatud riigi töörühma defineeritud nõuded, mille täidetuse korral võib andmehulga andmeid lugeda avaandmeteks sh pärinemine algallikast. Andmed on töötluseta kogutud algallikast, säilitades oma originaalkuju ja detailsuse. Nagu andmekogudegi puhul pole lubatud andmete võtmine teisest andmekogust. Teiste andmekogude andmeid on seega lubatud lisada andmehulka, kuid neid ei tohi kolmas osapool teist korda taaskasutada.

Riigi infosüsteemi koosvõime raamistik (2011) toob ühe koosvõime põhimõttena välja haldusprotsesside lihtsustamise. Sealjuures andmekogude dokumentatsiooni kooskõlastusel riigi infosüsteemi haldussüsteemis (RIHA) kontrollitakse andmete kogumise eesmärgipärasust ja andmete ühekordse küsimise põhimõtte järgimist.

Infoühiskonna arengukava 2020 üheks meetmeks on paremate avalike teenuste arendamine IKT abil ning selleks arendatakse avalike teenuste osutamise protsesse ja neid toetavaid IKT-lahendusi (nt rahvastikuregister) nii, et kaotada andmete korduva esitamise vajadus: eri ametkonnad peavad isikute esitatud andmeid nende esitamise sihtotstarbe piires taas- ja riskasutama, juhul kui isikud seda ei keela.

## Mitteliiasus

Liiasus võib kaasa tuua vastuolud andmetes, ebaefektivse andmetöötluse, ebavajaliku andmekogumise ja muid ebasoovitavaid tagajärgi. Mitteliiasus ongi ebasoovitav pigem oma kaudsete tagajärgede tõttu ja seepärast seda õigusaktides enamasti otseselt välja toodud ei ole. Ka ühekordsuse põhimõtet võib vaadata kui andmekogude vahelise mitteliiasuse nõude ühte näidet. Kaudselt võib mitteliiasusega seostada muuhulgas järgnevaid õigusakte ja dokumente.

Avaliku teabe seaduse § 43<sup>5</sup> lg 1 järgi sätestatakse andmekogu põhimääruses andmekogu pidamise kord, sealhulgas andmekogu vastutav töötaja (haldaja), andmekogusse kogutavate andmete koosseis, andmeandjad ja vajaduse korral muud andmekogu pidamise seotud korralduslikud küsimused.

Andmekaitse Inspektsiooni koostatud Andmekogude juhend (2013) seletab kogutavate andmete koosseisu sätestust. Nimelt ei pea andmekogusse kogutavate andmete täielik loetelu olema seaduses. Seadusega võib täpse andmete koosseisu kehtestamise delegeerida täitevvõimule. Kuid täitevvõim ei saa väljuda seaduse ja volitusnormi raamidest. Täitevvõim peab olema suuteline ammendavalt põhjendama, millise eesmärgi saavutamiseks mingeid andmeid möödapääsmatult on vaja koguda. Põhimäärus peab sisaldama ammendavat loetelu antud andmekogusse kogutavatest andmetest. Seejuures peab loetelust ka selguma, kelle kohta (st andmesubjektide kategooriad) neid andmeid kogutakse. Kui andmete koosseis on väga mahukas, võib loetelu esitada põhimääruse lisana. Põhimäärusesse ei ole soovitatav panna tehnilisi andmeid, mida on vaja infosüsteemi toimimiseks. Näiteks andmevälja või kande liigi süsteemi-sisene tunnus või andmete muutmise aeg. Ulatuslik andmekoosseis on soovitatav põhimääruses esitada süstematiseeritud kujul, nt menetluste kaupa. Seega on liiasus andmestruktuurides sõltuv seaduse ja volitusnormi raamist ning peab täitma nende eesmärgi.

Riigi Infosüsteemi Amet on koostanud relatsiooniliste andmemudelite koostamise juhendi<sup>1</sup>, milles kirjeldatakse andmemudelite koostamiseks soovituslikud struktuuriliste mustrite kogumid, mida andmekogu omanik peab andmekogu disainides ja arendades järgima.

## Andmekvaliteedi tagamise meetmed

Järgnevates peatükkides on välja toodud iga tunnuse osas andmekvaliteedi tagamise meetmed, millega andmekogu omanik saab ja peab arvestama.

### Õigsus

Õigsuse vastavusse viimiseks õigusaktidest tulenevate nõuetega tuleb täpsustada õigusaktidest tulenevaid nõudeid, hinnata andmete õigsuse vastavust fikseeritud nõuetele ja vajadusel rakendada meetmed andmete parendamiseks.

Õigusaktidest tulenevaid nõudeid tuleb täpsustada, arvestades lepinguid, põhitegevuse protsesse ning tagajärgede kaalukust, näited:

<sup>1</sup> [https://www.ria.ee/public/Arhitektuur/RELATSIOONILISTE\\_ANDMEMUDELITE\\_KOOSTAMISE\\_JUHEND\\_ver.1.0.pdf](https://www.ria.ee/public/Arhitektuur/RELATSIOONILISTE_ANDMEMUDELITE_KOOSTAMISE_JUHEND_ver.1.0.pdf)

Vastavalt ISKE raamistikule tuleb määrata andmete tervikluse turvaosaklass ning rakendada selle turvaosaklassiga seoses nõutud õigsuse kontrollid. Näiteks turvaosaklassi T2 puhul peab info allikas, selle muutmise ja hävitamise fakt olema tuvastatavad; vajalikud on perioodilised info õigsuse, täielikkuse, ajakohasuse kontrollid.

Andmete tervikluse kontrolli, sealhulgas veakäsitluse selliseid aspekte nagu pääsukontroll, sünkroonimise kontroll ja tervikluse kontroll käsitleb ISKE meede M 2.130 "Andmebaasi tervikluse tagamine". Haldusülesannete lahusust ja andmete ülekannet käsitlevad ISKE meetmed M 2.131 Haldusülesannete lahusus andmebaasisüsteemides ja M 2.135 Andmete turvaline teisaldus andmebaasi. Kõrgete tervikluse nõuete puhul tuleb ka veakäsitluse puhul rakendada meetet HT.10 Andmebaasi kannete krüptoaheldamine.

Õigusaktides ja poliitikates ei ole üldjuhul sätestatud õigsuse arvulisi näitajaid. Suuremahuliste andmekogude puhul ei ole enamasti võimalik tagada andmete absoluutset õigsust, sest andmetes on tavaliselt ikka vigu. Seepärast tuleb lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes sätestada täpsemad õigsuse nõuded. Selleks võib muuhulgas kehtestada protseduurid vigade avastamiseks ja ebaõigete andmete parandamiseks, määratleda ebaõigete andmete lubatava osakaalu jne.

Täpsustamise tulemusena tuleb õigsuse nõuded fikseerida ning esitada need lepingutes, teenustaseme lepetes, juhendites või muul sobival moel.

Tuleb hinnata, kas andmete õigsus vastab fikseeritud nõuetele, viies läbi andmete õigsuse ja/või rakendatavate protseduuride valikulisi, perioodilisi või reaaliajase hindamisi vastavalt ISKE tervikluse turvaosaklassile ja võrreldes tulemusi nõutud tasemega.

Kui andmete õigsust on vaja parendada, siis tuleb analüüsida kogu andmete elutsükli, et leida võimalused vigaste andmete osakaalu vähendamiseks vastavalt vajadustele ja nõuetele. Näiteks võib liiga suur ebaõigete andmete osakaal olla tingitud andmete kogumise ebapiisavatest protseduuridest, ebapiisavatest andmekontrollidest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest.

Rakendatavad meetmed sõltuvad konkreetse andmekogu organisatsioonist, protsessidest ja tehnilistest vahenditest ning need tuleb iga andmekogu puhul eraldi määratleda ning rakendada. Näited õigsuse tagamise meetmete kohta:

- Klassifikaatorite kasutamine
- Õigsuse automaatsete kontrolliprotseduuride rakendamine, sealhulgas andmete vaheliste loogiliste seoste kohta
- Liiasuse vältimine andmekogudes
- Ebaõigete andmete kohene parandamine.
- Perioodilised õigsuse kontrollid
- Kasutajate tagasiside arvestamine

## Täielikkus

Täielikkuse vastavusse viimiseks õigusaktidest tulenevate nõuetega tuleb täpsustada õigusaktidest tulenevaid nõudeid, hinnata andmete täielikkuse vastavust fikseeritud nõuetele ja vajadusel rakendada meetmed andmete parendamiseks.

Õigusaktidest tulenevaid nõudeid tuleb täpsustada, arvestades lepinguid, põhitegevuse protsesse ning tagajärgede kaalukust, näited:

- Vastavalt ISKE raamistikule tuleb määrata andmete tervikluse turvaosaklass ning rakendada selle turvaosaklassiga seoses nõutud täielikkuse kontrollid. Näiteks turvaosaklassi T2 puhul peab info allikas, selle muutmise ja hävitamise fakt olema tuvastatavad; vajalikud on perioodilised info õigsuse, täielikkuse, ajakohasuse kontrollid.

- Õigusaktides ja poliitikates ei ole üldjuhul sätestatud täielikkuse arvulisi näitajaid. Suuremahuliste andmekogude puhul ei ole enamasti võimalik tagada andmete absoluutset täielikkust, sest andmetes on tavaliselt ikka lünki. Seepärast tuleb lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes sätestada täpsemad täielikkuse nõuded. Selleks võib muuhulgas kehtestada protseduurid vigade avastamiseks ja mittetäielike andmete parandamiseks, määratleda mittetäielike andmete lubatava osakaalu jne.

Täpsustamise tulemusena tuleb täielikkuse nõuded fikseerida ning esitada need lepingutes, teenustaseme lepetes, juhendites või muul sobival moel.

Tuleb hinnata, kas andmete täielikkus vastab fikseeritud nõuetele, viies läbi andmete täielikkuse ja/või rakendatavate protseduuride valikulisi, perioodilisi või reaalselt hindamisi vastavalt ISKE tervikluse turvaosaklassile ja võrreldes tulemusi nõutud tasemega.

Andmete täielikkuse parendamisel, tuleb analüüsida kogu andmete elutsüklit, et leida vajadustele ja nõuetele vastavad võimalused mittetäielikkude andmete osakaalu vähendamiseks. Näiteks võib liiga suur mittetäielikkude andmete osakaal olla tingitud andmete kogumise ebapiisavatest protseduuridest, ebapiisavatest andmekontrollidest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest.

Rakendatavad meetmed sõltuvad konkreetse andmekogu organisatsioonist, protsessidest ja tehnilistest vahenditest ning need tuleb iga andmekogu puhul eraldi määratleda ning rakendada. Näited täielikkuse tagamise meetmete kohta:

- Täielikkuse automaatsete kontrolliprotseduuride rakendamine, sealhulgas aruanded kohustuslikult täidetavate, kuid tegelikult mittetäidetud andmeväljade kohta
- Muude protseduuride rakendamine, näiteks juhendid või koolitused andmeandjatele
- Mittetäielikkude andmete kohene lisamine
- Perioodilised täielikkuse kontrollid
- Kasutajate tagasiside arvestamine

## Kooskõla

Kooskõla vastavusse viimiseks õigusaktidest tulenevate nõuetega tuleb täpsustada õigusaktidest tulenevaid nõudeid, hinnata kooskõla vastavust fikseeritud nõuetele ja vajadusel rakendada meetmed kvaliteedi parendamiseks.

Õigusaktidest tulenevaid nõudeid tuleb täpsustada, arvestades lepinguid, põhitegevuse protsesse ning tagajärgede kaalukust. Näiteks määrab infosüsteemi eripära selle, kuidas ta peab toetama mitmekeelseid ja rahvusvahelisi semantikavarasid, millise ulatusega on infosüsteemi koosvõime arhitektuuri dokument ning kui ulatuslik on selle semantiliselt koosvõimet käsitlevat osa.

Samuti sõltub konkreetsest organisatsioonist, milliseid töövahendeid ja standardeid kasutatakse semantikavarade loomisel ja infovarade semantiliselt rikastamisel. Igal juhul tuleb semantikavarad ja annoteeritud infovarad avaldada RIHAs semantika juhise nõuete kohaselt.

Tuleb analüüsida, kas on võimalik ja asjakohane võimaldada juurdepääsu avaandmetele masinloetaval kujul ning avatud vormingus. Andmekogu andmete omavaheliste sisemiste vasturääkivuste vähendamise üks võimalus on liiasuse vältimine andmekoosseisudes.

Täpsustamise tulemusena tuleb kooskõla nõuded fikseerida ning esitada need lepingutes, teenustaseme lepetes, juhendites või muul sobival moel. Tehtud valikud tuleb ellu viia. Tuleb hinnata, kas andmete kooskõla vastab fikseeritud nõuetele, viies läbi andmete kooskõla ja/või rakendatavate protseduuride valikulisi, perioodilisi või reaalselt hindamisi.



Rakendatavad meetmed sõltuvad konkreetse andmekogu organisatsioonist, protsessidest ja tehnilistest vahenditest ning need tuleb iga andmekogu puhul eraldi määratleda ning rakendada. Näited andmekogu sisemise kooskõla tagamise meetmete kohta:

- Kooskõla automaatsete kontrolliprotseduuride rakendamine, sealhulgas aruanded vasturääkivate andmeväljade kohta
- Juhendid või koolitused andmeandjatele
- Vasturääkivuste kohene lahendamine
- Perioodilised kooskõla kontrollid
- Kasutajate tagasiside arvestamine

### Usaldusväarsus

Usaldusväarsuse vastavusse viimiseks õigusaktidest tulenevate nõuetega tuleb täpsustada õigusaktidest tulenevaid nõudeid, hinnata andmete usaldusväarsuse vastavust fikseeritud nõuetele ja vajadusel rakendada meetmed andmete parendamiseks.

Õigusaktidest tulenevaid nõudeid tuleb täpsustada, arvestades lepinguid, põhitegevuse protsesse ning tagajärgede kaalukust, näited:

- Vastavalt ISKE raamistikule tuleb määrata andmete tervikluse turvaosaklass ning rakendada selle turvaosaklassiga seoses nõutud autentsuse kontrollid. Nii nõutakse mitmetes ISKE meetmetes, sealhulgas M 2.162 “Krüptoprotseduuride ja -toodete vajaduse määramine”, M 2.163 “Krüptoprotseduure ja -tooteid mõjutavate tegurite määramine”, M 2.243 “Arhiveerimiskontseptsiooni väljatöötamine” andmete autentsuse nõuete väljaselgitamist ja vastavate protseduuride rakendamist.
- Vastavalt Riigi Infosüsteemi Ameti peadirektori käskkirjaga kinnitatud dokumendile “Kooskõlastajate ühine RIHAs dokumentatsiooni hindamise juhend” tuleb hinnata, kas andmed kogutakse võimalikult autentselt allikast.
- Õigusaktides ja poliitikates ei ole üldjuhul sätestatud usaldusväarsuse arvulisi näitajaid. Suuremahuliste andmekogude puhul ei ole enamasti võimalik tagada andmete absoluutset usaldusväarsust, sest usaldusväarsuse kadu võib tuleneda paljudest teguritest, mis ei pruugi olla andmekogu omaniku kontrolli all. Seepärast tuleb lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes sätestada täpsemad usaldusväarsuse nõuded. Selleks võib muuhulgas kehtestada protseduurid mitteusaldusväärsete andmete avastamiseks ja probleemide lahendamiseks, määratleda mitteusaldusväärsete andmete lubatava osakaalu jne.

Täpsustamise tulemusena tuleb usaldusväarsuse nõuded fikseerida ning esitada need lepingutes, teenustaseme lepetes, juhendites või muul sobival moel.

Tuleb hinnata, kas andmete usaldusväarsus vastab fikseeritud nõuetele, viies läbi andmete usaldusväarsuse ja/või rakendatavate protseduuride valikulisi, perioodilisi või reaajas hindamisi vastavalt ISKE tervikluse turvaosaklassile ja võrreldes tulemusi nõutud tasemega.

Kui andmete usaldusväarsust on vaja parendada, siis tuleb analüüsida kogu andmete elutsüklit, et leida võimalused mitteusaldusväärsete andmete osakaalu vähendamiseks vastavalt vajadustele ja nõuetele. Näiteks võib liiga suur mitteusaldusväärsete andmete osakaal olla tingitud andmete kogumise ebapiisavatest protseduuridest, ebapiisavatest andmekontrollidest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest.

Rakendatavad meetmed sõltuvad konkreetse andmekogu organisatsioonist, protsessidest ja tehnilistest vahenditest ning need tuleb iga andmekogu puhul eraldi määratleda ning rakendada. Näited usaldusväarsuse tagamise meetmete kohta:

- Andmeandjate usaldusväarsuse arvestamine ja võimalusel kontrollimine
- Teise andmekogu põhiandmete kasutamise puhul nende andmete usaldusväarsuse piisavuse analüüs, võimalusel kontrollimine ja mõjutamine.

- Andmete kogumise protseduuride analüüs ja täiendamine.
- Perioodilised andmete usaldusväärsuse kontrollid
- Kasutajate tagasiside arvestamine

## Ajakohasus

Ajakohasuse vastavusse viimiseks õigusaktidest tulenevate nõuetega tuleb täpsustada õigusaktidest tulenevaid nõudeid, hinnata andmete õigsuse vastavust fikseeritud nõuetele ja vajadusel rakendada meetmed andmete parendamiseks.

Õigusaktidest tulenevaid nõudeid tuleb täpsustada, arvestades lepinguid, põhitegevuse protsesse ning tagajärgede kaalukust, näited:

- Vastavalt ISKE raamistikule tuleb määrata andmete tervikluse turvaosaklass ning rakendada selle turvaosaklassiga seoses nõutud ajakohasuse kontrollid. Näiteks turvaosaklassi T1 puhul peab info allikas, selle muutmise ja hävitamise fakt olema tuvastatavad; tuleb rakendada info õigsuse, täielikkuse, ajakohasuse kontrollid erijuhtudel ja vastavalt vajadusele.
- Vastavalt Riigi Infosüsteemi Ameti peadirektori käskkirjaga kinnitatud dokumendile “Kooskõlastajate ühine RIHAs dokumentatsiooni hindamise juhend” tuleb aadressiandmete süsteemi kasutamisel hinnata, kas andmeid hoitakse ajakohasena, st uuendatakse regulaarselt ning kui andmeid ajakohastatakse, siis milliseid X-tee teenuseid ja kui sageli kasutatakse andmete ajakohastamiseks.
- Õigusaktides ja poliitikates ei ole (väheste eranditega) sätestatud ajakohasuse arvulisi näitajaid. Kuna absoluutset ajakohasust ei ole (andmed on alati mingil määral ajaliselt tegelikkusest maas), siis tuleb lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes sätestada täpsemad ajakohasuse näitajad. Eesmärgiks on seada ajalise viite jaoks selline piir, mis rahuldab kõiki andmete tarbijaid (kasutavaid teenuseid). Seejuures tuleb arvesse võtta, et ajalise viite piir võib olla erinevate andmete jaoks erinev.
- Vastavalt IKS § 24 on isikuandmete töötaja isikuandmete töötlemisel kohustatud tagama, et isikuandmed on õiged ja, kui see on eesmärkide saavutamiseks vajalik, viimases seisus. See, kas andmete viimases seisus olek on tegelikult vajalik, sõltub seega andmekogu eripärast.
- Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsiooni artikkel 5 sätestab, et automatiseeritud töödeldavad isikuandmed peavad olema säilitatud vormis, mis lubab andmesubjekti teha kindlaks mitte kauemaks, kui perioodiks, mis on vajalik andmete kogumise eesmärkide saavutamiseks. See periood on soovitatav kehtestada lähtudes põhitegevuse protsesside nõuetest.
- Eesti avaliku teabe masinloetava avalikustamise roheline raamat toob ära avatud riigi tööühma poolt defineeritud nõuded, mille täidetuse korral võib andmehulga andmeid lugeda avaandmeteks. Üks nõuetest on ajakohasus - andmehulk on avaldatud võimalikult kiirelt, et säilitada selle ajakohasus. Seega kui andmed peavad olema esitatud avaandmetena, siis tuleb sätestada ajakohasuse nõue.

Täpsustamise tulemusena tuleb ajakohasuse nõuded fikseerida ning esitada need lepingutes, teenustaseme lepetes, juhendites või muul sobival moel.

Tuleb hinnata, kas andmete ajakohasus vastab fikseeritud nõuetele, viies läbi andmete ajakohasuse valikulisi, perioodilisi või reaalselt hindamisi vastavalt ISKE tervikluse turvaosaklassile ja võrreldes tulemusi nõutud tasemega.

Kui andmete ajakohasust on vaja parendada, siis tuleb analüüsida kogu andmete elutsüklit, et leida võimalused andmeid tekitanud sündmuse ja andmete vahelise liiga suure ajalise viite vähendamiseks vastavalt vajadustele ja nõuetele. Näiteks võib liiga suur ajaline viide olla tingitud sündmusest teavitamise ebapiisavast kiirusest, andmete kogumise aeglusest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest.

Rakendatavad meetmed sõltuvad konkreetse andmekogu organisatsioonist, protsessidest ja tehnilistest vahenditest ning need tuleb iga andmekogu puhul eraldi määratleda ning rakendada.

## Reeglipärasus

Andmete reeglipärasuse saavutamiseks tuleb analüüsida, millised riigi infosüsteemi reguleerivad õigusaktid ning nende alusel välja antud dokumendid on antud andmekogu jaoks olulised. Muuhulgas on soovitatav reeglipärasuse nõuete väljaselgitamisel arvesse võtta teiste andmekvaliteedi tunnuste juures märgitud õigusakte.

Lähtudes sellest analüüsist kavandatakse andmete struktuur nii, et andmete atribuudid vastaksid andmekvaliteeti puudutavatele standarditele, kokkulepetele või kehtivatele õigusaktidele ning samalaadsetele normidele.

Näiteks sätestab IKS § 6 lõige 5 andmete kvaliteedi põhimõtte isikuandmete töötlemisel – isikuandmed peavad olema ajakohased, täielikud ning vajalikud seatud andmetöötluse eesmärgi saavutamiseks.

Andmekogude dokumentatsiooni kooskõlastusel riigi infosüsteemi haldussüsteemis (RIHA) kontrollitakse andmete kogumise eesmärgipärasust ja andmete ühekordse küsimise põhimõtte järgimist (koosvõime raamistik). Avaliku ja erasektori infosüsteemid peavad järgima teenusepõhise arhitektuuri reegleid (koosvõime raamistik).

## Konfidentsiaalsus

Konfidentsiaalsuse vastavusse viimiseks õigusaktidest tulenevate nõuetega tuleb neid nõudeid täpsustada, hinnata andmete konfidentsiaalsuse vastavust fikseeritud nõuetele ja vajadusel rakendada meetmed andmete parendamiseks. Üldjuhul on riigi ja kohaliku omavalitsuse andmekogude konfidentsiaalsuse meetmed kasutusele võetud seoses ISKE rakendamisega.

Õigusaktidest tulenevaid nõudeid tuleb täpsustada, arvestades lepinguid, põhitegevuse protsesse ning tagajärgede kaalukust, näited:

- Vastavalt ISKE raamistikule tuleb määrata andmete konfidentsiaalsuse turvaosaklass ning rakendada selle turvaosaklassiga seoses nõutud kontrollid. Näiteks konfidentsiaalsuse turvaosaklassi S1 puhul on info asutusesiseseks kasutamiseks ning juurdepääs teabele on lubatav juurdepääsu taotleva isiku õigustatud huvi korral.
- Õigusaktides ja poliitikates ei ole üldjuhul sätestatud konfidentsiaalsuse arvulisi näitajaid. Vajadusel tuleb seepärast lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes sätestada täpsemad konfidentsiaalsuse nõuded. Selleks võib muuhulgas kehtestada protseduurid konfidentsiaalsuse tagamiseks, määratleda aktsepteeritavad infoturbe riskid, luua hädaolukorraks valmisoleku plaanid ja protseduurid jne.

Täpsustamise tulemusena tuleb konfidentsiaalsuse nõuded fikseerida ning esitada need lepingutes, teenustaseme lepetes, juhendites või muul sobival moel.

Tuleb hinnata, kas andmete konfidentsiaalsus vastab fikseeritud nõuetele, viies läbi andmete konfidentsiaalsuse ja/või rakendatavate protseduuride valikulisi, perioodilisi või reaalajas hindamisi vastavalt ISKE konfidentsiaalsuse turvaosaklassile ja võrreldes tulemusi nõutud tasemega.

Rakendatavad meetmed sõltuvad konkreetse andmekogu organisatsioonist, protsessidest ja tehnilistest vahenditest ning need tuleb iga andmekogu puhul eraldi määratleda ning rakendada. Näited konfidentsiaalsuse tagamise meetmete kohta:

- ISKE meede M 2.128 Andmebaasisüsteemi pääsu reguleerimine
- ISKE meede M 2.132 Andmebaasi kasutajate ja kasutajagruppide configureerimise reeglid
- ISKE meede M 2.129 Andmebaasiinfo pääsu reguleerimine
- ISKE meede M 2.134 Andmebaasipäringute suunised
- ISKE meede M 2.127 Tuletamise vältimine andmebaasis.

- Perioodilised konfidentsiaalsuse kontrollid
- Kasutajate tagasiside arvestamine

## Ühekordsus

Ühekordsuse põhimõtte rakendamine eeldab selle põhimõtte arvestamist andmekogude kavandamisel ja realiseerimisel, ühekordsuse põhimõttest kõrvalekaldumiste põhjuste analüüsi ja kõrvalekaldumiste lahendamist.

Iga andmekogu kavandamisel on soovitatav analüüsida, kas vajalikke andmeid on võimalik saada teistest riigi infosüsteemi kuuluvatest andmekogudest, selmet neid inimestelt koguda. Kui see on võimalik, siis tuleb kasutada andmeid teistest andmekogudest, sõlmides vajadusel vastavad kokkulepped ning realiseerides tehnilised lahendused.

On soovitatav kontrollida, kas olemasolevates andmekogudes on ühekordsuse põhimõttest kinni peetud.

Kui ühekordsuse põhimõtet on rikutud, on soovitatav algatada tegevused ühekordsuse põhimõtte sisseviimiseks, näiteks andmete hõive protseduuride muutmiseks.

## Mitteliiasus

Liiasused andmestruktuurides, s.o samade andmete hoidmine andmestruktuuri erinevates osades võib kaasa tuua eelmises jaotises iseloomustatud ebasoovitavaid tagajärgi. Seepärast on soovitatav üldjuhul sellist liiasust võimalusel vältida. Sama põhimõtte kehtib liiasuste kohta andmetöötlusprotsessides. Juhiseid andmete korrektseks projekteerimiseks annab Riigi Infosüsteemi Ameti välja antud relatsiooniliste andmemudelite koostamise juhend.

Lisaks eelmises alajaotises toodud ebasoovitavatele mõjudele võivad andmeliiasusel olla ka positiivsed efektid. Oluline on, et positiivsed ja negatiivsed mõjud oleksid analüüsitud ja esimesed ületaksid teisi. Järgnevalt on toodud näiteid liiasuse, selle mõjude ja vastavate meetmete kohta.

Andmekogudes võidakse efektiivsuse huvides hoida põhiandmetele lisaks ka agregeeritud andmeid. Sellisel juhul peavad need olema põhiandmete põhjal üle arvutatavad ning sellist ülearvutamist on soovitatav kontrolli mõttes perioodiliselt teha.

Andmete haldamise ja ärioloogika eraldamine ning vastavate liideste kokkuleppimine võimaldab andmete füüsilise esituse optimeerimist, andmete tervikluse paremat kontrolli, andmepäringute paremat jõudlust jne. Samas võib see kaasa tuua liiasuse andmetes. Selline liiasus on üldjuhul ajutine ning selle positiivsed küljed kaaluvad üles negatiivsed.

Andmete turvalisuse ja hädaolukordadeks valmisoleku huvides võib olla vajalik liiasuse projekteerimine nii süsteemidesse (nt serverid, võrgukomponendid), andmetesse (nt varukoopiate hoidmine ning andmete säilitamine ja töötlemine erinevates lokatsioonides) kui ka protseduuridesse (nt lisaprotseduurid hädaolukordade puhuks). Sellise liiasuse vajalikkust on soovitatav analüüsida ning seda tuleks rakendada vaid vastavalt vajadusele.

Liiasuste andmete ja protseduuride puhul on soovitatav jälgida järgmisi üldisi reegleid.

- Hinnata liiasuse positiivseid ja negatiivseid mõjusid (teha nt SWOT analüüs) ning minimeerida liiasuse kahjulikke mõjusid.
- Rakendada analoogilisi reegleid ja kvaliteedinõudeid kõikide liiasuste andmete ja protseduuride puhul. Näiteks võivad varukoopiad kaasa tuua suuri turvariske, kui neid kaitstakse vähem kui põhiandmeid. Samasugused ohud on seotud avariiprotseduuridega.
- Perioodilised liiasuste andmete ja protseduuride kontrollid ning vajadusel ebavajaliku liiasuse kõrvaldamine.

## Andmekvaliteedi nõuete kontrollküsimused

Andmekvaliteedi nõuete väljaselgitamine on esimene oluline samm andmekvaliteedi tagamise suunas. Selleks, et tagada andmete kvaliteet, peab teadma andmekvaliteedi nõudeid, millele vastavust kontrollides saab teada, kas andmed on kvaliteetsed või mitte. Tunnuste nõuete kontrollküsimused on nimekiri küsimustest, mis aitavad veenduda, kas antud tunnuse osas on andmekvaliteedi nõuete väljaselgitamisel erinevate võimalike regulatsioonidega arvestatud või mitte. Samuti aitavad need veenduda selles, kas kvaliteedi loomiseks on kasutusele võetud tagamise meetmeid. Järgnevalt on kirjeldatud kontrollküsimused tunnuste kaupa:

### Õigsus

- Kas on määratud andmete tervikluse turvaosaklass ning rakendatud selle turvaosaklassiga seoses nõutud õigsuse kontrollid vastavalt ISKE raamistikule?
- Kas lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes on vajadusel sätestatud täpsemad õigsuse nõuded (vajadusel arvilised)?
- Kas isikuandmete töötlemisel on tagatud andmete õigsus vastavalt õigusaktidele?
- Kas õigsuse nõuded on fikseeritud ja esitatud lepingutes, teenustaseme lepetes, juhendites vms?
- Kas andmete õigsuse vastavust nõuetele on hinnatud?
- Kas on analüüsitud andmete elutsükli, et leida võimalused vigaste andmete osakaalu vähendamiseks vastavalt vajadustele ja nõuetele?
- Kas on rakendatud meetmeid, et vähendada vigaste andmete osakaalu, mis on tingitud andmete kogumise ebapiisavatest protseduuridest, ebapiisavatest andmekontrollidest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest?

### Täielikkus

- Kas on määratud andmete tervikluse turvaosaklass ning rakendatud selle turvaosaklassiga seoses nõutud täielikkuse kontrollid vastavalt ISKE raamistikule?
- Kas lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes on vajadusel sätestatud täpsemad täielikkuse nõuded (vajadusel arvilised)?
- Kas isikuandmete töötlemisel on tagatud andmete täielikkus vastavalt õigusaktidele?
- Kas täielikkuse nõuded on fikseeritud ja esitatud lepingutes, teenustaseme lepetes, juhendites vms?
- Kas andmete täielikkuse vastavust nõuetele on hinnatud?
- Kas on analüüsitud andmete elutsükli, et leida võimalused mittetäielike andmete osakaalu vähendamiseks vastavalt vajadustele ja nõuetele?
- Kas on rakendatud meetmeid, et vähendada mittetäielike andmete osakaalu, mis on tingitud andmete kogumise ebapiisavatest protseduuridest, ebapiisavatest andmekontrollidest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest?

### Kooskõla

- Kas on põhjendatud infosüsteemi tugi mitmekeelsetele ja rahvusvahelistele semantikavaradele, millise ulatusega on infosüsteemi koosvõime arhitektuuri dokument ning kui ulatuslik on selle semantilist koosvõimet käsitlevat osa?
- Kas on kindlaks määratud, milliseid töövahendeid ja standardeid kasutatakse semantikavarade loomisel ja infovarade semantilisel rikastamisel?
- Kas semantikavarad ja annoteeritud infovarad on avaldatud RIHAs vastavalt juhiste?
- Kas on võimaldatud juurdepääs avaandmetele masinloetaval kujul ning avatud vormingus, kui see on võimalik ja asjakohane?
- Kas andmekogu sisemise kooskõla tagamise vajadust on analüüsitud ning on rakendatud selleks meetmed?

### Usaldusväarsus

- Kas on määratud andmete tervikluse turvaosaklass ning rakendatud selle turvaosaklassiga seoses nõutud usaldusväärse kontrollid vastavalt ISKE raamistikule?
- Kas lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes on vajadusel sätestatud täpsemad usaldusväärse nõuded (vajadusel arvilised)?
- Kas isikuandmete töötlemisel on tagatud andmete usaldusväärse vastavalt õigusaktidele?
- Kas usaldusväärse nõuded on fikseeritud ja esitatud lepingutes, teenustaseme lepetes, juhendites vms?
- Kas andmete usaldusväärse vastavust nõuetele on hinnatud?
- Kas on analüüsitud andmete elutsükli, et leida võimalused usaldusväärse probleemidega andmete osakaalu vähendamiseks vastavalt vajadustele ja nõuetele?
- Kas on rakendatud meetmeid, et vähendada mitteusaldusväärse andmete osakaalu, mis on tingitud andmete kogumise ebapiisavatest protseduuridest, ebapiisavatest andmekontrollidest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest?

### Ajakohasus

- Kas on määratud andmete tervikluse turvaosaklass ning rakendatud selle turvaosaklassiga seoses nõutud ajakohasuse kontrollid vastavalt ISKE raamistikule?
- Kas lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes on sätestatud täpsemad ajakohasuse näitajad (vajadusel arvilised)?
- Kas isikuandmete töötlemisel on analüüsitud andmete viimases seisus oleku vajalikkust ja andmete töötlemise perioodi?
- Kas avaandmete puhul on kehtestatud ajakohasuse nõue?
- Kas andmete elutsükli on analüüsitud, et leida võimalused andmeid tekitanud sündmuse ja andmete vahelise liiga suure ajalise viite vähendamiseks vastavalt vajadustele ja nõuetele?
- Kas ajakohasuse nõuded on fikseeritud ja esitatud lepingutes, teenustaseme lepetes, juhendites vms?
- Kas andmete ajakohasuse vastavust nõuetele on hinnatud?
- Kas on analüüsitud andmete elutsükli, et leida võimalused ajakohasuse parendamiseks?
- Kas on rakendatud meetmeid, et vähendada ajalisi viiteid, mis on tingitud sündmusest teavitamise ebapiisavast kiirusest, andmete kogumise aeglusest, andmetöötlusprotsesside halvast korraldusest ja muudest teguritest?

### Reeglipärasus

- Kas on analüüsitud, millised riigi infosüsteemi reguleerivad õigusaktid ning nende alusel välja antud dokumendid on antud andmekogu jaoks olulised?
- Kas on arvesse võetud ka teiste andmekvaliteedi tunnuste juures märgitud õigusakte?
- Kas andmete struktuur on kavandatud selliselt, et andmete atribuudid vastaksid andmekvaliteeti puudutavatele standarditele, kokkulepetele või kehtivatele õigusaktidele ning samalaadsetele normidele?

### Konfidentsiaalsus

- Kas on määratud andmete konfidentsiaalsuse turvaosaklass ning rakendatud selle turvaosaklassiga seoses nõutud konfidentsiaalsuse meetmed vastavalt ISKE raamistikule?
- Kas lepingutest, põhitegevuse protsessidest, tagajärgede kaalukusest, teenustaseme lepetest ja muudest allikatest lähtudes on vajadusel sätestatud täpsemad konfidentsiaalsuse nõuded?
- Kas isikuandmete töötlemisel on tagatud andmete konfidentsiaalsus vastavalt õigusaktidele?
- Kas konfidentsiaalsuse nõuded on fikseeritud ja esitatud lepingutes, teenustaseme lepetes, juhendites vms?
- Kas andmete konfidentsiaalsuse vastavust nõuetele on hinnatud?

### Ühekordsus

- Kas andmekogu kavandamisel on analüüsitud võimalust andmehõiveks teistest andmekogudest ja kui see on võimalik, siis kas selline hõive on realiseeritud?
- Kas olemasolevates andmekogudes on ühekordsuse põhimõttest kinnipidamist analüüsitud?
- Kas on algatatud tegevused ühekordsuse põhimõtte sisseviimiseks?

## Mitteliiasus

- Kas on välditud ebavajalikku liiasust andmestruktuurides ja andmetöötlusprotsessides?
- Kas on analüüsitud liiasuse positiivseid ja negatiivseid mõjusid?
- Kas agregeeritud andmed on põhiandmete põhjal üle arvutatavad ning kas sellist ülearvutamist tehakse perioodiliselt?
- Kas andmete turvalisuse ja hädaolukordadeks valmisoleku huvides loodud liiasuse vajalikkust on analüüsitud ning kas seda rakendatakse vaid vastavalt vajadusele?
- Kas kõikide liiasuste andmete ja protseduuride puhul rakendada analoogilisi reegleid ja kvaliteedinõudeid?
- Kas viiakse läbi regulaarseid liiasuste andmete ja protseduuride kontrollide ning kõrvaldatakse ebavajalik liiasus?

## Andmekvaliteedi juhtimise uuring

Andmekvaliteedi juhtimise uuringu eesmärgiks oli andmekvaliteedi mõõtmise meetodika loomine riigi infosüsteemi andmekogude jaoks. Uuringu ülesandeks oli analüüsida kasutatavaid praktikaid andmekvaliteedi protsesside mõõtmiseks organisatsioonides ja välja töötada meetodika andmekvaliteedi mõõtmiseks riigi infosüsteemi andmekogudes.

Lähteülesande kohaselt tuli andmekvaliteedi hindamise ja parendamise võimalusi analüüsida vähemalt 4 aspektist kattes kuut rühma andmekvaliteedi mõjutavaid tegureid/valdkondi, milledeks olid:

1. andmekogu andmetega varustavad infosüsteemid ja nende kvaliteedist tulenevad tegurid;
2. andmekvaliteedi juhtimise meetoditest ja kasutatavatest vahenditest tulenevad tegurid;
3. riigi juhtimisest ja vastu võetavatest seadustest tulenevad tegurid;
4. andmekogu haldava asutuse juhtimisest ja sisemistest regulatsioonidest tulenevad tegurid;
5. organisatsiooni ülesehitusest ja organisatsiooni sisemisest kultuurist tulenevad tegurid;
6. välised tegurid (kõik tegurid, mis ei kattu eelmise viie rühma teguritega), mis on olulised, aga antud hetkel täiesti kirjeldamata.

Esimese sammuna veenduti selles, et loetletud valdkonnad on aspektidega kaetud. Selleks koostati tabel, milles hinnati iga aspekti kuuluvust kindlasse mõju valdkonda (vt Lisa 3: Hinnang tegurite kaetuse osas aspektidega). Hindamise tulemusena veenduti uuringu juhrühmas, et lähteülesandes loetletud tegurid on kaetud loetletud aspektidega. Aspektide selgemaks klassifitseerimiseks lepitati kokku, et edasiseks jaotatakse loetletud andmekvaliteedi aspektide 5-de põhilisse andmekvaliteedi kategooriasse, milleks olid:

1. **Juhtimine ja planeerimine.** Selle kategooriaga on seotud kogu andmekvaliteedi juhtimise ja planeerimisega seotud tegevuste ja vahendite aspektid, sh. reeglite, poliitikate, põhikirjade ja plaanide haldamine, andmekvaliteedi käsitlemine organisatsiooni strateegias.
2. **Organisatsioon ja vastutused.** Selles kategoorias käsitletakse andmekvaliteedi juhtimise eest vastutava organisatsiooni moodustamise ja vastutuste määratlemisega seotud aspekte.
3. **Protsessid.** Selles kategoorias käsitletakse andmekvaliteedi plaanide elluviimisega seotud aspekte, sh. monitooring, mõõtmine, tagamine.
4. **Teadmus ja kompetentsid.** Selles kategoorias käsitletakse andmekvaliteedi teadmuse juhtimisega (sh. teadmuse levitamise ja kompetentside juurutamisega) seotud aspekte.
5. **Tehnilised töövahendid.** Selles kategoorias käsitletakse andmekvaliteedi protsesse toetavate IT vahendite juurutamisega ja arendamisega seotud aspekte.



Järgmise sammuna tuli leida sobilik metoodika, mis võimaldaks hinnata andmekvaliteedi olukorda igas üksikus kategoorias eraldi. Püstitati hüpotees, et selliseks sobilikuks vahendiks võiks olla valdkonnale vastav küpsusmudel, kuna küpsusmudel on levinud töövahend inimeste/kultuuri, protsesside/struktuuride ja objektide/tehnoloogiate hindamiseks. Küpsusmudel on töövahend, mis võimaldab astmelisel skaalal hinnata vaadeldavas valdkonnas hetke olukorda ja seada astmeliselt eesmärged edasiseks arenguks.

Uuringu juhtrühmale tehti ettepanek hüpoteesi tõestamiseks vajaliku uuringu läbiviimiseks ja andmekvaliteedi küpsusmudeli koostamiseks, millega juhtrühm ka nõustus.

Küpsusmudeli koostamiseks viidi läbi otsing avalikus informatsioonis leidmaks praktilist rakendust leidnud andmekvaliteedi raamistikke, milles on kasutatud küpsusmudeleid. Uuringus vaadeldi järgnevaid raamistikke:

1. [Practitioners guide to data quality](#);
2. [Information Quality Management \(IQM\)- Trillium software data governance maturity assessment](#);
3. [The IBM Data Governance Council Maturity Model](#);
4. [The Storage Networking Industry- ILM Maturity Model](#);
5. [Process-Centric Organisation Maturity Model](#).

Raamistike küpsusmudelitest otsiti välja küpsustasemetele vastavaid väiteid ja klassifitseeriti need väited andmekvaliteedi kategooriatesse, et saada ülevaadet, millised andmekvaliteedi kategooriad on uuritud küpsusmudelites käsitletud. Tulemusi iseloomustab alljärgnev tabel:

Tabel 6: Andmekvaliteedi kategooriate kaetus küpsusmudelite väidetega

Kategooria	<i>Practitioners guide to data quality</i>	Information Quality Management (IQM)- Trillium software data governance maturity assessment	The IBM Data Governance Council Maturity Model	The Storage Networking Industry- ILM Maturity Model	Process-Centric Organisation Maturity Model
<b>Juhtimine</b>	x	x	x	x	x
<b>Organisatsioon</b>	x	x	x	x	x
<b>Teadmus</b>	x	x	x	puudub	x
<b>Protsessid</b>	x	x	x	x	x
<b>Töövahendid</b>	x	x	x	x	x

Välja valitud küpsusmudeli väited tõlgiti eesti keelde, konsolideeriti ja jaotati 5 küpsusmudeli astme vahel, milleks olid:

1. **Reaktiivne tase.** Andmekvaliteedi protsessid toimivad ettearvatult, on nõrgalt kontrollitud ja protsessidega tegelemine on reaktiivse iseloomuga.
2. **Kontrollitud tase.** Andmekvaliteedi juhtimise vajadust on teadvustatud, sisse on seatud korratavad protseduurid.
3. **Standardiseeritud tase.** Andmekvaliteedi protsessid on standardiseeritud ja andmete kvaliteet on nõuetele vastavuse osas kontrollitud.
4. **Juhitud tase.** Saavutatud on protsesside jätkusuutlikkus. Andmekvaliteedi mõõtmistulemusi kasutatakse olemasolevate protsesside juhtimisel.
5. **Optimeeritud tase.** Andmekvaliteedi protsesse vaadatakse üle ja täiendatakse regulaarselt.

Uuringu tulemusena loodi andmekvaliteedi küpsusmudel, mis sisaldas iga andmekvaliteedi kategooria küpsuse hindamiseks väiteid jaotatuna 5-s küpsuse taseme vahel. Loodud mudel võimaldab selle rakendajal mõõta andmekvaliteedi juhtimise küpsust 5-astmelisel skaalal. Väljatöötatud küpsusmudel on toodud aruande lisa Lisa 4: Andmekvaliteedi küpsusmudel.



## Rakendamise metoodika

Järgnevalt on kirjeldatud küpsustaseme hindamise rakendamise metoodika.

## Kategooriate küpsustasemete hindamine

Andmekvaliteedi juhtimise kindla kategooria küpsustaseme hindamist viiakse läbi küpsusmudelis vastava kategooria tasemetele kuuluvate väidete valideerimise abil. Selleks valitakse vaadeldav kategooria ja esitatakse kategooria väidete osas küsimus: „Kas vaadeldava andmekogu andmekvaliteedi juhtimise kontekstis vastav väide on vastab tõele (on täidetud) või mitte?“. Küpsusmudeli loogika kohaselt ollakse alati vähemalt 1.-l tasemel, seega esimese taseme väidete osas küsimust ei esitata.

Kategooria küpsuse hindamiseks liigutakse madalamast tasemest ülespoole kontrollides väidete täidetust kuni esimese väiteni, kus väide polnud täidetud. Küpsus määratakse selle väite tasemest ühe võrra madalam. Sama protseduuri korratakse kõikide kategooriate küpsustasemete hindamiseks.

Kategooriate küpsustasemete hinnangu oodatud tulemiteks on küpsustaseme hinnang iga andmekvaliteedi küpsusmudeli kategooria osas.

Äärmiselt oluline on väidetele vastamisel kirja panna konkreetse väitega seotud kontekst ja põhjendused, mille alusel saab väidet tõeseks lugeda. Kui tekib vajadus vastata „osaliselt“, siis tegelikult on vastus „ei“ ja dokumenteerida tuleb „osalise“ vastavuse kontekst. Kirja pandud konteksti kasutatakse hiljem planeerimise ja tegevuskava loomise juures parendamist vajava lähtekohana.

## Üldise küpsustaseme hindamine

Üldise küpsustaseme hindamine toimub kategooriate küpsustasemete hinnangute põhjal.

Hinnangu saamiseks koondatakse kõik küpsusmudeli kategooriate hinnangud kokku ja üldiseks andmekvaliteedi juhtimise küpsustaseme hindeks määratakse kõige madalam hinne, mis kategooriatele määratud on.

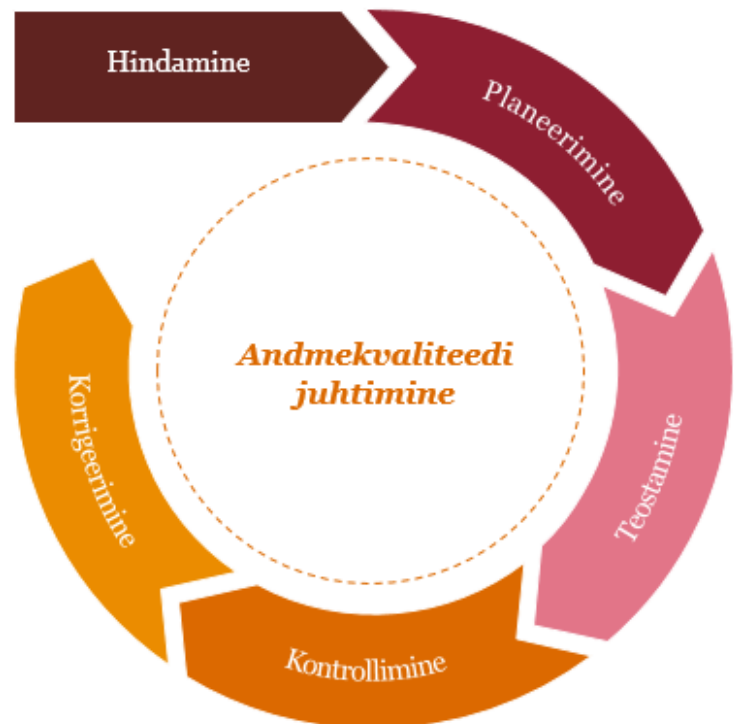
## Andmekvaliteedi juhtimise raamistik

Andmekvaliteedi juhtimise raamistiku eesmärgiks on pakkuda metoodika andmekvaliteedi parendamiseks riigi infosüsteemi andmekogudes.

Raamistiku aluseks on valitud nn. Demingi ratas, mis on protsesside parandamiseks mõeldud 4-astmeline korduv probleemilahendustsükkel (teatud ka kui PDCA-tsükkel – *plan, do, check, act*). Demingi rattal põhineb suur osa erinevaid kvaliteedijuhtimissüsteeme (nt. Kaizen, Six Sigma, Quality Circle) ja on sobilik vahend andmekvaliteedi juhtimisega seotud aspektide parendamise juhtimiseks.

PDCA-raamistik võimaldab järjepidevat protsesside tsüklilist parendamist, kus iga tsükkel sisaldab 4 tegevuste etappi:

1. **Planeerimine.** Etapi käigus seatakse parendustsükli eesmärgid ja määratletakse tegevused, mis on vajalikud nende eesmärkide saavutamiseks.
2. **Teostamine.** Etapi käigus viiakse ellu koostatud parendusplaan.
3. **Kontrollimine.** Etapi käigus uuritakse, kas teostamise etapis teostatud tegevused tagasid seatud eesmärkide saavutamise.
4. **Korrigeerimine.** Juhul kui plaan viidi ellu edukalt, siis uuest saavutatud tasemest saab standard (baastase) järgnevale parendustsüklile. Juhul kui plaani elluviimine loodetud muutusi kaasa ei toonud, siis jääb standardina kehtima parendustsükli alguses standardiks (baastasemeks) olnud tase.



Kuna riigi infosüsteemi andmekogudes pole siiani ühtsetel alustel andmekvaliteeti mõõdetud, siis puudub ka ülevaade olemasolevast andmekvaliteedi olukorrast. Parendamistsükli esimene etapp (planeerimine) aga eeldab, et olukord oleks mõõdetud ja seeläbi baastase seatud. Et seda

lünka täita, otsustati kasutada raamistikus laiendatud PDCA-mudelit – ODPCA, mis võimaldab mudelis kasutada lisaetappi „**hindamine**“ (*observe*). Hindamise etapis teostatakse olukorra väljaselgitamisega seotud tegevusi (sh. olemasoleva andmekvaliteedi olukorra hindamine, andmekvaliteedi skoobi määramine).

Kombineerides ODPCA parendamistsükli ja andmekvaliteedi küpsusmudeli, loodi andmekvaliteedi juhtimise raamistik, mis võimaldab teostada andmekvaliteedi tsüklilist parendamist andmekvaliteedi üksikute kategooriate küpsustasemete tõstmise läbi.

## Kokkuvõte

Andmekvaliteedi tagamise uuringu tulemusena töötati välja vajalikud töövahendid riigi infosüsteemi andmekogudes andmekvaliteedi mõõtmiseks ja tagamiseks. Uuringu tulemiteks olid:

1. Andmekvaliteedi tunnused, mis võimaldavad hinnata andmekogude andmete kvaliteedi olukorda;
2. Andmekvaliteedi küpsusmudel, mis võimaldab hinnata andmekvaliteedi juhtimisega seotud protsesse;
3. Andmekvaliteedi juhtimise raamistik, mis võimaldab rakendada tsüklilist parendamisprotsessi andmekvaliteedi juhtimisega seotud protsesside küpsuse tõstmiseks.

Uuringu tulemustest koostati Andmekvaliteedi tagamise juhend andmekogu omanikele (eraldi dokument), mis sisaldab juhiseid selle osas, kuidas andmekvaliteedi juhtimise raamistikku kasutada. Juhendis kirjeldatakse tegevusi:

1. andmekvaliteedi olemasoleva olukorra hindamiseks (sh. andmete ulatuse määramine, organisatsiooni tsoneerimine, küpsustasemete hindamine);
2. parendustsükli planeerimiseks (sh. küpsuse sihttaseme planeerimine, parendamise plaani koostamine);
3. parendustsükli teostamiseks (küpsuse tõstmine 1.-lt 3.-nda tasemeni);
4. saavutatud taseme kontrollimiseks;
5. andmekogu andmekvaliteedi taseme korrigeerimiseks.

# Andmekvaliteedi uuring andmekogudes

Andmekogu omanikele mõeldud andmekvaliteedi tagamise juhendi valmimise järel teostati andmekvaliteedi uuring andmekogudes, mille eesmärkideks olid testida juhendi rakendatavust ja täiendada juhendit praktilise rakendamise tulemusel saadava tagasiside abil.

Uuringu ülesandeks oli rakendada juhend kolmel Riigi infosüsteemi andmekogul nende andmekogude omanike abiga, kaardistades juhendi abil nende andmekogude andmekvaliteedi olukorra ja seades eesmärgid edasiseks andmekvaliteedi parendamiseks. Uuringu tulemina pidi iga andmekogu osas valmima andmekvaliteedi parendusplaan, mis sisaldab tegevusi andmekvaliteedi juhtimise küpsuse tõstmiseks.

Uuringus osales kolme riigi infosüsteemi andmekogu:

1. Riigi infosüsteemi haldussüsteem (RIHA);
2. Rahvastikuregister (RR);
3. Aadressiandmete süsteem (ADS).

Uuringu käigus seati eesmärgiks viia läbi igas andmekogus järgnevad tegevused:

1. Hindamine:
  - a. Andmete ulatuse määratlemine;
  - b. Andmekvaliteedi organisatsiooni tsoneerimine;
  - c. Andmekvaliteedi küpsustaseme hindamine;
2. Planeerimine
  - a. Küpsuse sihttaseme planeerimine;
  - b. Parendamise plaani koostamine.

Uuringu tulemusena selgitati olemasolev andmekvaliteedi olukord uuritud andmekogudes ja täiustati andmekvaliteedi juhendit andmekogu omanikelt juhendi rakendamisel laekunud tagasiside alusel. Uuringu käigus töötati lisaks välja andmekvaliteedi küpsustaseme määramise tööriist, mis täienes uuringu lõpuks.

Uuringu tulemusena koostati uuringu aruanne (eraldi dokument: „Andmekvaliteedi uuring andmekogudes“). Aruanne sisaldab ülevaadet uuringus teostatud tegevuste osas, uuringute tulemusi ja soovitusi andmekogu omanikele juhendi rakendamisel.

# *Andmekvaliteedi tagamise juhend riigile*

Uuringu raames töötati välja andmekvaliteedi tagamise juhend riigile, mis sisaldab metoodikaid riigi infosüsteemi kui terviku andmekvaliteedi mõõtmiseks ja tagamiseks. Juhend on kirjeldatud eraldi dokumendina.

Juhend kirjeldab andmekogude kvaliteedi juhtimise seire ja järelevalve metoodika tehes ettepanekuid andmekvaliteedi seire meetrika osas. Meetrika toetub andmekogu omanikele mõeldud andmekvaliteedi juhendis kirjeldatud metoodikatele.

Riigi tasemel soovitatakse rakendada vähemalt järgnevat kolme mõõdikut:

1. Andmekvaliteedi juhtimise osakaal riigi andmekogudes;
2. Andmekvaliteedi juhtimise osakaal RIHA kasutajate seas;
3. Riigi andmekvaliteedi juhtimise küpsus.

Juhendis kirjeldatakse rakendusettepanekud andmekvaliteedi taseme tõstmise metoodika rakendamiseks riigis. Rakendusettepanekutena tuuakse välja järgnevad soovitused:

1. Meetmed andmekvaliteedi seiramiseks ja mõõtmisandmete kogumiseks;
2. Andmekvaliteedi tagamise juhendi turundamine;
3. Juhendile tagasiside kogumine ja juhendi täiendamine;
4. Koostöö teiste institutsioonidega.

Lisaks antakse juhendi soovitusi seadusandluse muutmise osas, mis võimaldaks andmekvaliteeti efektiivsemalt juhtida.

# Lisad

## Lisa 1: Tunnuste allikate loetelu

Nr	Allika nimi	Autor(id)	Avaldaja	Allika viide	Tunnuste arv mudelis	Allika kategooria
1	EVS-ISO/IEC 25012:2014	Eesti Standardikeskus	Eesti Standardikeskus	<a href="https://www.evs.ee/tooted/evs-iso-iec-25012-2014">https://www.evs.ee/tooted/evs-iso-iec-25012-2014</a>	15	Rahvusvaheline standard
2	"Open Data and Metadata Quality"	Euroopa Komisjon	Euroopa Komisjon	<a href="http://www.slideshare.net/OpenDataSupport/open-data-quality-29248578">http://www.slideshare.net/OpenDataSupport/open-data-quality-29248578</a>	9	Rahvusvaheline praktika
3	Data quality framework	Bank of England	Bank of England	<a href="http://www.bankofengland.co.uk/statistics/Documents/about/dqf.pdf">http://www.bankofengland.co.uk/statistics/Documents/about/dqf.pdf</a>	6	Panga praktika
4	"A conceptual framework for data quality management"	Banco de Mexico	Banco De Mexico	<a href="http://www.cemla.org/actividades/2014/2014-05-necesidadesinformacionfinanciera/2014-05-NecesidadesInformacionFinanciera-18.pdf">http://www.cemla.org/actividades/2014/2014-05-necesidadesinformacionfinanciera/2014-05-NecesidadesInformacionFinanciera-18.pdf</a>	9	Panga praktika
5	"The Challenges of Data Quality and Data Quality Assessment in the Big Data Era"	Li Cai , Yangyong Zhu	Data Science Journal	<a href="http://datascience.codata.org/article/10.5334/dsj-2015-002/">http://datascience.codata.org/article/10.5334/dsj-2015-002/</a>	9	Artikkel
6	"Providing assurance on data quality"	Carina K. Wangwe	ISACA	<a href="http://www.isaca.org/Knowledge-Center/Blog/Lists/Posts/Post.aspx?List=ef7cbc6d-9997-4b62-96a4-a36fb7e171af&amp;ID=483">http://www.isaca.org/Knowledge-Center/Blog/Lists/Posts/Post.aspx?List=ef7cbc6d-9997-4b62-96a4-a36fb7e171af&amp;ID=483</a>	8	Blogipostitus
7	"Data Quality Assessment"	Leo L. Pipino, Yang W. Lee, and Richard Y. Wang	MIT	<a href="http://web.mit.edu/tdqm/www/tdqmpub/PipinoLeeWangCAMApro2.pdf">http://web.mit.edu/tdqm/www/tdqmpub/PipinoLeeWangCAMApro2.pdf</a>	16	Artikkel
8	"A Scorecard approach to improving Data Quality"	Phil Nousak, Rob Phelps, PWC Consulting, Chapel Hill, NC	SAS	<a href="http://www2.sas.com/proceedings/sugi27/p158-27.pdf">http://www2.sas.com/proceedings/sugi27/p158-27.pdf</a>	7	Artikkel/Soovitus ed
9	"Monitoring Data Quality Performance Using Data Quality Metrics"	David Loshin	Informatica Corporation	<a href="https://it.ojp.gov/documents/informatica_whitepaper_monitoring_dq_using_metrics.pdf">https://it.ojp.gov/documents/informatica_whitepaper_monitoring_dq_using_metrics.pdf</a>	8	Ettevõtte praktika
10	"Conducting Data Quality Assessments"	U.S. Automated Directive System	U.S. Automated Directive System	<a href="http://pdf.usaid.gov/pdf_docs/Pnadw118.pdf">http://pdf.usaid.gov/pdf_docs/Pnadw118.pdf</a>	5	Artikkel
11	"Data Quality Audit Tool: Guidelines for implementation"	U.S. Agency for International Development	U.S. Agency for International Development	<a href="http://www.cpc.unc.edu/measure/resources/tools/monitoring-evaluation-systems/data-quality-assurance-tools/dqa-auditing-tool-implentation-guidelines.pdf">http://www.cpc.unc.edu/measure/resources/tools/monitoring-evaluation-systems/data-quality-assurance-tools/dqa-auditing-tool-implentation-guidelines.pdf</a>	7	Rahvusvaheline praktika
12	"Data on the Web Best Practices: Data Quality Vocabulary"	W3C	W3C	<a href="https://www.w3.org/TR/vocab-dqv/">https://www.w3.org/TR/vocab-dqv/</a>	9	W3C Working Draft

## Lisa 2: Tunnuste sünteesi tabel

Eestikeelne nimetus	Tunnuse kirjeldus	Koondnimetus	Nimetus allikas	Definitsioon allikas	Pärinemine allikast															
					1	2	3	4	5	6	7	8	9	10	11	12				
Kättesaadavus	Näitab kuivõrd lihtsalt on kasutajate infovajadusi rahuldavad andmed kättesaadavad. Arvestatakse kasutajapõhiseid erisusi andmetele ligipääsuks, andmete saadavust ajas, saamise kiirust, andmete kättesaadavaks tegemise protsessi lihtsust.	Accessibility	Accessibility	Määr, milleni andmete poole saab spetsiifilises kasutuskontekstis pöörduda, eriti inimeste poolt, kes vajavad mingi puude tõttu toetatavat tehnoloogiat või erilist konfiguratsiooni.	x															
		Accessibility	Accessibility	Whether a data access interface is provided. Data can be easily made public or easy to purchase.					x											
		Accessibility	Accessibility	Data can be obtained with ease, is comprehensible and usable.						x										
		Accessibility	Accessibility	The extent to which data is available, or easily and quickly retrievable.							x									
		Accessibility	Accessibility	Accessibility refers to the physical conditions in which users can obtain data: where to go, how to order, delivery time, clear pricing policy, convenient marketing conditions (copyright, etc), availability of micro or macro data, formats, etc.			x													
		Accessibility	Availability	Määr, milleni andmetel on atribuudid, mis võimaldavad autoriseeritud kasutajatel ja/või rakendustel neid spetsiifilises kasutuskontekstis kätte saada.	x															
		Accessibility	Availability	The availability of data is the extent to which it can be accessed; this also includes the long-term persistence of data.		x														
		Accessibility	Availability	Can the data be accessed now and over time?																x
Täpsus	Näitab kuivõrd täpselt kajastavad andmeväärtused tegeliku elu nähtust.	Accuracy	Accuracy	Määr, milleni andmeatribuudid esitavad spetsiifilises kasutuskontekstis korrektselt mõiste või sündmuse kavatsatud atribuutide tõeseid väärtusi. Sellel on kaks põhilist aspekti: Süntaktiline õigsus (syntactic accuracy) on määratletud kui andmeväärtuste lähedus väärtuste hulgale, mis on määratletud süntaktiliselt korrektseks peetavas valdkonnas. Semantiline õigsus (semantic accuracy) on määratletud kui andmeväärtuste lähedus väärtuste hulgale, mis on määratletud semantiliselt korrektseks peetavas valdkonnas.	x															
		Accuracy	Accuracy	The accuracy of data is the extent to which it correctly represents the characteristics of the real-		x														

				world object, situation or event.															
		Accuracy	Accuracy	Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values. Statistics must accurately and reliably portray reality.		x													
		Accuracy	Accuracy	Data provided are accurate. Data representation (or value) well reflects the true state of the source information. Information (data) representation will not cause ambiguity.				x											
		Accuracy	Accuracy	Distance between data reported and data real.			x												
		Accuracy	Accuracy	Data reflects reality.					x										
		Accuracy	Accuracy	Data accuracy refers to the degree with which data correctly represents the “real-life” objects they are intended to model. In many cases, accuracy is measured by how the values agree with an identified source of correct information (such as reference data). There are different sources of correct information: a database of record, a similar corroborative set of data values from another table, dynamically computed values, or perhaps the result of a manual process. An example of an accuracy rule might specify that for healthcare providers, the Registration Status attribute must have a value that is accurate according to the regional accreditation board. If that data is available as a reference data set, and automated process can be put in place to verify the accuracy, but if not, a manual process may be instituted to contact that regional board to verify the accuracy of that attribute.												x			
		Accuracy	Accuracy	Also known as validity. Accurate data are considered correct: the data measure what they are intended to measure. Accurate data minimize errors (e.g., recording or interviewer bias, transcription error, sampling error) to a point of being negligible.															x
		Accuracy	Accuracy	Is the data correctly representing the real-world entity or event?															x
		Accuracy	Accurate	Data element values are properly assigned .											x				
		Accuracy	Validity	Do the data clearly and adequately represent the intended result? Some issues to consider are: Face Validity. Would an outsider or an expert in the field agree that the indicator is a valid and logical measure for the stated result? Attribution. Does the indicator measure the contribution of the project?													x		

				Measurement Error. Are there any measurement errors that could affect the data? Both sampling and non-sampling error should be reviewed.																		
Vormisobivus	Näitab andmete esitusformaadi vastavust kasutajate vajadustele.	Appropriateness	Appropriateness	One format is more appropriate than another if it is more suited to user needs.				x														
Koherentsus	Näitab statistiliste andmetöötluse kontekstis, kas erinevatest allikatest pärinevad algandmed on omavahel kokkupanadavad.	Coherence	Coherence	Coherence of statistics is their adequacy to be reliably combined in different ways and for various uses. It is, however, generally easier to show cases of incoherence than to prove coherence. When originating from a single source, statistics are normally coherent in the sense that elementary results derived from the concerned survey can be reliably combined in numerous ways to produce more complex results. When originating from different sources, and in particular from statistical surveys of different nature and/or frequencies, statistics may not be completely coherent in the sense that they may be based on different approaches, classifications and methodological standards.				x														
Võrreldavus	Näitab erinevatest allikatest pärinevate samade statistiliste muutujate võrreldavust klassifitseerivate muutujate lõikes.	Comparability	Comparability	Comparability aims at measuring the impact of differences in applied statistical concepts, definitions, and measurement tools/procedures when statistics are compared between geographical areas, non-geographical domains, or over time.				x														
Täielikkus	Näitab vaadeldava andmekomplekti täielikkust temas sisalduvate andmete väärtuste esinemise kontekstis.	Completeness	Complete	Data element is (1) always required or (2) required based on the condition of another data element																x		
		Completeness	Completeness	Määr, milleni spetsiifilises kasutus kontekstis on olemiga seotud teemakohaste andmete kõigil eeldatud atribuutidel ja nendega seotud olemieksemplaridel väärtused.	x																	
		Completeness	Completeness	The completeness of data is the extent to which it includes the data items or data points that are necessary to support the application for which it is intended.		x																
		Completeness	Completeness	Whether the deficiency of a component will impact use of the data for data with multi-components. Whether the deficiency of a component will impact data accuracy and integrity.																		x
		Completeness	Completeness	Degree to which values are present in a data collection																		x
		Completeness	Completeness	All the necessary data is present.																		





Reeglipärasus	Näitab vastavust kokkulepitud reeglitele ning standarditele andmete hõivel, avalikustamisel ja kirjeldamisel.	Conformance	Conformance	The conformance of data is the extent to which it follows a set of explicit rules or standards for capture, publication and description.	x																
		Conformance	Compliance	Määr, milleni andmetel on atribuudid, mis vastavad spetsiifilises kasutuskontekstis andmekvaliteeti puudutavatele standarditele, kokkulepetele või kehtivatele õigusaktidele ning samalaadsetele normidele.	x																
		Conformance	Conformance	This dimension refers to whether instances of data are either store, exchanged, or presented in a format that is consistent with the domain of values, as well as consistent with other similar attribute values. Each column has numerous metadata attributes associated with it: its data type, precision, format patterns, use of a predefined enumeration of values, domain ranges, underlying storage formats, etc.																x	
		Conformance	Conformance	Is the data following accepted standards ?																	x
Metoodiline järjepidevus	Näitab vaadeldavas andmehulgas andmetevaheliste vastuolude esinemist (või nende puudumist).	Consistency	Consistency	Määr, milleni andmeatribuudid on spetsiifilises kasutuskontekstis vabad vasturääkivustest ja kooskõlalised teiste andmetega. Kooskõla võib olla seoses ühe olemi andmetega, seoses võrreldavate olemite sarnaste andmetega või mõlemal juhul.	x																
		Consistency	Consistency	The consistency of data is the extent to which it does not contain contradictions that would make its use difficult or impossible.	x																
		Consistency	Consistency	After data have been processed, their concepts, value domains, and formats still match as before processing. During a certain time, data remain consistent and verifiable. Data and the data from other data sources are consistent or verifiable.																	
		Consistency	Consistency	Coherence of the same datum, represented in multiple copies, or different data to respect integrity constraints and rules																	
		Consistency	Consistency	There are no contradictions in the data.																	
		Consistency	Consistency	In its most basic form, consistency refers to data values in one data set being consistent with values in another data set. A strict definition of consistency specifies that two data values drawn from separate data sets must not conflict with each other, although consistency does not necessarily imply correctness. Even more complicated is the notion of consistency with a set of predefined constraints. More formal consistency constraints can be encapsulated as a set of rules that specify																	



				processes and analysis methods over time?																	
		Credibility	Reliability	The data generated by a program's information system are based on protocols and procedures that do not change according to who is using them and when or how often they are used. The data are reliable because they are measured and collected consistently.																x	
		Credibility	Reputation	The extent to which data is highly regarded in terms of its source or content.																x	
Objektiivsus	Näitab kuivõrd objektiivselt andmed kajastavad kirjeldatavat nähtust.	Objectivity	Integrity	Data have integrity when the system used to generate them is protected from deliberate bias or manipulation for political or personal reasons.																x	
		Objectivity	Objectivity	The extent to which data is unbiased, unprejudiced, and impartial.																x	
Ajakohasus	Näitab kuivõrd ajakohased on andmed reaalelulise nähtuse osas.	Currency	Currency	Degree to which a datum is up-to-date.																x	
		Currency	Currency	Currency refers to the degree to which information is current with the world that it models. Currency can measure how "up-to-date" information is, and whether it is correct despite possible time-related changes. Data currency may be measured as a function of the expected frequency rate at which different data elements are expected to be refreshed, as well as verifying that the data is up to date. This may require some automated and manual processes. Currency rules may be defined to assert the "lifetime" of a data value before it needs to be checked and possibly refreshed. For example, one might assert that the contact information for each customer must be current, indicating a requirement to maintain the most recent values associated with the individual's contact data.																	x
		Currency	Currentness	Määr, milleni andmeatribuutide väärtustel on spetsiifilises kasutuskontekstis õige vanus.	x																
		Currency	Timeliness	The timeliness of data is the extent to which it correctly reflects the current state of the entity or event and the extent to which the data (in its latest version) is made available without unnecessary delay.																x	
		Currency	Timeliness	Data values represent the most current information for the specific use.																x	
		Currency	Timeliness	The extent to which data is sufficiently up-to-date for the task at hand.																x	
		Currency	Timeliness	Is the data representing the actual situation and it is published soon enough ?																x	



				processes.																		
		Processability	Processability	Is the data machine readable ?																x		
Taastatavus	Määr, milleni spetsiifilises kasutuskontekstis on andmetel atribuudid, mis võimaldavad hallata ja alal hoida spetsifitseeritud käitus- ja kvaliteedi taset, sealhulgas tõrgete korral.	Recoverability	Recoverability	Määr, milleni spetsiifilises kasutuskontekstis on andmetel atribuudid, mis võimaldavad hallata ja alal hoida spetsifitseeritud käitus- ja kvaliteedi taset, sealhulgas tõrgete korral.	x																	
Asjakohasus	Näitab andmete asjakohasust äriülesannete täitmise kontekstis.	Relevance	Appropriate amount of data	The extent to which the volume of data is appropriate for the task at hand.																		
		Relevance	Relevance	The relevance of data is the extent to which it contains the necessary information to support the application.	x																	
		Relevance	Relevance	Relevance is the degree to which statistics meet current and potential users' needs. It refers to whether all statistics that are needed are produced, the extent to which concepts used (definitions, classifications etc) reflect user needs, and the extent to which users are satisfied with the disseminated statistics.			x															
		Relevance	Relevance	The extent to which data is applicable and helpful for the task at hand.																		
		Relevance	Relevance	Does the dataset include an appropriate amount of data?																		x
		Relevance	Value-Added	The extent to which data is beneficial and provides advantages from its use.																		x
Õigeaegsus	Näitab, kas andmed jõuavad kohale õigeaegselt nende vajatavuse kontekstis.	Timeliness	Timeliness	Within a given time, whether the data arrive on time. Whether data are regularly updated. Whether the time interval from data collection and processing to release meets requirements.																		
		Timeliness	Timeliness	Timeliness refers to the time expectation for accessibility and availability of information. Timeliness can be measured as the time between when information is expected and when it is readily available for use. For example, in the financial industry, investment product pricing data is often provided by third-party vendors. As the success of the business depends on accessibility to that pricing data, service levels specifying how quickly the data must be provided can be defined and compliance with those timeliness constraints can be measured.																		
		Timeliness	Timeliness	Are data timely enough to influence management																		x









## Lisa 3: Hinnang tegurite kaetuse osas aspektidega

Lähteülesandes loetletud aspektid:	Tegurid					
	1	2	3	4	5	6
<b>1) ANDMEKVALITEEDI JUHTIMINE</b>						
seadusloome mõju andmekvaliteedile;			X			
riigi juhtimise mõju andmekvaliteedile;			X			
juhtkonna arusaamine andmekvaliteedi tähtsusest;				X		
keskastme juhtide arusaamine andmekvaliteedi tähtsusest;				X		
andmekvaliteedi haldamisega seotud haridus- ja koolitusprogrammide tähtsustamine ja kasutamine;		X				
andmekvaliteedi visiooni olemasolu, detailsus ja rakendatavus;		X				
andmekvaliteedi juhtimine ja kontroll;		X				
andmete sisendvoo kvaliteedi juhtimine ja kontroll; „andmete tootjate“ kvaliteedi juhtimine;	X	X	X	X	X	X
kasutajate iseloom, hulk ja varieeruvus (avalik kasutaja vs ametkondlik kasutaja) ning selle mõju andmekvaliteedile;	X	X	X	X	X	X
andmeid muutvate infosüsteemide iseloom/iseärasused, hulk ja varieeruvus (avalikud infosüsteemid vs erakasutajate infosüsteemid), mis võivad mõjutada andmekvaliteeti;	X					
infosüsteemide muudatuste juhtimine ja kontroll;				X		
tööplaanide olemasolu ja nende teostatavus.				X	X	
<b>2) ORGANISATSIOONI VÕIMEKUS</b>						
personali arusaamine andmekvaliteedi tähtsusest;				X		
andmekvaliteedi mõõtmise metoodika olemasolu ja reaalne mõõtmine;		X				
protsesside pidev parendamine;				X		
koostöö organisatsiooni erinevate osapoolte vahel (juhid, arhitektid, projektijuhid, arendajad, haldajad);				X		
andmekvaliteedi kajastumine eelarves – andmekvaliteedi tagamisele eraldatud summad eelarves; tulude-kulude hindamine;				X		
riskijuhtimine – riskihalduse poliitikate, metoodikate, meetodite, kordade ja reeglite olemasolu ja nende rakendamine;				X		
andmekvaliteedi tagamise metoodikate ja meetodite mõistmine;		X		X	X	
infosüsteemide auditeerimiste läbi viimine ja ülevaatete esitamine; mahud, sagedused ja informatsiooni levitamine;				X		
äriprotsesside parendamine ja kohandamine kasutajate/teenindatavate vajadusele ja muutuvatele oludele				X		
sisekontrolli mõju andmekvaliteedile.				X		
<b>3) ORGANISATSIOONI STRUKTUUR</b>						
organisatsiooniline struktuur andmekvaliteedi tagamisel;					X	
andmekvaliteedi juhi asukoht organisatsioon struktuuris, tema roll ja volitused;					X	
andmekvaliteedi poliitikate, meetodite, metoodikate, kordade, standardite ja reeglite olemasolu;				X		
organisatsiooni sisemine kultuur.					X	
<b>4) VAHENID JA PERSONAL</b>						
personali kompetentsus;				X	X	
personali mõjutav töökeskkond;				X	X	
kasutatavad (materiaalsed) vahendid.				X	X	

## Lisa 4: Andmekvaliteedi kipsusmudel

Kategooria	Tase	Väide
Juhtimine ja planeerimine	1	Andmekogu andmekvaliteedi nõuded ei ole määratletud ega dokumenteeritud.
	2	Andmekogu andmekvaliteedi nõuded on määratletud ja dokumenteeritud.
	2	Määratletud ja dokumenteeritud on kõik poliitikatele alluvad kriitilised andmed.
	3	Andmekvaliteet on nõuetele vastavuse osas kontrollitud.
	3	Andmekvaliteedi poliitikad on dokumenteeritud ja avaldatud.
	4	Andmekvaliteedi juhtimine on vaadeldavas tsoonis ühtsetel alustel juurutatud.
	4	Andmekvaliteeti mõõdetakse ja parendatakse mõõtmistulemuste analüüsi alusel.
	4	Andmekvaliteedi mõõdikud on seotud kõrgemate tasemete strateegiatega või poliitikatega.
	5	Andmekvaliteedi protsesside osas toimub järjepidev parendamine.
Organisatsioon ja vastutused	5	Andmekvaliteeti juhitakse poliitikate regulaarse ülevaatus, muutmise ja avaldamisega.
	1	Andmekvaliteedi protsessi omanik ei ole tuvastatav.
	1	Andmete omanikud on määratlemata.
	2	Andmekvaliteedi protsessi omanik on tuvastatav.
	2	Osade andmete omanikud on määratletud.
	3	Andmekvaliteedi protsesside omaniku vastutused on selgelt määratud ametijuhendiga.
	3	Andmete omanikud on määratletud.
	4	Loodud on andmekvaliteedi eest vastutav juhtivatest ametikohtadest moodustatud rühmitus.
	4	Rühmitusse on kaasatud ka esindajad teistest andmekoguga seotud osapooltest.
Protsessid	5	Rühmituses toimub regulaarne andmekvaliteedi juhtimise vastutuste ülevaatus ja täiendamine.
	1	Andmekvaliteedi profiil ei ole kirjeldatud ja seda ei kasutata protsessi käigus puuduste avastamiseks.
	1	Mõõdetavad andmekvaliteedi tunnused ei ole määratletud ega dokumenteeritud.
	1	Ei otsita võimalusi probleemide ennetamiseks.
	1	Tuvastatud sündmuseid ei registreerita ja nende lahendamine ei ole jälgitav.
	1	Andmekvaliteedi tagamise protsessid ei ole määratletud ega dokumenteeritud.
	2	Andmekvaliteedi profiil on kirjeldatud ja seda kasutatakse protsessi käigus puuduste avastamiseks.
	2	Mõõdetavad andmekvaliteedi tunnused on määratletud ja dokumenteeritud.
	2	Otsitakse võimalusi probleemide ennetamiseks.
	2	Tuvastatud sündmused registreeritakse ja nende lahendamine on jälgitav.
	2	Andmekvaliteedi tagamise protsessid on määratletud ja dokumenteeritud.
	2	Andmekvaliteedi puuduste mõju on tuvastatud.
	3	Andmekvaliteeti mõõdetakse ja puudujäägid avastatakse protsessi varajases staadiumis.
	4	Andmekvaliteedi mõõtmine toimub regulaarselt.
	4	Mõõtmistulemused on juhtkonnale kättesaadavad.
Teadmus ja kompetentsid	4	Andmekvaliteedi parendamise tegevusplaan on määratletud ja dokumenteeritud.
	4	Andmekvaliteedi puuduste juurpõhjuste uurimine on levinud praktika.
	5	Toimub regulaarne andmekvaliteedi mõõdikute ülevaatus protsesside parendamisvõimaluste leidmise eesmärgil.
	1	Andmekvaliteedi teadlikkuse tõstmiseks ei ole läbi viidud andmekvaliteedi teemalisi koolitusi.
	1	Võtmeisikud ei ole tuvastatavad.
	2	Andmekvaliteedi teadlikkuse tõstmiseks on läbi viidud andmekvaliteedi teemalisi koolitusi.
	2	Võtmeisikud kavandavad ja annavad soovitusi andmekvaliteedi juhtimise osas.
	3	Andmekvaliteeti mõjutavad töötajad on läbinud andmekvaliteedi teemalise koolituse.
	3	Eksisteerib mitteametlik mentorlus andmekvaliteedi teadlikkuse tõstmise osas.
Töövahendid	3	Jagatud teadmusbaasina dokumenteeritakse parimaid praktikaid andmekvaliteedi tagamisel.
	4	Andmekvaliteedi teemalised koolitused toimuvad regulaarselt.
	5	Andmekvaliteedi koolituste sisu vaadatakse regulaarselt üle ja täiendatakse vastavalt vajadusele.
Töövahendid	1	Andmekvaliteedi standardid infotehnoloogiliste vahendite jaoks ei ole välja töötatud ega dokumenteeritud.
	1	Töövahendid andmekvaliteedi mõõtmiseks ja parendamiseks ei ole juurutatud.
	2	Andmekvaliteedi standardid infotehnoloogiliste vahendite jaoks on välja töötatud ja dokumenteeritud.

Kategooria	Tase	Väide
	2	On tehtud pingutusi "keskse tõeallika" andmekogude loomiseks.
	2	Töövahendid andmekvaliteedi mõõtmiseks ja parendamiseks on osaliselt juurutatud.
	3	Juurutatud on töövahendid andmekvaliteedi mõõtmiseks ja parendamiseks.
	3	Andmekvaliteedi standarditele vastavust jälgitakse uute projektide ettepanekute staadiumis.
	4	Andmekvaliteedi raporteerimise vahendid võimaldavad teostada analüüsi andmekvaliteedi juhtimisotsuste langetamiseks.
	4	Andmekvaliteedi juhtimist toetavate infosüsteemide arendustegevuste sihtportfell on kokku lepitud.
	4	Rakendatud on andmekvaliteedi profiilist lähtuvad automaatsed andmeparenduse või -kontrolli protseduurid.
	5	Andmekogude arendamisel on rakendatud põhiandmete põhimõtteid.
	5	Andmekvaliteedi juhtimist toetavate infosüsteemide arendustegevuste sihtportfelli vaadatakse regulaarselt üle ja täiendatakse vastavalt vajadusele.

---

***www.pwc.ee***