

TALLINNA PEDAGOOGIKAÜLIKOOL

Matemaatika-loodusteaduskond

Informaatika osakond

Martin Luts

INFOKVALITEET JA SELLE JUHTIMINE ORGANISATSIOONIS

Magistritöö IT juhtimise erialal

Juhendaja: PhD Katrin Niglas

Autor: " _____ " _____ 2004. a.

Juhendaja: " _____ " _____ 2004. a.

Osakonnajuhataja: " _____ " _____ 2004. a.

Tallinn 2004

Autorideklaratsioon

Deklareerin, et käesolev magistritöö on minu töö tulemus ja seda ei ole kellegi teise poolt varem kaitsmisele ega mingil muul puhul esitatud ega avaldatud.

26.07.2004

ANNOTATSIOON

Luts, M. (2004) Infokvaliteet ja selle juhtimine organisatsioonis. Magistritöö IT juhtimise erialal, TPÜ.

Infoühiskonnas on organisatsioonide juhtimise ja klientide teenindamise üheks kriitiliseks teguriks informatsiooni optimaalne kvaliteet. Infokvaliteedi käsitlusi on loodud mitmeid, käesolevas töös on esitatud infokvaliteet koosnevana neljast võtmemõõdikust: asjakohasus, õigsus, kasutatavus, andmete kirjeldatus; ning ligi kahekümnest alammõõdikust. Iga mõõdik on kirjeldatud eesti- ja ingliskeelse nimetusega ning defineeritud. Aga ainuüksi teadmisest, mis on infokvaliteet, ei piisa selle hindamiseks ja parendamiseks. Infokvaliteedi juhtimiseks on välja pakutud algupärane metoodika ning kirjeldatud selle rakendamist Riiklikus ehitisregistris. Toodud metoodikast võivad kasu leida nii andmete omanikud, IT-juhid kui ka infosüsteemide arendajad.

Märksõnad: informatsiooni kvaliteet, infokvaliteet, andmete kvaliteet, andmekvaliteet, infokvaliteedi juhtimine.

Viimane versioon tööst on saadaval Internetis aadressil

<http://www.itcollege.ee/~martin/infokvaliteet/>

SISUKORD

ANNOTATSIOON.....	3
SISUKORD.....	4
SISSEJUHATUS	6
Teema aktuaalsus	6
Probleemide sõnastus.....	8
Magistritöö eesmärkide püstitus	10
Magistritöö struktuur ja kasutatud meetodika	10
1. INFOKVALITEEDI TEOREETILINE KÄSITLUS	12
1.1. Andmed, informatsioon, teave.....	13
1.2. Info elutsükli käsitlused	14
1.2.1. CRUD-maatriks	15
1.2.2. Andmete elutsükkel Eesti seaduste mõistes.....	16
1.2.3. Informatsiooni elutsükkel organisatsioonis	16
1.3. Andmekvaliteet ja infokvaliteet.....	18
1.4. Olemasolevate infokvaliteedi käsitluste võrdlus	19
1.4.1. Gartner Groupi infokvaliteedi käsitlus.....	20
1.4.2. H. Milleri infokvaliteedi käsitlus	21
1.4.3. COBITi infokvaliteedi käsitlus	21
1.4.4. Olemasolevate käsitluste analüüs	23
1.5. Infokvaliteedi moodsuse raamistik	24
1.5.1. Asjakohasus	26
1.5.2. Õigsus	27
1.5.3. Kasutatavus	31
1.5.4. Andmete kirjeldatus	34
2. INFOKVALITEEDI JUHTIMISE METOODIKA.....	35
2.1. Olemasolevate meetodikate võrdlus	36
2.1.1. AS Resta kasutatav meetodika.....	36
2.1.2. Net Group OÜ kasutatav meetodika.....	36
2.1.3. Muud meetodikad	37
2.2. Meetodika loomise alused	38
2.3. Infokvaliteedi juhtimise protsessi kirjeldus	39

Samm 1. Infokvaliteedi strateegia koostamine	41
Samm 2. Oluliste andmeolemite määratlemine	43
Samm 3. Kvaliteedinõuete kehtestamine	43
Samm 4. Mõõdikute hindamistehnika koostamine	44
Samm 5. Infokvaliteedi mõõtmise läbiviimine	45
Samm 6. Mitteoptimaalse infokvaliteedi mõju hindamine	45
Samm 7. Infokvaliteedi parendusprojektide planeerimine	46
Samm 8. Infokvaliteedi parendusprojektide läbiviimine	49
Samm 9. Tulemuste hindamine.....	50
2.4. Metoodika rakendamise juhised	50
3. INFOKVALITEEDI JUHTIMISE METOODIKA RAKENDUSNÄIDE ...	51
KOKKUVÕTE	52
RESÜMEE.....	54
KASUTATUD KIRJANDUS	58
INFOKVALITEEDIALASTE TERMINITE REGISTER.....	62

SISSEJUHATUS

Magistritöö sissejuhatava osa eesmärkideks on:

- selgitada lugejale infokvaliteedi olulisust ja aktuaalsust Eesti organisatsioonide kontekstis;
- põhjendada infokvaliteedi teema sobivust magistritööks;
- sõnastada magistritöö eesmärk;
- kirjeldada eesmärgini jõudmise viisi ja magistritöö struktuuri.

Teema aktuaalsus

Organisatsioonidel ja indiviididel on mitmeid olulisi varasid. Läbi ajaloo on nende osatähtsus muutunud: kui agraarühiskonnas olid esmasteks rakendatavateks ressursideks loodusjõud ja lihased ning strateegiliseks ressursiks toorained; tööstusühiskonnas finantskapital ja energia; siis info- ehk teabeühiskonnas on organisatsiooni oluliseks varaks töötajate, klientide ning kaubamärgi kõrval informatsioon [Eesti Tulevikuuringute Instituut 1997]. Oluline on teha vahet, et konkurentsieelist loovaks varaks loetakse mitte *infotehnoloogiat* [Carr 2003] vaid infot ennast – *infovara* ehk *infokapitali*¹. IT juhi üheks ülesandeks on seda vara hallata, juhtida [Cobit Executive Summary 2000]. Üldteada on juhtimistarkus, et ei ole võimalik juhtida seda, mida ei saa mõõta. Kuidas siis infovara hinnata, mõõta? Shannoni jt tööde alusel on võimalik mõõta informatsiooni hulka ehk mahtu. Eesti suurettevõtete andmemahud on suurusjärgus kümnetes terabaitides, andmeaitade mahud sadades megabaitides; kogu maailmas ,toodeti' aastal 2002 suurusjärgus 10^8 – 10^9 baiti uut infot [Lyman 2003]. Paraku ei iseloomusta need arvud organisatsiooni hallatava info kõiki külgi, lisaks kvantiteedile võib ja tuleb infot iseloomustada kvaliteedi vaatevinklist. Aga mis on informatsiooni kvaliteet? Magistritöö autor on seda paljudelt IT-inimestelt küsinud, vastused on enamasti Augustinuse mis-on-aeg tõdemuse stiilis². Osutub, et info kvaliteet – erinevalt kvantiteedist – on mitmetahuline

¹ Andmed ise on väärtuslikumad kui andmete haldamise vahendid.

² *What then is time? If no one asks me, I know what it is. If I wish to explain it to him who asks, I do not know* [BrainyQuote 2004].

suurus [Friedman 2003, 2; Miller 1996]. Seda, mis on infokvaliteet ning kuidas seda juhtida, käsitlebki käesolev magistritöö.

Infokvaliteedi teema on oluline ja aktuaalne nii Eesti avaliku sektori asutustes kui ka erasektoris: autori andmetel mõõdavad mitmed omavalitsused (nt Tallinna Linnavalitsus), riigiasutused (nt Justiitsministeerium [Laks 2004]) ja firmad nende hallatava informatsiooni kvaliteeti; dokumendis „Eesti infopoliitika põhialused aastateks 2004-2006” on deklareeritud valdkonnad, millistes infokvaliteedi küsimused on tõstetud põhialuste tasemele:

- arendada andmekogusid, eesmärgiga tagada *andmete kvaliteet, käideldavus ja koostalitusvõime*³ [Infopoliitika põhialused 2003, 9];
- juhtimise ja majandamise tõhustamiseks arendada välja ministeeriumide ja riigiasutuste tegevusstatistikat ning finantsinformatsiooni koondavad operatiivsed infosüsteemid, mis võimaldaksid tagada *juhtimisotsuste kvaliteedi* [samas, 9];
- maaga seotud ja geograafilist asukohta omava informatsiooni (riiklike ruumiandmete) infrastruktuuri arendamine, Maainfosüsteemi edasiarendamine, (sh. täiendamine ja *andmekvaliteedi parandamine*) võimaldamaks pakkuda ruumiandmete avalikustamise ja haldamisega seotud e-teenuseid riigiasutustele ja kohalikele omavalitsustele [samas, 13];
- haldussuutlikkuse tõhustamine elanike arvestuse alal: rahvastikuregistri⁴ *andmekvaliteedi parandamine ja registrile parema juurdepääsu kindlustamine* [samas, 14].

Autori tõlgenduse järgi on Eesti riigi infosüsteemide arengus läbitud infrastruktuuri – so riistvara ja operatsioonisüsteemide – esmase soetamise tase, samuti on loodud erinevaid registreid ja andmekogusid ning kehtestatud nende asutamise ja pidamise

³ Autori rõhutused.

⁴ Rahvastikuregistri andmekvaliteedile ja sellel põhinevate e-teenuste kvaliteedile viitavad ka eri versioonid dokumendist „Infopoliitika tegevuskava 2005. a” [Infopoliitika tegevuskava 2004a; Infopoliitika tegevuskava 2004b], samuti BNSi uudised „Eesti riigi sajad registrid ei ühti omavahel” 13.10.2003 ja „Rahvastikuregistris pole umbes 17.000 inimese aadressi” 11.05.2004.

põhimäärused. Eesti riigi viimasel ajal püstitatud infopoliitilised eesmärgid sisaldavad juba viiteid infokvaliteedile, seadustesse (registrite asutamise ja pidamise aktid jm, vt allpool) ei ole need veel jõudnud. Põhitähelepanu liigub andmekogude põhjal pakutavate teenuste kvaliteedile. Viimase tagamiseks on aga lisaks infrastruktuurile ja registrite olemasolule vajalikuks eelduseks infosüsteemide kasutajate oskused ja motivatsioon ning teenindamiseks ja juhtimiseks tarviliku informatsiooni kvaliteet⁵.

Infokvaliteedi juhtimise teema on sobiv IT juhtimise eriala tööna, kuna tüüpiliselt on kaasajal see tööloik organisatsioonides usaldatud IT struktuuriüksustele.

Probleemide sõnastus

Vaatamata infokvaliteedi olulisusele on autori arvates infokvaliteedi teema Eestis varjul võrreldes muude infoarhitektuuri teemadega.

Olulistest andmekogusid reguleerivates dokumentides:

- Andmekogude seadus (viimane redaktsioon 28. 04. 2004),
- Riigi infosüsteemi ja andmeteenuste seaduse eelnõu [RISAT 2004],
- Avaliku teabe seadus,
- riigi (põhi)registrite aktid: Rahvastiku registri seadus, Ehitisregistri seadus,
- Andmekogude riikliku registri asutamise ja pidamise põhimäärus,
- Arhiiviseadus,
- infotehnoloogiaalaseid riigihankeid sätestavad aktid,
- Riigi infosüsteemide juhtimise korrastamine

ei viidata infokvaliteedile. Riigiteataja (<http://www.riigiteataja.ee>) märksõna „andmekvaliteet*“ otsingu ainukene tulem on akt „Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsioon“, mille artikkel 5 „Andmekvaliteet“ kõlab järgmiselt:

„Automatiseeritult töödeldavad isikuandmed:

- a. tuleb hankida ja töödelda ausal ning seaduslikul teel;

⁵ Infokvaliteet ei ole eesmärk omaette vaid vahend andmetel ja infol tuginevate otsustuste ja teenuste kvaliteedi tagamiseks.

- b. tuleb koguda seaduspärasel ja täpselt määratletud eesmärkidel ning kasutada vastavalt nendele eesmärkidele;
- c. peavad olema adekvaatsed, asjakohased, piisavad vastavalt kogumise eesmärkidele;
- d. peavad olema õiged ja vajadusel täiendatavad;
- e. peavad olema säilitatud vormis, mis lubab andmesubjekti teha kindlaks mitte kauemaks, kui perioodiks, mis on vajalik andmete kogumise eesmärkide saavutamiseks” [Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsioon 1981, artikkel 5].

Konventsiooni andmekvaliteedi põhimõtte leiab kajastust Isikuandmete kaitse seaduses, nimelt §5 lõige 5 ütleb „andmete kvaliteedi põhimõtte – isikuandmed peavad olema ajakohased, täielikud ning vajalikud antud andmetöötamise eesmärgi saavutamiseks”.

Euroopa seadustes käsitletakse infokvaliteedi küsimusi peamiselt järgmistes valdkondades: isikuandmete kaitse, põllumajanduslike andmete käsitlemine ja statistika kvaliteet [Euroinfo 2004; Eesti Õiguskeele Keskus 2004].

Infokvaliteedi küsimusi ei käsitleta autori andmetel⁶ Eesti ülikoolide ja rakenduskõrgkoolide infosüsteemide kursustes.

Autor näeb infokvaliteediga seotult järgmisi probleeme:

- infokvaliteet ja selle mõõdikud ei ole rangelt defineeritud, arusaamine infokvaliteedi olemusest on ebamäärane. Eksisteerib mitmeid definitsioone ja osaliselt kattuvaid loetelusid infokvaliteedi teguritest;
- infokvaliteedi auditeerimiseks, parendamiseks ja juhtimiseks ei ole väljakujunenud ühtset, laialdaselt heakskiidetud, piisavalt laia ja huvilistele vabalt kättesaadavat meetodilist käsitlust. Infokvaliteedi

⁶ Väite aluseks on vestlused IT Kolledži rektori ja õppejõududega, TTÜ ainekavade analüüs ja vestlused EBSi IT eriala tudengitega. Ainukeseks autorile teadaolevaks erandiks on P. Parmaksoni TPÜs loetavad infosüsteemide kursused.

hindamist ja auditeerimisteenust pakkuvad firmad ei avalda üldjuhul oma detailset metoodikat, pidades selle omamist ja unikaalsust konkurentsieeliseks;

- infokvaliteedialased teadmised on Eestis vähelevinud ning eestikeelsed terminid lõplikult välja kujunemata.

Magistritöö eesmärkide püstitus

Tulenevalt eelnevalt määratletud probleemidest seadis magistrant magistritööle järgmised eesmärgid:

1. eestikeelsete infokvaliteedi ja infokvaliteediga seotud terminite väljapakkumine ja infokvaliteedi mõõdikute defineerimine;
2. metoodika loomine infokvaliteedi juhtimiseks ning selle testimine ja täiendamine reaalses projektides;
3. infokvaliteedialaste teadmiste levitamiseks aluse loomine ja sellega eelduse loomine infokvaliteedi kasvuks Eesti organisatsioonides.

Magistritöö struktuur ja kasutatud metoodika

Magistritöö koosneb sissejuhatausest, kolmest peatükist, kokkuvõttest, ingliskeelsest resümeeist, kasutatud kirjanduse loetelust ja registrist.

Juhtimiseks on vaja teada, mida juhitakse [IT juhtimise käsiraamat 1999, ptk 11.1 lk 4]. Magistritöö esimene peatükk – „Infokvaliteedi teoreetiline käsitlus” – käsitleb infokvaliteeti üldisel tasemel, üritades vastata küsimusele „mis on infokvaliteet?”. Valitud ja võrreldud on kolme ingliskeelsetest allikatest pärit infokvaliteedi käsitlust, kohendatud eestikeelseid vasteid ning sünteesitud infokvaliteedi struktuur.

Teine peatükk – „Infokvaliteedi juhtimise metoodika” – kirjeldab autori koostatud algupärast metoodikat infokvaliteedi hindamiseks ja organisatsiooni infokvaliteedi juhtimiseks. Toodud metoodika loomisel on aluseks võetud praktilised kogemused: autor on metoodikat kasutanud ja täiendanud mitmetes infokvaliteedi projektides, millistest ühe läbiviimist ja tulemusi on kirjeldatud kolmandas peatükis – „Infokvaliteedi haldamise metoodika rakendusnäide”.

Infokvaliteediga seotud terminid on defineeritud tekstis nende sissetoomisel. Eestikeelsetele terminitele on lisatud ingliskeelsed vasted, seda eelkõige eesmärgiga hõlbustada valdavalt ingliskeelse infokvaliteedialase kirjanduse kasutamist. Töö lõpus asub infokvaliteediga seotud terminite register.

Käesoleva töö kirjutamine toimus samaaegselt infokvaliteedi mõisteraamistikku tudengitele õpetades ning infokvaliteedialaseid teoreetilisi teadmisi erinevate organisatsioonide infokvaliteedi projektides rakendades. Saadud tagasiside põhjal täiendas autor oluliselt esimest ja teist peatükki. Kasutades infosüsteemide arendusmetoodikate termineid, võib öelda, et antud magistritöö loodi iteratiivset ja inkrementaalset tehnikat kasutades.

Autor kaalus ühe teadmiste hankimise vahendina ka küsitluste ja intervjuude läbiviimist Eesti organisatsioonides: nii andmeid haldavates kui ka infokvaliteedi hindamist teenusena pakkuvates firmades. Viimane neist sai teoks, kuid esimesest loobus autor peale tuttavate IT-töötajate pilootküsitlust, millest selgus, et infokvaliteedi teema on Eestis niivõrd uus, et küsitlusest ei ole oodata abi magistritöö eesmärkide saavutamiseks.

Magistritöö loomise ajenditeks olid autori tööalane huvi infoarhitektuuri ja infokvaliteedi teema vastu ning soov valdkonna olukorda paremaks muuta.

1. INFOKVALITEEDI TEOREETILINE KÄSITLUS

Käesolev peatükk käsitleb infokvaliteeti üldisel tasemel, üritades vastata küsimusele „mis on infokvaliteet?”. Käsitlus on universaalne, st rakendatav suvalisele infokogumile – olgu see siis analoog- või digitaalkujul, paberipõhises asjaajamissüsteemis või objektorienteeritud andmebaasis, kui ei ole märgitud teisiti.

Peatüki eesmärgiks on defineerida infokvaliteet kujul, milline oleks kasutatav:

- infokvaliteedi juhtimisel:
 - infokvaliteedi auditi projektides lepingu objekti ja ulatuse määratlemisel ning projektide läbiviimisel;
 - infokvaliteedi probleemide klassifitseerimisel ja kvaliteediparenduse eesmärkide seadmisel;
- didaktilistel eesmärkidel – infokvaliteedi käsitlemisel andmebaaside ja infosüsteemide õppeainetes.

Seega definitsiooni nõutavad omadused on selle hõlmavus, detailsus ja arusaadavus.

Eesmärgi täitmiseks on valitud ja võrreldud kolme ingliskeelsetest allikatest pärit autorile kättesaadavat infokvaliteedi käsitlust, kohendatud eestikeelseid vasteid ning sünteesitud ja esitatud infokvaliteedi struktuur.

Peatüki ülesehitus on järgmine:

- esmalt lähenetakse infokvaliteedile elementaristliku käsitlusviisi kohaselt – antakse ülevaade mõistetest ‚kvaliteet’ ja ‚info’ ning viimasega seotud terminitest ‚andmed’ ja ‚teave’;
- seejärel vaadeldakse informatsiooni elutsüklite mudeleid. Vajadus info elutsüklite tundmiseks ja kasutamiseks tekib järgmises peatükis ‚Infokvaliteedi juhtimine’;
- peatüki põhiosa esimeses pooles on analüütiline-võrdlev ülevaade kolmest infokvaliteedi käsitlusest: Cobiti juhtimiseesmärgid [Cobit Control Objectives 2000], H. Milleri artikkel infokvaliteedi dimensioonidest [Miller 1996] ja Gartner Groupi artiklitele [Friedman 2003; jmt] toetudes;

- peatüki põhiosa teises pooles on esitatud magistritöö autori väljapakutud infokvaliteedi definitsioon ja infokvaliteedi mõõdikute grupeerimise üks võimalik viis.

Peatüki tulemid on:

- infokvaliteedi mõõdikute loetelu;
- infokvaliteedi mõõdikute definitsioonid;
- infokvaliteedi mõõdikute üks võimalik grupeerimisviis.

1.1. Andmed, informatsioon, teave

Andmed (ingl. k *data*) kirjeldavad midagi toimunut – näiteks peegeldades reaalse maailma objektide seisundeid [Pearlson 2001, 190] ja sündmuseid; või esitades teise andmehulga omadusi. Andmeteks „Andmekogude seaduse” mõttes loetakse igasuguseid üksteisest eraldatavaid informatsiooniühikuid [Andmekogude seadus 2004, §2 lg 2]. Andmed on ‚külmad’ faktid, objektiivsed, nii analoog- kui ka digitaalkujul esitatavad; salvestatud andmed täidavad mälu funktsiooni.

Mõningate allikate järgi [nt SNIA 2004] on umbes pool ettevõtetes säilitatavatest andmetest kas:

- korduvad ehk dubleeritud või
- mittevajalik ettevõtte eesmärkide seisukohast.

Strategic Research Corporationi uurimuse kohaselt umbes 70% andmetest jääb kasutuseta 90 päeva möödudes nende esmasest salvestamisest [SRC 2004].

Andmetest saab kvalitatiivselt järgmine tase – informatsioon ehk info (ingl. k *information*) – siis, kui andmeid kasutatakse. Siin tuleb mängu subjekt, (andmete või infosüsteemi) kasutaja, andmete tarbija kelle jaoks andmetel on tähendus, kes interpreteerib andmeid. Termin ‚informatsioon’ tuleneb ladinakeelsest sõnast *informatio* – kujutamine, selgitamine. Kaasaegse juhtimise isa Peter F. Drucker defineerib informatsiooni järgmiselt: „omistades andmetele eesmärgi ja asjakohasuse (või suhestades andmed millegagi), saame informatsiooni” [Drucker 1988, 45]. Infot

võib defineerida ka kui sõnumit, mis esineb dokumendi või audiovisuaalses vormis oleva kommunikatsioonina. Nagu igal sõnumil, on ka infol saatja ja vastuvõtja. Info funktsiooniks on mõjutada vastuvõtja hinnanguid ja/või käitumist [IT juhtimise käsiraamat 1999, ptk 4.2.1], ehk info funktsiooniks on olla millegi aluseks – näiteks otsuste tegemisel, tegutsemisel, käitumisel.

Andmetesse võib piltlikult väljendudes uppuda ja ikkagi olla infonäljas. Liiga suurest infohulgast, info liigkiirest esitusest või vasturääkivustest infos tekkida psüühiline pingeseisund mida nimetatakse infostressiks [ENE 1998, 622].

Väärtuslik info inimese teadvuses, sisaldades refleksiooni, sünteesi, konteksti, on teave (ingl. k *knowledge*) [Pearlson 2001, 190]. Teave on väärtustatud info, sisaldades subjektikohaseid mälestusi, emotsioone jm.

Huvitav on siinkohal märkida, kuidas on ajapikku muutunud infoga ja infotehnoloogiaga tegelevate organisatsiooni allüksuste nimetused: algselt nimetati neid enamasti *andmetöötlusosakondadeks*, kaasajal on levinud *infosüsteemide* nimetus. Kas on oodata nende ümbernimetamist *teabeteenistusteks*, seda eriti andmeladude ja andmekaevanduse levides?

1.2. Info elutsükli käsitlused

Infokvaliteeti vaadeldes on oluline käsitleda informatsiooni dünaamilist poolt järgmistel põhjustel:

- informatsiooni dünaamika käsitlemine annab ühe vaate avastamiseks ja mõistmaks põhjusi, miks tekib infokvaliteedi langus ning kuhu organisatsiooni infoarhitektuuris projekteerida infokvaliteeti monitoorivaid ja/või tõstvaid meetmeid;
- info dünaamika käsitlemise tehnikad on ühtlasi kasutatavad mitmete infokvaliteedi mõõdikute hindamisel⁷.

⁷ Näiteks CRUD-analüüsi abil saab hinnata andmete asjakohasust.

Traditsiooniline viis info dünaamika käsitlemiseks on info elutsüklite abil. Järgnevalt vaadeldakse lühidalt kahte neist:

- CRUD-maatriks,
- Eesti seadustes [Andmekogude seadus 2004] kasutatav.

Ülevaatest on välja jäetud infokvaliteedi juhtimisel vähem huvipakkuvad info elutsüklite kontseptsioonid, näiteks andmete arhiveerimismetoodikatega seotud ILM⁸ (ingl. k *Information Lifecycle Management*).

1.2.1. CRUD-maatriks

Antud andmete elutsükli käsitlemise tehnika seisneb CRUD-maatriksi koostamises ja analüüsis. Lühend CRUD tuleneb järgmistest andmeoperatsioonide ingliskeelsete nimetuste esitähtedest:

- *Create* – looma, tekitama;
- *Retrieve, Read või Reference* – lugema, kasutama;
- *Update* – uuendama, muutma;
- *Delete* – kustutama.

CRUD-maatriksi koostamisel loetletakse andmeolemid⁹ ridades ja (äri)funktsioonid veergudes ning maatriksi lahtrites näidatakse, milliseid eelnimetatud operatsioone iga funktsionaalne element iga andmeolemiga läbi viib [Open Group 2004]. Lähenemine on andmebaasi operatsioonide ja SQL keele keskne.

Maatriksi analüüsimisel kontrollitakse muuhulgas, et iga andmeolemi jaoks eksisteeriks vähemalt üks funktsionaalne üksus, mis andmeolemit loob ning vähemalt üks üksus, mis andmeolemit kasutab.

⁸ Vt näiteks PrincetonSoftech.com.

⁹ Andmeolem (ingl. k *data entity*) – abstraktsioon samaliigilistest andmetest.

1.2.2. Andmete elutsükkel Eesti seaduste mõistes

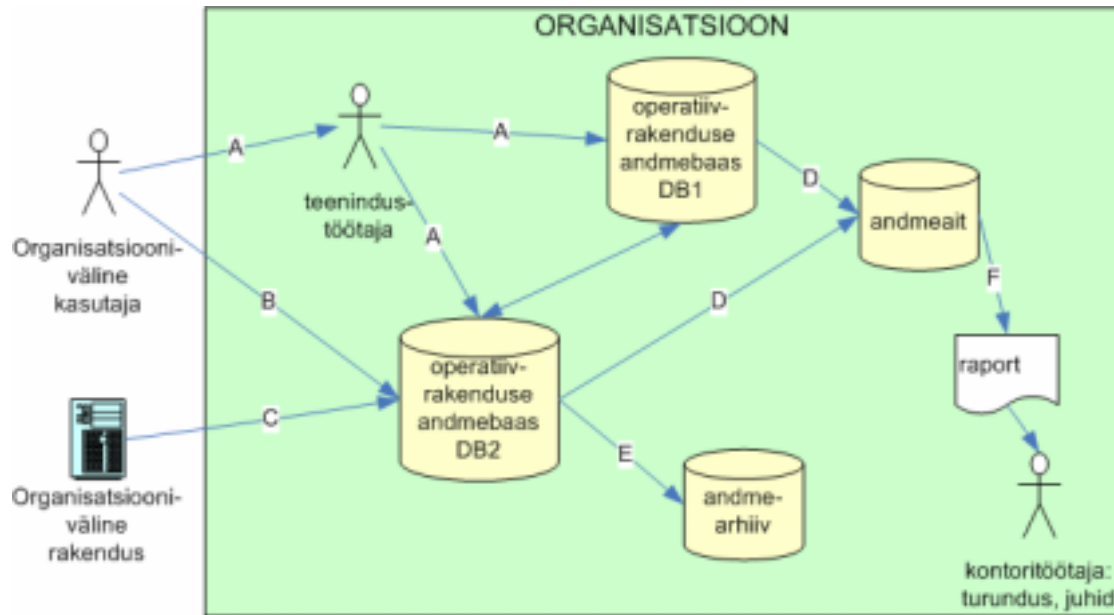
Andmekogude seaduse [Andmekogude seadus 2004, §2 lg 3] järgi on võimalikud järgmised toimingud andmetega:

- kogumine,
- salvestamine,
- korrastamine,
- säilitamine,
- muutmine,
- päringute teostamine,
- väljavõtete tegemine,
- kasutamine,
- üleandmine,
- ühendamine,
- juurdepääsu sulgemine,
- kustutamine,
- hävitamine;

kokku seega 13 operatsiooni. Põhimõtteliselt võib toodud andmeoperatsiooni kasutada eelmises punktis kirjeldatud viisil funktsionaalsusega seostamiseks ja kvaliteedikontrolliks.

1.2.3. Informatsiooni elutsükkel organisatsioonis

Eelnevalt toodud kaks andmete elutsükli käsitlemist on loodud erinevate eesmärkidega – CRUD-matriks eelkõige süsteemianalüütiku töövahendina andmestruktuuride projekteerimiseks ja funktsionaalsuse seostamiseks andmeolemitega ning infosüsteemi mudeli kvaliteedikontrolliks; teine juriidilise konstruktsioonina. Käesolevas alapunktis pakub autor välja eelnevatele käsitlemistele tugineva ja neid laiendava visuaalse pildi info elutsükli käsitlemiseks organisatsioonis ning organisatsioonisisese ja väliskeskkonnaga andmevahetuse käsitlemiseks (vt Joonis 1).



Joonis 1 – Informatsiooni elutsüklil organisatsioonis.

Joonisel on kujutatud:

- organisatsiooni andmevahetus organisatsioonivälise andmeallikatega, millised on automatiseerituse alusel jaotatavad:
 - organisatsiooni töötaja ja välise kasutaja osalusega liidesed (Joonis 1, tähis A), nt traditsiooniline letiteenindus;
 - organisatsiooni töötaja osalusega, kuid välise kasutaja osalusega liidesed (tähis B). Siia alla kuuluvad iseteenindusbürood;
 - töötaja osalusega, kuid välispartneri poolt automatiseeritud liides, näiteks organisatsiooni teenindava pangaga suhtlemine läbi panga teleteenuse, kus organisatsiooni raamatupidamis-finantsüsteemi ja panga infosüsteemi andmevahetust vahendab raamatupidaja;
 - organisatsiooni ja välispartneri poolt automatiseeritud andmevahetus (tähis C). Siia alla kuuluvad muuhulgas B2B (ingl. k *business-to-business*) automatiseeritud andmevahetusliidesed;
- organisatsioonisisestest rakendustevahelistest andmevahetustest:
 - automatiseeritud andmevahetusliidesedga infosüsteemid, sh:
 - operatiivrakenduste andmebaasidest andmete ETL (ingl. k *extraction-transformation-loading* ehk kogumine, kohandamine ja laadimine) protsessiga andmeaita viimine (tähis D);

- operatiivrakenduse andmebaasidest arhiveerimisele kuuluvate andmete kustutamine operatiivrakenduse andmebaasist ja sisestamine andmearhiivi (tähis E);
- inimsekkumist ja käsitsitööd nõudvad andmevahetused süsteemide vahel.
- andmete kasutamised (tähis F), näiteks turunduse ja juhtimisinfosüsteemide raportite väljastamine andmeajada poolt.

Jälgides andmete liikumist (nooled joonisel), saab tuua sisse andmete allavoolu liikumise mõiste (ingl. k *downstream*). Teises osas seostatakse andmete allavoolu liikumisega andmekvaliteedi langus, samuti leiab II osas kasutust eelnevas loetelus toodud organisatsiooni andmeoperatsioonide seostamine infokvaliteedi juhtimisega.

1.3. Andmekvaliteet ja infokvaliteet

Vastavalt rahvusvahelise standardimisorganisatsiooni ISO terminoloogiale on kvaliteet määr, mille olemuslike omaduste kogum täidab eeldatavaid või seatud nõudeid. Rakendades kvaliteedi mõistet infole, saame ühe võimaliku informatsiooni kvaliteedi definitsiooni – info on kvaliteetne siis kui ta täidab erinevate osapoolte ootused ja vajadused antud info suhtes.

Kui vajadused on esitatud nõudmiste vormis, siis on kvaliteet vastavus nõudmistele [IT juhtimise käsiraamat 1999, ptk 11.1]. Nõudmised võib tuletada järgmistest allikatest:

- ülesande püstitustest;
- lepingutest,
- spetsifikatsioonidest,
- standarditest, normidest,
- heast tavast [samas, ptk 11.1].

Kvaliteeti saab juhtida, kui vajadused on määratletud. Kvaliteedi tagamine (ingl. k *quality assurance*) on tegevuste kogum, mis on vajalik piisava usaldatavuse tagamiseks, et kvaliteedi objekt – siin siis info – vastaks esitatud kvaliteedinõuetele.

Tulenevalt sellest, et andmekvaliteeti ja infokvaliteeti mõõdetakse erinevalt ja saavutatakse erinevalt [Information Technology Standards Guidance 1999, 273] on antud töös oluline suhestada terminid 'andmekvaliteet' ja 'infokvaliteet'. K. E. Pearlson kasutab andmete, informatsiooni ja teabe suhete illustreerimiseks järgmist mõttekäiku – andmed on alamtüüp infost ja info omakorda alamtüüp teabest [Pearlson 2001, 191]. Seega andmekvaliteet sisaldub infokvaliteedis, infokvaliteet on laiem mõiste – sisaldades tegureid, millised ei ole kasutatavad andmeid ja andmekvaliteeti käsitledes. V. Eilenfieldi ja S. Luhrmani järgi saab kvaliteetsetest andmetest kvaliteetse informatsiooni siis, kui on täidetud järgmised tingimused:

- andmetöötlusalgoritmid on õiged,
- kasutaja interpreteerib andmeid õieti [Eilenfield 2000].

Käesolevas töös on vastavalt alapunktis 1.1 toodud andmete ja informatsiooni eristusele kasutatud terminit 'andmekvaliteet' siis, kui käsitletakse kitsamalt andmeid ja andmete suhteid reaalsusega; terminit 'infokvaliteet' siis, kui käsitletakse andmeid koos andmete kasutajaga ja tema vajadustega. Erandiks on tsitaadid ja viited, kus on säilitatud algupärane terminikasutus.

1.4. Olemasolevate infokvaliteedi käsitluste võrdlus

Käesolevas alapunktis on esitatud analüütiline-võrdlev ülevaade valikust¹⁰ autorile kättesaadavatest infokvaliteedi käsitlustest:

- Cobiti juhtimiseesmärgid [Cobit Control Objectives 2000],
- H. Milleri artikkel infokvaliteedi dimensioonidest [Miller 1996] ja
- Gartner Groupi uurimused [Friedman 2003; jmt].

Valikukriteeriumitest lähtuvalt on võrdlusest välja jäetud mitmed infokvaliteeti käsitlevad allikad, nimetades:

- ISO standardeid (ISO12207, ISO9126:2000 jmt) kui tarkvara-, mitte niivõrd infokvaliteedi keskseid;

¹⁰ Valiku alusteks olid: (i) tunnustatus IT alase teabe allikana ja (ii) käsitluse universaalsus.

- C. Foxi [Fox 1994] kirjeldatud nelja andmekvaliteedi faktorit, mis on alamosa H. Milleri käsitlusest, so sellele mitte juurde andev;
- spetsiifilistel eesmärkidel loodud infokvaliteedi käsitlusi, nt meditsiinilistele andmetele keskendunud [Long 2004]¹¹, geoinfo kvaliteedi käsitlused jmt.

Alapunkti lõpus on toodud refereeritud materjali sisuline ja süstemaatiline analüüs.

1.4.1. Gartner Groupi infokvaliteedi käsitlus

Gartner Group on juhtiv rahvusvaheline IT uuringu- ja konsultatsioonifirma. Gartneri analüütikute poolt läbiviidud uuringute tulemusi pakutakse firma klientidele süstematiseeritud veebiraamatukogus, infokvaliteedile on pühendatud alamrubriik „Software Infrastructure\Database Management Software\Data Quality”. Seisuga 27.06.2004 oli infokvaliteedi rubriigis kokku 56 artiklit, põhiliselt neljalt autorilt: Ted Friedman, Mary Knox, Kevin H. Strange, Annemarie Early.

Friedmani järgi [Friedman 2003, 2] on infokvaliteedi tahkudeks (ingl. k *facets*):

a. tehnilised aspektid:

1. olemasolu (ingl. k *existence*),
2. kehtivus (ingl. k *validity*),
3. kooskõlalatus (ingl. k *consistency*),
4. õigsus (ingl. k *accuracy*);

b. „pehmed”, ärilised aspektid:

5. ajakohasus (ingl. k *timeliness*),
6. asjakohasus (ingl. k *relevance*).

Gartneri artiklite kogumik ei moodusta ühtset käsitlust, samuti ei leia artiklitest infokvaliteedi definitsiooni.

¹¹ Siin on nimetatud ligi 100 infokvaliteedi tegurit.

1.4.2. H. Milleri infokvaliteedi käsitlus

Holmes Miller loetleb artiklis „The Multiple Dimensions of Information Quality” [Miller 1996] kümme infokvaliteedi faktorit:

1. asjakohasus (ingl. k *relevance*). Asjakohane info vastab kasutaja vajadustele;
2. õigsus (ingl. k *accuracy*). Tõene info peegeldab reaalsust piisava täpsusega;
3. ajakohasus (ingl. k *timeliness*). Ajakohane info on igal hetkel aktuaalne
4. täielikkus (ingl. k *completeness*). Täielik info kajastab reaalsust paraja mitmekülgusega, detailsusega;
5. sisemine sidusus (ingl. k *coherence*). Sisemiselt sidus info on vastuoludeta;
6. formaat (ingl. k *format*). Info esitamine kasutajale sobivas vormis ja kontekstis;
7. kättesaadavus (ingl. k *accessibility*). Info peab olema kasutajale kättesaadav vajalikul hetkel;
8. väline sidusus (ingl. k *compatibility*). Väline sidusus iseloomustab võimaluste olemasolu seostada infot teise infohulgaga;
9. turvalisus (ingl. k *security*). Turvalisuse tagamine seisneb loogilises ja füüsilises infoturbes;
10. kehtivus (ingl. k *validity*). Kehtivus seisneb põhimõttelises võimaluses kontrollida info õigsust ja infokvaliteedi vastavust eelnimetatud dimensioonidele.

Milleri artikli eesmärk on teadvustada infokvaliteedi struktuursus, mitte defineerida iga infokvaliteedi mõõdikut või esitada selle mõõtmise skaala ja tehnika. Samuti on esitatud loetelu haaramiseks liiga pikk ning nõuaks grupeerimist, osade dimensioonide kokkuliitmist või diferentseerimist.

1.4.3. COBITi infokvaliteedi käsitlus

Cobit Quickstart nimetab infokvaliteeti kahes kohas. Esiteks Cobiti raamstruktuuri tutvustuses, loetledes tegurid millistele peab vastama informatsioon äritegevuse seisukohalt [Cobit Quickstart 2003, 6]:

- toimivus (ingl. k *effectiveness*),

- tõhusus (ingl. k *efficiency*),
- konfidentsiaalsus (ingl. k *confidentiality*),
- terviklus (ingl. k *integrity*),
- käideldavus (ingl. k *availability*),
- kooskõlalisus¹² (ingl. k *compliance*) ja
- usaldatavus (ingl. k *reliability*);

ning teiseks Cobit Quickstart Etalonis Cobiti IT-ala „tarnimine ja tugi” IT-protsessis TT11 – Hallata andmeid [Cobit Quickstart 2003, 29]. Protsessi TT11 äriliseks nõudeks on seatud: kõik andmed on nende loomisel, kasutamisel ja talletamisel:

- täielikud (ingl. k *complete*),
- õiged (ingl. k *accurate*) ning
- kehtivad (ingl. k *valid*).

Cobitis puuduvad infokvaliteedi faktorite ranged definitsioonid, samuti tõdevad Cobiti autorid, et toodud tegurid – kombineerituna mitme erineva käsitluse alusel – on kattuvad [Cobit Control Objectives 2000, 14].

¹² Cobiti varasem, teine redaktsioon [Cobit juhtimiseesmärgid 1998, 11] kasutab termini ‚kooskõlaline’ asemel terminit ‚vastavus’ töömääratlusega „puudutab vastavust neile seadustele, eeskirjadele ja lepinguaktidele, millele allub äriprotsess, st väljastpoolt tulenevaid ärikriteeriume.”.

1.4.4. Olemasolevate käsitluste analüüs

Eelnevalt käsitletud infokvaliteedi tegurite mitterange võrdlus on esitatud alljärgnevas tabelis (Tabel 1). Ühte ritta on seatud eri käsitluste sisuliselt kattuvad mõisted, tühi lahter märgib antud infokvaliteedi teguri puudumist vastavas käsitluses. Infokvaliteedi teguritele on lisatud järjekorranumbrid mis märgivad antud teguri nimetamise järjekorda kasutatud allikas.

Cobit	Gartner	Miller	Eestikeelne vaste
	5. <i>timeliness</i>	3. <i>timeliness</i>	ajakohasus
2. <i>efficiency</i>			tõhusus
3. <i>confidentiality</i>		9. <i>security</i> (turvalisus)	konfidentsiaalsus
4. <i>integrity</i>			terviklus
5. <i>availability</i>			käideldavus
8. <i>completeness</i>		4. <i>completeness</i>	täielikkus
6. <i>compliance</i>			vastavus
7. <i>reliability</i>			usaldatavus
	1. <i>existence</i>		olemasolu
10. <i>validity</i>	2. <i>validity</i>	10. <i>validity</i>	kehtivus
1. <i>effectiveness</i> (toimivus), 9. <i>accuracy</i>	3. <i>consistency</i>	5. <i>coherence</i>	kooskõlalisus, sisemine sidusus
	4. <i>accuracy</i>	2. <i>accuracy</i>	õigsus (täpsus)
		7. <i>accessibility</i>	õigeaegsus
	6. <i>relevance</i>	1. <i>relevance</i>	asjakohasus
		6. <i>format: form, context</i>	formaat, vorming, kontekst
		8. <i>compatibility</i>	sidusus

Tabel 1. Infokvaliteedi tegurite võrdlus Gartneri, Milleri ja Cobiti käsitluses.

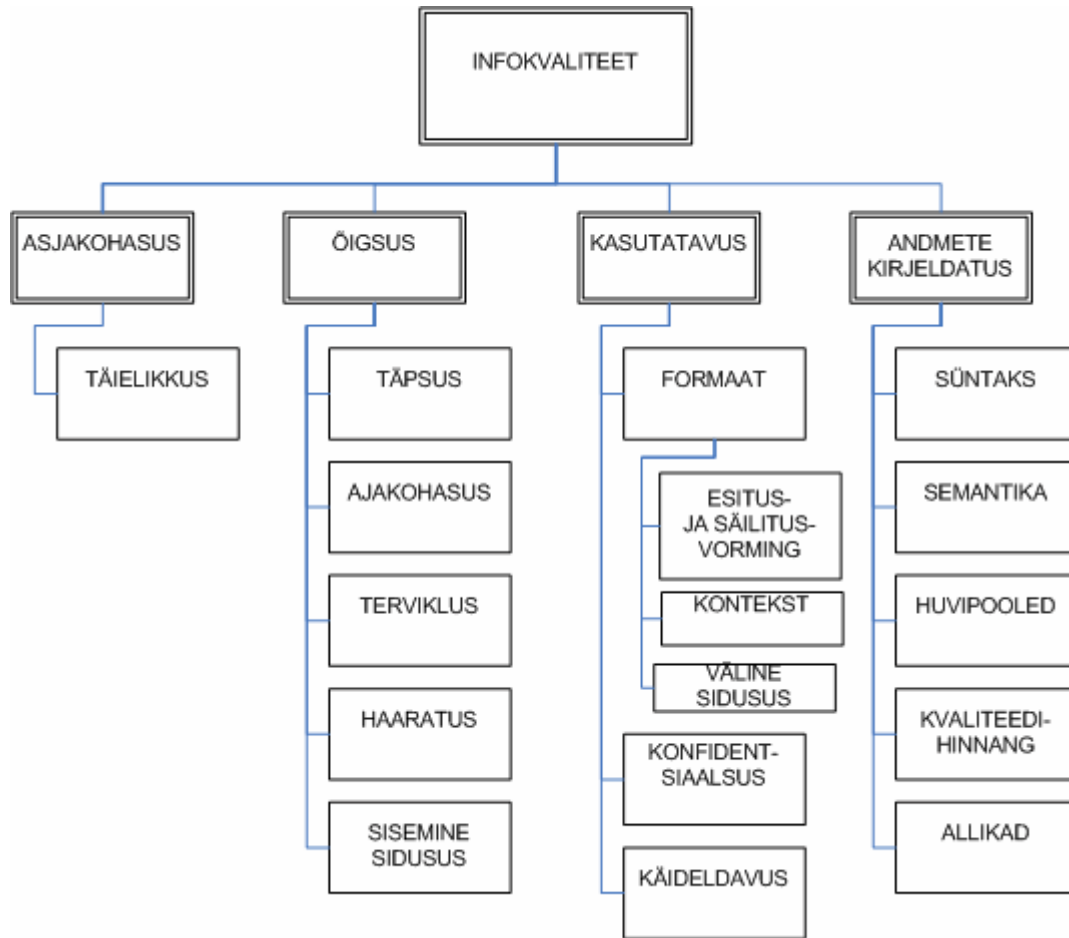
Gartneri, Milleri ja Cobiti käsitlusi võrreldes võime täheldada, et infokvaliteedi mõiste struktuuri osas ei ole ühest arusaama ei (inglisekeelsetes) terminites ega infokvaliteedi dimensioonides:

- tegurite loetelud on osaliselt kattuvad, kuid ei eksisteeri kõikehõlmavat, teisi käsitlusi endas haaravat käsitlust;
- käsitlustest puuduvad infokvaliteedi ranged definitsioonid ja mõõtmistehnikad, enamasti 'defineeritakse' kirjelduste ja näidete abil;
- mõõdikuid on suurem hulk (Gartner 6, Cobit 7, Miller 10, neid kombineerides umbes 15-16). Psühholoogiast tuntud 7+/-2 reegli järgi ei suuda inimese teavus sellist hulka objekte korraga hõlmata [Miller 1956]. Siit tekkis magistritöö autoril idee infokvaliteedi tegureid loodavas käsitluses grupeerida;
- eksisteerib infokvaliteedi tegureid, millised ei sisaldu võrreldud käsitlustes.

1.5. Infokvaliteedi mõõdikute raamistik

Käesolevas alapunktis on esitatud magistritöö autori väljapakutud infokvaliteedi definitsioon¹³ ja üks võimalik viis infokvaliteedi mõõdikute grupeerimiseks hierarhiasse (vt joonis 2). Mõõdikuks on siin nimetatud iga infokvaliteedi tegurit, võtmemõõdikuks hierarhia ülemise taseme mõõdikuid ning alammõõdikuteks hierarhia alamate tasemete mõõdikuid.

¹³ Kuna infokvaliteet on mitmemõõtmeline suurus, on sobivaks defineerimise viisiks tema struktuuri (alamsüsteemide) kirjeldamine ehk infokvaliteedi mõõdikute loetelu ja nende definitsioonide esitamine.



Joonis 2. Magistrandi väljapakutud infokvaliteedi mõõdikud ja üks mõõdikute grupeerimise võimalusi.

Toodud mõõdikute loetelu ei ole kindlasti ei lõplik ega ammendav, vaid on arendatav ja täiendatav vastavalt konkreetsete projektide vajadustele. Kui mõõdikute hulk ja grupeerimistasemete sügavus muutub mõõdikute lisandumisel raskesti haaratavaks, tuleb luua struktuur mõõdikutele metaandmete lisamiseks ja algoritm mõõdikute valikuks vastavalt ülesandele.

Alljärgnevalt igast mõõdikust detailsemalt. Iga mõõdiku juures on lisaks eestikeelsele nimetusele toodud selle ingliskeelne vaste, mõõdiku definitsioon ning võimalusel ka sünonüümid, antonüümid, mõõdiku seletusi, näiteid jm.

1.5.1. Asjakohasus

Mõõdiku eestikeelne nimetus	Asjakohasus
Ingliskeelne nimetus	<i>Relevance</i>
Definitsioon	Vastavuse määr kasutaja(te)le vajaliku ja kogutava/säilitatava info vahel. Kitsamalt, objektorienteeritud paradigmas, võib asjakohasust mõista kui objektide/klasside vastavust kasutaja(te) vajadustele
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Asjakohane – asja kohta käiv, sobiv, asjalik [ÕS 1999]; relevantsus, otstarbekus, kohane, kohasus, olulisus, kompaktsus, konkreetsus, vajalikkus, efektiivsus, ökonoomsus, optimaalsus
Antonüümid	Irrelevantsus
Seletus, kontrollküsimused	Kas kogutavad ja andmekogus talletatavad andmed on tarbijale vajalikud? Kas salvestatud info rahuldab kõikide kasutajate kõik olulised/tähtsad infovajadused? Niipalju kui vaja, nii vähe kui võimalik. <i>Fitness for purpose</i> . Kogu vajalik ja ainult vajalik info
Miks on oluline, millised on suboptimaalse kvaliteedi võimalikud tagajärjed	Puuduolev info toob kaasa organisatsiooni tööprotsesside efektiivsuse languse
Võimalikud ohud liigse (optimaalset ületava) kvaliteediga	Üleliigne info tekitab infostressi ja mittevajaliku info kogumine ning säilitamine raiskab ressursse
Alamvõtmemõõdikute nimetused	<ul style="list-style-type: none"> • Täielikkus
Miks ja millal tekivad antud infokvaliteedi teguri probleemid?	Andmestruktuuride projekteerimise etapis ning nende hilisemalt arendustööl
Muud olulist	Lisaks andmete kasutajate vajadustele arvestada

	seadusandlikke jm piirangud (nt delikaatsed isikuandmed)
Seos teiste mõõdikutega	Asjakohasuse mõõdik ei kirjelda andmete õigsust ega seda osa kvaliteedist, kuidas tarbija andmeid kätte saab, st need võtmemõõdikud on sõltumatud. Asjakohasus on seotud andmete kirjeldatuse võtmemõõdikuga – viimane loob eelduse asjakohasuse juhtimiseks

1.5.1.1. Täielikkus

Mõõdiku eestikeelne nimetus	Täielikkus
Ingliskeelne nimetus	<i>Completeness</i>
Definitsioon	vastavuse määr kasutaja(te)le vajaliku ja kogutava info mitmekülgsuse vahel. Kitsamalt, objektorienteeritud paradigmas, võib täielikkust mõista kui objekti/klassi atribuutide vastavust kasutaja(te) vajadustele
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Mitmekülgsus, detailsus
Antonüümid	Ühekülgsus, pinnapealsus
Näited	Kliendiandmebaasis on info raamatupidamisele vajalike kliendikohaste atribuutide kohta: kliendi nimi, aadress ja firma käibemaksu kohuslase number. See ei pruugi aga katta turunduse vajadusi, kes vajab aadressi, kliendi sugu, vanust

1.5.2. Õigsus

Mõõdiku eestikeelne nimetus	Õigsus
Ingliskeelne nimetus	<i>Accuracy; conformity; correctness</i>

Definitsioon	Vastavuse määr andmete ja andmete poolt esitatava reaalsuse vahel ning selle määra ja kasutaja(te)le optimaalse õigsuse määra suhe
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	tõesus, usaldusväarsus, vastavus; korrektsus, <i>correct</i> – korrektne, õige, veatu; sünnis, viisakas [Silvet 1989]
Seletus, kontrollküsimused	Kas andmed vastavad reaalsusele?
Miks on oluline, millised on suboptimaalse kvaliteedi võimalikud tagajärjed	Andmetel põhinevate tegevuste kvaliteet sõltub andmete õigsusest
Võimalikud ohud liigse (optimaalset ületava) kvaliteediga	Ressursside raiskamine andmete kogumisel ja säilitamisel
Alamvõtmemõõdikute nimetused	<ul style="list-style-type: none"> • Täpsus, • ajakohasus, • terviklus, • haaratus, • sisemine sidusus
Miks ja millal tekivad antud infokvaliteedi teguri probleemid ja parendamise tehnikad	Enamjaolt andmete kogumisel ja salvestamisel
Seos teiste mõõdikutega	Antud dimensioon ei kirjelda andmete ajakohasust ega seda osa kvaliteedist, kuidas tarbija andmeid kätte saab, st need mõõdikud on sõltumatud

1.5.2.1. Täpsus

Mõõdiku eestikeelne nimetus	Täpsus
Ingliskeelne nimetus	<i>Exactness</i>
Definitsioon	Andmete ja reaalsuse vastavuse määr on kasutaja seisukohalt optimaalne
Antonüümid	Ebatäpsus

Näited	GIS koordinaadid; komakohtade arv rahalises atribuudis
--------	--

1.5.2.2. Ajakohasus

Mõõdiku eestikeelne nimetus	Ajakohasus
Ingliskeelne nimetus	<i>Timely, timeliness</i>
Definitsioon	Info vastavus igal ajahetkel (või määratud hetkedel) tegelikkuse seisule fikseeritud ajalise hilinemisega
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Ajakohane – moodne, mingi aja nõuetele vastav [ÕS 1999]. Aktuaalsus, aktuaalne – päevakohane [ÕS 1999]. <i>Timely</i> – ajaldine ¹⁴ , õigeaegne; ajakohane [Silvet 1990]. Kehtivus, ajalisus
Antonüümid	Mahajäämus, vananenud
Seletus, kontrollküsimused	Kuidas toimub andmete uuendamine: ühekordselt, juhuslikult; pidevalt, lausaliselt ehk totaalse uuendamisenä kampaania korras, süstemaatiliselt vm moel? Milline on mahajäämus? Millise ajavahemiku tagant uuendatakse andmeid (andmete pidevus)? Millise aja jooksul jõuab info reaalse elu sündmusest andmetalletuseni?
Näited	Kliendiandmete korral võib probleemiks olla surnud (füüsilised isikud) või pankrotistunud ning sundlõpetatud (juriidilised isikud) klientide väljaarvamine aktiivsest kliendihulgast

1.5.2.3. Terviklus

Mõõdiku eestikeelne nimetus	Terviklus
Ingliskeelne	<i>Integrity</i>

¹⁴ Antud tähendusvarjund on rohkem seotud käideldavuse alammooduliga.

nimetus	
Definitsioon	Andmete terviklus on andmete pärinemine autentsest allikast ning veendumine, et need pole hiljem muutunud ja/või neid pole hiljem volitamatu muudetud. Andmeterviklus on standardi ISO 7498-2:1989 kohaselt omadus, mis näitab, et andmeid ei ole volitamatu viisil muudetud ega hävitatud
Seletus, kontrollküsimused	Kas info on kaitstud inimeste ja loodusõnnetuste eest? Kas andmed on seotud selle loojaga, loomisajaga jm kontekstiga (auditatribuutidega)?

1.5.2.4. Haaratus

Mõõdiku eestikeelne nimetus	Haaratus
Ingliskeelne nimetus	<i>Conformity</i>
Definitsioon	Andmekogusse on kaasatud kõik ja ainult kõik asjakohased objektid reaalsusest ühekordselt
Seletus, kontrollküsimused	Kas igale asjakohasele objektile reaalsusest vastab üks unikaalne kirjeldus andmekogus, st puuduvad topeltkirjed? Kas andmekogus on andmeid objektidest, mida reaalsuses ei eksisteeri?

1.5.2.5. Sisemine sidusus

Mõõdiku eestikeelne nimetus	Sisemine sidusus
Ingliskeelne nimetus	<i>Coherence, accordance to business rules</i>
Definitsioon	Vasturääkivuste puudumine andmekogus; andmete vastavus kasutaja defineeritud reeglitele
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Kooskõllalisus, koherentsus, kokkusobivus, vasturääkivusteta, vastavus ärireeglitele
Antonüümid	Vastuolulisus
Seletus, kontrollküsimused	Kas andmestruktuurid on normaliseeritud kujul? Kontrollnumbrite rakendamine (nt isikukoodis,

	viitenumbrites jm)
Näited	Isikukoodi seos sootunnusega; lepingu kuupäev on suurem kui firma asutamisaeg

1.5.3. Kasutatavus

Mõõdiku eestikeelne nimetus	Kasutatavus
Ingliskeelne nimetus	<i>Usability</i>
Definitsioon	Vastavuse määr info esituse ja sellekohaste tarbija vajaduste vahel
Antonüümid	Kasutuskõlbmatu
Seletus, kontrollküsimused	Kas info on personaliseeritud? Kas info on õigele kasutajaskonnale suunatud ja edastatud sobivas kommunikatsioonikanali vahendusel? Kas info on arusaadav (sh keel)?
Miks on oluline, millised on suboptimaalse kvaliteedi võimalikud tagajärjed	Mittekasutatav info toob kaasa infol baseeruvate tegevuste efektiivsuse languse ja ressursside raiskamise
Kui on, siis alamvõtmemõõdikute nimetused	<ul style="list-style-type: none"> • Vorming, • käideldavus, • konfidentsiaalsus
Miks ja millal tekivad antud infokvaliteedi teguri probleemid ja parendamise tehnikad	Infosüsteemide kasutajaliideste projekteerimisel

1.5.3.1. Vorming

Mõõdiku eestikeelne nimetus	Vorming
Ingliskeelne nimetus	<i>Format</i>
Definitsioon	Andmete salvestuse ja esituse vastavuse määr kasutaja seisukohalt vajalikule
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Formaat

Kui on, siis alamvõtmemõõdikute nimetused	<ul style="list-style-type: none"> • esitusvorming, • säilitusvorming, • kontekst, • väline sidusus
---	---

Esitusvorming

Mõõdiku eestikeelne nimetus	Esitusvorming
Ingliskeelne nimetus	<i>Presentation format</i>
Definitsioon	Andmed on kasutajale esitatud sobivas vormingus
Seletus, kontrollküsimused	Kas kasutatakse standardset andmete esituse vormi (ekraanipildid tuttavad, menüü samas kohas)?
Näited	Teksti dubleerimine graafikute, tabelite jms

Säilitusvorming

Mõõdiku eestikeelne nimetus	Säilitusvorming
Ingliskeelne nimetus	<i>Data structure format</i>
Definitsioon	Andmed on säilitatud vormingus mis võimaldab kasutajale vajaliku funktsionaalsuse realiseerimist
Seletus, kontrollküsimused	Kas andmed on säilitatud piisavalt granulaarselt?
Näited	XMLi kasutamine andmete struktureerimisel – aadressi salvestamisel piisab ümbrikule trükkimiseks ühes atribuudis hoitavast aadressist, ruumilisteks operatsioonideks on tarvilik aadressi komponendid lahus hoida

Kontekst

Mõõdiku eestikeelne nimetus	Kontekst
Ingliskeelne nimetus	<i>Context</i>
Definitsioon	Andmed on esitatud kasutajale koos sobivate, interpreteerimist võimaldavate andmetega, taustsüsteemis
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Taustainfo, sidusinfo, võrdlusandmed
Näited	Rahalistele andmeühikutele on lisatud valuutaühik

Väline sidusus

Mõõdiku eestikeelne nimetus	Väline sidusus
Ingliskeelne nimetus	<i>Compatibility</i>
Definitsioon	Info ristseostatavus (ühilduvus) teise infohulgaga
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Kokkusobivus, ühtesobivus ehk ühilduvus, ühitatavus, kombineeritavus, seostatavus
Seletus, kontrollküsimused	Kas objektidel on standardsed ja unikaalsed atribuudid?
Näited	Isikukood, firma äriregistri number

1.5.3.2. Käideldavus

Mõõdiku eestikeelne nimetus	Käideldavus
Ingliskeelne nimetus	<i>Availability, accessibility</i>
Definitsioon	Info õigeaegne kättesaadavus selleks volitatud isikutele ning subjektidele
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Ligipääsetavus, kättesaadavus, õigeaegselt edastatud (ingl. k <i>push</i>)
Antonüümid	Hilinenud info
Seletus, kontrollküsimused	Kas info on vajaduse tekkimisel kasutajale aktsepteeritava aja jooksul kättesaadav?
Seosed teiste mõõdikutega	Seotud ajakohasusega – kasutajale on tähtis nende mõõdikute kvaliteeti iseloomustavate ajategurite summa

1.5.3.3. Konfidentsiaalsus

Mõõdiku eestikeelne nimetus	Konfidentsiaalsus
Ingliskeelne nimetus	<i>Confidentiality</i>
Definitsioon	Andmete kättesaadavus ainult selleks volitatud isikutele ning kättesaamatus kõikidele ülejäänutele
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Piiratus, salastatus

kaugemad) sünonüümid	
Antonüümid	Avalikkus, avalikustatus

1.5.4. Andmete kirjeldatus

Möödiku eestikeelne nimetus	Andmete kirjeldatus
Ingliskeelne nimetus	<i>Metadata, data description, data comments</i>
Definitsioon	Andmed on piisavalt kirjeldatud metaandmetega
Eestikeelsed (lähedasemad ja kaugemad) sünonüümid	Metaandmed, dokumenteeritud andmestruktuur
Seletus, kontrollküsimused	Kas andmebaasi tabelite ja veergudele on lisatud kommentaarid? Kas eksisteerib organisatsiooni andmesõnastik?
Miks on oluline, millised on suboptimaalse kvaliteedi võimalikud tagajärjed	Arendaja saab metaandmete olemasolul efektiivsemalt infosüsteemi edasi arendada; aitab kaasa jätkusuutliku infoarhitektuuri loomisele
Võimalikud ohud liigse (optimaalset ületava) kvaliteediga	Ressursimahukus ajakohasena hoidmisel
Alamvõtmemöödikute nimetused	<ul style="list-style-type: none"> • Süntaks ehk tehniline kirjeldatus: andmetüüp, pikkus, kodeering jt, • semantika ehk tähendus, • kvaliteedihinnang, • huvipooled: omanik, valitseja, kasutajad jt, • allikad, • jm vajalik vastavalt olukorrale
Miks ja millal tekivad antud infokvaliteedi teguri probleemid ja parendamise tehnikad	Kogu infoarhitektuuri projekteerimise ja edasiarenduse käigus. Lahenduseks on vastava haldusprotsessi sisseviimine ja kontroll

2. INFOKVALITEEDI JUHTIMISE METOODIKA

Eelmises peatükis käsitleti informatsiooni kvaliteeti eesmärgiga luua juhtimisobjekti määratlus. Käesoleva peatüki ülesanne on kirjeldada instrumentarium mille abil infokvaliteeti juhtida.

Autorile teadaolevalt ei ole infokvaliteedi hindamise ning juhtimise metoodikate vallas *de facto* standardit. Eestis infokvaliteediga seotud teenuseid pakkuvate firmade teenusekirjeldusi uurides ja töötajaid küsitledes selgus, et kasutatakse andmeanalüüsi tarkvara tootjate ideoloogiast lähtuvaid ning omaenda praktika käigus täiendatud infokvaliteedi hindamise läbiviimise metoodikaid.

Tegurid, millest lähtuvalt otsustas töö autor koostada uue metoodika infokvaliteedi juhtimiseks, on järgmised:

1. puuduvad infokvaliteedi hindamise ja juhtimise metoodika standardid;
2. kasutatavad firmasisesed juhised ei ole detailselt ja täpselt kirjeldatud või ei ole juhised vabalt kättesaadavad – teenusepakkujad peavad neid konkurentsieeliseks;
3. olemasolevate metoodikate ja vahendite keskendumine:
 - mingile ärivaldkonnale (nt finantsandmete kvaliteedikontrollile, kliendiandmete puhastamisele);
 - kitsamalt andmekvaliteedile või mõnele mõõdikule (nt dubleerivate kirjete eraldamine);
 - infokvaliteedi juhtimise konkreetsele etapile (nt audit);
4. vajadus luua infokvaliteedi juhtimise metoodika, mis oleks kasutatav nii konsulteerimistöös infokvaliteedi hindamisel ja parendamisel kui ka IT tudengite õppeprogrammides ning mida oleks võimalik vajaduse korral täiendada.

Peatüki ülesehitus on järgmine. Esmalt kirjeldatakse ja antakse hinnang olemasolevatele infokvaliteedi auditeerimise ja juhtimise metoodikatele, seejärel esitatakse magistritöö autori koostatud metoodika koos läbiviimise protsessi sammude

ja rakendamise juhistega. Väljapakutud metoodika rakendusnäide on järgmises, kolmandas peatükis.

2.1. Olemasolevate metoodikate võrdlus

Magistritöö autoril õnnestus leida kaks Eesti firmat kes pakuvad infokvaliteediga seotud teenuseid: AS Resta ja Net Group OÜ. Lisaks neile uuris magistrant IT- ja infokvaliteediga seotuid ISO ja IEC standardeid ning IT Governance Institute'i väljaannet Cobit Framework.

2.1.1. AS Resta kasutatav metoodika

AS Resta on Tartu Ülikooli matemaatikateaduskonna inimjõule tuginev, ettevõtte algusaastatel SAS Institute'i ja SAS tarkvaraga seotud andmeanalüüsi- ja konsultatsioonifirma. Ühe teenusena pakub Resta andmekvaliteedi auditit. Firma kodulehe [Resta 2004] ja firma töötaja Riin Tillmanni bakalaureusetöö [Tillmann 2004] järgi kasutab Resta andmekvaliteedi auditeerimiseks SAS Institute'i tööde ja firma Eesti kogemuse põhjal arendatud viie-etapilist andmekvaliteedi auditi protsessi:

1. ettevõtte valitud andmebaaside kohta info kogumine:
 - a. andmebaaside tabelite nimekirja koostamine;
 - b. andmebaaside tabelite struktuuri esitamine (väljade nimed, tüübid, pikkused);
 - c. ärioloogikast lähtuvate väljadevaheliste seoste määramine;
2. auditi töökeskkonna loomine:
 - a. tarkvara konfigureerimine;
 - b. programmide ja kontrollitavate andmebaaside vastavusse viimine;
3. andmebaaside sisu analüüs ja testkontrollide läbiviimine;
4. andmekvaliteedi aruande koostamine;
5. töö tulemuste presentatsioon.

2.1.2. Net Group OÜ kasutatav metoodika

OÜ Net Group on äriintelligentsi, IT ja veebifirma, ärianalüüsi tarkvaratootja SAS Institute'i Consulting Partner. Ühe teenusena pakub Net Group andmekvaliteedi

auditit. Net Group kodulehel [Net Group 2004] on andmekvaliteedi protsessi kirjeldus järgmine – „Andmekvaliteedi protsessi võib jagada 5 järjestikusse etappi:

1. andmekvaliteedi auditi tegemine. Siin etapis hinnatakse andmekvaliteedi probleemide taset ja olemust erinevates andmeallikates. See osa peaks andma vastused küsimustele „kui kehvad andmed on?“ ning „kui palju mõjutab see organisatsiooni majandustegevust (saavutusi/ebaõnnestumisi)?“. Andmekvaliteedi audit põhineb 6 andmekvaliteedi võtmemõõdikul:

- a. täielikkus,
- b. vastavus,
- c. pidevus,
- d. täpsus,
- e. dubleerimine ja
- f. terviklikkus.

Tegevus põhineb üldistel ja kasutaja (ehk siis kliendi) defineeritud ärireeglitel.

Auditi võib jagada kolmeks järgmiselt:

- a. struktuuri analüüs;
- b. andmete analüüs;
- c. seoste avastamine.

2. andmekvaliteedi reeglid ja sihtmärgid;
3. andmekvaliteedi täiustamine;
4. andmekvaliteedi rakendamine;
5. andmekvaliteedi aruandlus ja jälgimine.

Sealjuures võib esimene tegevus olla ühekordne ja ülejäänud korduvad.”

2.1.3. Muud meetodikad

2.1.3.1. COBIT

COBIT, mille missiooniks on uurida, arendada, avaldada ja edendada asjakohast rahvusvahelist üldtunnustatud infotehnoloogia *juhtimise eesmärgistikku*, nimetab, ühena kolmekümneneljast, IT-protsessi „TT11 Hallata andmeid” [Cobit Framework 2000, 20]. Protsess on seatud rahuldama ärinõuet tagada kõikide andmete täielikkus,

õigsus ja kehtivus nende loomisel, kasutamisel, uuendamisel ja talletamisel. Ärinõude täitmise peavad tagama kolmekümne¹⁵ detailse juhtimiseesmärgi saavutus.

COBIT annab ainult juhtimiseesmärgid, mitte metoodikat ega protsessi kirjeldust mis vastaks küsimustele „miks?“, „kes?“ ja „kuidas?“ viia läbi infokvaliteedi auditit või juhtida infokvaliteeti.

2.1.3.2. Rahvusvahelised standardid

Standard „EVS-ISO/IEC 9126-1:2003 Tarkvaratehnika. Toote kvaliteet. Osa 1: Kvaliteedimudel“ kirjeldab tarkvaratoote kvaliteedi kaheosalist mudelit: a) sisekvaliteeti ja väliskvaliteeti ning b) kasutuskvaliteeti. Määratletud näitajad on kohaldatavad iga liiki tarkvarale, sealhulgas püsivaras sisalduvatele programmidele ja andmetele.

ISO 9000 seeria standard „EVS-EN ISO 9000-3:2003 Quality management and quality assurance standard“ sisaldab suuniseid ISO 9001:1994 kohaldamiseks tarkvara väljatöötamisele, tarnimisele, installeerimisele ja hooldusele.

Vaadeldud standardid käsitlevad mitte niivõrd andmeid ja infokvaliteedi juhtimist kui võrd andmete haldamise tehnoloogiat, sh tarkvara ja tarkvara arendus- jt tarkvara elutsükli etappide läbiviimise kvaliteeti.

2.2. Metoodika loomise alused

Kvaliteeti saab juhtida mitmeti – alates lihtsate põhimõtete rakendamisest ja lõpetades keerukate kvaliteedisüsteemidega. Magistrant esitas loodavale metoodikale järgmised nõuded:

- rakendatavus Eesti avalikus halduses ja erasektoris;
- arusaadavus äripoolele;

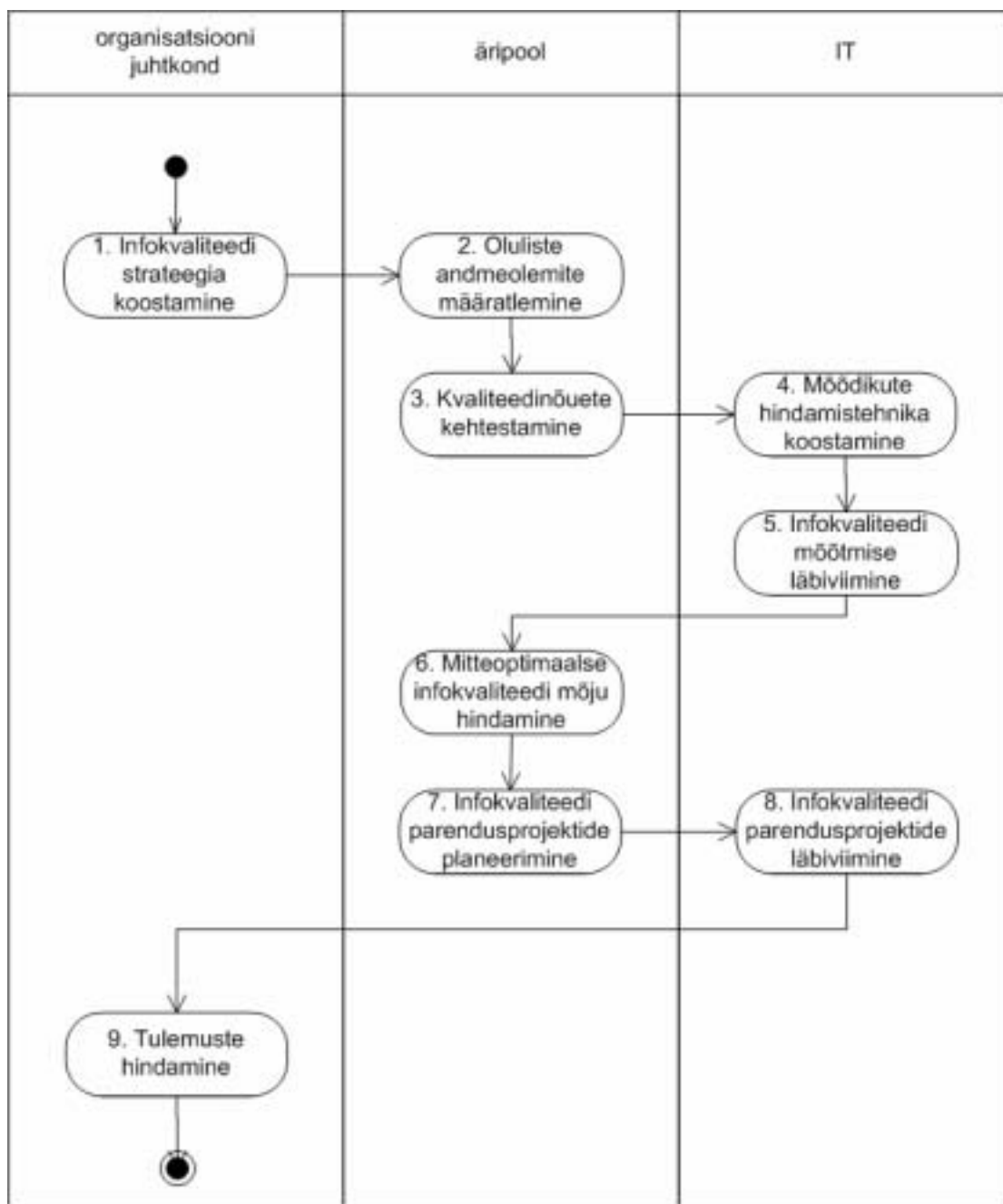
¹⁵ Võrdluseks – Cobit Quickstart [Cobit Quickstart 2003] järgi on andmehaldusprotsessi TT11 detailseid juhtimiseesmärke neli.

- läbiviidavus väikeste meeskondade poolt (üks-kaks konsultanti, lisaks äripoole esindajad).

Metoodika loomisel on aluseks võetud organisatsiooniteooria valdkonna juhtimismudelid [Lorents 2001, 145-150]. Autor on metoodikat kasutanud ja täiendanud mitmetes infokvaliteedi projektides, samuti mitmete Gartneri artiklite põhjal. Metoodika ei ole kindlasti valmis ja lõplik, vaid areneb infokvaliteedi projektides edasi. Hetkel kasutab magistrant seda metoodikat kahes erinevas organisatsioonis infokvaliteedi projektide läbiviimisel.

2.3. Infokvaliteedi juhtimise protsessi kirjeldus

Loodud infokvaliteedi juhtimise protsess on jagatud üheksaks järjestikuseks sammuks. Iga protsessi samm on tegevusdiagrammil (Joonis 3) liigitatud organisatsiooni üksuse alla, milline on vastava sammu läbiviimisel peamine tegutseja ja otsustaja.



Joonis 3. Infokvaliteedi juhtimise protsessi tegevusdiagramm UML notatsioonis¹⁶.

Järgnevalt on iga infokvaliteedi juhtimise protsessi sammu detailsemalt kirjeldatud.

¹⁶ UML – lahtikirjutatuna *Universal Markup Language* – on *de facto* standard äriprotsesside ja infosüsteemide modelleerimiseks ja kirjeldamiseks visuaalses keeles. UMLi kohta vaata näiteks <http://www.uml.org/> või [Booch 1999].

Samm 1. Infokvaliteedi strateegia koostamine

Milliselt juhtimistasandilt läheneda infokvaliteedile? T. Friedmani arvates tuleb infokvaliteedi parendamisel edu saavutamiseks tegeleda sellega mitte taktikaliselt, vaid strateegiliselt tasandilt [Friedman 2002, 1], so juhtkonna tasemel. Infokvaliteedi haldus peaks olema seotud organisatsiooni strateegiliste eesmärkidega – toetades nende saavutamist.

Organisatsiooni juhtkonna otsus on, kas ja kui ulatuslikult infokvaliteedi juhtimise protsess organisatsioonis juurutada. Võimalikud põhjused on järgmised:

- kahtlused informatsiooni kvaliteedis,
- organisatsioonide ühinemisel,
- üleminekul uuele infosüsteemile,
- andmelao loomisel.

Infokvaliteedialase tegevuse aluseks võib võtta – kui eksisteerib – organisatsiooni kvaliteedipoliitika, so kvaliteedialased üldised eesmärgid ja juhtnöörid.

Kvaliteedi juhtimiseks on vaja organisatsioonilist struktuuri, vastutuse määranguid, protseduure ja vahendeid, mis kokku moodustavad kvaliteedisüsteemi [IT juhtimise käsiraamat 1999, ptk 11.1].

Teadvusta huvigruppidele infokvaliteedi olulisus

Infokvaliteet on primaarselt 'äripoole', mitte IT-valdkonda kuuluv juhtimisobjekt, seetõttu on organisatsiooni ja ärivaldkonna juhtidele vajalik ülevaade ja arusaamine ebapiisava kvaliteediga info tagajärgedest organisatsioonile kui tervikule, sh finantstulemustele ja klienditeenindusele [Friedman 2002, 1].

Miks on infokvaliteet oluline?

Alljärgnev loetelu on magistritöö autori kogemuste ja mitmete infokvaliteedialaste artiklite ([Tambaum 2000; Allekand 2004]; Gartner uuringutulemused jmt) refereering. Loetelu võib kasutada lisaks infokvaliteedi strateegia koostamise etapile nii infokvaliteedi projektide tasuvuse arvutamisel (vt samm 7) kui ka

müügiargumentidena infokvaliteedi konsultatsiooniteenuse pakkumisel ja huvigruppides infokvaliteedi olulisust teadvustades.

Niisiis:

- ebapiisava kvaliteediga informatsioonile toetuva analüüsi tulemuste kasutamine võib viia ebaadekvaatse olukorra kirjelduseni ning see omakorda kehvade (juhtimis)otsusteni [Friedman 2002, 1]. Näiteks analüütilise kliendisuhete juhtimise (ingl. k lühend CRM) rakenduse määrang kliendi segmendi ja tasuvuse kohta mõjutavad operatiivset klienditeenindust, hinnapakumisi jms ning võivad halvematel juhtudel viia kasuliku kliendi kaotuseni. Parem on omada informatsiooni kvaliteedihinnangut ja võimalusel kahtlase kvaliteediga andmete juures vastavat märget;
- ebapiisava kvaliteediga andmed võivad endaga kaasa tuua nendele andmetele toetuva rakenduste või infosüsteemide alakasutuse või koguni hülgamise kasutajate poolt. Ehk, kasutajate rahulolematus infokvaliteediga väljendub rahulolematusest rakenduse vastu, usaldamatus andmete ja raportite vastu pöördub näiteks andmeida projekti vastu¹⁷;
- mitteoptimaalse, so ebapiisava või optimaalsust ületava kvaliteediga andmed toovad kaasa ressursside raiskamise. Näiteks vale kliendi aadress infosüsteemis toob kaasa selle, et arve ei jõua õigel ajal kliendini, nõuab mõlema poole raamatupidamiselt lisatööaega erikäsitusel, lisa trüki- ja postikulud, võimalikud probleemid kassavooga jm;
- ettevõtted, mis pakuvad informatsioonil baseeruvaid tooteid ja teenuseid, võivad infokvaliteedi parendamist kasutada selleks, et tõsta oma konkurentsieeliseid¹⁸ nagu tootmisettevõtted kasutavad selleks tootekvaliteedi juhtimist [Miller 1996].

¹⁷ Kirjeldatud olukord on üks andmeida loomise suurimaid projekti riske kuna andmeida projektid on tihti keskendunud raportitele ja andmete allikatele, mitte andmete kvaliteedile.

¹⁸ Tihti nähakse infokvaliteediga seoses ainult probleeme, näiteks andmete ebapiisavast kvaliteedist tingitud organisatsiooni tööprotsesside efektiivsuse langust. Selline lähenemine võib viia huvipoolte suhtumiseni infokvaliteeti kui millessegi „vastikult paratamatusesse”, mitte võimalusteni. Seetõttu on

Kõrvalefektina võib infokvaliteedi auditeerimine ja juhtimine:

- anda omanikele, klientidele, regulaatoritele, töötajatele, avalikkusele jt huvipooltele kinnituse organisatsiooni toimivuse ja teenuste kvaliteedist;
- aidata avastada võimalikke andmete kuritarvitamisi, nt tahtlikke andmete rikkumisi.

Samm 2. Oluliste andmeolemite määratlemine

Alustada korruga kogu ettevõtte hallatavate andmete parendamisega on ebaefektiivne [Friedman 2002, 2]. Suundi ja vihjeid fookuseerumiseks võib leida organisatsiooni missioonist, strateegiast, visioonist ja nendest alamates eesmärkides, millistest annab tuletada seoseid infokvaliteedi eesmärkidega. Määratleda tuleks organisatsiooni olulised eesmärgid ja nende täitmist tagavad protsessid (kui organisatsiooni strateegias on otseselt rõhutatud *klienditeeninduse* kvaliteedi olulisusele, siis tuleb selgitada protsessid, milles klient on osaline), ja nende protsesside infovajadused (näiteks: kliendi nimi, kontaktandmed, krediidilimiit, segmendimäärang), seejärel järjestada¹⁹ andmeolemid äri jaoks olulisuse alusel ning keskenduda esmajärjekorras olulisimale.

Samm 3. Kvaliteedinõuete kehtestamine

Selle sammu sisuks on iga sammus 2 valitud andmeolemi jaoks leida peatükis I toodud infokvaliteedi mõõdikutest olulised, samuti seada igale andmeolemi ja mõõdiku kombinatsioonile nõutavad kvaliteeditasemed.

Sobiv tehnika, mille alusel valida mõõdikud ja seada nõutavad tasemed, on iga konkreetse projekti jaoks unikaalne. Kaaluda võiks osapoolte intervjuerimist ja

oluline suuta näha ja sõnastada infokvaliteedi parendamises mitte ainult „kulude maailma” poolt, vaid ka „tulede maailma”.

¹⁹ Sobivad Eisenhoweri printsiip ja selle üldistused, samuti MoSCoW tehnika jmt.

küsitluste läbiviimist leidmaks olulised tegurid, millised mõjutavad infokvaliteedile seatavaid nõudeid.

Siinjuures tuleb arvestada, et:

- kvaliteedinõuded on dünaamilised. Ärikeskkond ja siit tulenevalt ärinõuded infokvaliteedile on ajas muutuvad, kõrgekvaliteetne informatsioon täna võib olla madalakvaliteetne homme. Kord juba kehtestatud kvaliteedinõuded tuleb vajadusel üle vaadata ja kohendada, vastasel juhul tegeletakse ‚eilsete‘ probleemidega ning kulutatakse ressursse ebaolulise peale [Miller 1996];
- kvaliteedinõuded on relatiivse loomuga. Igale infokvaliteedi mõõdikule nõutavat taset määrates tuleb arvestada, et eesmärk ei ole saada maksimaalset, vaid tarbijale piisavat kvaliteeti vastuvõetava ressursikuluga, so optimaalsust. Näiteks tulemust „98,5% kliendiandmetest on täpsed” ei saa konteksti teadmata hinnata skaalal hea-halb ega suboptimaalne-optimaalne-optimaalset ületav;
- kvaliteedi tagamiseks on soovitatav määratleda kvantifitseeritud kvaliteedinõuded;
- kvaliteedinõuete kehtestamisel arvestada kogu informatsiooni elutsükli (vt alapunkt 1.2.3. „Informatsiooni elutsükkel organisatsioonis”) jooksul andmeolemiga kokkupuuduvate huvipoolte nõudmiste tagamisega – punktis, kus andmed tekivad, võidakse andmeolemile esitada madalamaid kvaliteedinõudeid kui allavoolu vajalik on.

Samm 4. Mõõdikute hindamistehnika koostamine

Selle sammu sisuks on iga sammus 2 valitud andmeolemi ja sammus 3 valitud mõõdiku kombinatsiooni jaoks valida, kohendada või koostada sobiv hindamisprotseduur, samuti vajadusel määratleda hinnatavate andmete allikad. Hindamisprotseduurina võib kaaluda järgnevaid tehnikaid:

- andmekogu võrdlemine samasisulise, kuid kõrgema kvaliteediga andmekoguga. Näiteks võib võrrelda firma kliendiandmete registrit rahvastikuregistri ja äriregistriga;

- statistilised tehnikad. Näiteks korrelatsiooni analüüsimine andmeväljade täidetuse ja kirje sisestaja vahel võib osutada töötajatele, kelle andmesisestuskvaliteet ei ole piisav;
- nn projektspetsiifiliste ärireeglite koostamine.

Samm 5. Infokvaliteedi mõõtmise läbiviimine

Selle sammu sisuks on sammus 2 valitud andmeolemite infokvaliteedi mõõtmine vastavalt sammus 4 plaanitud protseduurile ja tulemuste dokumenteerimine; vajadusel sammu 4 tagasilikumine ja protseduuri täiustamine.

Samm 6. Mitteoptimaalse infokvaliteedi mõju hindamine

Infokvaliteedi auditi tulemusi võib võrrelda (ingl. k *benchmark*) nt teiste organisatsioonide tulemustega, kuid üldjuhul on olulisem hinnata mitteoptimaalse kvaliteediga informatsiooni mõju organisatsioonile endale. Kontekstist väljavõetuna ei ütle fakt „98,5% kliendiandmetest on täpsed” kas tegu on ebapiisava, piisava või liigse kvaliteediga, samuti ei aita eri andmeolemite kvaliteeditegurite mõõtmistulemuste aritmeetiline võrdlemine järjestada kvaliteedi parendusvajadusi. Seetõttu tuleks hinnata puudujäägid infokvaliteedis ning sellest tulenevaid võimalikke kahjusid ja – mis tihti arvestamata jääb – saamatajäänud võimalikke tulusid, mõlemaid ideaalis rahaliselt väljendatult.

Kuid ka selle sammu tulemuste alusel ei saa veel otsustada millise andmeolemite parendamisele tuleks esmajärjekorras asuda – selleks on vaja läbi viia parendamisprojektide kulukuse hindamine (kirjeldatud sammus 7) ning kalkuleerida parenduste tasuvus.

Samm 7. Infokvaliteedi parendusprojektide planeerimine

Selle sammu sisuks on hinnata lahenduste kulukus, võrrelda seda sammus 6 hinnatud võimaliku kulude kokkuhoiu ja tuludega, selgitada optimaalne kvaliteeditase²⁰, teadvustada see osapooltele koos argumentatsiooniga²¹ ning järjestada ja planeerida infokvaliteedi parendusprojektid.

Rahvatarkus ütleb – mõjusam on tegeleda põhjustega kui tagajärgedega. Infokvaliteedi kontekstis tähendab see järgmist – esmalt tuleb selgitada vigade tekkimiskohad ning seejärel:

- juba tekkinud vead andmetes võimaluste piires parandada ning
- hoiduda sarnaste vigade tekkimise kordumise eest või nende mõju vähendada.

Infokvaliteedi langemise põhjused (ingl.k *DQ failure*)

Kasutades alapunktis 1.2.3. käsitletud informatsiooni elutsükli mudelit, vaatleme põhjuseid, miks informatsiooni kvaliteet ei ole 100% – kus, millal ja kelle tõttu tekivad vead andmetes – ning kuidas andmete allavoolu liikudes infokvaliteet halveneb.

Vead võivad tekkida:

- andmete sisestamisel:
 - sisestusvead inimliku hooletuse tõttu. Näiteid: vead sünniaja sisestamisel, suur- ja väiketähtede valestikasutamine, kirillitsa mittekorrektne transkriptsioon;
 - sensorite vm vigade tõttu;
 - andmete vastuvõtmisel väljastpoolt organisatsiooni, näiteks info partneritel, klientidelt;

²⁰ Infokvaliteedi parendusprojektide planeerimisel arvestada, et taotleda tuleks soovitusel kõrgemat kvaliteeti – „Tabamaks sihtmärki tuleb sihtida kõrgemale”.

²¹ Sobiv tehnika on ROI-analüüs

- andmete liikumisel²² ja konverteerimisel (ingl.k *after capture point*). Näiteks andmete operatiivbaasist andmeaita laadimisel andmeatribuutide erinevate semantiliste tähenduste tõttu; rakendust vahetades²³;
- andmete hajusas ja heterogeenses keskkonnas töötlemisel ja säilitamisel. Hajusus ja heterogeensus on infoarhitektuuriliselt paratamatu, seda eelkõige valmiskonstruktsioonide, pärandrakenduste, kohandatud ja spetsiaalrakenduste samaaegse kasutamise tõttu. See aga toob paratamatult kaasa andmete liiasuse ja dubleerimise ning potentsiaalseks ohuks saab andmete sünkronisatsiooniseisundist väljaminek ehk „mitmese töö teke”;
- suure andmehulga tõttu, nii et lihtsalt ei piisa ressurssidest infokvaliteedi juhtimiseks, vigade avastamiseks, uurimiseks ja parendamiseks.

Infokvaliteedi parendusviise (ingl.k *DQ improvement, treatment, enhancement*) võib jagada mitmel viisil, nt infokvaliteedi mõõdikute kaupa. Levinud on siiski käsitleda neid grupeerituna järgmistesse valdkondadesse²⁴:

- tehnoloogilised²⁵,
- organisatsioonilised ja
- protsessilised parendusviisid.

Järgnevalt on vaadeldud igauht neist täpsemalt.

Organisatoorsed ning inimestele suunatud parendusviisid

- andmekäsitlejate, sh sisestajate:
 - motiveerimine – siduda infokvaliteedi mõõdikud nt tulemusjuhtimise mõõdikutega ja motiveerimissüsteemiga²⁶;

²² Info „liikumine” organisatsioonis on ülimalt vajalik, kuid võib kõrvalefektina tuua endaga kaasa infokvaliteedi languse.

²³ Eestis on mitmeid probleeme esinenud riiklike registrite tarkvara vahetades, nt Hooneregistri väljavahetamisel Riikliku ehitisregistri vastu.

²⁴ Praktikas on levinud infokvaliteedi parendusprojektides eri valdkondade tehnikate samaaegne kasutamine, nt võivad protsessilised ümberkorraldused vajada tehnoloogilist tuge.

²⁵ Lähenedes infokvaliteedile kui „IT-probleemile” võib tehnoloogiline infokvaliteedi parendusviis olla ainukeseks nähtavaks vahendiks, kuid mitte alati mõjusaimaks ja tõhusaimaks.

- koolitus;
- väljavahetamine;
- sisemiste regulatsioonide: ametijuhendite, protseduurireeglite jm; täiendamine;
- kultuurilised muudatused, sh suhtumise muutmine infokvaliteeti;
- andmeomanike määramine (ingl. k *stewardship* – valitsejad);
- kontrollprotseduuride kehtestamine.

Protsessilised infokvaliteedi parendusviisid

- protsesside ümberkorraldused (ingl.k *business process reengineering*), näiteks dubleerides andmete sisestustöö.

Tehnoloogilised infokvaliteedi parendusviisid

- andmete sisestusvigade vähendamiseks:
 - lubatud väärtuste sisestuskontrollid. Andmete sisestuskontrolli andmete tekkimise punktis loetakse efektiivsemaks kui allavoolu andmete parendamisel. Sisestuskontrollide rakendamisel tuleb arvestada eespool mainitud organisatsiooniliste meetmete juures toodud konfliktiohtudega: allavoolu rakenduste kvaliteedinõuded võivad olla erinevad selle organisatsiooniüksuse, kes andmeid sisestab, kvaliteedinõuetest;
 - välisosapooltega andmevahetuse automatiseerimine, mis võimaldab kohest andmekvaliteedi kontroll ja tagasisidet. Näiteks võtavad mitmed teenindustevõtted vastu maksepikenduse avaldusi iseteenindusbüroost, kaotades nii fakside jm raskesti loetamatute andmekandjate dešifreerimise ja ühtlasi sisestamise vajaduse;
- andmete organisatsioonisisel liikumisel tekkivate vigade vähendamiseks:

²⁶ Siin peitub konfliktioht – andmesisestaja ja andmesisestajale mõõdikute seadja (tüüpiliselt tema otsene ülemus) ei pruugi olla andmekvaliteedist huvitatud kuna nende tekitatud andmed ei mõjuta nende (osakonna) tööd, samas kulutab ressursse: raha, aega. Näiteks kõnekeskuse operaatorit hinnatakse mõõdikuga vastuvõetud kõnede ning lahenduste arv, mitte kontaktinfo detailsus mis kontakti kohta üles täheldati. Viimane on aga oluline nt turundusele ja tootearendusele.

- perioodiline 'andmepesu' (ingl.k *data cleansing*), so vastavuse kontrollimine etteantud ärireeglitele;
- 'andmetulemüüri' (ingl. k *data firewall*) kasutamine. Andmetulemüür on loogiline element mille eesmärk on mitte lasta näiteks operatiivandmebaasist andmeaita laadimisel läbi vigaseid andmeid²⁷ või märgistada need 'kahtlasteks';
- välise andmeomanike kasutamine (ingl. k *service-bureaus*) ja andmete võrdlemine. Eestis pakuvad viimasel ajal aktiivsemalt vastavaid teenuseid rahvastikuregistri volitatud töötaja AS Andmevara, samuti äriregister. Kontroll võib olla nii pisteline (infokvaliteedi hindamiseks), ainult kahtlaste kirjete osas kui ka totaalne; ning intervalli alusel ühekordne, perioodiline ja reaalajas;
- andmete võrdlus andmedoonori ja andmeretsipiendi andmebaasides. Kontrollitakse ja parandatakse järgmisi andmete lahknevusi: andmete liiasust, andmete puuduolekut, valestitõlgendatud andmeid.

Infokvaliteedi tehnoloogilise parendamise punkti alla kuulub ka tehnilise arhitektuuri mõiste (ingl. k *technical architecture*). Tehniline arhitektuur ei ole ette nähtud otseselt andmete parendamiseks, vaid andmete *käitlemise* tehnoloogia parendamiseks. Üldlevinud viisid on lisada objektidele unikaalsed koodid (nt füüsilistele isikutele isikukoodid, juriidilistele isikutele äriregistri registreerimisnumbrid), andmete tsentraliseerimine, normaliseerimine, XML kasutamine, käsitsitöö vähendamine.

Samm 8. Infokvaliteedi parendusprojektide läbiviimine

Sammus 8 viiakse läbi sammus 7 planeeritud infokvaliteedi parendusprojektid, st nii andmete 'ajaloo' parendamine kui ka andmete haldamise vahendite parendamine. Sobivaks tehnikaks on projektijuhtimine meetodil mis on organisatsioonis juba kasutusel, näiteks lähtudes piirangute teooriast²⁸.

²⁷ Võrdle võrgutulemüüri, millise tööpõhimõtte on leida kirjeldatud muustritele (so IP aadressid, pordid jms) vastavad paketid ning need seejärel blokeerida.

²⁸ Piirangute teooria kohta vaata näiteks <http://www.goldratt.ee>

Samm 9. Tulemuste hindamine

Aja möödudes on tarvilik hinnata, kas rakendatud parendamisprojektid on edukad ja mil määral. Vastavalt vajadusele tuleb siirduda tagasi sammuni 1-7: ümber defineerida olulisi andmeolemeid, laiendada nende loetelusid, võtta ette vähemprioriteetseid alasid, rakendada teisi kvaliteedi parendusmeetodeid jms.

2.4. Metoodika rakendamise juhised

Magistritöö autori senise kogemuse järgi ligi kümne infokvaliteedi projekti põhjal on toodud metoodika rakendatav nii era- kui ka avaliku sektori organisatsioonides. Metoodika on kohandatav nii ühekordse infokvaliteedi auditeerimis- ja/või parendusprojekti läbiviimiseks kui ka protsessina kasutamiseks. Viimasel juhul ei tähenda toodud metoodika sammudeks jagatus seda, et korraga võib olla käsil üks ja ainult üks samm – paralleelselt võib töös olla erinevate andmeolemite kvaliteediparendusprotsesse eri faasides, kriitiline on siinjuures organisatsiooni mitme projekti haldamise võimelisus.

3. INFOKVALITEEDI JUHTIMISE METOODIKA RAKENDUSNÄIDE

Käesolevas peatükis vaadeldakse eelnevalt käsitletud infokvaliteedi mõõdikute (peatükk I) ja väljapakutud infokvaliteedi juhtimise meetodika (peatükk II) ühte rakendusnäidet. Kirjeldatud on magistritöö kirjutamise ajal käimasolevat Riikliku ehisregistri infokvaliteedi projekti, milles magistritöö autor osaleb konsultandi rollis. Projekti teostusest on magistrandi kanda umbes 50%.

Magistritöö avalikustatud versioonist on kärbitud kolmanda peatüki sisu.

KOKKUVÕTE

Informatsioon on infoühiskonnas üheks olulisimaks ressursiks ning kriitiliseks teguriks organisatsioonide ellujäämise seisukohast. Info haldamiseks organisatsioonis on otstarbekas sisse tuua infokvaliteedi mõiste. Infokvaliteeti on võimalik mõõta ja juhtida ning infokvaliteeti saab parendada. Infokvaliteedi parendamisse on otstarbekas suhtuda kui võimalusse tõsta informatsioonil baseeruvate teenuste ja otsuste kvaliteeti.

Magistrant seadis käesolevale tööle järgmised eesmärgid:

1. eestikeelsete infokvaliteedi ja infokvaliteediga seotud terminite väljapakumise ja infokvaliteedi mõõdikute defineerimine;
2. meetodika loomine infokvaliteedi juhtimiseks ning selle testimine ja täiendamine reaalsetes projektides;
3. infokvaliteedialaste teadmiste levitamiseks aluse loomine ja sellega eelduse loomine infokvaliteedi kasvuks Eesti organisatsioonides.

Töö käigus koostas magistrant infokvaliteedi struktuuri kirjeldava mudeli, mis koosneb neljast võtmemõõdikust: asjakohasus, õigsus, kasutatavus ja andmete kirjeldatus; ning paarikümnest alamõõdikust. Mõõdikutele on leitud eestikeelsed nimetused, samuti on mõõdikud defineeritud.

Infokvaliteedi juhtimiseks on loodud algupärane ja universaalne meetodika ning seda on testitud mitmetes organisatsioonides, millest ühe – Riikliku ehtisregistri infokvaliteedi projekti – läbiviimist ja tulemusi on töös pikemalt kirjeldatud.

Esitatud tulemused infokvaliteedi struktuuri ja infokvaliteedi juhtimise meetodika osas ei ole kindlasti ammendavad ega lõplikud, vaid vajavad edasist täiendamist. Siit lähtuvalt pakub töö autor välja edasiarenduse võimalikud suunad:

- rakenduslikud teemad:
 - infokvaliteedi juhtimise meetodika rakendamine organisatsioonides ja tagasiside põhjal meetodika täiendamine;

- üksikute infokvaliteedi mõõdikute mõõtmistehnikate kasutamine ja täiendamine;
- teoreetilised teemad:
 - infokvaliteedi ja infosüsteemi kvaliteedi seoste uurimine ning infokvaliteeti arvestavate lähenemiste integreerimine infosüsteemide arendusmetoodikatesse;
 - IT juhtimise käsitlemine läbi infokvaliteedi juhtimise;
 - infokvaliteedi standardite ja infokvaliteedi klasside väljatöötamine või olemasolevate standardite ülevõtmine Eesti organisatsioonide tarbeks.

RESÜMEE

Information Quality and Its Management in Organization

Martin Luts

Summary

The goal of the summary is to give a brief review of the aims, solutions and results of this Master's thesis.

In the information society data is a valuable asset in any organization besides employees, clients and brand. The duty of CIO is to administrate this asset and the technology used to process it. It is common knowledge that 'you can't manage what you don't measure'. But how does one value and measure data? Data can and must be described taking into account both quantity and quality. Unlike quantity quality has multiple facets. This Master's thesis deals with two subjects: firstly what information quality is and secondly how to manage it.

The subject of information quality is important and actual both in public and private sectors of Estonia. To assure the quality of services based on information systems it is equally important to value the skills and motivation of information system's users and also the quality of data.

1. Problems

Although the subject of information quality is important, it has in the author's opinion been underestimated: there is no reference to the quality of data in relevant documents regulating databases (including state registers) and the subject is not discussed in the information systems' courses of Estonian universities.

The author sees the following problems concerning information quality:

- the quality of information and its characteristics are not tightly defined. Understanding the nature of information quality is vague. There are many and partly overlapping definitions of the characteristics of information quality;
- there is no universal, public, and commonly accepted methodology to audit, improve and manage information quality. Consulting companies offering

information quality evaluation and auditing services regard their unique methodology as competition advantage;

- knowledge of information quality and the subjects concerned is not widely spread in Estonia and the corresponding terms are not yet fully developed.

2. The aims of the Master's thesis

Taking into account the abovementioned problems, the author set the following aims:

- to define information quality characteristics and to propose terminology in Estonian;
- to develop an information quality managing method; to test and improve it in real-life projects;
- to develop a basis for spreading knowledge of information quality and thereby creating preconditions for improvement of information quality in Estonian organizations.

3. Solutions

In order to manage information quality, it is important to have a solid knowledge of the object under management. The First Part of the Master's thesis' (Theoretical Treatment of Information Quality) tries to find an answer to the question 'What is information quality?', deals with information quality at general level, and so is applicable to data of any kind and form. Three approaches to information quality are chosen and compared, namely:

- Cobit's Control Objectives,
- H. Miller's article 'The Multiple Dimensions of Information Quality',
- Gartner Group's research notes.

The author defines information quality characteristics, proposes terminology in Estonian, and presents the structure of information quality.

In the Second Part (Control Method of Information Quality) the author introduces the original method to manage information quality in organization. The author has improved the method with experience gained from real-life information quality projects and from students whom the author has taught the subject. One of such

projects carried out following the new method is described in the Third Chapter of the thesis.

4. Results

Information quality is defined by the following means:

- list of information quality's characteristics;
- definitions of information quality's characteristics;
- one possible way of grouping characteristics of information quality.

The definition can be used:

- managing information quality:
 - when defining the object and scope of the contract of information quality audit projects;
 - when analyzing information quality problems and setting required quality levels;
- teaching subjects covering information systems and database management systems.

The proposed method of managing information quality has been put into practice at the National Construction Register (NCR) and in other information quality projects.

The two most important outcomes of the NCR information quality projects are:

- to improve the information quality characteristic of relevance a list of data entities – that are lacking at the current version of NCR database but which several stakeholder groups need – has been compiled. To implement this plan two solutions are suggested:
 - to extend the database of NCR;
 - to re-establish Local Construction Register. The local database would contain the lacking data entities;
- to improve the information quality characteristic of accuracy a method has been created for total inventory of constructions within the municipal borders of Tallinn.

The results concerning the structure of information quality and the method to manage information quality are certainly not thorough and final and thus need improvement. Taking this into account, the author proposes possible ways of further development:

- practical subjects:
 - putting the method of information quality management into use in different organisations and improving the method based on feedback;
 - using and improving single information quality metric measuring techniques;
- theoretical subjects:
 - examination of connections between information quality and the quality of information system and integrating those approaches to the development methods of information systems that take into account information quality;
 - treatment of IT management through information quality management;
 - development of information quality standards or adapting the ones already existing for use in Estonian organisations.

KASUTATUD KIRJANDUS

Allekand, I. (2004) Andmete kvaliteet tagab usaldusväarsuse. Äripäev, 06.05.2004

Andmekogude Riiklik Register (2004) <http://www.riik.ee/arr/register/ehitus.html> seisuga 26.06.2004

Andmekogude seadus (2004) Redaktsioon 28.04.2004.
<https://www.riigiteataja.ee/ert/act.jsp?id=745339> seisuga 16.06.2004

Booch, G., Rumbaugh, J., & Jacobson, I. (1999) The Unified Modeling Language User Guide. Addison-Wesley

BrainyQuote (2004)
<http://www.brainyquote.com/quotes/quotes/s/saintaugus108119.html> seisuga 14.06.2004

Carr, N. G. (2003) IT Doesn't Matter. Harvard Business Review 05.2003.
<http://www.nicholasgcarr.com/articles/matter.html> seisuga 14.06.2004

Cobit juhtimiseesmärgid (1998) IT Governance Institute <http://www.itgi.org> seisuga 12.07.2004

Cobit Executive Summary (2000) IT Governance Institute <http://www.itgi.org> seisuga 12.07.2004

Cobit Control Objectives (2000). IT Governance Institute <http://www.itgi.org> seisuga 12.07.2004

Cobit Framework (2000) IT Governance Institute <http://www.itgi.org> seisuga 12.07.2004

Cobit Quickstart (2003) IT Governance Institute <http://www.itgi.org> seisuga 12.07.2004

- Drucker, P. F. (1998) The Coming of the New Organization. Harvard Business Review Jan-Feb 1998 pp 45-53.
- Eesti Õiguskeele Keskus (2004) Euroopa Liidu õigusaktide kogu.
<http://www.legaltext.ee/> seisuga 18.06.2004
- Ehitisregistri põhimäärus (2002) „Riikliku ehitisregistri” asutamine ja pidamise põhimäärus. Vabariigi Valitsuse 17. detsembri 2002. a määrus nr 405.
<https://www.riigiteataja.ee/ert/act.jsp?id=230771>
- Eilenfielt, V. & Lührman, S. (2000) Data Quality - The EI/DS Approach
- ENE (1998) Eesti Nõukogude Entsüklopeedia. Osa 3. Tallinn
- Eesti Tulevikuuuringute Instituut (1997) Maailm ja Eesti. Tulevikutrendid. Tartu/Tallinn
- Euroinfo (2004) Käesoleva töö autori küsimus ja Euroinfo töötaja Merle Tõniste vastus andmekvaliteedi õigusliku regulatsiooni kohta Euroopa Liidus.
<http://www.eurotelefon.ee/?kysimus> seisuga 19.06.2004
- FocusIT (2003) TLV registreeritavate ehitisandmete loetelu. Lõpparuanne. Tallinn
- Fox C., Levitin A., & Redman T. (1994) The Notion of Data and Its Quality Dimensions. Information Processing & Management 30 (1), pp. 9-19
- Friedman, T. (2002) A Strategic Approach to Improving Data Quality. Research Note. Gartner Group <http://www.gartner.com> seisuga 11.06.2004
- Friedman, T. (2003) Client Issues for Data Quality. Research Note. Gartner Group <http://www.gartner.com> seisuga 11.06.2004
- Infopoliitika põhialused (2003) Eesti infopoliitika põhialused aastateks 2004-2006.
<http://www.riso.ee/et/Infopoliitika2.doc> seisuga 08.06.2004

Infopoliitika tegevuskava (2004a) Infopoliitika tegevuskava 2005.a.

<http://www.riso.ee/et/Atg2005.doc> seisuga 08.06.2004

Infopoliitika tegevuskava (2004b) Infopoliitika tegevuskava 2005.a.

http://www.riso.ee/et/Atg2005_14jun.doc seisuga 16.06.2004

Information Technology Standards Guidance (1999). US DON CIO ITSG Integrated Product Team, <http://www.doncio.navy.mil/itsgpublic/> seisuga 11.07.2004

Isikuandmete automatiseeritud töötlemisel isiku kaitse konventsioon (1981)

Strasbourg, 28. jaanuar 1981

IT juhtimise käsiraamat (1999) Äripäeva kirjastus

Laks, M. (2004) Andmekvaliteedi hindamine kohtulahendite registri näitel.

Magistritöö. Juhendaja P. Parmakson. TPÜ 2004.

http://www.cs.tpu.ee/osakond/opilaste_tood/magistri_tood/2004_kevad/Mihkel_Laks/Mihkel_Laks_Magistri_Too_Lisadeta.pdf seisuga 12.07.2004

Long J. A., Richards J. A., & Seko C. E. (2004) The Canadian Institute for Health Information (CIHI) Data Quality Framework. Version 1: A Meta-Evaluation and Future Directions.

http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page=quality_framework_e seisuga 29.06.2004

Lorents, P. (2001) Süsteemse käsitluse alused. Tallinn

Lyman, P. and Varian, H. (2003) How Much Information.

<http://www.sims.berkeley.edu/research/projects/how-much-info-2003> seisuga 14.06.2004

Miller, G. (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, 1956, vol. 63, pp. 81-97. <http://www.well.com/user/smalin/miller.html> seisuga 30.06.2004

- Miller, H. (1996) The Multiple Dimensions of Information Quality. Information Systems Management, pp 79-82
<http://www.muhlenberg.edu/depts/abe/business/miller/mdiqual.html> seisuga 16.06.2004
- Netgroup (2004) OÜ Net Group andmekvaliteedi auditi teenuse kirjeldus firma kodulehel. <http://www.netgroup.ee> seisuga 30.06.2004
- Open Group (2004) The Open Group Architectural Framework. TOGAF "Enterprise Edition" Version 8.1. <http://www.opengroup.org/architecture/togaf8-doc/arch/> seisuga 27.06.2004
- Pearlson, K. E. (2001) Managing and Using Information Systems
- Resta (2004) AS Resta andmekvaliteedi auditi teenuse kirjeldus firma kodulehel. <http://www.resta.ee> seisuga 30.06.2004
- RISAT (2004) Riigi Infosüsteemi ja andmeteenuste seaduse eelnõu. http://www.riso.ee/et/RISAT_21_04.doc seisuga 16.06.2004
- Silvet, J. (1989) Inglise-Eesti sõnaraamat. I osa. 1989
- Silvet, J. (1990) Inglise- Eesti sõnaraamat. II osa. 1990
- SNIA (2004) Storage Networking Industry Association. <http://www.snia.org/about/> seisuga 18.06.2004
- SRC (2004) Strategic Research Corporation. <http://www.sresearch.com/> seisuga 18.06.2004
- Tambaum, T. (2000) Tee andmebaasid korda. Äripäev, 04.12.2000
- Tillmann, R. (2004) Bakalaureusetöö mustand andmekvaliteedist
- ÕS (1999) Õigekeelsussõnaraamat. Tallinn

INFOKVALITEEDIALASTE TERMINITE REGISTER

- ajakohasus, 29
- alammõõdik, 24
- andmed, 13
- andmekvaliteet, 19
- andmeolem, 15, 43, 45
- andmeomanik, 48
- andmepesu, 49
- andmete allavoolu liikumine, 18, 44, 46, 48
- andmete kirjeldatus, 34
- andmetulemüür, 49
- asjakohasus, 26
- esitusvorming, 32
- haaratus, 30
- info elutsükkel, 15
- infokvaliteedi langemise põhjused, 46
- infokvaliteedi parendusviisid, 47
- infokvaliteet, 24
- informatsioon ehk info, 13
- informatsiooni kvaliteet, 18
- käideldavus, 33
- kasutatavus, 31
- konfidentsiaalsus, 33
- kontekst, 32
- kvaliteedi tagamine, 18
- kvaliteet, 18
- mõõdik, 24
- õigsus, 27
- organisatoorsed infokvaliteedi
 parendusviisid, 47
- protsessilised infokvaliteedi
 parendusviisid, 48
- säilitusvorming, 32
- sisemine sidusus, 30
- täielikkus, 27
- täpsus, 28
- teave, 14
- tehnoloogilised infokvaliteedi
 parendusviisid, 48
- terviklus, 29
- väline sidusus, 33
- vorming, 31
- võtmemõõdik, 24