

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Iris Nael 204267IABM

**E-KAUBANDUSETTEVÕTTE X  
ANDMEKVALITEEDI ANALÜÜS JA  
PARENDAMINE**

Magistritöö

Juhendaja: Erki Eessaar

PhD

Tallinn 2023

## **Autorideklaratsioon**

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Iris Nael

06.05.2023

## **Annotatsioon**

Väike e-kaubandusettevõtte X peab e-poodi kasutades populaarsed sisuhaldussüsteemi WordPress ja selle pistikprogrammi WooCommerce. Käesoleva töö eesmärgiks oli saavutada ettevõtte X andmekvaliteedi püsiv paranemine.

Eesmärgi saavutamiseks uuriti tänapäeval kasutuses olevaid andmekvaliteedi raamistike et valida ja rakendada sobivaim andmekvaliteedi hindamise ja parendamise raamistik väikesele eraettevõttele. Samuti koostati ettevõtte MySQL andmebaasi disaini mudel koos selle andmebaasi tabelite ja veergude vabatekstilise sõnalise selgitusega.

Töös kasutatakse ettevõtte X andmekvaliteedi parandamiseks Eesti riigi infosüsteemi andmekogude andmekvaliteedi juhtimise raamistikku, mida on kohandatud väikese eraettevõtte jaoks. Töö tulemusena viidi ettevõtte andmekvaliteet esimeselt (kõige madalamalt) küpsustasemelt kolmandale küpsustasemele (viietasemelisel skaalal). Teiste sõnadega viidi andmekvaliteet reaktiivselt tasemelt standardiseeritud tasemele.

Magistritöö sisaldab muuhulgas ettevõtte andmekvaliteedi analüüsi ja andmebaasi disainiprobleemide analüüsi.

Püstitatud andmekvaliteedi nõuetele mittevastavate andmete leidmiseks loodi ja juurutati veebipõhise andmekvaliteedi hindamistöölaua prototüüp. See programm teeb andmebaasi põhjal päringuid, et leida andmekvaliteedi nõuetele mittevastavaid andmeid.

Andmekvaliteedi hindamistöölaua valideerimiseks kasutati A/B testimist ettevõtte X töötajatega, kelle igapäevane tööülesanne on uute kaupade, toodete ja teenuste kohta informatsiooni kogumine, selle töötlemine ja e-poe andmebaasi sisestamine.

Töö tulemuseks saadi vastuseid kõigile püstitatud uurimisküsimustele, leiti ja rakendati sobiv andmekvaliteedi juhtimise raamistik ning loodi andmekvaliteedi hindamistöölaua prototüüp. Andmekvaliteedi püsiva paranemise tagab andmekvaliteedi regulaarne kontrollimine ja parandustegevuste läbiviimine vastavalt kindlaksmääratud juhistele.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 108 leheküljel, 10 peatükki, 17 joonist, 29 tabelit.

## **Abstract**

# **Analysis and Improvement of Data Quality of an E-commerce Enterprise X**

For organizations, the quality of the data it uses is important, as it is the basis on which decisions are made that affect all stakeholders in the organization. Therefore, it is important for an organization to have high-quality data at its disposal – this can give them a competitive advantage and enable them to make more accurate business decisions.

The main business of a small private e-commerce enterprise X is to keep an e-shop by using a popular content management system WordPress and its plugin WooCommerce. The main problem that motivates this thesis was the insufficient quality of the data managed by the enterprise. The goal of the thesis was to persistently improve the data quality of the enterprise X.

To achieve this objective, the data quality frameworks in use today were studied to find the most suitable framework for data quality assessment and improvement for a small private enterprise. Moreover, the database design model of the enterprise was constructed that includes free-form descriptions of tables and columns. The database was created in MySQL database management system.

The work uses the framework for managing data quality in Estonian national information system databases. The work adapts it for a small private enterprise. The framework specifies five levels of data quality. As the result of the work the data quality of the enterprise was taken from the lowest level (reactive level) to the third level (standardized level).

Among other things the work includes an analysis of data quality in the database of the enterprise X and an analysis of design problems of its database.

A prototype of a web-based data quality assessment dashboard was created and deployed to identify data that does not meet the set data quality requirements. This program queries

the database to find data elements that does not conform to the data quality requirements. The program was created in PHP.

The validation of the dashboard prototype was carried out with the employees of enterprise X, whose daily task is to collect information on new goods and services, process it, and enter it into the e-shop database. The validation was carried out through A/B testing of the evaluation dashboard prototype, comparing the enterprise's initial solution, i.e., manual search, with the prototype.

The work resulted with answers to all research questions, finding, and implementing a suitable data quality management framework, and creating a prototype data quality assessment dashboard. The additional results include an analysis of database design problems in WordPress and WooCommerce database, an evaluation of how well the data quality requirements are fulfilled in the current database, a case study of using the framework for managing data quality in Estonian national information system databases, and its adaptation to a small private enterprise. The amount of changes that the author made in the framework was small, i.e., in general it is suitable for the small private enterprises.

The persistent improvement of data quality is achieved through the regular checking of data quality and fixing the identified data quality problems based on the specified guidelines.

The thesis is in Estonian and contains 108 pages of text, 10 chapters, 17 figures, 29 tables.

## Lühendite ja mõistete sõnastik

Andmekvaliteedi profiil	Andmekontrollide ja ärireeglite hulk, mille abiga on võimalik välja selgitada andmete kvaliteeti [1], [2].
API	<i>Application Programming Interface</i> – rakendusliides, programmikoodi hulk, mille abil on võimalik edastada andmeid ühest tarkvaratootest teise ja vastupidi [3].
CASE	<i>Computer Aided Software Engineering</i> . Tarkvara, mis on mõeldud info- ja tarkvarasüsteeme kirjeldavate mudelit loomiseks ja nendest mudelitest uute tulemite genereerimiseks.
GMT	<i>Greenwich Mean Time</i> – keskmine päikeseaeg Greenwichi meridiaanil [4].
Kasutaja	Organisatsiooni e-poodi registreerunud töötaja või klient.
Kaup	Organisatsiooni poolt müüdavad asjad, mida toodavad organisatsiooni partnerid.
Klient	Kliendi all mõistetakse organisatsiooni kaupade ja toodete ostjaid.
Makse <i>tokenid</i>	Võtmed, mida kasutatakse suhtlemiseks erinevate maksete töötajatega (nt. PayPal) [5].
Metakirjeldus	Lühike kokkuvõttev lause lehekülje, kauba, toote, teenuse ja ajaveebi postituse kohta, mida kasutajale otsingumootoris otsides kuvatakse.
MIME	<i>Multipurpose Internet Mail Extensions</i> – protokoll, mille abiga kasutajad saavad vahetada omavahel erinevaid andmefaile (nt pildid, videod, heli) [6].
PayPal	Elektroonne maksesüsteem, mis võimaldab võtta vastu ja saata raha [7].
REST API	Rakendusliides, mis vastab REST ( <i>Representational State Transfer</i> ) disainipõhimõtetele [8].
RIA	Riigi Infosüsteemi Amet.

Semantiline õigsus	Kontrollib, et andmed oleksid sisuliselt korrektsed [9].
Sisesed lingid	Veebisaidi sisesed lingid ehk link, mis ettevõtte veebisaidilt viitab ettevõtte teisele veebilehele.
SKU	Tootekood.
Slug	URL-i osa, mis jääb peale protokolliga ja domeeni nime. Näiteks <a href="https://domeen.ee/slug">https://domeen.ee/slug</a> .
SQL	<i>Structured Query Language</i> – valdkonnapõhine programmeerimiskeel, mille abiga on võimalik salvestada ja töödelda andmeid andmebaasis ja hallata andmebaasiobjekte [10].
Stripe	Elektroonne maksesüsteem, mille abiga on võimalik saata arveid, võtta vastu raha ja töödelda makseid [11].
Süntaktiline õigsus	Näitab, et andmed on andmekogus kirjavigadeta [12].
Teenus	Ettevõtte X poolt osutatav tegevus, näiteks koolitus või konsultatsioon.
Toode	Organisatsiooni enda toodang, mida klientidele müüakse.
TSD	Tulu- ja sotsiaalmaksu, kohustusliku kogumispensionini makse ja töötuskindlustusmakse deklaratsioon [13].
TÖR	Töötamise register [14].
UML	<i>Unified Modeling Language</i> – modelleerimiskeel, mida kasutatakse info- ja tarkvarasüsteemide visuaalsete mudelite loomiseks [15].
URI	<i>Uniform Resource Identifier</i> – tähemärkide jada, mis identifitseerib ressursi [16]. URI näide: <a href="https://domeen.ee/pildid/pilt1.png">https://domeen.ee/pildid/pilt1.png</a>
URL	<i>Uniform Resource Locator</i> – tähemärkida jada, mis toimub kui võrguaadress ressursi laadimiseks veebis [16]. URL näide: <a href="https://domeen.ee">https://domeen.ee</a>
VAT	<i>Value Added Tax</i> ehk käibemaks.
Võtmesõna	Sõna või lause, mille järgi kliendid võivad veebilehti, kaupu, tooteid, teenuseid ja ajaveebi postitusi leida (nii veebilehel endal, kui ka üldiselt otsingumootorites).
Välised lingid	Lingid, mis suunavad teisele veebisaidile ehk ettevõtte veebilehel on link, mis viib teise ettevõtte veebileheküljele.



## Sisukord

Autorideklaratsioon .....	2
Annotatsioon.....	3
Abstract Analysis and Improvement of Data Quality of an E-commerce Enterprise X...	5
Lühendite ja mõistete sõnastik .....	7
Sisukord.....	9
Jooniste loetelu .....	14
Tabelite loetelu .....	15
1 Sissejuhatus .....	16
1.1 Taust ja probleem .....	16
1.2 Töö kirjeldus.....	18
2 Metoodika.....	20
2.1 Ülevaade objektist .....	20
2.2 Tööriistade kirjeldus .....	21
2.3 Tööprotsessi kirjeldus.....	22
3 Teoreetiline taust .....	26
3.1 Andmekvaliteet.....	26
3.1.1 Andmekvaliteedi üldine mõiste .....	26
3.1.2 Andmekvaliteedi tähtsus organisatsioonidele .....	28
3.1.3 Andmekvaliteedi aspektid .....	28
3.2 Andmekvaliteedi raamistikud.....	33
3.3 Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistik.....	38
4 Andmebaasi mudel .....	41
4.1 Andmebaasi alamosad .....	41

4.1.1 WordPress.....	41
4.1.2 E-poe pistikprogramm .....	42
4.1.3 Kodulehe optimeerimise pistikprogramm .....	42
4.1.4 Hinnapakumiste pistikprogramm .....	42
4.2 Põhiobjektid .....	42
4.3 Ettevõtte X andmebaasi andmemudel .....	43
4.3.1 Klassifikaatorite register .....	44
4.3.2 Kasutajate register .....	45
4.3.3 Klientide register .....	46
4.3.4 Kommentaaride register .....	46
4.3.5 Konfiguratsioonide register .....	47
4.3.6 Maksete register.....	49
4.3.7 Märkmete register .....	50
4.3.8 Müügiobjektide register.....	50
4.3.9 Sündmuste register .....	51
4.3.10 Tellimuste register .....	52
5 Andmekvaliteedi juhtimise raamistiku rakendamine .....	54
5.1 Hindamine .....	54
5.1.1 Andmete ulatuse määratlemine ja tsoneerimine .....	54
5.1.2 Andmekvaliteedi küpsustaseme hindamine.....	54
5.2 Planeerimine .....	55
5.2.1 Küpsuse sihttaseme planeerimine.....	56
5.2.2 Parendamise plaani koostamine.....	56
5.3 Teostamine.....	57
5.3.1 Juhtimine ja planeerimine.....	57
5.3.2 Organisatsioon ja vastutused .....	58
5.3.3 Protsessid .....	58
5.3.4 Teadmused ja kompetentsid .....	59
5.3.5 Töövahendid .....	59
5.4 Kontrollimine.....	59
5.5 Korrigeerimine.....	61

6 Andmekvaliteedi nõuded ja nende täidetud .....	62
6.1 Õigsus .....	62
6.1.1 Olemasoleva andmebaasi vastavus .....	63
6.1.2 Vastavuse tagamise meetmed .....	66
6.1.3 Kriitiliste nõuete põhjendus .....	67
6.2 Täielikkus .....	70
6.2.1 Olemasoleva andmebaasi vastavus .....	70
6.2.2 Vastavuse tagamise meetmed .....	71
6.2.3 Kriitiliste nõuete põhjendus .....	72
6.3 Kooskõla .....	73
6.3.1 Olemasoleva andmebaasi vastavus .....	73
6.3.2 Vastavuse tagamise meetmed .....	74
6.3.3 Kriitiliste nõuete põhjendus .....	75
6.4 Usaldusväarsus .....	75
6.4.1 Olemasoleva andmebaasi vastavus .....	75
6.4.2 Vastavuse tagamise meetmed .....	76
6.4.3 Kriitiliste nõuete põhjendus .....	77
6.5 Ajakohasus .....	77
6.5.1 Olemasoleva andmebaasi vastavus .....	77
6.5.2 Vastavuse tagamise meetmed .....	78
6.5.3 Kriitiliste nõuete põhjendus .....	78
6.6 Reeglipärasus .....	78
6.6.1 Olemasoleva andmebaasi vastavus .....	79
6.6.2 Vastavuse tagamise meetmed .....	79
6.6.3 Kriitiliste nõuete põhjendus .....	80
6.7 Konfidentsiaalsus .....	80
6.7.1 Olemasoleva andmebaasi vastavus .....	80
6.7.2 Vastavuse tagamise meetmed .....	81
6.7.3 Kriitiliste nõuete põhjendus .....	81
6.8 Ühekordsus .....	82
6.8.1 Olemasoleva infosüsteemi vastavus .....	82

6.8.2 Vastavuse tagamise meetmed .....	82
6.8.3 Kriitilised nõuded .....	83
6.9 Mitteliiasus .....	83
6.9.1 Olemasoleva andmebaasi vastavus.....	83
6.9.2 Vastavuse tagamise meetmed .....	84
6.9.3 Kriitilised nõuded .....	84
7 Protsessid .....	85
7.1 Rollid .....	85
7.2 Andmete haldamise protsessid .....	86
7.3 Andmekvaliteedi tagamise protsess.....	89
7.4 Andmekvaliteedi nõuete lisamise ja muutmise protsessid .....	90
8 Andmekvaliteedi profiil ja andmete sellele vastavuse hindamise töölaua prototüüp..	92
8.1 Andmekvaliteedi profiilid.....	92
8.1.1 Kauba ja toote andmekvaliteedi profiil .....	92
8.1.2 Teenuse andmekvaliteedi profiil .....	94
8.2 Kontrollpäringud .....	96
8.2.1 Õigsuse kontrollpäringud .....	96
8.2.2 Täielikkuse kontrollpäringud.....	96
8.2.3 Kooskõla kontrollpäringud .....	97
8.2.4 Konfidentsiaalsuse kontrollpäringud .....	98
8.3 Tagarakendus .....	98
8.4 Kasutajaliides.....	100
9 Tulemuste analüüs ja järeldused .....	104
9.1 Andmebaasi disainivigade analüüs.....	104
9.1.1 Andmebaasiobjektide nimetamine .....	105
9.1.2 Pole arvestatud isikuandmete kultuurilise erinevusega .....	106
9.1.3 EAV antimuster .....	106
9.1.4 Ebajärjekindel tõeväärtuste esitamine .....	107
9.1.5 Puuduvad kitsendused .....	107
9.1.6 Sama nimega veergudel on erinevad tüübid või väljapikkused .....	108

9.1.7 Andmebaasi disainiprobleemide mõju andmekvaliteedile .....	108
9.2 Andmekvaliteedi tagamise parimad praktikad .....	110
9.2.1 Andmehaldus .....	111
9.2.2 Andmekvaliteedi koolitused .....	111
9.2.3 Andmekvaliteedi probleemide tekkimise põhjused.....	112
9.2.4 Andmekvaliteedi kontrollid .....	113
9.2.5 Andmete profileerimine ja kontroll .....	113
9.2.6 Kriitiliste valdkondade määramine.....	113
9.3 Andmekvaliteedi juhtimise raamistiku kohandused väikesele eraettevõttele ....	114
9.4 Andmekvaliteedi juhtimise raamistiku rakendamise ja selle tulemuste analüüs	114
9.5 Andmekvaliteedi hindamise töölaua prototüübi analüüs ja täiendusvõimalused	116
9.5.1 Hindamistöölaua prototüübi valideerimine .....	117
9.5.2 Prototüübi täiendusvõimalused.....	118
9.6 Tehtud töö nõrkused .....	118
9.7 Ettepanekud edasiseks arenduseks ja uuringuteks.....	119
9.7.1 Andmekvaliteet küpsustasemele neli ja viis .....	119
9.7.2 Andmekvaliteedi nõuete täiendamine .....	121
9.7.3 Andmekvaliteedi hindamistöölaua prototüübi täiendamine .....	121
10 Kokkuvõte .....	123
Kasutatud kirjandus .....	125
Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks .....	133
Lisa 2 – Ettevõtte X wiki .....	134
Lisa 3 – Andmekvaliteedi hindamistöölaua prototüübi esialgne kasutajaliidese kavand .....	135
Lisa 4 - WordPressi andmebaasi disaini diagramm.....	136
Lisa 5 – Müügiobjektide register .....	137
Lisa 6 – Andmekvaliteedi küpsusmudel.....	141

## Jooniste loetelu

Joonis 1. Andmekvaliteedi moodsikud [53].....	27
Joonis 2. Kauba andmete haldamise protsess.....	87
Joonis 3. Teenuse andmete haldamise protsess.....	88
Joonis 4. Andmekvaliteedi tagamise protsess.....	89
Joonis 5. Andmekvaliteedi nõuete lisamise protsess.....	91
Joonis 6. Õigsuse andmekvaliteedi nõue „Ühegi kauba, toote ja teenuse nimetuse pikkus ei tohiks ületada 60 tähemärki“ kontrollpäring.....	96
Joonis 7. Täielikkuse andmekvaliteedi nõude „Igal kaubal, tootel ja teenusel peab olema kirjeldus“ kontrollpäring.....	97
Joonis 8. Kooskõla nõude „Igal avalikustatud kaubal, tootel ja tootega komplekti pandud teenusel, peavad olema kooskõlas lao staatus ja kogus“ kontrollpäring.....	98
Joonis 9. Konfidentsiaalsuse nõude “Kõigil veebisaidil registreerinud klientidel on staatus „customer“ ja neil puudub ligipääs veebipoe administraatoripaneelile” kontrollpäring.....	98
Joonis 10. Hindamistöölaua prototüübi projekti struktuur.....	99
Joonis 11. Hindamistöölaua kasuajaliideses kasutatud värvitoonid.....	100
Joonis 12. Andmekvaliteedi hindamistöölaua prototüübi veebilehe päis.....	100
Joonis 13. Andmekvaliteedi hindamistöölaua prototüübi veebilehe jalus.....	101
Joonis 14. Õigsuse andmekvaliteedi kontrollimise veebilehe fragment.....	101
Joonis 15. Täielikkuse andmekvaliteedi kontrolli veebilehe fragment.....	102
Joonis 16. Kooskõla andmekvaliteedi kontrollimise veebilehe fragment.....	102
Joonis 17. Konfidentsiaalsuse andmekvaliteedi kontrollimise veebilehe fragment.....	103

## Tabelite loetelu

Tabel 1. Andmekvaliteedi aspektid kirjanduse alusel. ....	28
Tabel 2. Mistahes valdkonnale rakendatavad andmekvaliteedi raamistikud. ....	33
Tabel 3. Andmekvaliteedi aspektide esinemine andmekvaliteedi raamistik. ....	37
Tabel 4. Klassifikaatorite registri tabelid. ....	44
Tabel 5. Kasutajate registri tabelid. ....	45
Tabel 6. Klientide registri tabelid. ....	46
Tabel 7. Kommentaaride registri tabelid. ....	46
Tabel 8. Konfiguratsioonide registri tabelid. ....	47
Tabel 9. Maksete registri tabelid. ....	49
Tabel 10. Märkmete registri tabelid. ....	50
Tabel 11. Müügiobjektide registri tabelid. ....	50
Tabel 12. Sündmuste registri tabelid. ....	51
Tabel 13. Tellimuste registri tabelid. ....	52
Tabel 14. Parentustegevused andmekvaliteedi 2. küpsustaseme saavutamiseks. ....	56
Tabel 15. Parentustegevused andmekvaliteedi 3. küpsustaseme saavutamiseks. ....	57
Tabel 16. Õigsuse nõuded ja nende täidetud ning kriitilisus ettevõttes X. ....	63
Tabel 17. Õigsuse nõuded, mille täidetust on võimalik IT-süsteemil kontrollida. ....	64
Tabel 18. Täielikkuse nõuded ja nende täidetud ning kriitilisus ettevõttes X. ....	70
Tabel 19. Koosõla nõuded ja nende täidetud ning kriitilisus ettevõttes X. ....	73
Tabel 20. Usaldusväarsuse nõuded ja nende täidetud ning kriitilisus ettevõttes. ....	76
Tabel 21. Ajakohasuse nõuded ja nende täidetud ning kriitilisus ettevõttes. ....	78
Tabel 22. Reeglipärasuse nõuded ja nende täidetud ning kriitilisus ettevõttes. ....	79
Tabel 23. Konfidentsiaalsuse nõuded ja nende täidetud ning kriitilisus ettevõttes. ....	80
Tabel 24. Ühekordsuse nõuded ja nende täidetud ning kriitilisus ettevõttes. ....	82
Tabel 25. Mitteliiasuse nõuded ja nende täidetud ning kriitilisus ettevõttes. ....	83
Tabel 26. Kauba ja toote andmekvaliteedi profiil. ....	92
Tabel 27. Teenuse andmekvaliteedi profiil. ....	94
Tabel 28. Parentustegevused andmekvaliteedi küpsustaseme 4 saavutamiseks. ....	119
Tabel 29. Parentustegevused andmekvaliteedi küpsustaseme 5 saavutamiseks. ....	120

# 1 Sissejuhatus

Selles peatükis esitatakse ülevaade magistritöö taustast ja probleemist. Peatükis on toodud välja töö kirjeldus, milles käsitletakse antud magistritöö jaotust.

## 1.1 Taust ja probleem

Magistritöös käsitletakse andmekvaliteedi olukorda ja parendamist e-kaubandusettevõtte X näitel. See ettevõtte tegeleb õppevahendite ja koolituste müümisega WordPressi pistikprogrammi WooCommerce ja teiste väiksemate pistikprogrammide baasil realiseeritud e-poe kaudu.

W3techs andmete järgi on 2022. aasta oktoobrikuu seisuga sisuhaldussüsteemi WordPressi turuosa sisuhaldussüsteemide hulgas 64.2% [17]. 43% maailma veebisaitidest on loodud WordPressi põhjal [18]. WordPressil põhineva WooCommerce tarkvaral on suurim – 29% suurune – turuosa e-poe tarkvarade hulgas [19]. Allika [19] kohaselt on 2022. aastal võrreldes 2021. aastaga see populaarsus tõusnud kolm protsendipunkti. 2022. aasta seisuga leidub vähemalt 3.58 miljonit veebisaiti, mis kasutavad WooCommerce pistikprogrammi [20].

Töö põhiprobleemiks on organisatsiooni hallatavate andmete ebapiisav kvaliteet, mille tulemusena tuleb andmeid täiendavalt kontrollida ja teha muudatusi vastavalt nõuetele. Probleemi esinemise tulemuseks on ressursside (nagu näiteks töötajate aeg, mida saaks kasutada millegi muu jaoks) kulutamine ettevõtte poolt. Probleemi loodetakse lahendada, võttes organisatsioonis kasutusele sobiv andmekvaliteedi raamistik ja luues andmekvaliteedi hindamise veebipõhise töölaua prototüübi, mis aitaks senisest kiiremini tuua välja andmete probleemsed kohad ning vähendada tulevikus probleemide tekkimist. Andmekvaliteedi parandamiseks plaanitakse rakendada RIA andmekvaliteedi tagamise juhendit riigi infosüsteemi andmekogude omanikele [12].

Käesoleva töö eesmärgiks oli saavutada ettevõtte X andmekvaliteedi püsiv paranemine.

Eesmärgi saavutamiseks on vaja saada vastus järgmistele uurimisküsimustele ja lähtudes vastustest astuda samme andmekvaliteedi parandamiseks.



- Milline on väikeettevõttele sobiv andmekvaliteedi raamistik? Millised on alternatiivid?
- Kui RIA raamistik sobib Eesti riigi infosüsteemi andmekogudele, siis kas võiks sobida ka eraettevõtete andmekogudele?
- Millised on kõige sagedasemad kvaliteediprobleemid ettevõtte X andmetes?
- Millised on organisatsiooni andmebaasi skeemi probleemid? Kas need võivad tingida andmekvaliteedi probleeme?

Magistritöö käigus soovitakse analüüsida olemasolevat olukorda sh uurida ettevõtte X SQL-andmebaasi nii seal olevate andmete kui struktuuri ja jõustatud piirangute suhtes. Olemasoleva olukorra analüüsimiseks plaanitakse kõigepealt koostada andmebaasi disaini mudel, mis organisatsioonil on töö alustamise seisuga puudu. Olemasolev andmebaas koosneb WordPressi enda ja ka erinevate pistikprogrammide tabelitest. Programmide arendajad pakuvad nende tabelite dokumentatsiooni [21], kuid see on organisatsiooni jaoks ebapiisav. Organisatsiooni töötajate jaoks, kellel kokkupuude andmebaasidega on üsna väike või puudub üldse, on olemasolevas dokumentatsioonis esitatud kirjeldused raskesti mõistetavad või ei sisalda need vajalikku infot. Näiteks ei ole pistikprogrammide tabelite dokumentatsioonis välja toodud veergude sõnalisi kirjeldusi ja kergesti mõistetavaid väärtuste näiteid. Olemasolevad kirjeldused on võõrkeeles ja asuvad erinevates kohtades, mis omakorda ei ole töötajate jaoks mugav. Näiteks WordPress dokumentatsioonis [21] on toodud välja andmebaasi kirjeldus, mis koosneb tabelite lühikirjeldusest ja tabelite veergude nimetustest, kuid tabeli veergude tähenduse kirjeldused puuduvad. Uus loodav dokumentatsioon sisaldab kasutatavate pistikprogrammide tabelite ja tabeli veergude detailset kirjeldust. Disaini mudeli koostamise näidet pakuvad Matkur ja Saluvere [22] ning Siitan [23].

Järgnevalt tuleb valida organisatsioonile sobiv andmekvaliteedi raamistik.

Peale sobivaima andmekvaliteedi raamistiku valimist ja kohandamist plaanitakse valitud raamistik organisatsioonis kasutusele võtta. Andmekvaliteedi raamistiku valimise ja kasutuselevõtu näidet ettevõttes pakub Oja [24]. Samuti plaanitakse luua töötav andmekvaliteedi hindamistoolaua prototüüp kasutades PHP programmeerimiskeelt. Töölaua teeb päringuid ettevõtte andmebaasist, leidmaks nii andmekvaliteeti iseloomustavaid koondandmeid, kui ka detailseid parandamist vajavaid andmeid.

Käesoleva töö oodatavateks tulemusteks on ettevõtte andmebaasi andmemudel, olemasoleva olukorra analüüs sh olemasolevate andmekvaliteedi probleemide ja andmebaasi disainiprobleemide analüüs, andmekvaliteedi raamistiku valimine ja rakendamine, teostatud andmekvaliteedi parendused ja hindamistöölaua prototüüp.

Tulemused, mis on huvipakkuvad laiemalt kui ainult ettevõttele X, on järgmised.

- WordPressi ja selle pistikprogrammide tabelitest moodustunud andmebaasi disainivigade analüüs.
- Analüüs konkreetse ettevõtte näitel, kui palju ja milliseid andmekvaliteedi vigu tekib WordPress ja WooCommerce tarkvara reaalse kasutamise tulemusel.
- WordPress ja WooCommerce tarkvara kasutajate andmekvaliteedi probleemide lahendamise võimaluste analüüs koos konkreetse näitega, mis on kvaliteedi parandamiseks ette võetud.
- WordPressi ja selle pistikprogrammide tabelitest moodustunud andmebaasi täiendatud dokumentatsioon.
- RIA raamistiku kohandus väikesele eraettevõttele ja selle raamistiku kasutamise näide.

Kuna WordPress ja WooCommerce on väga populaarsed süsteemid, siis on võimalike huviliste ring lai. Tulemusi saavad kasutada äriomanikud, et teha otsuseid andmekvaliteedi ja WordPressi kasutuse suhtes.

Töö valideerimine toimub läbi hindamistöölaua prototüübi A/B testimise ettevõtte X töötajatega [25]. Testimise käigus võrreldakse algset lahendust, kus töölauda ei olnud ja vigu tuli otsida käsitsi ja uut versiooni. Uuritakse kumb lahendus on organisatsiooni töötajatele meeldivam, kumma kasutamine võtab vähem aega ja kummaga leitakse ning parandatakse rohkem vigu. Uue versiooni all mõistetakse hindamistöölaua prototüübi kasutamist ja vigade parandamist.

## **1.2 Töö kirjeldus**

Käesolev töö jaguneb kaheksaks põhiosaks: meetodika, teoreetiline taust, andmebaasi mudel, andmekvaliteedi juhtimise raamistiku rakendamine, andmebaasi andmekvaliteedi nõuded, protsessid, andmekvaliteedi hindamise töölaua prototüüp ning tulemuste analüüs ja järeldused.

Ülevaade töö objektist, tööprotsessist ja kasutatud tarkvaralistest tööriistadest tuuakse välja teises peatükis. Andmekvaliteedi ja andmekvaliteedi raamistike kohta käiv teoreetiline taust on leitav kolmandast peatükist. Ettevõtte X andmebaasi mudel koos tabelite vabatekstilise kirjeldusega esitatakse neljandas peatükis. Andmekvaliteedi juhtimise raamistiku rakendamist ettevõttes X käsitletakse viiendas peatükis. See on kokkuvõtlik peatükk. Rakendamise olulisi tulemusi käsitletakse töö teistes peatükkides. Kuuendas peatükis tuuakse välja ettevõtte X andmebaasi andmekvaliteedi nõuded ja analüüsitakse nende täidetust ettevõtte andmebaasis. Andmete haldamise ja andmekvaliteedi tagamise protsessidest ning andmekvaliteedi nõuete lisamise ja muutmise protsessidest ettevõttes X kirjutatakse seitsmendas peatükis ning andmekvaliteedi hindamise töölaua prototüüpi esitletakse kaheksandas peatükis. Töö tulemuste analüüs ja vastavad järeldused esitatakse üheksandas peatükis.

## 2 Metoodika

Selles peatükis esitatakse ülevaade töös kasutatud metoodikast.

### 2.1 Ülevaade objektist

Töö aluseks on ettevõtte X veebisait, läbi mille toimub onlain müük Eestis, Lätis ja Leedus. Veebisaidil saavad kasutajad osta erinevaid kaupu, tooteid ja teenuseid. Lisaks on võimalik veebisaidil lugeda ettevõtte ajaveebi postitusi ja osaleda erinevates kampaaniates.

Ettevõtte kasutab erinevates riikides müümiseks ühte ja sama veebipoodi, mis on ehitatud WordPress platvormile. Veebipoodi on võimalik kasutada eesti, inglise, läti, leedu ja vene keeles. Andmete hoiustamiseks kasutatakse ühte SQL-andmebaasi, mis on loodud MySQL andmebaasisüsteemi põhjal.

Veebisaidi vajaliku funktsionaalsuse saavutamiseks kasutatakse lisaks WordPress võimalustele veel ka pistikprogramme, millest igaüks vastutab kindla funktsionaalsuse täitmise eest. Näiteks e-poe kataloogi loomiseks ja tellimuste haldamiseks kasutatakse pistikprogrammi WooCommerce. Ettevõtte andmebaasi struktuur on ära määratud WordPressi ja selle pistikprogrammide poolt, st andmebaasis ei ole ettevõtte enda töötajate poolt tehtud struktuurimuudatusi.

Kaupu, tooteid ja teenuseid lisavad ettevõtte töötajad e-poe kataloogi käsitsi. Kõik ettevõtte töötajad pääsevad veebisaidile ligi vastavalt tööülesannetele, et uurida tellimuste statistikat, lisada tooteid/teenuseid/kaupu ja täiendada ajaveebi postitusi või veebisaidil esitatud muud informatsiooni.

Magistritöö raames võttis autor vaatluse alla veebisaidi e-poe osa ehk kaupade, toodete ja teenuse kataloogi ning klientide võimaluse registreerida end e-poe kasutajaks ja esitada tellimusi või küsida hinnapakumisi.

Töös ei tõmmata selget piiri andmete ja informatsiooni ja käsitletakse neid sünonüümidena. Mis ühele on töötlemist vajavad lähteandmed - märkide hulgas mingis

keskkonnas – see võib teisele olla kasulik sisendinfo, mis võimaldab tal midagi järeldada või otsustada.

## 2.2 Tööriistade kirjeldus

Materjalide otsimiseks kasutati otsingukeskkondi Google ja Google Scholar.

Töö käigus loodavate dokumentide hoidmiseks kasutati Microsoft OneDrive'i [26]. Dokumentatsiooni ja lõputöö kirjutamiseks kasutati onlain Word dokumenti, et magistritöö juhendajal ja ettevõtte juhatusel oleks ligipääs dokumentatsioonile ja võimalust jätta parandusi ja kommentaare. Lisaks kasutati tabelite ja jooniste tegemiseks Exceli töövihikut.

Suhtlemiseks juhendajaga kasutati e-kirju ja videokoosolekute tegemiseks Microsoft Teamsi [27]. Ettevõtte töötajate ja autori vaheline suhtlemine toimus enamasti läbi Slack keskkonna [28].

Lokaalse andmebaasi loomiseks töö autori arvutis kasutati Docker Desktop rakendust [29], milles loodi kaks konteinerit – phpMyAdmin [30] ja MySQL [31] jaoks.

PhpMyAdmin konteinerit kasutati selleks, et oleks võimalik vaadelda ettevõtte X andmebaasis olevaid tabeleid ja andmeid [32]. PhpMyAdminis käivitas autor SQL-päringuid, et leida vastuseid püstitatud küsimustele. Neid vastuseid läks vaja andmekvaliteedi vigade leidmiseks, andmebaasi struktuurist arusaamiseks ja andmekvaliteedi töölaual kasutavate andmekontrolli päringute koostamiseks.

Ettevõtte X andmekvaliteedi parendamiseks kasutati Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistiku, mis loodi Riigi Infosüsteemi Ameti projekti „Andmekvaliteedi tagamise juhend“ raames [2]. Rohkem infot andmekvaliteedi raamistiku kohta esitatakse jaotises „3.3 Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistik“.

Andmekvaliteedi tagamise protsesside modelleerimiseks kasutati UML keelt ja Enterprise Architect [33] (ver 12) CASE vahendit.

Andmebaasi disaini kontrollimiseks kasutati Tallinna Tehnikaülikooli õppeaine „Andmebaasid II“ jaoks loodud kontrollpäringuid [34]. Kuna päringud on mõeldud

PostgreSQL andmebaasisüsteemile ja ettevõttes on andmebaas loodud MySQL andmebaasisüsteemis, siis tuli autoril teisendada ettevõtte andmebaas PostgreSQL. Esialgu laaditi PhpMyAdminist alla tabelite loomise SQL-laused (ilma andmete tabelitesse lisamise lauseteta). Selleks kasutati rakendust Sequel Ace [35]. Järgmise sammuna võeti kasutusele RazorSQL [36], milles kasutades MySQL andmebaasi teiseldamistöööriista [37] teisendati MySQL jaoks mõeldud SQL-laused PostgreSQL andmebaasi tabelite loomise lauseteks. Saadud laused käivitati ülikooli pakutavas serveris ning nende abil loodud andmebaasis käivitati andmebaasi disaini kontrollpäringud.

Andmekvaliteedi töölaua prototüübi loomiseks kasutati JetBrains PhpStorm rakendust [38] ja PHP programmeerimiskeelt [39]. Töölaua prototüübi struktuuri loomiseks kasutati HTMLi [40] ja kujundamiseks CSSi [41].

Andmebaasi disaini kirjeldava mudeli, andmekvaliteedi raamistiku ja selle rakendandamise käigus loodud dokumentatsiooni hoiustamiseks ettevõtte siseselt kasutati ettevõtte X wikit, mis esialgu loodi Microsoft OneNote'is [42] ja edaspidi plaanitakse üle viia MediaWiki platvormile [43]. Ettevõtte X wiki struktuur Microsoft OneNote'is on leitav Lisas 2. Ettevõtte töötajad kasutavad nõuetele mittevastavate andmete registreerimiseks Trello rakendust [44], milles on võimalik luua ja jälgida pileteid.

### **2.3 Tööprotsessi kirjeldus**

Välja toodud probleemide lahendamiseks ja püsitatud uurimisküsimustele vastutuste leidmiseks alustas töö autor andmekvaliteedi ja andmekvaliteedi raamistike kohta käivate artiklite otsimisest ja lugemisest. Lisaks artiklitele pööras autor teemast parema ülevaate saamiseks tähelepanu ka teemaga seotud postitustele foorumites ja andmekvaliteediga tegelevate ettevõtete veebilehtedel olevale informatsioonile.

Autor kasutas tööks lokaalset koopiat ettevõtte andmebaasist, kus olid ettevõtte reaalsed andmed. Andmebaasile mugavaks ligipääsu saamiseks võttis autor kasutusele Docker Desktop tarkvara, milles ta lõi konteinerid MySQL ja PhpMyAdmin programmide jaoks. Konteinerite installeerimisel tegi autor muudatusi vaikimisi sätetes, et oleks võimalik ettevõtte andmebaasi üleslaadimine koopias (vaikimisi oli üleslaaditavate failide suurus liiga väike) ja võimalus edaspidi lokaalselt avada andmebaasi kasutades PhpMyAdmini ja Sequel Ace rakendusi.

Autor otsustas oma töös käsitleda ettevõtte andmebaasi 52 tabelit, mis on seotud ettevõttes kõige enam kasutatavate pistikprogrammidega, millest omakorda sõltub kaupade, toodete ja teenuste müük. Järgmise sammuna jaotati andmebaas alamosadeks e registriteks, et paremini aru saada andmebaasi struktuurist ja paremini hallata andmekvaliteedi uurimisega seotud keerukust. Andmebaas jaotati kümneks registriks. Töö käigus analüüsiti tabeleid ja nende dokumentatsioon lisati ettevõtte wikisse, tehes selle kaudu dokumentatsiooni kõigile ettevõtte töötajatele kättesaadavaks.

Analüüsi käigus uuriti andmebaasis sisalduvaid andmeid ja nende vastavust ettevõtte ootustele ja reeglitele seoses klientidele antava informatsiooniga. Andmekvaliteedi probleemide kiiremaks tuvastamiseks loodi SQL-päringuid, et leida reeglitele mittevastavaid andmeid. Edaspidi kasutati neid SQL-päringuid andmekvaliteedi hindamistöölaua prototüübis.

Peale andmebaasi analüüsimist ja dokumenteerimist valiti ettevõtte vajadustele sobiv andmekvaliteedi raamistik ja kohandati seda, et sobitada see väikeettevõtte vajaduste ja võimalustega. Andmekvaliteedi raamistiku rakendamisel hinnati ettevõtte X andmebaasi andmekvaliteeti, tehti kindlaks olemasolev küpsustase ja määrati soovitud küpsustase. Seejärel planeeriti ja viidi läbi tegevused, mida oli vaja teha, et jõuda olemasolevalt küpsustasemelt eesmärgiks seatud tasemele.

Andmekvaliteedi küpsustaseme tõstmiseks koostati andmekvaliteedi nõuded, mida edaspidi kasutati andmekvaliteedi hindamise töölaua prototüübi loomisel. Mitmete nõuete puhul saab koostada SQL-päringuid, millega saab otsida nõuetele mittevastavaid andmeid. Andmekvaliteedi nõuete koostamisega paralleelselt loodi andmekvaliteedi profiilid, mida ettevõtte töötajad saavad kasutada kaupade, toodete ja teenuste andmete kogumisel ja töötlemisel, et edaspidi need kaubad, tooted ja teenused e-poodi lisada ja avalikustada.

Andmekvaliteedi kontrollimiseks loodi andmekvaliteedi hindamise töölaua prototüüp JetBrains PhpStorm rakendust kasutades. Loodud prototüüpi saab ettevõtte X edaspidi kasutada, et leida kiiremini probleemseid ja kvaliteedinõuetele mittevastavaid andmeid.

Enne prototüübi loomist visandati paberile kasutajaliidese kavand – millised leheküljed võiksid prototüübis olla ning milline informatsioon ja millisel kujul esinema peaks (v.t. Lisa 3). Edaspidi loodi vastavad .php, .css ja .html failid.

Prototüübi valideerimiseks tehti A/B testimine ettevõtte töötajatega, kellel võib oma tööülesannete täitmiseks sellest töölauast kasu olla. Testimine toimus etappidena [25], [45], [46]:

- Etapp 1: Määrata testimise hüpotees. Selle etapi käigus esitati viis hüpoteesi:
- Õigsuse andmekvaliteedi nõuete tulemuste leidmine ei vaja rohkem kui kahte klikki.
- Täielikkuse andmekvaliteedi nõuete tulemuste leidmine ei vaja rohkem kui kahte klikki.
- Kooskõla andmekvaliteedi nõuete tulemuste leidmine ei vaja rohkem kui kahte klikki.
- Konfidentsiaalsuse andmekvaliteedi nõuete tulemuste leidmine ei vaja rohkem kui kahte klikki.
- Töötajad mõistavad kergesti, milline andmekvaliteedi nõue on täidetud ja milline ei ole.
- Etapp 2: Koostada testimisülesanded. Selle etapi käigus loodi ülesanded, mis jagati ettevõtte töötajatele A/B testimise ajal e kolmanda testimise etapi ajal. Igal kaustajal paluti täita kolm ülesannet, mis olid sarnased järgmistele ülesannetele.
- Palun too välja kõik eestikeelsed tootekaardid, millel puudub hind.
- Palun leia üles kolm eestikeelset tootekaarti, millel lao staatus ja lao kogus ei klapi omavahel.
- Palun leia üles üks eestikeelne tootekaart, millel puudub metakirjeldus ja lisa see.
- Etapp 3: Testida varasemat ja uut lahendust. Varasem lahendus on see lahendus, kus töölauda ei olnud ja ettevõtte töötajatel tuli andmekvaliteedi probleemid käsitsi üles otsida. Uus lahendus on andmekvaliteedi hindamistöölaua prototüübi kasutamine.
- Etapp 4: Küsida tagasisidet.
- Etapp 5: Analüüsida testi tulemusi.

Selgitamaks välja, kas andmebaasi skeemi probleemid võivad tingida andmekvaliteedi probleeme, analüüsiti andmebaasi disainivigu. Esimese tegevusena laaditi alla andmebaasi tabelite koostamise SQL-laused MySQL jaoks. Lauseid, mis lisaksid andmeid tabelitesse ei laaditud alla, kuna eesmärk oli analüüsida andmebaasi struktuuri, mitte tabelite sees olevaid andmeid. Järgmise tegevusena teisaldati saadud SQL-laused



PostgreSQL lauseteks, et edasi käivitada neid ülikooli õppeserveril, millel omakorda hiljem käivitati andmebaasi disaini kontrollpäringud.

## 3 Teoreetiline taust

Käesolevas peatükis esitatakse andmekvaliteedi, andmekvaliteedi raamistike ja töös rakendatava Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistiku teoreetiline taust.

### 3.1 Andmekvaliteet

Selles jaotises tuuakse välja andmekvaliteedi mõiste, selle tähtsus organisatsioonile, andmekvaliteedi iseloomustamiseks kasutatavad mõõdikud ning millised on andmekvaliteedi tagamiseks mõeldud raamistikud.

#### 3.1.1 Andmekvaliteedi üldine mõiste

Andmekvaliteet on suhteline mõiste [47], mille määratlus sõltub vaadeldavatest eesmärkidest [48] ja mida seletatakse läbi teiste mõistete ning definitsioonide.

Üldisteks kirjanduses kasutatavateks definitsioonideks andmekvaliteedile on „kasutamiseks sobiv“ ja „lõppkasutaja vajaduste rahuldamiseks sobiv“. Andmeid loetakse kvaliteetseteks siis, kui need sobivad rakenduste, otsuste tegemise ja planeerimise jaoks. Sobivuse määramiseks kasutatakse erinevaid andmekvaliteedi mõõdikuid, mis on üksteisega sõltuvuses, st nende vahel on otsene või kaudne seos. Järgnev nimekiri on kõigest väike valik andmekvaliteedi mõõdikutest – täpsus (*accuracy*), täielikus (*completeness*), värskendatavus (*update*), järjepidevus (*consistency*), kättesaadavus (*availability*), õigeaegsus (*timeliness*) ja lisandväärtus (*value added*) [48].

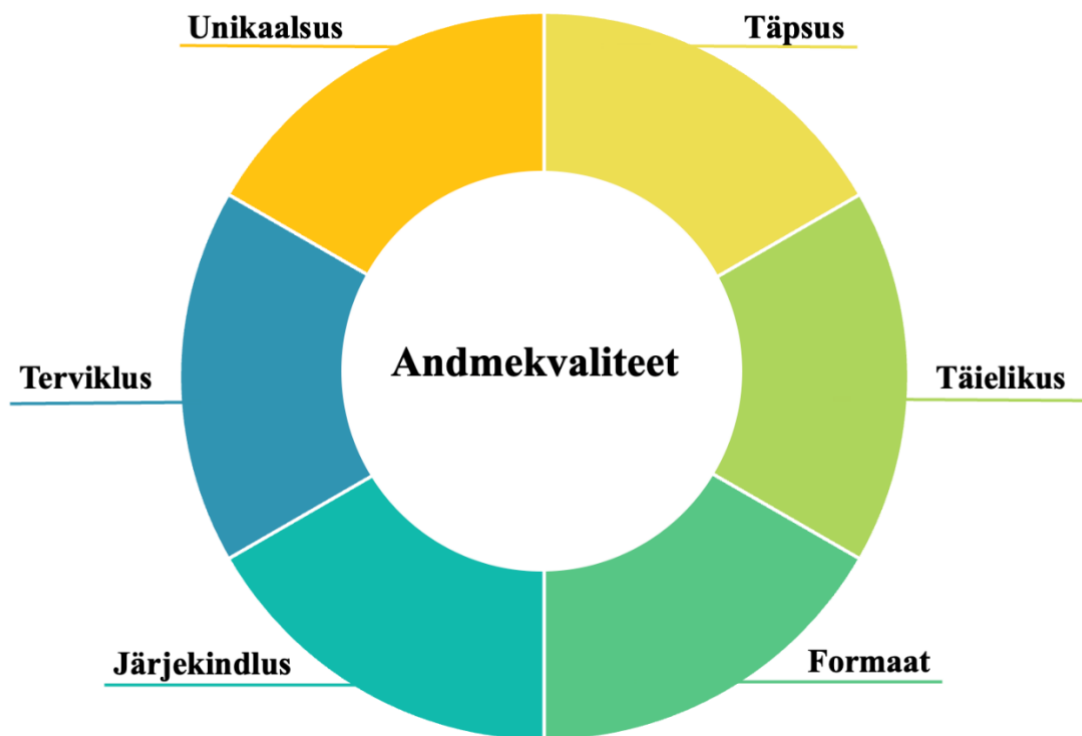
Giri Kumar Tayi ja Donald P. Ballou [47] defineerivad artiklis andmekvaliteedi mõiste läbi mõiste „kasutuskõlblikus“, mis rõhutab seda, et andmed peavad olema andmeid koguva üksuse (organisatsioon või eraisik) kontekstis kasutamiseks sobivad. Kasutuskõlblikkuse definitsioon võib üksuste vahel ja ka organisatsiooni alamüksuste vahel erineda ja seetõttu on andmeid kogaval üksusel tähtis määrata kvaliteet vastavalt vajadustele ning mõelda tavaprobleemidest kaugemale [47].

Howard Veregin seletab andmekvaliteeti kui immateriaalsete omaduste funktsiooni. Põhjus on selles, et võrreldes toodetud kaubaga ei ole andmetel füüsikalisi omadusi, mille

järgi saab määrata kvaliteeti. Immateriaalsete omadustena nimetab ta usaldusväarsust, täielikkust ja järjepidevust. [49]

Li Cai ja Yangyong Zhu [50] seletavad andmekvaliteeti kui „kasutuskõlblikust“, „nõuetele vastavust“ ja „astet, milles olemuslike omaduste kogum vastab nõuetele“.

TechTargeti järgi mõistetakse andmekvaliteedi hinnangu all kogumit mõõdikute väärtuseid, mis näitavad andmete seisukorda [51]. Kvaliteedi määramisel kontrollitakse, kas andmed vastavad ettenähtud kuuele mõõdikule – täpsus, täielikkus, formaat, järjekindlus, terviklus ja unikaalsus [51], [52]. Joonisel 1 on illustratiivne näidis andmekvaliteedi mõõdikute kohta.



Joonis 1. Andmekvaliteedi mõõdikud [53].

Kokkuvõttes saab öelda, et allikates nagu [51], [52], [53] defineeritakse andmekvaliteeti kui andmete vastavust teatud hulga aspektidele, kusjuures erinevates artiklites on see aspektide hulk erinev. Täpsemalt kirjutatakse andmekvaliteedi aspektidest jaotises „3.1.3 Andmekvaliteedi aspektid“.

### 3.1.2 Andmekvaliteedi tähtsus organisatsioonidele

Andmekvaliteet on organisatsioonidele tähtis, sest organisatsiooni huvitatud osapooled (nt omanikud, töötajad, kaastöölised ja kliendid) võtavad vastu otsuseid tuginedes organisatsiooni käsutuses olevatele andmetele.

Organisatsioon saab oma töös eeliseid, sh ettevõtted konkurentsieeliseid, kui selle andmekvaliteet on heal tasemel. Näiteks hoides andmekvaliteeti heal tasemel on organisatsioonil võimalik vähendada täiendavatele tegevustele tehtuid kulutusi ja suurendada müüki (õiged andmed => täpsemate tulemustega ärianalüütika => paremad äriotsused) [51]. Andmete kvaliteedi tagamine on kulu, kuid õigel ajal õiges kohas kasutada olnud andmete alusel tehtud õiged otsused on tulu, mis aitab neid kulusid kompenseerida.

Tehes otsuseid tuginedes halvale andmekvaliteedile võib organisatsioon end kahjustada [52], [54]. Halva kvaliteediga andmed võivad vähendada klientide ja töötajate rahulolu ning tekitada organisatsioonis usaldamatust [54].

### 3.1.3 Andmekvaliteedi aspektid

Andmed on kvaliteetsed kui need on tööülesannete täitmisel, sh otsustamisel, kasutamiseks sobivad [54]. Kvaliteetseteks andmeteks nimetakse neid andmeid, mis vastavad standarditele ja on kasutuskõlblikud [54]. ISO/IEC 25012 standardi järgi on andmed kvaliteetsed, kui need vastavad organisatsiooni poolt määratletud nõuetele [55].

Tabelis 1 on kokkuvõtvalt välja toodud autori poolt loetud artiklites nimetatud andmekvaliteedi aspektid koos lühikirjelduste ja nende esinemiste arvuga artiklites. Tabelis on read sorteeritud artiklites esinemiste arvu järgi kahanevalt ja selle sees ingliskeelse nime järgi kasvavalt.

Tabel 1. Andmekvaliteedi aspektid kirjanduse alusel.

Andmekvaliteedi aspekt	Lühikirjeldus	Viitavad artiklid	Art. arv
Täielikkus ( <i>Completeness</i> )	[54] järgi nimetatakse andmebaase täielikuks, kui ükski kirje neist ei puudu ja ühelgi neist kirjetest ei ole puuduvaid andmelemente. ISO 25012 standardi [55] järgi mõistetakse andmete täielikkuse all määra, mil andmetel on	[48], [50], [51], [52], [54], [55], [56], [57]	8

<b>Andmekvaliteedi aspekt</b>	<b>Lühikirjeldus</b>	<b>Viitavad artiklid</b>	<b>Art. arv</b>
	olemas kõik eeldatavad atribuudid ja väärtused konkreetses kasutuskontekstis. Corinna Cichy ja Stefan Rass oma artiklis [56] kirjeldavad andmete täielikkust määrana, kuivõrd andmed on konkreetse kasutuskonteksti jaoks piisava ulatuse, sügavuse ja ulatusega.		
Järjepidevus ( <i>Consistency</i> )	Erinevates andmekogumites samade olemite ja seoste kohta esitatud faktide vahel puuduvad konfliktid [51]. Näiteks kui ühes andmekogumis on kliendi sünnipäevaks märgitud „01. juuni“ ja teises andmekogumis „01. juuli“, siis on tegemist andmete konfliktiga [58]. Järjepidev esitus näitab kuivõrd on andmed esitatud ühesuguses formaadis [57].	[48], [50], [51], [52], [55], [56], [57]	7
Õigeaegsus ( <i>Timeliness</i> )	Andmeid puudutava teabe ajakohasus [54]. Näiteks võidakse hinnata, et teave, mis on kogutud viimase kahe tunni jooksul ja millele pole uuendusi tulnud, õigeaegseks [59].	[48], [50], [51], [52], [54], [56] [57]	7
Täpsus ( <i>Accuracy</i> )	Andmete täpsust puudutavad küsimused. Iga andmete valdaja (isik/organisatsioon) peab enda jaoks ise määrama, millised on talle peamised huvipakkuvad andmed ja kui täpsed need andmed peavad olema. Näiteks kui vaadata e-kaubandusettevõtet, siis ettevõtte jaoks on tähtsad müüdavad kaubad ja nende kaupade kohta käiv müügi informatsioon – näiteks kui mitu korda seda kaupa osteti. [54].	[48], [50], [52], [54], [55], [56]	6
Kättesaadavus ( <i>Accessibility</i> )	Allikas [55] defineerib kättesaadavust kui ligipääsetavuse taset ehk millisel määral on andmed konkreetses kasutuskontekstis kättesaadavad. Eelkõige rõhutab allikas [55] andmete kättesaadavust puuetega inimestele. Allika [56] järgi on kättesaadavus määr, mis näitab, kui kiiresti ja kergesti on andmed leitavad.	[50], [54], [55], [56], [57]	5
Kättesaadavus ( <i>Availability</i> )	Allika [48] järgi määrab kättesaadavus, kui kiiresti ja kergesti on andmed leitavad. Allikas [50] on ühendanud kättesaadavuseks ( <i>availability</i> ) kaks teist andmekvaliteedi	[48], [50], [55]	3

Andmekvaliteedi aspekt	Lühikirjeldus	Viitavad artiklid	Art. arv
	aspekti – ligipääsetavus ( <i>accessibility</i> ) ja õigeaegsus ( <i>timeliness</i> ).		
Usaldusväarsus (Credibility/ Believability)	Allika [55] kohaselt tähendab see, et kasutajad peavad andmeid tõeseks ja usutavaks konkreetses kasutuskontekstis. Allika [50] järgi kasutatakse usaldusväarsust, kui andmeid on vaja hinnata mittedumbriliselt. Näiteks on andmed usaldusväärased, kui nad [50]: <ul style="list-style-type: none"> <li>▪ pärinevad riigi, valdkonna või tööstusharu organisatsioonidelt,</li> <li>▪ on ekspertide või auditite abil regulaarselt kontrollitavad,</li> <li>▪ on teadaolevate või vastuvõetavate väärtuste vahemikus.</li> </ul>	[50], [55], [57]	3
Ajakohasus (Currentness/ Relevance)	Andmete vastavus põhivajadustele (kas vastab sellele, milleks andmeid koguti), andmete kasutuse võimalus täiendavatel eesmärkidel (kas neid andmeid saab kasutada millekski veel), kas andmebaasi saab kasutada teistel eesmärkidel (andmebaasi teisene kasutamine) [54].	[54], [55], [57]	3
Asjakohane andmete hulk ( <i>Appropriate amount of data</i> )	Näitab kuivõrd andmete hulk on ülesande jaoks sobiv [57].	[56], [57]	2
Arusaadavus (Understandability)	Arusaadavus näitab, millisel määral on andmetel omadused, mis võimaldavad neid kasutajate poolt lugeda ja tõlgendada ning mis on väljendatud konkreetses kasutuskontekstis sobivas keeles, sümbolites ja ühikutes [55].	[55], [57]	2
Unikaalsus ( <i>Uniqueness</i> )	Andmekogumites puuduvad üksteist dubleerivad andmekirjed [51].	[55], [57],	2
Kehtivus ( <i>Validity</i> )	Allika [51] kohaselt tähendab see, et andmed on õigesti struktureeritud ja sisaldavad õigeid väärtusi. Allikas [52] märgib, et kehtivad andmed peavad olema esialgu korjatud õigesti vastavalt ärireeglitele ning need peavad vastama vormingule ja kuuluma õigesse vahemikku.	[51], [52]	2

<b>Andmekvaliteedi aspekt</b>	<b>Lühikirjeldus</b>	<b>Viitavad artiklid</b>	<b>Art. arv</b>
Lisandväärtus ( <i>Value added</i> )	Lisandväärtus ehk kasumäär [48], mis näitab millisel määral on andmed kasulikud [57].	[48], [57]	2
Sidusus ( <i>Coherence</i> )	Andmete loogilised ühendused andmekogumi sees või andmekogumite vahel [60].	[54]	1
Võrreldavus ( <i>Comparability</i> )	Andmeid saab lugeda võrreldavateks siis, kui nende kasutajad saavad neid kasutada vaatamata eesmärkide erinevusele [61].	[54]	1
Vastavus ( <i>Compliance</i> )	Vastavus näitab, millisel määral vastavad andmed oma omaduste poolest kehtivatele eeskirjadele ja standarditele [55].	[55]	1
Konfidentsiaalsus ( <i>Confidentiality</i> )	Konfidentsiaalsus näitab, millisel määral on tagatud, et andmed on konkreetses kontekstis kättesaadavad ja tõlgendatavad ainult volitatud kasutajatele [55].	[55]	1
Sisutihedus ( <i>Concise representation</i> )	Sisutihedus näitab kuivõrd kompaktselt on andmed esitatud [57].	[57]	1
Töötlemise lihtsus ( <i>Ease of Manipulation</i> )	Töötlemise lihtsus näitab kuivõrd lihtne on andmeid erinevate ülesannete puhul käsitleda [57].	[57]	1
Efektiivsus ( <i>Efficiency</i> )	Efektiivsus näitab, millisel määral on andmetel omadusi, mis aitavad neid töödelda ja mis tagavad piisava töötlemise jõudluse arvestades kasutuskontekstist tulenevate ressursside mahuga [55].	[55]	1
Sobivus ( <i>Fitness</i> )	Sobivus näitab, millisel määral vastavad andmed kasutajate vajadustele [50].	[50]	1
Veavaba (Free-of-Error)	Veavabadus näitab kuivõrd usaldusväärsed ja õiged on andmed [57].	[57]	1
Terviklikkus ( <i>Integrity</i> )	Allika [50] järgi on terviklikeks andmeteks need, millel on terviklik struktuur ehk andmeväärtused on standardiseeritud vastavalt andmemudelile ja/või andmetüübile (andmete vorming on selge ja vastab kriteeriumitele).	[50]	1
Tõlgendatavus ( <i>Interpretability</i> )	Tõlgendatavus näitab kuivõrd esitatud andmed on esitatud asjakohastes keeltes, kasutades sobivaid sümboleid ja ühikuid [57].	[57]	1
Objektiivsus ( <i>Objectivity</i> )	Objektiivsus näitab kuivõrd andmed on eelarvamuste vabad ja erapooletud [57].	[57]	1

<b>Andmekvaliteedi aspekt</b>	<b>Lühikirjeldus</b>	<b>Viitavad artiklid</b>	<b>Art. arv</b>
Kaasaskantavus ( <i>Portability</i> )	Kaasaskantavus näitab, millisel määral on andmetel omadusi, mis võimaldavad neid olemasolevat kvaliteeti säilitades paigaldada, asendada või ühest süsteemist teise üle viia [55].	[55]	1
Täpsus ( <i>Precision</i> )	Täpsus näitab andmete omaduste täpsustaset [55].	[55]	1
Loetavus ( <i>Readability</i> )	Loetavus näitab, kas andmed on selged ja arusaadavad, kas neid on lihtne hinnata ja need on kergesti mõistetavad [50].	[50]	1
Taastatavus ( <i>Recoverability</i> )	Taastatavus näitab, millisel määral on andmetel omadused, mis võimaldavad neil konkreetses kasutuskontekstis rikke korral säilida ja tagada selle kaudu kindlaksmääratud toimingute ja kvaliteedi taseme [55].	[55]	1
Maine (Reputation)	Maine näitab, millisel määral hinnatakse andmeid nende allika või sisu poolest [57].	[57]	1
Turvalisus ( <i>Security</i> )	Turvalisus näitab kuiõrd on juurdepääs andmetele piiratud [57].	[57]	1
Jälgitavus ( <i>Traceability</i> )	Jälgitavus näitab, millisel määral on andmetel omadusi, mis tagavad andmetele juurdepääsu ja andmete muutmise kohta kontrolljälje säilimise konkreetses kasutuskontekstis [55].	[55]	1
Värskendatavus ( <i>Update</i> )	Värskendatavus näitab kui kiiresti andmeid värskendatakse [48].	[48]	1

Tabelis 1 on toodud välja 33 andmekvaliteedi aspekti kaheksast erinevast allikast, mis puudutavad andmekvaliteeti ja andmekvaliteedi raamistikke. Võrdluseks kasutas Riigi Infosüsteemi Amet (RIA) 12 allikat, et korjata kokku 28 andmekvaliteedi aspekti [1]. Autori kokku koondatud aspektid langesid enamasti kokku RIA välja toodud andmekvaliteedi aspektidega. 22 andmekvaliteedi aspekti langesid kokku RIA andmekvaliteedi uuringus väljatoodud aspektidega. Sellist kattuvust toetas kolme ühise allika kasutamine [55], [50], [57]. Andmekvaliteedi aspektid, mis esinesid RIA uuringus, kuid jäid välja autori uuringust, on detailsus, vormisobivus ja vormipaindlikus. Aspektid, mis olid autori uuringutes, kuid ei esinenud RIA uuringutes, on lisandväärtus, tõlgendatavus ja värskendatavus.



Kokkuvõttes võib öelda, et andmekvaliteedi aspekte on palju ja neid kirjeldatakse erinevates allikates. Mõnikord kirjeldatakse erinevates allikates sama nimega aspekti erinevalt ja mõnikord kasutatakse erinevates allikates sisuliselt samatähenduslikule aspektile viitamiseks erinevaid nimesid.

### 3.2 Andmekvaliteedi raamistikud

Andmekvaliteedi probleemide tähelepanuta jätmise võib tekitada ärile kahju ning mõningate selliste probleemide lahendamata jätmise võib organisatsioonidele pikas plaanis maksta minna miljoneid või miljardeid dollareid [62]. Selle probleemi lahendamiseks on juba pikemalt tegeletud andmekvaliteedi raamistike/tagamise ja parandamise meetodite väljatöötamisega [62]. Seega võib täheldada, et erinevaid andmekvaliteeti parendamiseks mõeldud raamistike on töö kirjutamise ajaks (2023. aasta kevad) välja pakutud üpris palju.

Allikas [56] tuuakse välja 25 andmekvaliteedi raamistikku, millest 12 on teoreetiliselt võimalik kohandada mistahes valdkonna andmete suhtes. Neid saab rakendada kõigi andmekvaliteedi juhtimise aspektides. Ülejäänud 13 andmekvaliteedi raamistikku on suunatud mingi kindla valdkonna andmekvaliteedi probleemide lahendamiseks (näiteks tootmisandmete ja tervishoiu andmete kvaliteediga seotud andmekvaliteedi raamistikud AMEQ ja CIHI). Allika [56] autorid rõhutavad oma töös eelkõige 12 andmekvaliteedi raamistikku, mis sisaldavad andmete kvaliteedi määramist, andmete kvaliteedi hindamist ja andmete kvaliteedi parandamise tegevusi. Tabelis 2 on lühidalt toodud välja üldotstarbeliste e mistahes valdkonnale rakendatavad andmekvaliteedi raamistike kirjeldused. Tabel on sorteeritud akronüümide järgi tähestiku järjekorras.

Tabel 2. Mistahes valdkonnale rakendatavad andmekvaliteedi raamistikud.

Akronüüm	Nimetus	Lühikirjeldus
AIMQ	A Methodology for Information Quality Assessment	Metoodika, mis hindab andmete kvaliteeti organisatsioonis ja võrdleb seda teise organisatsioonidega [63]. AIMQ metoodikat on rakendatud finants-, tervishoiu- ja toomisvaldkonnas. See metoodika on kasulik andmekvaliteedi probleemide tuvastamisel, andmekvaliteedi probleemide parandamise valdkondade tähtsuse järjekorra määramisel ja andmekvaliteedi parendamise jälgimisel [63].

Akronüüm	Nimetus	Lühikirjeldus
		AIMQ meetodika aga ei ole nii kasulik, et selle põhjal teha otsuseid andmekvaliteedi parandamiseks [63].
CDQ	Comprehensive Methodology for Data Quality Management	<p>Metoodika, mis pakub tuge optimaalse andmekvaliteedi parandamise protsessi valimisel. Selle meetodika eesmärk on saada hinnang, kuivõrd mõjutavad halvad andmed organisatsiooni äriprotsesse. [64]</p> <p>CDQ raamistiku rakendamisel on kolm faasi [64]:</p> <ol style="list-style-type: none"> <li>1. Oleku rekonstrueerimine, mille väljundiks on organisatsiooni kirjeldus koos andmete, teabevoogude ja protsessidega (toimub organisatsiooni nõuete väljaselgitamine).</li> <li>2. Hindamine, mille käigus analüüsitakse andmeid ja andmekvaliteedi nõudeid ning mõõdetakse kvaliteeti. Hindamise väljundiks on andmekvaliteeti iseloomustavad väärtused.</li> <li>3. Parandamine, mille käigus tehakse selgeks, mis olid andmekvaliteedi probleemide põhjused ja võetakse vastu otsus, milliseid strateegiaid ja tehnikaid organisatsioonis andmekvaliteedi parandamiseks kasutada.</li> </ol>
COLDQ	Cost-effect of Low Data Quality	<p>COLDQ pakub hindamiskaarti, mille kasutuselevõtt aitab organisatsioonil vältida kulusid, mis on seotud andmete halva kvaliteediga [64].</p> <p>COLDQ raamistiku rakendamisel on viis faasi [64]:</p> <ol style="list-style-type: none"> <li>1. Andmekvaliteedi keskkonna loomine ehk andmete ja andmekvaliteedi nõuete analüüs.</li> <li>2. Määratlemine, mille käigus modelleeritakse protsesse ja määratakse organisatsiooni kriitilised valdkonnad.</li> <li>3. Andmekvaliteedi mõõtmine.</li> <li>4. Andmekvaliteedi vigade põhjuste analüüs.</li> <li>5. Parandamine, mille käigus toimub andmekvaliteedi protsessi rakendamine, protsessi kontroll ja parandamise jälgimine.</li> </ol> <p>Sarnaselt CDQ raamistikule on ka COLDQ eesmärk saada hinnang selle kohta, kuidas halvad andmed mõjutavad organisatsiooni äriprotsesse [64].</p>
DQA	Data Quality Assessment	<p>DQA on andmekvaliteedi raamistik, mis põhineb järgmisel kolmel faasi läbimisel [64]:</p> <ol style="list-style-type: none"> <li>1. Subjektiivne ja objektiivne andmekvaliteedi mõõtmine.</li> <li>2. Tehtud kvaliteedi mõõtmiste tulemuste võrdlus.</li> </ol>

Akronüüm	Nimetus	Lühikirjeldus
		3. Parandamine, mille käigus tuvastatakse vigade põhjused ja valitakse sobiv strateegia andmekvaliteedi parandamiseks.
DQPA	A Data Quality Practical Approach	<p>DQPA meetodika põhineb äriküsimuste ja andmeallikate järjestamisel, et määrata kindlaks mõju ja vajalikud parandused. Meetod rõhutab andmete kvaliteediprobleemide ennetamise ja paranduste kulutõhususe tähtsust. [56]</p> <p>DQPA meetodika andmekvaliteedi hindamise etapid on [65] järgmised:</p> <ul style="list-style-type: none"> <li>▪ Andmete analüüs.</li> <li>▪ Andmekvaliteedi nõuete analüüs.</li> <li>▪ Kriitiliste valdkondade tuvastamine.</li> <li>▪ Andmekvaliteedi mõõtmine.</li> </ul> <p>DQPA meetodika kohaselt on andmekvaliteedi parandamise etapid [65] järgmised:</p> <ul style="list-style-type: none"> <li>▪ Kulude hindamine.</li> <li>▪ Tuvastada andmekvaliteeti mõjutavate vigade põhjused.</li> <li>▪ Protsessi ümberkujundamine.</li> </ul>
DQAF	Data Quality Assessment Framework	DQAF on meetodika, mis aitab luua organisatsioonis raamistiku, mis aitaks hinnata andmete kvaliteeti ühtselt ja toetudes standarditele (nii organisatsiooni enda püstitatud, kui ka maailmas kokku lepitutele). Meetodika on suunatud ka sellele, et täiustada andmete koostamise tavasid [65].
HDQM	A Data Quality Methodology for Heterogeneous Data	HDQM meetodika põhineb ressursside parandamistegevustel ja kulude-tulude analüüsil [56]. HDQM meetodika rakendamisel tuleb esialgu välja selgitada kõik organisatsioonis kasutatavad teabeallikad [66]. Selleks, et oleks võimalik kõigi allikate andmekvaliteeti üheselt hinnata, tuleb andmed viia ühisele kujule [66].
HIQM	Hybrid Information Quality Management	HIQM meetodika põhineb sellel, et vigade tuvastamine ja parandamine toimub andmekvaliteedi mõõtmise protsessi käigus. Võrreldes teiste andmekvaliteedi raamistikega, vaatab HIQM andmekvaliteedi ka kasutaja vaatepunktist. [67].
OODA DQ	The Observe – Orient – Decide – Act Methodology for Data Quality	OODA DQ meetodika põhineb tsüklil Jälgi-Orienteeru-Otsusta-Tegutse [56].

Akronüüm	Nimetus	Lühikirjeldus
		<ul style="list-style-type: none"> <li>▪ Orienteerumisfaasi käigus analüüsitakse algpõhjuseid ja hinnatakse andmekvaliteedi probleeme.</li> <li>▪ Otsustamisfaasi käigus võetakse vastu otsused, mis võivad ulatuda andmete puhastamisest kuni rakendussüsteemide muutmiseni.</li> <li>▪ Tegutsemise faasis viiakse määratud meetmed ellu.</li> </ul>
TBDQ	Task-Based Data Quality Method	<p>TBDQ raamistik on suunatud sellele, et lisada olemasolevate ülesannetele parandavaid tegevusi, mitte muuta olemasolevaid [62].</p> <p>TBDQ peamised komponendid on järgmised [56].</p> <ul style="list-style-type: none"> <li>▪ Andmekvaliteedi hindamine.</li> <li>▪ Arendus- ja parendustegevuste kavandamine.</li> <li>▪ Tehtud parenduste hindamine.</li> </ul> <p>Raamistik sobib kasutamiseks väiksemates ja keskmise suurusega organisatsioonides, kus suurt tähtsust omavad andmeid sisestavad töötajad [62].</p>
TDQM	Total Data Quality Management	<p>Metoodika põhineb tsüklil Määratle-Mõõda-Analüüsi-Paranda [56]. TDQM eesmärk on toetada andmete kvaliteedi parandamise protsessi alguses lõpuni ehk andmekvaliteedi nõuete analüüsist kuni nende nõuete rakendamiseni [64].</p>
TIQM/ TQdM	Total Information Quality Management	<p>TIQM e TQdM raamistik põhineb eesmärgil teha andmekvaliteedi parendustegevusi vaid siis, kui parendustegevustest tulenev kasu on suurem, kui parendustegevuste kulu [64]. Raamistiku rakendamise käigus määratletakse andmed, hinnatakse arhitektuuri kvaliteeti ja mõõdetakse kulusid.</p> <p>Metoodika etapid on andmekvaliteedi hindamine, andmekvaliteedi parendamine ja andmekvaliteedi parendamise juhtimine ning järelvalve [64].</p>

Tabelis 3 on toodud välja kokkuvõtte andmekvaliteedi aspektidele viitamisest üldotstarbeliste andmekvaliteedi raamistikutes. Tabeli lahtris on X, kui selles raamistikus viidatakse sellele aspektile. Tabelis on read sorteeritud raamistikutes viitamiste arvu järgi kahanevalt ja veerud on sorteeritud andmekvaliteedi raamiste akronüümide järgi tähestiku järjekorras.

Tabel 3. Andmekvaliteedi aspektide esinemine andmekvaliteedi raamistik.

Andmekvaliteedi raamistik	A I M Q	C D Q	C O L D Q	D Q A	D Q A F	D Q P A	H D Q M	H I Q M	O O D A D Q	T B D Q	T D Q M	T I Q M
Andmekvaliteedi aspekt												
<b>Täielikkus</b> ( <i>Completeness</i> )	X	X	X	X	X	X		X		X	X	X
<b>Õigeaegsus</b> ( <i>Timeliness</i> )	X		X	X	X	X		X		X	X	X
<b>Täpsus</b> ( <i>Accuracy</i> )		X	X			X	X	X		X	X	X
<b>Järjepidevus</b> ( <i>Consistency</i> )			X		X	X		X		X		
<b>Ajakohasus</b> (Currentness/Relevance)	X	X	X	X							X	
<b>Kättesaadavus</b> ( <i>Accessibility</i> )	X		X	X								X
<b>Turvalisus</b> ( <i>Security</i> )	X		X	X							X	
<b>Asjakohane andmete hulk</b> (Appropriate amount of data)	X			X							X	
<b>Sisutihedus</b> (Concise representation)	X			X							X	
<b>Usaldusväarsus</b> (Credibility/Believability)	X			X							X	
<b>Tõlgendatavus</b> ( <i>Interpretability</i> )	X			X							X	
<b>Objektiivsus</b> ( <i>Objectivity</i> )	X			X							X	
<b>Maine</b> (Reputation)	X			X							X	
<b>Arusaadavus</b> (Understandability)	X			X							X	
<b>Veavaba</b> (Free-of-Error)	X			X								
<b>Terviklikkus</b> ( <i>Integrity</i> )					X							X
<b>Unikaalsus</b> ( <i>Uniqueness</i> )						X						X
<b>Kehtivus</b> ( <i>Validity</i> )					X							X
<b>Lisandväärtus</b> ( <i>Value added</i> )				X							X	
<b>Töötlemise lihtsus</b> ( <i>Ease of Manipulation</i> )				X								
<b>Kaasaskantavus</b> ( <i>Portability</i> )			X									

Tabelist 3 jäeti välja 12 andmekvaliteedi aspekti (tabelist 1), mida allikate järgi valitud andmekvaliteedi raamistikud ei sisalda. 12 andmekvaliteedi aspekti, mis jäeti välja on

järgmised: Kättesaadavus (*Availability*), Sidusus (*Coherence*), Võrreldavus (*Comparability*), Vastavus (*Compliance*), Konfidentsiaalsus (*Confidentiality*), Efektiivsus (*Efficiency*), Sobivus (*Fitness*), Täpsus (*Precision*), Loetavus (*Readability*), Taastatavus (*Recoverability*), Jälgitavus (*Traceability*) ja Värskendatavus (*Update*).

### **3.3 Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistik**

Eesti e-riigi osaks olevate registreeritud infosüsteemide ja andmekogude arv on (2023. aasta kevade seisuga) juba üle 1300 [68] ja enamus nendest süsteemidest mõjutavad riigi toimimist [1]. Sellisteks süsteemideks on näiteks Maksu- ja Tolliametile kuuluvad TSD (tulu- ja sotsiaalmaksu süsteem), TÖR (töötamise register) ja Lubade süsteem [69]. Seega selleks, et riigi juhtimisel saaksid otsustajad võtta vastu täpseid otsuseid ja et riigi elanikkonnale oleks võimalik pakkuda mugavaid ja kättesaadavaid teenuseid, peavad nendes andmekogudes olevad andmed olema kvaliteetsed [1].

Tagamaks riigi andekogudes ühtlase andmekvaliteedi, käivitas Riigi Infosüsteemi Amet andmekvaliteedi tagamise juhendi loomise projekti. Selle raames loodi andmekvaliteedi tagamise juhend, testiti juhendi rakendatavaust ja vajadusel täiendati juhendit saadud tagasiside abil [1].

Projekti raames töötati välja andmekvaliteedi juhtimise raamistik, mis koosneb kolmest komponendist: andmekvaliteedi küpsusmudelist, andmekvaliteedi tunnuste komplektist ja andmekvaliteedi juhtimise raamistikust [1]. See raamistik põhineb pideva parendamise tsüklil, mis tähendab, et raamistiku rakendamine pole ühekordne akt, vaid raamistiku tegevusi tuleb pidevalt läbi viia.

RIA projekti raames loodud andmekvaliteedi juhtimise raamistik sisaldab viit tegevust, mis võimaldab ettevõtte andmekvaliteeti parendada ja juhtida andmekvaliteedi parendamise protsesse [12].

1. Andmekvaliteedi hindamine. Kõigepealt määratletakse andmete ulatus e skoop ehk tehakse selgeks andmete hulk, mille andmekvaliteedi soovitakse parandada. Järgmisena tzoneeritakse valitud andmete hulk ja määratakse iga tsoonile organisatsioonis vastutav töötaja, kes edaspidi vastutab selle tsooni andmekvaliteedi juhtimise eest. Peale tzoneerimist hinnatakse andmekvaliteedi

küpsustaset, mille käigus selgub milline on küpsustase kategooriate kaupa ja organisatsiooni andmekvaliteedi küpsustaseme koondhinnang. [12], [2].

2. Andmekvaliteedi parandamistegevuste planeerimine. Peale küpsustaseme väljaselgitamist tuleb organisatsioonil määrata küpsustase, milleni soovitakse jõuda ja koostada parendamise plaan [12].
3. Andmekvaliteedi parandustegevuste teostamine. Selles etapis toimub plaanis olevate parandamistegevuste tsüklite kaupa ellu viimine [12].
4. Parandustegevuste kontrollimine. Selle etapi käigus tehakse selgeks, kas saavutati sihiks seatud küpsustase või mitte. Küpsustaseme kontrollimiseks hinnatakse uuesti küpsustaset (v.t punkti „1. Andmekvaliteedi hindamine“) [12].
5. Andmekvaliteedi korrigeerimistööd. Peale parandustegevuste kontrollimist toimub selles etapis saavutatud muudatuste standardiseerimine. Juhul kui eelmise etapi käigus selgus, et mõni küpsustase ei ole saavutatud, siis korrigeerimistöde etapis tuleb analüüsida, millisel põhjusel antud küpsustase ei saavutatud [12].

RIA loodud andmekvaliteedi küpsusmudel koosneb viiest kategooriast ja iga kategoorias on võimalik saavutada tase ühest viieni [12].

1. Juhtimine ja planeerimine. Selles kategoorias hinnatakse andmekvaliteedi parendamise protsesside juhtimist ja planeerimist. Näiteks selles kategoorias 2. küpsustaseme saavutamiseks peavad organisatsioonil olema määratletud ja dokumenteeritud andmekogu andmekvaliteedi nõuded ja kõik poliitikatele alluvad kriitilised andmed. [12]
2. Organisatsioon ja vastutused. Selles kategoorias hinnatakse andmekvaliteedi parendamise protsessiga seotute töötajate määratlemist ja nende vastutusi. Näiteks selles kategoorias 3. küpsustase saavutamiseks peavad organisatsioonis olema määratletud andmete omanikud ja ametijuhendites selgelt välja toodud andmekvaliteedi protsesside omanike vastutused. [12]
3. Protsessid: Selles kategoorias hinnatakse organisatsiooni andmekvaliteedi parendamise protsesse. Näiteks selles kategoorias 3. küpsustaseme saavutamiseks tuleb organisatsioonil andmekvaliteeti mõõta ja avastava puudujääke protsessi varajases staadiumis. [12]
4. Teadmus ja kompetentsid. Selles kategoorias hinnatakse organisatsiooni andmekvaliteediga töötavate isikute teadmust ja kompetentsi. Näiteks selles kategoorias 2. küpsustaseme saavutamiseks tuleb organisatsiooni võtmeisikutel

läbida andmekvaliteedi teemalisi koolitusi ning tuleb kavandada ja anda soovitusi andmekvaliteedi juhtimise osas. [12]

5. Töövahendid. Antud kategoorias hinnatakse organisatsioonis kasutuses olevaid töövahendeid. Näiteks selles kategoorias 2. küpsustaseme saavutamiseks tuleb organisatsioonil töötada välja ja dokumenteerida standardid infotehnoloogiliste vahendite jaoks ja osaliselt juurutada töövahendid andmekvaliteedi mõõtmiseks. [12]

RIA andmekvaliteedi tunnuste komplektis on kokku üheksa tunnust: õigsus, täielikkus, kooskõla, usaldusväärsus, ajakohasus, konfidentsiaalsus, ühekordsus, mitteliiasus ja reeglipärasus [1]. Ettevõtte X andmebaasi analüüsi nende tunnuste suhtes leiab peatükis „6. Andmekvaliteedi nõuded ja nende täidetatus“. Tunnuste komplektid moodustati järgmisel viisil [1].

- Analüüsi 12 allikat ja toodi nendest allikatest välja kõik seal nimetatud andmekvaliteedi aspektid. Andmekvaliteedi aspekte kokku oli 108. [1]
- Konsolideeriti andmekvaliteedi aspekte – võeti kokku andmekvaliteedi aspektid, mille nimetused ja definitsioonid olid sarnased või samad ning võeti kokku ka need andmekvaliteedi aspektid, mille nimetused erinesid, kuid definitsioonid olid sarnased või samad. Konsolideerimise tulemusena jäi alles 53 unikaalset andmekvaliteedi aspekti. [1]
- Järgnevalt tõlgiti ingliskeelsed andmekvaliteedi aspektide nimetused eestikeelseteks ja korraldati konsolideerimist. Alles jäi 28 unikaalset andmekvaliteedi aspekti. [1]
- Järgmisena hinnati 28 andmekvaliteedi aspekti mõõdetavust, mõjutatavust ja valdkonna kaetust. Hinnangu tulemusena pandi kokku tunnuste komplekt andmekvaliteedi aspektidest. [1]



## 4 Andmebaasi mudel

Selles peatükis analüüsitakse ettevõtte X olemasolevat andmebaasi, selgitades selle tabelite tähendust. Mudel, mis kirjeldab SQL-andmebaasi tabeleid, on disainitaseme mudel.

### 4.1 Andmebaasi alamosad

Ettevõtte andmebaas on SQL-andmebaas, mis on loodud MySQL andmebaasisüsteemis. Andmebaas moodustub WordPressi ja selle pistikprogrammide tabelitest. Järgnevalt kirjeldatakse andmebaasi jaotust alamosadeks vastavalt programmidele, läbi mille toimub andmebaasis olevate andmete haldamine.

- WordPressi tabelite nimed algavad eesliitega “wp”. Töö kirjutamise ajal (2023. aasta kevad) kasutatakse WordPressi versiooni 6.1.1.
- E-poe pistikprogrammi WooCommerce kaudu toimub kaupade lisamine, müümine (hinnapakkumine, onlain müük) ja hinna kujundamine (käibemaks). Vastavate tabelite nimed algavad eesliitega “woocommerce” või “wc”. Töö kirjutamise ajal (2023. aasta kevad) kasutatakse pistikprogrammi versiooni 7.4.0.
- Kodulehe optimeerimiseks kasutatakse pistikprogrammi Rank Math, mis kontrollib, kas otsingumootorite jaoks on informatsioon (võtmesõnad, metaandmed jne) korrektselt sisestatud. Vastavate tabelite nimed algavad eesliitega “rank\_math”. Töö kirjutamise ajal (2023. aasta kevad) kasutatakse versiooni 3.0.3.1.
- Hinnapakkumiste pistikprogrammi kasutades saavad kliendid küsida ettevõttelt individuaalseid hinnapakkumisi. Töö kirjutamise ajal (2023. aasta kevad) kasutatakse versiooni 4.8.0.

#### 4.1.1 WordPress

Kasutuses olevas WordPressi andmebaasis on 12 tabelit. WordPressi andmebaasis on näiteks andmed kasutajate, postituste, kommentaaride ja linkide kohta. Nendes tabelites on ka informatsioon kaupade kohta (lühikirjeldus, spetsifikatsioon). Lisas 4 on WordPressi dokumentatsioonist pärit andmebaasi diagramm, millel on esitatud WordPressi tabelid ja nende vahelised seosed [21].

#### **4.1.2 E-poe pistikprogramm**

Kasutusel olevas WooCommerce pistikprogrammi andmebaasis on 30 tabelit. WooCommerce andmebaasis on näiteks andmed e-poes sätestatud maksude (nagu määratud käibemaksu suurus) ja maksevõimaluste kohta. Lisaks enda loodud tabelitele kasutab WooCommerce ka WordPressi looduid tabeleid – näiteks kauba kirjeldusi hoitakse WordPressi tabelis *wp\_posts*.

#### **4.1.3 Kodulehe optimeerimise pistikprogramm**

Kasutusel olevas kodulehe optimeerimise pistikprogrammi andmebaasis on 10 tabelit, milles hoitakse kodulehe siseseid ja väliseid linke ning veebilehe optimeerituse analüüsi tulemusi. Lisaks hoiab pistikprogramm andmeid WordPressi tabelites. Näiteks võtmesõnu hoitakse WordPress tabelis *wp\_posts*.

Optimeeritud veebileheks saab nimetada näiteks lehte, millel on olemas sisesed ja välised lingid ning millele on määratud võtmesõna, mis teatud arv kordi kordub. Optimaalne võtmesõnade kasutamine lehel peaks jääma 1–1.5% vahele kogu veebilehe teksti mahust – üle selle on juba liiga tihe võtmesõnade kasutamine ja otsingumootor peab seda halva kvaliteedi signaaliks [70]. Mida paremini on veebileht optimeeritud, seda paremini on see otsingumootorite poolt üles leitud.

#### **4.1.4 Hinnapakkumiste pistikprogramm**

Hinnapakkumise pistikprogramm kasutab E-poe pistikprogrammi WooCommerce tabeleid. Näiteks kasutatakse tabelit *wc\_customer\_lookup*, et salvestada hinnapakkumise soovi esitanud kliendi ees- ja perekonnanime ja *wc\_order\_product\_lookup*, et salvestada hinnapakkumises esitatud kaubad.

### **4.2 Põhiobjektid**

Keerukuse paremaks haldamiseks ja andmemudeli paremaks esitamiseks tükeldatakse andmebaas lähtuvalt põhiobjektidest loogilisteks alamosadeks ehk registriteks. Iga registris on kokku-kuuluvad – ühe asja kohta käivad – andmed. Iga registri realiseerimiseks on kasutusel üks või mitu tabelit, kusjuures need tabelid võivad kuuluda erinevate pistikprogrammide andmebaasidesse. Järgnevalt esitatakse nimekiri põhiobjektidest järjekorras, et kõigepealt on klassifikaatorid (kirjeldavad kõiki ülejäänud andmeid), siis põhiandmete (*master data*) vastavad põhiobjektid ja siis tehingutele (*transactional data*) vastavad põhiobjektid.

- Klassifikaator
- Kasutaja
- Klient
- Kommentaar
- Konfiguratsioon
- Makse
- Märkmed
- Müügiobjekt
- Sündmus
- Tellimus (sh hinnapakkumine)

### 4.3 Ettevõtte X andmebaasi andmemudel

Olemasolevas andmebaasis olevaid tabeleid saab liigitada järgnevalt.

1. Praeguse programmi versiooni tabel, kus on andmeid.
2. Praeguse programmi versiooni tabel, mis on tühi (sest vastavat funktsionaalsust ei kasutata).
3. Vana programmi versiooni tabel, mida praeguses versioonis enam ei ole ja kus on andmeid.
4. Vana programmi versiooni tabel, mida praeguses versioonis enam ei ole ja mis on tühi (sest vastavat funktsionaalsust ei kasutatud).

Järgnev andmebaasi kirjeldus hõlmab tabeleid, mis kuuluvad klassidesse 1 ja 2. Klassi 3 ei kuulu ükski tabel ja klassi 4 kuulub üks tabel (*wp\_links*). Iga registri kohta tuuakse välja sellesse kuuluvad tabelid. Iga tabeli kohta esitatakse selle nimetus ja vabatekstiline sõnaline kirjeldus. Samuti tuuakse välja, kas töö kirjutamise hetke seisuga (2023. aasta kevad) kuulub see klassi 1 või 2. Tabelid esitatakse tähestiku järjekorras. Lisaks koostati ka veergude sõnalised kirjeldused. Kuna nende täielik esitus käesolevas töös oleks liiga mahukas, siis käesoleva töö Lisas 5, esitatakse näitena ühe tabeli (*wp\_posts*) veergude sõnalised kirjeldused. Nii tabelite kui veergude kirjeldused on tehtud ettevõtte töötajatele kättesaadavaks wikis.

- Loogiliselt kokku kuuluvad tabelid on grupeeritud ühte registrisse.
- Töö käigus koostatavad tabelite ja veergude vabatekstilised kirjeldused on oluliselt põhjalikumad kui ametlikus dokumentatsioonis esitatud kirjeldused.

- Kirjeldus on ühes kohas koos ja selle alusel luuakse ettevõtte töötajatele mõeldud wiki. Praktikas on WordPressi ja pistikprogrammide dokumentatsioon laiali paljudel erinevatel veebilehtedel, kust on ka raske informatsiooni leida ja infot esitatakse erinevatel viisidel (piltidena, videotena, tekstina).
- Kirjeldus on eesti keeles, samas kui ametlik dokumentatsioon on inglise keeles.

#### 4.3.1 Klassifikaatorite register

Klassifikaatorite registrisse kuuluvad kõik ettevõtte andmebaaside tabelid, kus hoitakse veebipoe kohta käivat üldist informatsiooni veebipoe maksemäärade, kaupade kategooriate, tarne jm sarnase kohta. Tabelis 4 on toodud välja klassifikaatorite registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 4. Klassifikaatorite registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wc_category_lookup	Tabelis on informatsioon e-poe kategooriate suhete kohta e on märgitud, kas tegemist on vanem ( <i>parent</i> ) või laps ( <i>child</i> ) kategooriaga. Juhul, kui tegemist on laps-kategooriaga, siis selles tabelis on võimalik leida vanem-kategooria ID. Tabelis kuvatakse vaid kategooriate ID väärtused.	1
wc_tax_rate_classes	Veebipoes kehtestatud maksude suurused. Näiteks käibemaksu suurus (Eestis 2023. aasta alguses 20%).	1
woocommerce_attribute_taxonomies	Tabelis on informatsioon e-poe atribuutide kohta. Atribuutideks võib olla kauba/toote bränd ehk kaubamärk või kauba/toote tüüp. Tabelis ei ole mitte atribuutide väärtused, vaid nende nimetused. Atribuute võib e-poes kasutada näiteks kaupade/toodete järgi filtreerimisel – nii e-poe klient, kui ka organisatsiooni töötaja saavad kataloogis olevaid kaupu/tooteid filtreerida kasutades brändi või kauba/toote tüüpi.	1
woocommerce_shipping_zone_methods	Tabelis on tarneviisid vastavalt tsoonidele. Näiteks Eestis saab tulla ise tellimusele järgi või tellida kulleriga kohaletoimetamise.	1
woocommerce_shipping_zone_locations	Tarnetsoonide asukohad. Asukohtadeks saavad olla määratud nii riik üldiselt, kui ka kindlad piirkonnad nendes riikides. Tarnetsoonide näideteks on Eesti ja Helsinki.	1

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
woocommerce_shipping_zones	E-poe tarnetsoonid. Üldisel on siin tabelis toodud välja riigid, millesse kaupu tarnitakse. Riikide näideteks on Eesti, Soome, Läti ja Leedu.	1
woocommerce_tax_rate_locations	Tabelis on asukohad, kus määratud käibemaks kehtib.	2
woocommerce_tax_rates	E-poes määratud käibemaksud ja nende tähistus. Näiteks VAT ja 20%.	1
wp_term_taxonomy	Tabeli <i>wp_terms</i> kategooriate ja siltide kirjeldused säilitatakse selles tabelis [21]. Näiteks kui veebipoes on olemas kategooria STEM, siis selles tabelis on STEM kategooria kirjelduseks STEM kategooria tähendus ehk <i>Science Technology Engineering Mathematics</i> .	1
wp_termmeta	Seda tabelit kasutatakse kategooriate ja siltide kohta täiendava info hoidmiseks [21].	1
wp_terms	Tabelis hoitakse WordPressi lehel kasutatavaid kategooriaid ja silte [21]. Kategooriad veebipoes on näiteks bioloogia, keemia ja füüsika. Sildid veebipoe kaupadel on justkui märksõnad, mille õige kasutamine aitab kliendil otsingumootorit kasutades avada ettevõtte X kaupa kirjeldav leht. Näiteks kui klient otsib otsingumootoris teemat „osmoos“ ja ettevõtte X kaubal on sildina märgitud antud märksõna, siis võib otsingumootor pakkuda kliendile ka lingi ettevõtte X kauba lehele.	1
rank_math_analytics_keyword_manager	Tabel, milles hoitakse võtmesõnu ja infot selle kohta, millisesse kolleksiooni need kuuluvad. Ettevõtte antud tabelit ei kasuta.	2

### 4.3.2 Kasutajate register

Kasutajate registrisse kuuluvad kõik ettevõtte andmebaaside tabelid, kus hoitakse informatsiooni WordPress kasutajate kohta. Tabelis 5 on toodud välja kasutajate registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 5. Kasutajate registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
woocommerce_api_keys	Tabelis on mitte-administraator kasutajatele genereeritud API võtmed. REST API võtmed luuakse selleks, et mitte-administraator kasutajad saaksid lugeda ja kirjutada WooCommerce'i andmeid.	2

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
	WooCommerce'is kasutatavad andmed on tellimused, kaubad/tooted, kupongid, kliendid ja tarnepiirkonnad [71]. Allikas [72] on toodud näidisjuhend, kuidas WooCommerce'is genereerida REST API võti.	
wp_usermeta	Tabelis hoitakse informatsiooni kasutajate kohta. Näiteks salvestatakse kasutaja ees- ja perekonnanimi, aadress, e-posti aadress, viimane sisselogimise aeg ning muud WordPressi kasutajaprofiili sätteid puudutavat informatsiooni.	1
wp_users	Tabelis hoitakse kõiki WordPressi registreerunud kasutajaid ja põhiteavet nende kohta. Näiteks hoitakse siin kasutajanime ( <i>user_login</i> ), parooli räsi ja kasutaja enda valitud hüüdnime [73].	1

#### 4.3.3 Klientide register

Klientide registrisse kuuluvad kõik ettevõtte andmebaaside tabelid, kus hoitakse informatsiooni ettevõtte e-poes tehinguid teinud klientide kohta. Tabelis 6 on toodud välja klientide registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 6. Klientide registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wc_customer_lookup	Üldine informatsioon ettevõtte e-poes tehinguid teinud klientide kohta ehk klientide kohta, kes on esitanud tellimusi või küsinud hinnapakkumist. Näiteks hoitakse siin kliendi ees- ja perekonnanime ning e-posti aadressi.	1

#### 4.3.4 Kommentaaride register

Kommentaaride registrisse kuuluvad kõik ettevõtte andmebaaside tabelid, kus hoitakse WordPress kasutajate kommentaare ja nendega seotud lisateavet. Tabelis 7 on toodud välja kommentaaride registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 7. Kommentaaride registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wp_commentmeta	Tabelis hoitakse lisainformatsiooni kommentaari ja kommenteerija kohta. Näiteks leiab sellest tabelist informatsiooni, kas kommenteerija on WordPressi kasutaja või mitte.	1

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wp_comments	<p>Tabelis hoitakse kommentaare, koos kommenteerija andmetega. Näiteks on selles tabelis olemas veerud järgmiste andmete jaoks:</p> <ul style="list-style-type: none"> <li>▪ kommentaari autori nimi,</li> <li>▪ kommentaari autori e-posti aadress,</li> <li>▪ kommentaari autori IP-aadress,</li> <li>▪ kommenteerimise aeg,</li> <li>▪ kommentaar ise.</li> </ul>	1

### 4.3.5 Konfiguratsioonide register

Konfiguratsioonide registrisse kuuluvad kõik ettevõtte andmebaaside tabelid, kus hoitakse informatsiooni WordPress veebilehe otsingumootori optimeerimise ja sündmuste märguannete kohta. Tabelis 8 on toodud välja konfiguratsioonide registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 8. Konfiguratsioonide registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
rank_math_analytics_ga	<p>Tabelis hoitakse statistikat selle kohta, kui mitu korda veebisaidi postitusi/lehti kindlal kuupäeval ja ajal vaadati ja kui mitu erinevat kasutajat seda on teinud.</p> <p>Rank Math pistikprogrammi on võimalik siduda Google analüütikaga (Google Analytics) ja siit tuleb ka selle tabeli nimetusse „ga“. Ettevõtte X on seda teinud.</p>	1
rank_math_analytics_gsc	<p>Tabelis hoitakse statistikat selle kohta, millised kasutajate otsingulaused/sõnad andsid tulemuseks veebisaidi, mitmes oli veebisait otsingutulemuste järjekorras ja kui mitu korda veebisaiti avati.</p> <p>Rank Math pistikprogramm on ettevõttel X seotud Google Search Console'iga. Tabeli nimetuses „gsc“ tähendab Google Search Console.</p>	1
rank_math_analytics_inspections	<p>Tabelis hoitakse lehtede indekseerimise analüüsi. Näiteks leiab siit tabelist, et lehte X [74]:</p> <ul style="list-style-type: none"> <li>▪ võib indekseerida (<i>index_verdict = PASS</i>),</li> <li>▪ indekseerimine on lubatud (<i>indexing_state = INDEXING_ALLOWED</i>),</li> <li>▪ leht on indekseeritud ja saidikaart on loodud (<i>indexing_state = Submitted and indexed</i>).</li> </ul>	1
rank_math_analytics_objects	<p>Tabelis hoitakse üldist informatsiooni postituse otsingumootori optimeerimise analüüsi kohta.</p>	1

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
	Selles tabelis on veerud, mis vastavad sellele, millist tüüpi postitusega on tegu (ajaveebi postitus, kaup, leht), kas postitus on indekseeritav otsingumootorite poolelt (1 – jah, 0 - ei) ja milline on ettevõtte töötaja poolelt sisestatud otsingu lause/sõna. Ettevõtte töötajad lisavad otsingu lauseid/sõnu, et ettevõtte kaupade lehed oleksid otsingumootorite poolt paremini indekseeritavad.	
rank_math_internal_links	Tabel, milles hoitakse postituse siseseid ja väliseid linke. Välisteks linkideks on need lingid, mis suunavad kasutajat ettevõtte veebilehelt mõnele teisele veebisaidile (näiteks YouTube'i). Siseseid lingid on need lingid, mis suunavad kasutajat mõnele muule ettevõtte veebilehele (näiteks ühelt kaubalt teisele kaubale).	1
rank_math_internal_meta	Tabelis hoitakse informatsiooni selle kohta, kui mitu linki postituses on ja kui mitu linki teistes postituses antud postitusele on. Näiteks võidakse siin registreerida, et postitusel ID-ga 39611 on kaks veebisaidi sisest linki, null välist linki ja null linki teistelt postitustelt postitusele ID-ga 39611.	1
rank_math_redirections	Tabelis hoitakse veebisaidis seadistatud ümbersuunamiste andmeid: <ul style="list-style-type: none"> <li>▪ milliselt lehelt soovitakse ümbersuunata,</li> <li>▪ millisele lehele soovitakse suunata.</li> </ul>	2
wc_rate_limits	Tabel, milles hoitakse kiiruse piiranguid Store API lõpp-punktidele [75]. Näiteks võib selles tabelis olla hoitud informatsioon, et X päringut on võimalik teha Y korda Z sekundi jooksul [75]. Store API pakub avalike REST API lõpp-punkte, et oleks võimalik arendada ostukorvi ja tootekaartide funktsionaalsust [76]. Vaikimisi on tabel lülitatud välja [77]. Ettevõtte X seda samuti ei kasuta.	2
wc_webhooks	Selles tabelis hoitakse veebihaake, mille kasutamine võimaldab saata automaatselt andmeid kolmandatele osapooltele kui veebisaidil toimus kindel sündmus [78], [79]. Näiteks saab kasutades veebihaake selleks, et saata kliendile automaatse teavituse meilile, et nende tellimus on esitatud [79].	2
wp_options	Tabelis on kõikvõimalikke andmeid, mida WordPressil põhinev veebisait toimimiseks vajab [80]. Näiteks on antud tabelis salvestatud kodulehekülje URL e link ja veebipoes kasutatavate valuutade andmed.	1



### 4.3.6 Maksete register

Maksete registrisse kuuluvad kõik ettevõtte andmebaaside tabelid, kus hoitakse informatsiooni e-poes tehtavate maksete kohta. Tabelis 9 on toodud välja maksete registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 9. Maksete registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
woocommerce_payment_tokenmeta	Metaandmed makse lubade ( <i>tokenite</i> ) kohta. Täpsemalt, kui klient on sooritanud ostu ja maksnud onlaineis, siis selles tabelis säilitatakse täiendavaid andmeid makse kohta. Näiteks säilitatakse maksja ees- ja perekonnanimi ning tema Stripe kasutajanimber.	2
woocommerce_payment_tokens	E-poes saavad kliendid maksta kahel viisil – kohe onlaineis pangalaulekandega või hiljem esitatud arve alusel pangalaulekandega. Kui klient maksab kohe onlaineis, siis kasutatakse tehingu tegemiseks spetsiaalseid maksesüsteeme. Nende maksesüsteemide integreerimiseks e-poodi kasutatakse vastavaid lubasid, mis sisaldavad näiteks makse lüüsi (nt. Apple Pay), kelle arvele tellimuse summa liigub. Tabelis säilitatakse makse lubasid ja nendega seotud informatsiooni, et e-poe kasutaja saaks näiteks maksta soovitud kauba eest PayPal või Stripe'ga.	2

### 4.3.7 Märkmete register

Märkmete registrisse kuuluvad kõik ettevõtte andmebaasi tabelid, kus hoitakse informatsiooni WordPress administraatori õigustes kasutajatele kuvatavate teadete kohta. Tabelis 10 on toodud välja märkmete registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 10. Märkmete registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wc_admin_note_actions	Tabelis säilitatakse WordPress administraatori õigustes kasutajate tegevusi seoses WooCommerce pistikprogrammi loojatelt tulnud uudiste, uuenduste, küsitluste ja muu reklaamiga. Näiteks kui mustal reedel kuvati WordPress administraator kasutajatele reklaami, et ta saab osta soodsamalt mõne täiendava WooCommerce rakenduse, siis sai administraator reklaami sulgeda (tegevus #1) või vajutada reklaamis olevale bannerile (tegevus #2) ja liikuda WooCommerce loojate veebilehele.	1
wc_admin_notes	Tabelis hoitakse WooCommerce pistikprogrammi loojatelt tulnud uudiseid, uuenduste, küsitlusi ja muud reklaami. Näiteks musta reede tähistamise puhul näidata WordPress administraator kasutajatele WooCommerce reklaami, et soetada mõni täiendav WooCommerce rakendust suure soodustusega.	1

### 4.3.8 Müügiobjektide register

Müügiobjektide registrisse kuuluvad kõik ettevõtte andmebaasi tabelid, kus hoitakse informatsiooni ettevõtte e-poes müüdavate kaupade, toodete ja teenuste kohta. Tabelis 11 on toodud välja müügiobjektide registri tabelid koos vabatekstilise sõnalise selgitusega. Tabeli *wp\_posts* veergude kirjeldused esitatakse Lisas 5.

Tabel 11. Müügiobjektide registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wc_product_attributes_lookup	Kaupade atribuudid ja nende väärsused. Atribuudid võivad olla näiteks kauba värvus, kui on võimalik valida kaupu erinevate värvidega.	1
wc_product_download_directories	Allalaaditavate kaupade ja toodete URL'id.	1
wc_product_meta_lookup	Antud tabelis hoitakse järgmist informatsiooni toote, kauba ja teenuse kohta:	1

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
	<ul style="list-style-type: none"> <li>▪ Tootekood (SKU). Näiteks AB-1234.</li> <li>▪ Kas toode on virtuaalne?</li> <li>▪ Kas toode on allalaaditav?</li> <li>▪ Tavahind, soodushind.</li> <li>▪ Laoseis (kogus, laos olemise staatus). Näiteks laoseis 0 ja staatus on <i>onbackorder</i>.</li> <li>▪ Reiting ehk milline hinnang toote, kaubal või teenusel on.</li> </ul>	
wp_postmeta	Tabelis hoitakse iga WordPressi lehel oleva postituse (nt: ajaveebi postitus, kaup, toode, teenus) kohta lisainfot. Näiteks üheks postituse lisainfoks, mida selles tabelis hoiustatakse, on postituse metakirjeldus ( <i>meta description</i> ).	1
wp_posts	Postituste tabelis hoitakse kõiki postitusi, veebilehti ja nende versioone. Samuti hoitakse selles tabelis ka menüü-üksusi, meediafaile, manuseid, postituse staatust jne [21]. Iga kauba, teenuse ja toote kohta on andmebaasis eraldi postitus iga riigi kohta, kus seda müügiks pakutakse. See postitus on selles keeles, mida seal riigis kasutatakse. Näiteks hoitakse tabelis konkreetse kauba kirjeldust koos spetsifikatsiooniga.	1
wp_term_relationships	Tabelis on seosed postituste ja kategooriate/siltide vahel [21]. <i>Wp_term_relationships</i> on seotud tabeliga <i>wp_terms</i> , milles hoitakse kategooriaid ja silte.	1

#### 4.3.9 Sündmuste register

Sündmuste registrisse kuuluvad kõik ettevõtte andmebaasi tabelid, milles hoitakse informatsiooni WordPressi veebilehel toimunud sündmuste kohta. Tabelis 12 on toodud välja sündmuste registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 12. Sündmuste registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wc_download_log	Logide tabel, milles hoitakse informatsiooni allalaaditavate kaupade allalaadimiste kohta. Näiteks hoitakse selles tabelis informatsiooni selle kohta, milline kasutaja laadis alla kauba ja millal ta seda tegi (kuupäev ja aeg).	2
woocommerce_log	Logide tabel, milles hoitakse üldist informatsiooni toimunud sündmuste kohta.	2

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
	Näiteks hoitakse selles tabelis informatsiooni selle kohta kui konkreetsele kaubale/või tootele lisati uus atribuut ja/või selle väärtus. Tabel ei sisaldab kasutajate logisid. Organisatsioonis seda tabelit ei kasutata – see on vaikimisi tühi.	
woocommerce_sessions	WooCommerce sessiooni tabelis salvestatakse klientide seansi andmeid, näiteks nende ostukorv [81].	1
rank_math_redirections_cache	Ümbersuunamiste spetsifikatsioonid on tabelis <i>rank_math_redirections</i> . Siin tabelis hoitakse veebisaidil toimunud tegelike ümbersuunamiste andmeid. Tabelist leiab infot: <ul style="list-style-type: none"> <li>▪ milliselt lehelt suunati ümber,</li> <li>▪ millist ümbersuunamist kasutati (seotud tabeliga <i>rank_math_redirections</i>),</li> <li>▪ kas sai ümbersuunatud.</li> </ul>	2
rank_math_404_logs	Tabelis hoitakse informatsiooni selle kohta, milliste URI-dega seoses olid veebisaidil <i>error 404</i> vead. <i>Error 404</i> vead võisid tekkida erinevatel põhjustel nagu näiteks [82]: <ul style="list-style-type: none"> <li>▪ kasutaja tegi veebiaadressi sisestamisel vea,</li> <li>▪ kellegi poolt oli edastatud vale veebiaadress (näiteks selles oli kirjaviga),</li> <li>▪ fail või leht, millele üritatakse saada ligipääsu, on kustutatud või avalikusele kättesaamatuks tehtud (näiteks postitus oli mustanditesse kantud).</li> </ul> Antud tabel aitab ettevõttel leida vigu veebisaidis ja need parandada, et kasutajatel jääks veebisaidi ja ettevõtte kohta hea mulje.	1

#### 4.3.10 Tellimuste register

Tellimuste registrisse kuuluvad kõik ettevõtte andmebaasi tabelid, kus hoitakse informatsiooni veebipoes tehtud hinnapakumiste taotluste ja tellimuste kohta. Tabelis 13 on toodud välja tellimuste registri tabelid koos vabatekstilise sõnalise selgitusega.

Tabel 13. Tellimuste registri tabelid.

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wc_order_stats	Tellimuste statistika tabel, milles on toodud välja informatsioon, milline oli müüdud kaupade arv, summa ja tasutud maksud.	1

<b>Tabeli nimi</b>	<b>Tabeli kirjeldus</b>	<b>Klass</b>
wc_order_coupon_lookup	Kupongide tabelis on toodud kõik kupongid ja nende saadavate allahindluse summad. WooCommercis on kupongiks sooduskood (näiteks võib sooduskood olla „HäidJõule2022“), mille kasutaja poolt rakendamisel tehakse talle kupongi suuruses soodustus.	2
wc_order_product_lookup	Tellitud kaubad, kauba kogus ja selle lõppsummad (käibemaksuga ja käibemaksuta).	1
wc_order_tax_lookup	Kokkuvõtlik tabel, milles on toodud välja tellimuse tasud tarne ja kaupade eest.	1
wc_reserved_stock	Reserveeritud kaubad ja tooted.	2
woocommerce_downloadable_product_permissions	E-poes saab müüa erinevat tüüpi kaupu – tavakaup ( <i>simple product</i> ), virtuaalne ( <i>virtual</i> ), allalaaditav ( <i>downloadable</i> ), grupeeritud kaup ( <i>grouped product</i> ), väline- või sidustoode ( <i>external/affiliate product</i> ) ja muutuv toode ( <i>variable product</i> ) [83]. Antud tabelis hoiustatakse juurdepääsulubasid allalaaditavatele kaupadele. Näiteks hoitakse selles tabelis allalaaditava kauba/toote identifikaatorit, loa saanud kasutaja e-posti aadressi, tellimuse identifikaatorit ja kui mitu korda on kaupa/toodet ostnud kasutaja allalaadimisi teinud. Organisatsioonis seda tabelit ei kasutata, kuna ei müüda e-poest otse allalaaditavaid kaupu/tooteid.	2
woocommerce_order_itemmeta	Metaandmed e täiendav informatsioon klientide tellimuste kohta. Metaandmete hulka kuulub näiteks <i>product_id</i> , mis tähendab tootekoodi ja <i>qty</i> , mis tähendab toote või kauba kogust.	1
woocommerce_order_items	Antud tabelis on klientide tellimuste kirjed [84]. Tellimuse kirjed võivad olla järgmised. <ul style="list-style-type: none"> <li>▪ Tellitud kaup, tood või teenus. Näiteks Robot A.</li> <li>▪ Kõikide tellitud kaupade, toodete või teenuste käibemaks kokku.</li> </ul> Iga tellimuse kirje on seotud kindla kliendi tellimusega ja kliendi tellimuses võib olla mitu kirjet.	1

## **5 Andmekvaliteedi juhtimise raamistiku rakendamine**

Selles peatükis kirjutatakse andmekvaliteedi juhtimise raamistiku (rohkem jaotises „3.3 Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistik“) rakendamisest ettevõtte X e-poes, et parandada selle ettevõtte hallatavas andmebaasis olevate andmete kvaliteeti.

### **5.1 Hindamine**

Hinnatakse ettevõtte X veebirakendust, mis koosneb e-poest, blogist ehk ajaveebist ja informatiivsetest lehtedest (kontakt, meist, koduleht jne).

Andmete kvaliteedi küpsustaseme hindamiseks kasutatakse RIA andmekvaliteedi tagamise juhendis kirjeldatud andmekvaliteedi mudelit. Selles mudelis esitatakse andmekvaliteedi küpsustaseme kindlaks tegemiseks kasutatavad väited. Need on rühmitatud küpsuse kategooriate (neid on kokku viis) kaupa ja sorteeritud kategooriate sees küpsustaseme järgi (tasemeid on kokku viis). Autor muutis seda mudelit veidi, et see vastaks paremini väikese eraettevõtte vajadustele. Muudatuse sisuks oli kolme väite kustutamine. Mudelis tehtud muudatused tuuakse välja jaotises „9.3 Andmekvaliteedi juhtimise raamistiku kohandused väikesele eraettevõttele“. Muudatuse tulemusena valminud ja töös kasutatav mudel on Lisas 6. Väidete sõnastust ei peetud vajalikuks muuta.

#### **5.1.1 Andmete ulatuse määratlemine ja tsoneerimine**

Andmekvaliteedi parandamine hõlmab kõiki registreid, mida kirjeldati peatükis „4. Andmebaasi mudel“. Andmeobjektid on tsoneeritud vastavalt peatükis 4 kirjeldatud registritele. Registrite eest vastutavad ettevõtte töötajad, kes tegelevad e-poe sätete haldusega, andmete sisestamisega (andmehaldurid) ja andmete analüüsimisega.

#### **5.1.2 Andmekvaliteedi küpsustaseme hindamine**

Seejärel tuleb hinnata ettevõtte X andmekvaliteedi küpsustaset. Järgnevalt esitatakse hindamise tulemused. Iga kategooria kohta tuuakse välja ainult need väited (koos vastava küpsustasemega), millele ettevõtte andmekvaliteedi tagamine enne käesoleva töö tegemise algust vastab. Väidete sõnastus on samasugune nagu RIA mudelis [12].

Kategooria: Juhtimine ja planeerimine

- Tase 1: Andmekogu andmekvaliteedi nõuded ei ole määratletud ega dokumenteeritud.

Kategooria: Organisatsioon ja vastutused

- Tase 1: Andmekvaliteedi omanik ei ole tuvastatav. Andmete omanikud on määratlemata.

Kategooria: Protsessid

- Tase 1: Andmekvaliteedi profiil ei ole kirjeldatud ja seda ei kasutata protsessi käigus puuduste avastamiseks. Mõõdetavad andmekvaliteedi tunnused ei ole määratletud ega dokumenteeritud. Tuvastatud sündmuseid ei registreerita ja nende lahendamine ei ole jälgitav. Andmekvaliteedi tagamise protsessid ei ole määratletud ega dokumenteeritud.

Kategooria: Teadmus ja kompetentsid

- Tase 2: Andmekvaliteedi teadlikkuse tõstmiseks on läbi viidud andmekvaliteedi temalisi koolitusi. Võtmeisikud ei ole tuvastatavad.

Kategooria: Töövahendid

- Tase 1: Andmekvaliteedi standardid infotehnoloogiliste vahendite jaoks ei ole välja töötatud ega dokumenteeritud. Töövahendid andmekvaliteedi mõõtmiseks ja parendamiseks ei ole juurutatud.

Hindamise tulemusel selgus, et kõikides kategooriates, peale ühe, on maksimaalne küpsustase enne andmekvaliteedi parandamisega tõsisemalt tegelema hakkamist 1 (reaktiivne tase). Kategoorias “Teadmus ja kompetentsid” on tase 2 (kontrollitud tase), sest töö autor on oma töökaaslaseid andmekvaliteedi osas järjepidevalt juhendanud. Järelikult on ka ettevõtte üldine andmekvaliteediga tegelemise küpsustase 1 ehk madalaim võimalik tase.

## **5.2 Planeerimine**

Selles jaotises räägitakse andmekvaliteedi parandamise planeerimisest.

### 5.2.1 Küpsuse sihttaseme planeerimine

Plaani järgselt, peale andmekvaliteedi raamistiku ettevõttes kasutusele võtmist, on ettevõtte X andmekvaliteedi küpsus 3. tasemel (standardiseeritud tase). Standardiseeritud taseme korral on andmekvaliteedi protsessid standardiseeritud ja samuti on teada, kui hästi vastab andmekvaliteet nõuetele. Selle saavutamiseks tuleb kõikides kategooriates tõsta küpsuse tase 1. tasemel (või ühe puhul 2. tasemelt) 3. tasemele. Töö analüüsi osas, jaotises „9.4 Andmekvaliteedi juhtimise raamistiku rakendamise ja selle tulemuste analüüs“, on toodud välja põhjendus, miks otsustati viia ettevõtte andmekvaliteedi küpsus 3. tasemele. Jaotises „9.7.1 Andmekvaliteet küpsustasemele neli ja viis“ on toodud välja visioon 3. tasemelt edasi liikumiseks.

### 5.2.2 Parendamise plaani koostamine

Parendamise plaani koostamise aluseks võeti jaotises “5.1 Hindamine” esitatud kategooriate kaupa küpsustaseme hinnang ja RIA küpsustaseme tõstmise juhend [12]. Ühegi väite sõnastust ei peetud vajalikuks muuta. Tabelis 14 on toodud välja andmekvaliteedi parendusvaldkond ja vastavad parendustegevused. Kõikidel tegevustel on üks ja sama vastutaja (töö autor rollis andmehaldur) ja täitmise tähtaeg (2023. aasta aprilli lõpp). Küpsustasemete saavutamine peaks toimuma järk-järgult. Kõigepealt tuleb kõikides kategooriates saavutada 2. küpsustase, siis on ka üldine andmekvaliteedi küpsustase 2. Seejärel tuleb kõikides kategooriates saavutada 3. küpsustase, et ka üldine küpsustase oleks 3. Tabelis 14 esitatakse parendustegevused, mida peaks tegema tasemelt 1 tasemele 2 jõudmiseks ja Tabelis 15 esitatakse parendustegevused, mida peaks tegema tasemelt 2 tasemel 3 jõudmiseks.

Tabel 14. Parendustegevused andmekvaliteedi 2. küpsustaseme saavutamiseks.

<b>Kategooria</b>	<b>Parendusvaldkond</b>	<b>Kategoorias 2. taseme saavutamiseks vajalikud parendustegevused</b>
Juhtimine ja planeerimine	Andmekvaliteedi nõuded	<ul style="list-style-type: none"><li>▪ Määratleda andmekvaliteedi nõuded ja need dokumenteerida.</li><li>▪ Määratleda ja dokumenteerida kriitilised andmed.</li></ul>
Organisatsioon ja vastutused	Andmete ja andmekvaliteedi omanikud	<ul style="list-style-type: none"><li>▪ Määrata andmete omanikud.</li><li>▪ Määrata andmekvaliteedi protsesside omanikud.</li></ul>
Protsessid	Andmekvaliteedi profiili loomine ja	<ul style="list-style-type: none"><li>▪ Kirjeldada andmekvaliteedi profiil ja see dokumenteerida.</li></ul>



Kategooria	Parendus-valdkond	Kategoorias 2. taseme saavutamiseks vajalikud parendustegevused
	protsesside määramine	<ul style="list-style-type: none"> <li>Määratleda andmekvaliteedi protsessid ja need dokumenteerida.</li> </ul>
Protsessid	Andmekvaliteedi probleemide jälgimine ja analüüs	<ul style="list-style-type: none"> <li>Võtta kasutusele süsteem andmekvaliteedi probleemide registreerimiseks ja lahendamise jälgimiseks.</li> <li>Analüüsida andmekvaliteedi probleeme.</li> </ul>
Töövahendid	Andmekvaliteedi nõuded	<ul style="list-style-type: none"> <li>Töötada välja andmekvaliteedi standardid süsteemidele ja dokumenteerida need.</li> </ul>
Töövahendid	Andmekvaliteedi töövahendite juurutamine	<ul style="list-style-type: none"> <li>Alustada töövahendite juurutamisega.</li> </ul>

Tabel 15. Parendustegevused andmekvaliteedi 3. küpsustaseme saavutamiseks.

Kategooria	Parendus-valdkond	Kategoorias 3. taseme saavutamiseks vajalikud parendustegevused
Juhtimine ja planeerimine	Andmekvaliteedi nõuded	<ul style="list-style-type: none"> <li>Teha ettevõtte töötajatele kättesaadavaks määratletud andmekvaliteedi nõuded.</li> </ul>
Juhtimine ja planeerimine	Andmekvaliteedi kontroll	<ul style="list-style-type: none"> <li>Kontrollida, kas andmed vastavad andmekvaliteedi nõuetele.</li> </ul>
Organisatsioon ja vastutused	Andmete ja andmekvaliteedi omanikud	<ul style="list-style-type: none"> <li>Dokumenteerida ametijuhenditesse andmekvaliteedi protsessi vastutused.</li> </ul>
Protsessid	Andmekvaliteedi profiili parendamine	<ul style="list-style-type: none"> <li>Rakendada andmekvaliteedi profiile, et tulevikus avastada puudujääke kohe alguses.</li> </ul>
Teadmus ja kompetentsid	Ettevõtte wiki	<ul style="list-style-type: none"> <li>Lisada ettevõtte wikisse andmekvaliteedi tagamise parimad praktikad</li> </ul>
Töövahendid	Andmekvaliteedi töövahendite juurutamine	<ul style="list-style-type: none"> <li>Juurutada töövahendid andmekvaliteedi mõõtmiseks.</li> </ul>

## 5.3 Teostamine

Selles jaotises kirjeldatakse küpsuse kategooriate kaupa, kuidas toimus parendusplaani elluviimine ettevõttes X.

### 5.3.1 Juhtimine ja planeerimine

2. küpsustaseme saavutamiseks kategoorias “Juhtimine ja planeerimine” võeti vaatluse alla üheksa andmekvaliteedi tunnust ja tehti kindlaks neile vastavad nõuded. Peale

andmekvaliteedi nõuete arutamist ja kinnitamist kolleegide ja ettevõtte juhatusega, dokumenteeriti ettevõtte wikisse 98 andmekvaliteedi nõuet, neist 53 kriitilise tähtsusega. Andmekvaliteedi nõuded on toodud välja peatükis „6 Andmebaasi andmekvaliteedi nõuded“.

3. küpsustaseme saavutamiseks avalikustati ettevõtte andmekvaliteedi nõuded wikis kõikidele töötajatele ja viidi läbi andmekvaliteedi kontroll. Ettevõtte andmete vastavust nõuetele kontrolliti kasutades peatükis 8 nimetatud andmekvaliteedi hindamise töölaua prototüüpi ja ettevõtte töötajaid. Kuna ettevõtte kasutab onlain müügiks WordPressi ja selle pistikprogramme, siis mõningate andmekvaliteedi nõuete täidetust tuli käsitsi kontrollida (nt jaotises „6.1 Õigsuse nõuded“).

### **5.3.2 Organisatsioon ja vastutused**

2. küpsustaseme saavutamiseks kategoorias „Organisatsioon ja vastutused“ määratleti andmete omanikud ja andmekvaliteedi protsesside omanikud. Andmete ja andmekvaliteedi protsesside omanikud on toodud välja jaotises „7.1 Rollid“.

3. küpsustaseme saavutamiseks tuli töö autoril täiendada töötajate ametijuhendeid. Tehti kindlaks töötajad, kelle tööülesanneteks on materjalide loomine või e-poodi kaupade, toodete ja teenuste lisamine. Nende ametijuhenditesse lisati andmete kvaliteeti ja selle ettevõttes tagamist puudutav peatükk.

### **5.3.3 Protsessid**

2. küpsustaseme saavutamiseks kategoorias „Protsessid“ loodi kaks andmekvaliteedi profiili – kauba/toote profiil ja teenuse profiil. Andmekvaliteedi profiilid leiab jaotises 8.1. Kirjeldati ka andmekvaliteedi tagamise protsessid ja määrati kindlaks nende eest vastutavad ettevõtte töötajad. Andmekvaliteedi protsessid on toodud välja peatükis „7 Protsessid“.

Ettevõtte wikisse lisati juhend, milles ettevõtte töötajatele tutvustatakse Trello keskkonda. Trello keskkonnas saavad töötajad registreerida andmekvaliteedi probleeme ja jälgida nende lahendamist. Trello keskkonnale said ligipääsu kõik ettevõtte töötajad. Peale andmekvaliteedi kontrollimist analüüsiti ettevõttes esinevaid andmekvaliteedi probleeme ja arutati nende tekkimise põhjuseid.

3. küpsustaseme saavutamiseks rakendati kahte eelnimetatud andmekvaliteedi profiili, mille alusel muudeti olemasolevate kaupade, tootete ja teenuste andmeid. Mõlemat andmekvaliteedi profiili plaanitakse tulevikus kasutada uute kaupade, toodete ja teenuste lisamisel.

#### **5.3.4 Teadmus ja kompetentsid**

Ettevõttel X oli enne töö algust kategoorias „Teadmus ja kompetentsid“ saavutatud 2. küpsustase, põhjusel, et ettevõtte töötajad läbivad käesoleva töö autori läbiviidavaid andmekvaliteedi teemalisi lühikoolitusi. 3. küpsustaseme saavutamiseks lisati ettevõtte wikisse andmekvaliteedi tagamise parimad praktikad. Jaotises 9.2 on toodud välja andmekvaliteedi tagamise parimad praktikad.

#### **5.3.5 Töövahendid**

2. küpsustaseme saavutamiseks kategoorias „Töövahendid“ töötati välja andmekvaliteedi standardid süsteemidele ja dokumenteeriti need ettevõtte wikis. Näiteks toodi ettevõtte wikis välja, et enne pistikprogrammide uuendamist tuleb töötajal kontrollida, kas pistikprogrammi versiooni on testitud kasutuses oleva WordPressi versiooniga. Organisatsiooni töötajatele õpetati wiki kasutamist, võeti kasutusele Trello keskkond ja alustati andmekvaliteedi prototüübi juurutamist. Edaspidi installeeriti andmekvaliteedi hindamistöölauda prototüüp koos lokaalse andmebaasiga ettevõtte X töötaja arvutisse. Töötaja saab käivitada enda arvutist kontrolli, kas ettevõtte X e-poes olevate kaupade, toodete ja teenuste tootekaardid vastavad ettevõttes püstitatud andmekvaliteedi nõuetele. 3. küpsustase saavutati peale andmekvaliteedi kontrollimise töövahendite juurutamist ja uute projektide ettepanekute staadiumis andmekvaliteedi standarditele vastavuse jälgimist. Töövahendist kirjutatakse peatükis „8 Andmekvaliteedi profiil ja andmete sellele vastavuse hindamise töölauda prototüüp“.

### **5.4 Kontrollimine**

Kontrollimise etapis toimub taas ettevõtte X küpsustaseme hindamine, aga juba eesmärgiga saada teada, kas planeeritud sihttase saavutati või mitte [12].

Järgnevalt esitatakse küpsustaseme kontrollhindamise tulemused. Iga kategooria kohta tuuakse välja ainult need väited (koos vastava küpsustasemega), millele ettevõtte andmekvaliteedi tagamine peale andmekvaliteedi juhtimise raamistiku rakendamist vastab. Väidete sõnastus on samasugune nagu RIA mudelis [12].

#### Kategooria: Juhtimine ja planeerimine

- Tase 2: Andmekogu andmekvaliteedi nõuded on määratletud ja dokumenteeritud. Kriitilised andmed on määratletud ja dokumenteeritud.
- Tase 3: Andmekvaliteet on nõuetele vastavuse osas kontrollitud.

#### Kategooria: Organisatsioon ja vastutused

- Tase 2: Andmekvaliteedi protsessi omanik on tuvastatav. Osade andmete omanikud on määratletud.
- Tase 3: Andmekvaliteedi protsesside omaniku vastutused on selgelt määratud ametijuhenditega. Andmete omanikud on määratletud.

#### Kategooria: Protsessid

- Tase 2: Andmekvaliteedi profiil on kirjeldatud ja seda kasutatakse protsessi käigus puuduste avastamiseks. Mõõdetavad andmekvaliteedi tunnused on määratletud ja dokumenteeritud. Otsitakse võimalusi probleemide ennetamiseks. Tuvastatud sündmused registreeritakse ja nende lahendamine on jälgitav. Andmekvaliteedi tagamise protsessid on määratletud ja dokumenteeritud. Andmekvaliteedi puuduste mõju on tuvastatud.
- Tase 3: Andmekvaliteeti mõõdetakse ja puudujäägid avastatakse protsessi varajases staadiumis.

#### Kategooria: Teadmus ja kompetentsid

- Tase 3: Andmekvaliteeti mõjutavad töötajad on läbinud andmekvaliteedi temalise koolituse. Eksisteerib mitteametlik mentorlus andmekvaliteedi teadlikkuse tõstmise osas. Jagatud teadmusbaasina dokumenteeritakse parimad praktikad andmekvaliteedi tagamisel.

#### Kategooria: Töövahendid

- Tase 2: Andmekvaliteedi standardid infotehnoloogilistele vahendite jaoks on välja töötatud ja dokumenteeritud. Töövahendid andmekvaliteedi mõõtmiseks ja parandamiseks on osaliselt juurutatud.

- Tase 3: Juurutatud on töövahendid andmekvaliteedi mõõtmiseks ja parandamiseks. Andmekvaliteedi standarditele vastavust jälgitakse uute projektide ettepanekute staadiumis.

Hindamise tulemusel selgus, et kõikides kategooriates, on saavutatud 3. küpsustase. Järelikult on ka ettevõtte üldine andmekvaliteediga tegelemise küpsustase 3 e standardiseeritud tase, mis viitab sellele, et andmekvaliteet on teada [12].

## **5.5 Korrigeerimine**

Korrigeerimise etapis toimub andmekvaliteedis tehtud muudatuste standardiseerimine, mis on järgmise tsükli sisendiks. Sellel etapil on tähtis dokumenteerida kõik andmekvaliteedi parandamiseks tehtud muudatused, sh ka need, mis ebaõnnestusid. Näiteks tuleb ära kirjeldada ja dokumenteerida kõik kategooriate küpsustasemed, mida õnnestus saavutada ja need mida ei õnnestunud. Kategooriates, milles ei õnnestunud küpsustaset saavutada, tuleb analüüsida ebaõnnestumise põhjuseid. [12]

Ettevõtte X korral dokumenteeriti tehtud muudatused ettevõtte X wikis, et ettevõtte töötajatel oleks võimalik tehtud muudatuste ja saavutatud küpsustasemetega tutvuda. Kuna töö käigus läbitud tsüklis õnnestus soovitud küpsustaseme saavutamine, siis ebaõnnestumiste analüüs puudus.

## 6 Andmekvaliteedi nõuded ja nende täidetud

Andmebaasi andmekvaliteedi nõuete loomisel lähtuti RIA loodud andmekvaliteedi tunnuste komplektist. Komplekt koosneb üheksast andmekvaliteedi tunnusest, mis RIA raamistiku loomisel pandi kokku nendest andmekvaliteedi aspektidest, mis vastavad RIA püstitatud kriteeriumitele olla unikaalne ja mõõdetav [12].

Käesolevas peatükis analüüsitakse iga eelnimetatud tunnuse korral e-poe andmebaasi andmete vastavust nende tunnustele. Samuti kirjutatakse, milliseid meetmeid plaanitakse kasutada, et tagada tulevikus andmete vastavus nendele tunnustele. Iga tunnuse kohta tuuakse välja ka ettevõtte kriitilised nõuded seoses selle tunnusega, st millised nõuded peavad olema kindlasti täidetud. Organisatsioon määrab ise, millised on tema jaoks kriitilised andmed ja millistele nõuetele nad peavad vastama. Nõuete kriitiliseks märkimisel lähtus autor enda arvamusest ja kolleegidega peetud aruteludest.

Selleks, et ettevõtte käsutuses olevad andmed vastaks tunnuse nõuetele, peavad selle andmed vastama kõigile tunnuse kohta välja toodud nõuetele.

### 6.1 Õigsus

Andmekvaliteedi tunnus „Õigsus“ iseloomustab seda, kui võrd tõesed e tegelikkusele vastavad on andmeväärtused [12]. Andmete õigsus on määratud nii organisatsiooni poolt kehtestatud reeglitega, kui ka riigi poolt kehtestatud seadusega. Näiteks avaliku teabe seaduses, mis küll ettevõttele X ei rakendu, sest ettevõtte ei täida avalikke ülesandeid, on märgitud, et teabevaldaja on kohustatud mitte andma teadvalt teavet, mis on eksitava, tegelikkusele mittevastav või ebaõiget. Teabevaldaja peab kahtluse korral kontrollima väljastatava teabe õigsust ja tegelikkusele vastavust. [12], [85] Tarbijakaitseseaduses kirjutatakse eksitavatest kauplemisvõtetest. "Kauplemisvõtet peetakse eksitavaks, kui sellega esitatud teave on ebaõige või kui faktiliselt õige teabe esitusviis petab või tõenäoliselt petab keskmist tarbijat ja kummalgi juhul teeb või tõenäoliselt teeb keskmise tarbija selle mõjul tehinguotsuse, mida ta muul juhul ei oleks teinud." [86] Seadus sätestab, et valeandmeid võidakse esitada näiteks kauba või teenuse olemuse, peamiste omaduse, kaupleja kohustuste, hinna või selle arvutamise aluse, hoolduse, parandamise vajaduse ja tarbija õiguste kohta [86].

### 6.1.1 Olemasoleva andmebaasi vastavus

Tabelis 16 tuuakse välja õigsuse nõuded ja hinnang nende täidetusele olemasolevas andmebaasis. Tabel on sorteeritud nõuete kriitilisuse järgi ja selle sees tähestikulises järjekorras.

Tabel 16. Õigsuse nõuded ja nende täidetus ning kriitilisus ettevõttes X.

ID	Nõue	Täidetus	On kriitiline
1	Iga kauba, toote ja teenuse võtmesõna on õige.	Ei	Jah
2	Iga tellimuse ja hinnapakkumise taotluse kliendil on õige e-posti aadress.	Jah	Jah
3	Iga tellimuse ja hinnapakkumise taotluses peab olema õige arv kaupu, tooteid ja teenuseid.	Jah	Jah
4	Igal e-poe kasutajal on registreeritud õige e-posti aadress.	Jah	Jah
5	Igal e-poe kasutajal on registreeritud õige kasutajanimi.	Jah	Jah
6	Igal kaubal, tootel ja teenusel on märgitud õige pilt.	Ei	Jah
7	Igal kaubal, tootel ja teenusel on õige nimetus.	Jah	Jah
8	Igal kaubal, tootel ja teenusel on õige SKU kood.	Ei	Jah
9	Igal kaubal, tootel ja teenusel on registreeritud õiged hinnad.	Ei	Jah
10	Igal tellimusel ja hinnapakkumise taotlusel peab olema õige klient.	Jah	Jah
11	Igal, kaubal, tootel ja teenusel on õige silt ( <i>tag</i> ).	Ei	Jah
12	Igal kaubal ja tootel ja tootega komplekti pandud teenusel (ühel tootekaardil nii teenus kui toode) on registreeritud õige lao kogus.	Ei	Ei
13	Igal kaubal, tootel ja teenusel on märgitud õige tootja.	Ei	Ei
14	Igal kaubal, tootel ja teenusel on õige kirjeldus.	Ei	Ei
15	Igal kaubal, tootel ja teenusel on õige lühikirjeldus.	Ei	Ei
16	Igal kaubal, tootel ja teenusel on õige metakirjeldus.	Ei	Ei
17	Igal kaubal, tootel ja teenusel on registreeritud õige lao staatus.	Ei	Ei
18	Igal kaubal, tootel ja teenusel on märgitud õiged kategooriad.	Jah	Ei

Kõik nõuded tabelis 16 on semantilise õigsuse kohta e andmed peavad olema tähenduse

poolest õiged. Nõuetest on täidetud 7 ja täitmata 11. Seega töö alustamise hetkel on andmebaasis olevad andmed ei vasta õigsuse kehtivatele nõutele.

Näiteks ettevõtte X e-poes on avalikustatud kaupu, mille hind e-poes on märgitud „0.00€“, kuid tegelikult peab nende hind olema „1500.00€“. Samuti leidub veebipoes kaupu, mille kogus laos on 1 ja staatus peaks olema „instock“, kuid nende staatuseks on märgitud „onbackorder“ ja kogus laos on 0.

Tarkvara (sh andmebaasisüsteem) ei saa kontrollida seda, kas e-poes olevad andmed on õiged (näiteks, kas kokkulepitud teenuse hind on 100 või 110 eurot). Küll aga on võimalik leida reeglid, millele õiged andmed igal juhul ei vasta. See tähendab, et kui registreerida andmeid, mis nendele reeglitele ei vasta, siis need on kindlasti valed andmed. Selliseid reegleid on võimalik tarkvara abil jõustada (näiteks deklareerides andmebaasis kitsendusi) ja reeglite täidetust kontrollida.

Tabelis 17 tuuakse välja piirangud andmetele, mis aitavad leida ebaõigeid andmeid. Samas vastavus nendele nõuetele ei tähenda veel tingimata, et andmed on õiged. Tabel on sorteeritud nõuete kriitilisuse järgi ja selle sees tähestikulises järjekorras.

Tabel 17. Õigsuse nõuded, mille täidetust on võimalik IT-süsteemil kontrollida.

ID	Nõue	Täidetud	On kriitiline
1	Iga avalikustatud kauba, toote ja teenuse hind peab olema nullist suurem arv.	Jah	Jah
2	Iga kauba, toote ja teenuse üks võtmesõna peab olema sama, mis tootekaardi nimetus.	Ei	Jah
3	Iga tellimusele ja hinnapakkumise taotlusele lisatud kauba/toote/teenuse identifikaator on andmebaasis registreeritud kauba/toote/teenuse identifikaatorite hulgas.	Jah	Jah
4	Iga tellimusele ja hinnapakkumise taotlusele märgitud registreeritud kliendi identifikaator on andmebaasis registreeritud kliendi identifikaatorite hulgas.	Jah	Jah
5	Iga tellimuses ja hinnapakkumise taotluses esitatud kogus peab olema suurem 0st.	Jah	Jah
6	Igal e-poe kasutajal on registreeritud e-posti aadress, milles on olemas „@“ märk.	Jah	Jah
7	Igal e-poe kasutajal on unikaalne kasutajanimi.	Jah	Jah
8	Igal kaubal, tootel ja teenusel olev nimetus peab olema unikaalne.	Jah	Jah



ID	Nõue	Täidetud	On kriitiline
9	Igal kaubal, tootel ja teenusel oleva koodi (SKU) ja keele kombinatsioon peab olema unikaalne.	Jah	Jah
10	Igal tellimust ja hinnapakumise taotlust esitanud kliendil peab olema märgitud e-posti aadress, milles on olemas „@“ märk.	Jah	Jah
11	Tootekaardi loomisel tuleb arvestada kodulehe optimeerimise pistikprogrammi soovitustega.	Ei	Jah
12	Ühegi kauba, toote ja teenuse nimetus ei tohiks sisaldada kirjavahemärke (v.a punkti ja sidekriipsu).	Ei	Jah
13	Iga kaubale, tootele ja teenusel märgitud lao staatus on lubatud lao staatuste hulgas ( <i>instock, onbackorder, outofstock</i> ).	Jah	Ei
14	Iga kauba, toote ja teenuse metakirjeldus on kirjavigadeta	Jah	Ei
15	Iga kauba, toote ja teenuse nimetus on kirjavigadeta	Jah	Ei
16	Iga kaubale, tootele ja teenusele registreeritud kategooria identifikaator on andmebaasis registreeritud kategooria identifikaatorite hulgas.	Jah	Ei
17	Iga kaubale, tootele ja teenusele registreeritud pildi identifikaator on andmebaasis registreeritud pildi identifikaatorite hulgas.	Ei	Ei
18	Iga kaubale, tootele ja teenusele registreeritud tootja identifikaator on andmebaasis registreeritud tootja identifikaatorite hulgas.	Jah	Ei
19	Iga teenuse SKU algab järgmiselt „DF-K-“.	Ei	Ei
20	Igal kaubal, tootel ja teenusel, millel on valitud mitu kategooriat, peab olema märgitud üks kategooria peamiseks ( <i>primary</i> ).	Ei	Ei
21	Kirjeldus kaubal tootel ja teenusel peab sisaldama vähemalt ühte võtmesõna päises ja ühte võtmesõna tavatekstis.	Ei	Ei
22	Kirjeldus kaubal, tootel ja teenusel peab olema vähemalt 560 tähemärki pikk.	Ei	Ei
23	Kõik kauba, toote ja teenuse kirjeldused on kirjavigadeta.	Ei	Ei
24	Kõik kauba, toote ja teenuse lühikirjeldused on kirjavigadeta	Ei	Ei
25	Kõik kauba, toote ja teenuse märksõnad on kirjavigadeta	Jah	Ei

ID	Nõue	Täidetud	On kriitiline
26	Kõik kauba, toote ja teenuse sildid on kirjavigadeta	Jah	Ei
27	Lühikirjeldus kaubal, tootel ja teenusel peab olema vähemalt 130 tähemärki pikk.	Ei	Ei
28	Lühikirjeldus kaubal, tootel ja teenusel peab sisaldama vähemalt ühte võtmesõna.	Ei	Ei
29	Ühegi kauba, toote ja teenuse nimetuse pikkus ei või olla lühem kui 9 tähemärki.	Ei	Ei
30	Ühegi kauba, toote ja teenuse nimetuse pikkus ei tohiks ületada 60 tähemärki.	Ei	Ei

Nõuetest on täidetud 16 ja täitmata 14. Seega töö alustamise hetkel andmebaasis olevad andmed ei vasta õigsuse kehtivatele nõutele. Kontrollpäringute abil saab kontrollida andmete vastavust vähemalt 19-le Tabelis 17 esitatud andmekvaliteedi nõudele. Nõudeid 9, 11, 17, 21, 28 tuleb (vähemalt esialgu) töötajal käsitsi kontrollida ja nõuete 14, 15, 23-26 korral tuleb kaasata programm, mis pakub õigekirja kontrolli.

### 6.1.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida andmete vastavust järgmiste meetmetega [1], [12]:

- Juhul, kui organisatsiooni töötaja on leidnud vea andmetes, siis tuleks tal see dokumenteerida ja kui tal on võimalus endal viga parandada, siis koheselt seda teha ja dokumenteerida Trello vea parandus. Juhul, kui töötajal ei ole paranduse tegemiseks piisavalt õigusi või töötaja ei ole teadlik, millised on õiged andmed, siis tuleb töötajal antud viga samuti Trello dokumenteerida ja märkida, et viga ootab parandamist.
- Andmete parandamise vajadus võib tekkida sellest, et:
  - töötaja vaatab ise perioodiliselt andmeid üle või märkab nõuetel mittevastavust muu töö käigus,
  - klient võtab organisatsiooniga ühendust ja teatab, et organisatsiooni veebisaidil on ebaõiged andmed,
  - partnerid saavad uudise selle kohta, et nende saadetud andmetes oli viga.
- Õigsust on keeruline/võimatu automaatselt kontrollida ja seega eeldab selle tunnuse tagamine täpsust, tähelepanelikkust ja kogemust nii andmete registreerimisel (et mitte teha hooletusvigu) kui kontrollimisel (et märgata

probleeme). Samuti eeldab see, et andmete registreerija on saanud registreerimiseks õiged andmed.

- Teksti süntaktilise õigsuse tagamiseks tuleks kasutada õigekirja kontrolli pakkuvate programmide abi.
- Tarkvara ei saa kontrollida andmete õigsust. Küll aga on võimalik leida reeglid, millele õiged andmed kindlasti peavad vastama ning kontrollida andmete puhul nende reeglite kehtivust. Kui andmed reeglitele ei vasta, siis on need kindlasti ebaõiged ja tuleb parandada. Kui andmed reeglitele vastavad, siis ei taga see veel andmete õigsust. Sellised reeglid on esitatud tabelis 17 ja nende alusel plaanitakse luua enamike päringuid andmekvaliteedi töölauale. Samas andmebaasi või rakenduse tasemel kontrollide jõustamine ei tule praegu kõne alla, sest ettevõtte kasutab valmistarkvara ja seda uuendatakse iga mõne aja tagant ning ettevõttel puudub kompetents selliste muudatuste sisseviimiseks ja hooldamiseks.

### **6.1.3 Kriitiliste nõuete põhjendus**

Tabeli 16 kriitiliste nõuete põhjendused.

Nõue 1 on kriitiline, sest tänu õigetele võtmesõnade olemasolule on e-pood otsingumootorite poolt paremini indekseeritav ja selle tulemusena paremini otsingus kuvatav. Näiteks enne võtmesõnade sisestamist võis toode olla Google otsingus viiendal leheküljel, kuid peale õigete võtmesõnade sisestamist võib juhtuda, et seda kuvatakse juba esimesel või teisel leheküljel. Siinkohal tasub ka märkida, et enamus inimesi ei pruugi kauba, toote või teenuse otsingut teisel, kolmandal või enamal leheküljel jätkata. Kõige tavapärasem on olukord, et kliendid leiavad neile vajaliku/sobiva informatsiooni esimeselt leheküljelt.

Nõue 2 on kriitiline, sest vastasel juhul ei ole ettevõttel võimalik saata kliendile arvet või hinnapakkumist.

Nõue 3 on kriitiline, sest vastasel juhul ei saa ettevõtte töötajad esitada arvet ja hinnapakkumist.

Nõue 4 on kriitiline, sest vastasel juhul ei ole kasutajal võimalik e-poodi sisse logida (üks viis sisselogimisel on e-posti aadressi kasutamine) ja vajadusel taastada parooli. Samuti võib ettevõtte saata informatsiooni valele inimesele.

Nõue 5 on kriitiline, sest üks e-poodi sisselogimise viise on kasutades kasutajanime.

Nõue 6 on kriitiline, sest mõningad kliendid esitavad tellimusi toetudes kaupade, toodete või teenuste piltidele, mitte lühikirjelduses või kirjelduses toodud informatsioonile. Pildilt vaadatakse, kuidas kaup või toode välja näeb või kuidas teenust osutatakse. Kaupade, toodete ja teenuste komplektide puhul võidakse lugeda pildilt välja selle, mis komplekti kuulub. Näiteks juhul kui klient näeb pildil mõnda kaupa või toodet, mida komplektis tegelikult ei ole, võib ta ikkagi arvata, et antud kaup või toode kuulub komplekti. Juhul kui klient saab kauba või toote kätte, siis on ta ebameeldivalt üllatunud.

Nõue 7 on kriitiline, sest juhul kui kaubal, tootel või teenusel on vale nimetus, siis ei oska kliendid ja ka ettevõtte töötajad seda üles leida.

Nõue 8 on kriitiline, sest sarnaste kaupade, toodete ja teenuste müümisel on nii kliendil kui ka ettevõtte töötajal nende vahel kergem vahet teha, kui on registreeritud õige SKU kood. Ka kliendi tellimust on ettevõtte töötajatel kergem ja kiirem komplekteerida, kui kaupu või tooteid saab otsida SKU koodi järgi. Samuti on SKU koodi järgi mugav otsida kaupa, toodet ja teenust e-poe või Google otsingumootori abil.

Nõue 9 on kriitiline, sest juhul kui ettevõtte lülitab sisse kohese kaupade, toodete ja teenuste eest tasumise, siis võib ta kanda kahjumit ja/või rikkuda vale hinna küsimisega ja hiljem raha juurde küsimise/tagasimaksmisega ettevõtte mainet. Näiteks oletame, et klient ostis e-poes endale kauba, mille hind e-poes oli 15€. Ettevõtte töötaja tegi eelnevalt kauba hinna sisestamisel vea ja kauba tegelik müügihind on 150€. Makstud ja tegeliku maksumuse hinna vahe on 135€. Nüüd on ettevõtte kohustatud müüma kaupa hinnaga 15€ või tellimuse tühistama ja kliendile raha tagasi kandma.

Nõue 10 on kriitiline, sest vastasel juhul võib esitada arveid, hinnapakumisi valele kliendile. Samuti võidakse tellimuse komplekteerimise järel saata kuller valele aadressile.

Nõue 11 on kriitiline, sest sildid aitavad ettevõtte veebipoes otsingu abil paremini kaupu, tooteid ja teenuseid temaatika järgi üles leida.

Tabeli 17 kriitiliste nõuete põhjendused.

Nõue 1 on kriitiline, sest ettevõtte ei tegele tasuta kaupade, toode ja teenuste pakkumisega. Igal kaubal, tootel ja teenusel on olemas oma maksumus.

Nõue 2 on kriitiline, sest juhul kui nimetus ei sisalda võtmesõna, muutub kauba, toote või teenuse veebilehe skoor madalamaks. Madal veebilehe skoor omakorda muudab kauba, toote või teenuse otsingumootorite poolt halvemini otsitavaks ja indekseeritavaks.

Nõue 3 on kriitiline, sest vastasel juhul tähendab see, et klient on tellinud e-poes kauba, toote või teenuse, mida e-poes ei ole olemas.

Nõue 4 on kriitiline, sest vastasel juhul ettevõtte ei tea, kellele saata arve või hinnapakkumine.

Nõue 5 on kriitiline, sest negatiivse koguse korral peaks toimuma justkui kauba/toote tagastusprotsess, milles klient tagastab kauba/toote ja ettevõtte tagastab tellimuse raha. Null koguse juures ei toimu justkui mingisugust ostu.

Nõue 6 on kriitiline sest, vastasel juhul ei ole tegemist korrektse e-posti aadressiga. Sellisel juhul ei õnnestu ettevõtte töötajatel saata arveid, hinnapakkumisi kasutaja e-postile. Samuti ei ole kasutajal võimalik e-posti kasutades logida sisse e-poodi ja vajadusel taastada enda parooli.

Nõue 7 on kriitiline, sest juhul kui kahel või enamal kasutajal on sama kasutajanimi, võib tekkida olukord, mil üks kasutaja võib saada ligi teise kasutaja andmetele.

Nõue 8 on kriitiline, sest juhul kui leiduvad kaubad, tooted ja teenused, mis kannavad ühesuguseid nimetusi, tekitab see arusaamatusi klienditele ja ka organisatsiooni töötajatele.

Nõue 9 on kriitiline, sest vastasel juhul esineb andmebaasis korduvaid kirjeid.

Nõue 10 on kriitiline sest, vastasel juhul ei ole tegemist korrektse e-posti aadressiga. Ettevõtte töötajatel ei õnnestu saata arveid, hinnapakkumisi kliendi e-postile.

Nõue 11 on kriitiline, sest vastasel juhul on veebileht otsingumootorite poolelt halvemini indekseeritav, mis omakorda tähendab, et veebileht ei pruugi otsingus olla konkurentides eespool ja ettevõtte võib kaotada oma kliente.

Nõue 12 on kriitiline, sest muud sümbolid, v.a tähed ja numbrid, ei ole lingis märgitud. Kauba, toote või teenuse veebilehe skoori määramisel arvestatakse ka märksõna sisaldust

lingis e juhul kui lingis ei sisaldu märksõna kukub veebilehe skoor automaatselt paar punkti.

## 6.2 Täielikkus

Andmekvaliteedi tunnus „täielikkus“ iseloomustab seda, kuivõrd täielikud on andmed ehk kas andmed sisaldavad selliste andmeelementide väärtuseid, mida organisatsioon on kohustuslikuks määranud [12].

### 6.2.1 Olemasoleva andmebaasi vastavus

Tabelis 18 tuuakse välja täielikkuse nõuded ja hinnang nende täidetusele olemasolevas andmebaasis. Tabel on sorteeritud nõuete kriitilisuse järgi ja selle sees tähestikulises järjekorras.

Tabel 18. Täielikkuse nõuded ja nende täidetud ning kriitilisus ettevõttes X.

ID	Nõue	Täidetud	On kriitiline
1	Iga tellimus ja hinnapakkumise taotlus peab sisaldama vähemalt ühte toodet, kaupa või teenust.	Jah	Jah
2	Igal avalikustatud kaubal, tootel ja teenusel on registreeritud vähemalt üks hind.	Ei	Jah
3	Igal avalikustatud kaubal, tootel ja teenusel on registreeritud vähemalt üks pilt.	Ei	Jah
4	Igal kaubal, tootel ja teenusel peab olema nimetus.	Jah	Jah
5	Igal kaubal, tootel ja teenusel peab olema SKU kood.	Ei	Jah
6	Igal kaubal, tootel ja teenusel peab olema vähemalt üks võtmesõna.	Ei	Jah
7	Igal tellimisel ja hinnapakkumise taotlusel on märgitud selle esitamise aeg.	Jah	Jah
8	Igal tellimisel ja hinnapakkumise taotlusel peab olema märgitud klient.	Jah	Jah
9	Igal kasutajal on märgitud e-poe kasutajaks registreerimise aeg.	Jah	Ei
10	Igal kauba, tootel ja teenusel on märgitud vähemalt üks <i>up-sell</i> e kaup, toode või teenus, mida soovitatakse selle kauba, toode või teenuse asemele.	Ei	Ei
11	Igal kaubal ja tootel ja tootega komplekti pandud teenusel (ühel tootekaardil nii teenus kui toode) on registreeritud lao kogus.	Ei	Ei

ID	Nõue	Täidetud	On kriitiline
12	Igal kaubal, tootel ja teenusel on märgitud vähemalt üks <i>cross-sale</i> e kaup, toode või teenus, mida soovitatakse lisaks sellele kaubale, tootele või teenusele.	Ei	Ei
13	Igal kaubal, tootel ja teenusel on märgitud vähemalt üks silt ( <i>tag</i> ).	Ei	Ei
14	Igal kaubal, tootel ja teenusel on märgitud vähemalt üks tootja.	Ei	Ei
15	Igal kaubal, tootel ja teenusel on registreeritud lao staatus.	Ei	Ei
16	Igal kaubal, tootel ja teenusel peab olema kirjeldus.	Ei	Ei
17	Igal kaubal, tootel ja teenusel peab olema lühikirjeldus.	Jah	Ei
18	Igal kaubal, tootel ja teenusel peab olema metakirjeldus.	Ei	Ei
19	Igal kaubal, tootel ja teenusel on märgitud vähemalt üks kategooria.	Ei	Ei
20	Igal teenusel, mis on tüüpi koolitus või kursus, on lahti kirjutatud koolituse/kursuse eesmärk, väljundid, teema, toimumise asukoht ja keel.	Ei	Ei
21	Kõik Eestis müüdavad avalikustatud kaubad, tooted ja teenused on kättesaadavad eesti, inglise ja vene keeles.	Ei	Ei

Nõuetest on täidetud 6 ja täitmata 15. Seega töö alustamise hetkel andmebaasis olevad andmed ei vasta täielikkuse tunnusele kehtivatele nõuetele.

Näiteks ettevõtte X korraldatavatel koolitustel ei ole märgitud tootjat, kuid peaks olema märgitud „Ettevõtte X“. Avalikustatud kaupadel, toodetel ja teenustel on registreeritud kategooriad. Samas avalikustamata kaupadel, toodetel ja teenustel ei ole registreeritud kategooriat. Mõningatel kaupadel ja teenustel ei ole märgitud hinda ja kliendid näevad neid veebipoes hinnaga 0€.

### 6.2.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida andmete vastavust järgmiste meetmetega [1], [12]:

- Juhul, kui organisatsiooni töötaja märkas, et andmed ei vasta täielikkuse nõuetele, siis tuleks tal see dokumenteerida ja kui tal on võimalus endal parandada andmed nõuetele vastavaks, siis koheselt seda teha ja dokumenteerida Trello vea parandus. Juhul, kui töötajal ei ole paranduste tegemiseks piisavalt õigusi või

töötaja ei ole teadlik, mida tal tuleb andmete nõuetele vastavaks muutmiseks teha, siis tuleb töötajal antud mittevastavus samuti Trellos dokumenteerida ja märkida, et see ootab parandamist.

- Korraldada koolitusi ja luua juhendeid. Õpetada organisatsiooni töötajaid, et nad järgiksid andmete sisestamisel juhendeid.
- Andmete parandamise vajadus võib tekkida sellest, et:
- töötaja vaatab ise perioodiliselt andmeid üle või märkab nõuetele mittevastavust muu töö käigus,
- klient võtab organisatsiooniga ühendust, et küsida täpsustavaid küsimusi kauba, toote või teenuse kohta,
- partnerid saadavad uudise selle kohta, et nende saadetud andmetes oli viga.
- Luua mittetäielike andmete otsimiseks päringud (osana töölaust). Andmebaasi või rakenduse tasemel piirangute jõustamine ei tule kõne alla samal põhjusel kui õigsuse tunnuse korral.

### **6.2.3 Kriitiliste nõuete põhjendus**

Nõue 1 on kriitiline, sest vastasel juhul tähendab see, et klient on esitanud tühja tellimuse ja ettevõtte töötajad ei tea, mida komplekteerida või millele esitada hinnapakumise.

Nõue 2 on kriitiline, sest juhul kui hind ei olnud registreeritud, siis näidatakse e-poe kasutajaliideses kaubale, tootele või teenusel automaatselt hinnaks 0. See tähendab, et klient näeb veebipoes, et kauba hind on 0€ (see on tasuta), mitte seda, et kaubal, tootel või teenusel olev hind on puudu.

Nõue 3 on kriitiline, sest mõningad kliendid esitavad tellimusi toetudes kaupade, toodete või teenuste piltidele, mitte lühikirjelduses või kirjelduses toodud informatsioonile.

Nõue 4 on kriitiline, sest juhul kui kaubal, tootel või teenusel puudub nimetus, siis ei ole ettevõtte klientidel ega töötajatel võimalik kaupa, toodet või teenust nimetuse järgi üles otsida.

Nõue 5 on kriitiline, sest SKU koodi puudumisel on raske eristada sarnaseid tooteid, kaupu ja teenuseid. SKU koodi on mugav kasutada nii komplekteerimisel kui ka kaupade, toodete ja teenuste otsimisel.



Nõue 6 on kriitiline, sest võtmesõnade puudumisel võib veebilehekülje indekseeritus otsingumootorite poolt olla halb ja seetõttu võib ettevõtte kaup, toode või teenuse asuda mitte esimesel otsingutulemuste leheküljel, vaid teisel või kolmandal. Tavaliselt kliendid isegi ei vaata nii kaugele, kui neil selleks just erilist põhjust ei ole. Seega sellises olukorras võivad konkurendid olla eespool ja kliendid võivad sattuda nende e-poodi ja sooritada ostu hoopis sealt.

Nõue 7 on kriitiline sest, tellimuste tagastamisel ja komplekteerimisel võib see mängida tähtsat rolli. Näiteks kui kaks klienti tellisid e-poest samaaegselt ühte ja sama kaupa või toodet ning laos ei ole mõlema kliendi soove arvestades piisavat kauba või toote kogust, siis komplekteeritakse algul esimesena tellimuse esitanud kliendi tellimus.

Nõue 8 on kriitiline, sest vastasel juhul on tegemist tellimusega või hinnapakkumisega, mis on esitatud tundmata kliendilt. Sellisel korral ei ole võimalik ei esitada arvet, ei saa saata hinnapakkumist ja kindlasti ei saa hakata komplekteerima tellimust.

## 6.3 Kooskõla

Andmekvaliteedi tunnus „kooskõla“ iseloomustab andmete vasturääkivuse määra ja kas andmed on kooskõlas teiste andmetega [12]. Kooskõla on seotud mitteliiasusega – mida suurem on andmete liiasus andmebaasis, seda rohkem võib nendes olla kooskõla probleeme.

### 6.3.1 Olemasoleva andmebaasi vastavus

Tabelis 19 tuuakse välja kooskõla nõuded ja hinnang nende täidetusele olemasolevas andmebaasis. Tabel on sorteeritud nõuete kriitilisuse järgi ja selle sees tähestikulises järjekorras.

Tabel 19. Kooskõla nõuded ja nende täidetud ning kriitilisus ettevõttes X.

ID	Nõue	Täidetud	On kriitiline
1	<p>Igal avalikustatud kaubal, tootel ja tootega komplekti pandud teenusel, peavad olema kooskõlas lao staatus ja kogus:</p> <ul style="list-style-type: none"> <li>▪ Kauba, toote ja tootega komplekti pandud teenuse staatus <i>instock</i> tähendab, et kogus peab olema suurem 0st.</li> <li>▪ Kauba, toote ja teenuse staatus <i>onbackorder</i> tähendab, et kogus peab olema 0.</li> </ul>	Ei	Jah

ID	Nõue	Täidetud	On kriitiline
	<ul style="list-style-type: none"> <li>▪ Kauba, toote ja teenuse staatus <i>outofstock</i> tähendab, et kogus peab olema 0.</li> </ul>		
2	Kaupade, toodete ja teenuste erinevate keelte jaoks loodud avalikustatud versioonides peavad olema samad kogused.	Jah	Jah
3	Kaupade, toodete ja teenuste erinevate keelte jaoks loodud avalikustatud versioonides peavad olema samad tootjad.	Jah	Jah
4	Kaupade, toodete ja teenuste erinevate keelte jaoks loodud avalikustatud versioonides peavad olema üldjuhul samad hinnad.	Jah	Ei
5	Kaupade, toodete ja teenuste erinevate keelte jaoks loodud avalikustatud versioonides peavad olema samad pildid.	Jah	Ei
6	Kaupade, toodete ja teenuste erinevate keelte jaoks loodud avalikustatud versioonides peavad olema samad kategooriad.	Jah	Ei

Nõuetest on täidetud 5 ja täitmata 1. Seega töö alustamise hetkel andmebaasis olevad andmed ei vasta kooskõla tunnusele kehtivatele nõuetele.

Olemasolevas andmebaasis on kõik avalikustatud kaubad, tooted ja teenused samade hindadega eesti-, vene-, inglise-, läti- ja leedukeelsetes veebisaidi versioonides. Kõik avalikustatud kaubad, tooted ja teenused on sama lao staatuse ja kogusega eesti-, vene-, inglise-, läti- ja leedukeelsetes veebisaidi versioonides. Samas leidub näiteks kaupu, kus on märgitud staatuseks „instock“, aga kogus on „NULL“ ehk pole määratud. Nõude 4 puhul tuleb märkida, et ettevõtte võib otsustada müüa erinevatesse riikidesse erineva hinnaga kaupu, tooteid ja teenuseid.

### 6.3.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida andmete vastavust järgmiste meetmetega [1], [12]:

- Juhul, kui organisatsiooni töötaja märkas, et andmebaasis olevad andmed on vasturääkivad, siis tuleb need andmed dokumenteerida Trellos ja võimalusel vasturääkivused parandada ning märkida Trellos, et probleem lahendati. Juhul, kui töötajal ei ole paranduste tegemiseks piisavalt õiguseid, siis tuleb

dokumenteerida Trellos vasturääkivad andmed ja märkida, et antud probleemiga on vaja tegeleda.

- Andmete parandamise vajadus võib tekkida sellest, et:
- töötaja vaatab ise perioodiliselt andmeid üle või märkab nõuetele mittevastavust muu töö käigus,
- klient võtab organisatsiooniga ühendust, et teatada mittekooskõlalistest andmetest,
- partnerid saavad uudise selle kohta, et nende saadetud andmetes oli viga.
- Koolitada organisatsiooni töötajaid, et nad teaksid, milliseid juhendeid kasutada ja millistest allikatest tuleb andmeid võtta.
- Luua mittekooskõlaliste andmete otsimiseks päringud (osana töölaust).

### **6.3.3 Kriitiliste nõuete põhjendus**

Nõue 1 on kriitiline, sest tihtipeale vaatavad ettevõtte töötajad, kellel ei ole ligipääsu ettevõtte laosüsteemile, kaupade ja toodete olemasolu laos e-poest nende staatuse järgi.

Nõue 2 on kriitiline, sest ettevõttel X on vaid üks ladu, milles hoitakse kaupu. Seega puudub võimalus tellida kaupa ühest ettevõtte laost teise lattu. Seega, kui kaupa ostetakse eestikeelsest e-poe versioonist, siis muutub automaatselt kauba kodus ka vene-, inglise-, läti- ja leedukeelsetest e-poe versioonides, sest kauba kogus on kõigil neil versioonidel ühesugune.

Nõue 3 on kriitiline, sest juhul kui klient vaatab ühte toodet mitmes keeles ja märkab, et samal kaubal, tootel või teenusel on erinevad tootjad, siis tekitab see kliendis kahtlust (kahtlus ettevõtte usaldusväärsuse ja maine suhtes).

## **6.4 Usaldusväärsus**

Andmekvaliteedi tunnus „usaldusväärsus“ iseloomustab seda, kuivõrd tõesteks ja usaldatavateks võib pidada andmeid [12].

### **6.4.1 Olemasoleva andmebaasi vastavus**

Tabelis 20 tuuakse välja usaldusväärsuse nõuded ja hinnang nende täidetusele olemasolevas andmebaasis. Tabel on sorteeritud nõuete kriitilisuse järgi ja selle sees tähestikulises järjekorras.

Andmed jagunevad:

- Tootjalt pärinevad andmed (hind, nimetus, kirjeldus, pilt), mida sisestab ettevõtte. Neid andmeid saab lugeda kõige usaldusväärsemaks.
- Kauba, toote, teenuse kasutajatelt pärinevad kogemuslikud andmed, mida sisestab ettevõtte. Need andmed sisestatakse ka kirjeldusena ja ettevõtte üritab neid enne sisestamist üle kontrollida. Nende andmete usaldusväärsus on väiksem.
- Klientide kommentaarid. Need on kõige vähem usaldusväärsed. Kliendi kommentaarid avalikustatakse alles peale ülevaatamist.

Tabel 20. Usaldusväärsuse nõuded ja nende täidetus ning kriitilisus ettevõttes.

ID	Nõue	Täidetus	On kriitiline
1	Kliendi kasutajakogemuse kirjeldus, mis toote, kauba või teenuse juures registreeritakse, on eelnevalt üle kontrollitud.	Jah	Jah
2	Kliendi kasutajakogemuse kirjelduse juures on võimalusel viide allikale, kust see pärineb.	Jah	Jah
3	Kliendi kommentaar on enne avalikustamist üle kontrollitud.	Jah	Jah

Kõik nõuded on täidetud. Seega töö alustamise hetkel andmebaasis olevad andmed vastavad usaldusväärsuse tunnusele kehtivatele nõuetele.

Näiteks kõik klientide käest saadud kasutuskogemused, millest kauba või toote kaartidel e-poes on kirjutatud, on eelnevalt läbinud ühe/kahe ettevõtte töötaja poolse kontrolli. Juhul, kui ettevõtte töötajad ei ole kindlad millegi kasutuskogemuses, siis pöörduv ettevõtte tootja poole, et täpsustada kasutusvõimalusi ja teemasid.

#### 6.4.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida andmete vastavust järgmiste meetmetega [1], [12]:

- Andmete parandamise vajadus võib tekkida sellest, et töötaja vaatab ise perioodiliselt andmeid üle või märkab nõuetele mittevastavust muu töö käigus.
- Juhul, kui organisatsiooni töötajas tekib kahtlus andmete usaldusväärsuse suhtes, siis tuleb need andmed dokumenteerida Trellos ning võimalusel parandada ja märkida Trellos, et usaldusväärsuse probleem parandati. Juhul, kui töötajal ei ole

piisavalt õigusi, siis tuleb andmed Trellos dokumenteerida ja märkida, et antud probleemiga on vaja tegeleda.

- Analüüsida kogutud andmeid ja võrrelda neid usaldusväärsete teabeandjate andmetega. Selleks tuleb esialgu valida välja teabeandjad, keda organisatsioon usaldab. Sellisteks teabeandjateks on e-poe puhul näiteks kaupade tootjad – kellel, kui mitte neil, peaks olema usaldusväärne teave nende enda toodete kohta.
- Pidada kinni reeglitest, et klientide kasutuskogemuse andmed ja kommentaarid vajavad enne avalikustamist üle kontrollimist. Koolitada töötajaid nendest reeglitest kinni pidama.
- Luua roppsõnu või kahtlaseid linke sisaldavate kommentaaride otsimiseks päringud (osana töölauast). Sellised päringud võiksid vaadelda ka avalikustatud kommentaare, et leida kahe silma vahele jäänud ebasobivaid kommentaare.

### **6.4.3 Kriitiliste nõuete põhjendus**

Nõue 1 on kriitiline, sest juhul kui ettevõtte e-poes olevad andmed ei ole usaldusväärsed, võib see rikkuda ettevõtte mainet ja olla otsustav põhjus, et kliendid ostaksid sarnast kaupa, toodet või teenust konkurentidelt.

Nõue 2 on kriitiline, sest viide allikale võib tõsta klientide silmis usaldusväärset ettevõtte enda ja kauba, toote või teenuse suhtes. Lisaks saab e-poe klient sellest allikast lugeda juurde informatsiooni kauba ja toote kasutamise või teenuse osutamise kohta.

Nõue 3 on kriitiline, sest ettevõtte veebilehele ei ole oodatud kommentaarid, mis sisaldavad roppsõnu või kahtlaseid linke.

## **6.5 Ajakohasus**

Andmekvaliteedi tunnus „ajakohasus“ iseloomustab seda, kui võrd andmed vastavad vajadustele ja nõuetele ajaliselt [12]. Näiteks hinnapakkumise loomise ja kliendile saatmise korral on tähtis, millisel päeval oli hinnapakkumine loodud ja mitu päeva hinnapakkumine kehtib. Kui klient reageerib hinnapakkumisele peale kehtivusaja lõppu, siis ei kasuta ta ajakohaseid andmeid.

### **6.5.1 Olemasoleva andmebaasi vastavus**

Tabelis 21 tuuakse välja ajakohasuse nõuded ja hinnang nende täidetusele olemasolevas andmebaasis. Tabel on sorteeritud nõuete kriitilisuse ja tähestikulises järjekorras.

Tabel 21. Ajakohasuse nõuded ja nende täidetud ning kriitilisus ettevõttes.

ID	Nõue	Täidetud	On kriitiline
1	Kaupade, toodete ja teenuste andmed muudetakse niipea kui andmed muutusid (nt hind, lao kogus, kirjeldus).	Ei	Jah

Nõue ei ole täidetud. Seega töö alustamise hetkel andmebaasis olevad andmed ei vasta ajakohasuse tunnusele kehtivatele nõuetele.

Näiteks ettevõtte X veebipoes leidub kaupasid, mille informatsioon on aegunud, kuid ei ole vale – näiteks ajaga on lihtsalt juurde tulnud informatsiooni, mida konkreetse kaubaga on võimalik teha.

### 6.5.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida andmete vastavust järgmiste meetmetega [1], [12]:

- Juhul, kui organisatsiooni töötaja märkas, et andmed ei ole ajakohased, siis tuleb need andmed dokumenteerida Trellos ning võimalusel parandada ja märkida Trellos, et ajakohasuse nõuete probleem parandati. Juhul, kui töötajal ei ole parandamiseks piisavalt õigusi, siis tuleb andmed Trellos dokumenteerida ja märkida, et antud probleemiga on vaja tegeleda.
- Andmete parandamise vajadus võib tekkida sellest, et:
- töötaja vaatab ise perioodiliselt andmeid üle või märkab nõuetele mittevastavust muu töö käigus,
- partneritel tuleb organisatsiooni informatsioon, et kaupade sisseostu hind muutub.

### 6.5.3 Kriitiliste nõuete põhjendus

Nõue 1 on kriitiline, sest selle mittetäitmine võib tuua endaga kaasa ettevõtte maine rikkumise või kahju kandmise. Näiteks, kui tootja muutis oma kauba sisseostuhinna – see tõusis 20%, siis tuleb ka ettevõttel määrata kaubale uus müügihind arvestades 20% suurust tõusu.

## 6.6 Reeglipärasus

Andmekvaliteedi tunnus „reeglipärasus“ näitab määra, kuidas andmed vastavad kehtestatud seadustele, formaadile ja struktuurile [1], [9]. E-poe puhul on olulisteks seadusteks Isikuandmete kaitse seadus (IKS) ja Tarbijakaitseseadus (TKS).

### 6.6.1 Olemasoleva andmebaasi vastavus

Tabelis 22 tuuakse välja reeglipärasuse nõuded ja hinnang nende täidetusele olemasolevas andmebaasis. Tabel on sorteeritud nõuete kriitilisuse ja tähestikulises järjekorras.

Tabel 22. Reeglipärasuse nõuded ja nende täidetus ning kriitilisus ettevõttes.

ID	Nõue	Täidetus	On kriitiline
1	Ettevõtte e-poes on kättesaadavad privaatsuspoliitika, ostu- ja tarnetingimused, milles on märgitud e-poe kasutaja ja kliendi õigused ja kohustused.	Jah	Jah
2	Iga tellimuse puhul registreeritakse tellitud tooted/kaubad, nende kogus ja toote/kauba väljastamise aeg, et saaks täita nõuet, mille kohaselt Internetis sõlmitud lepingust on kliendil õigus taganeda 14 päeva jooksul alates toote kättesaamisest.	Jah	Jah
3	Iga teenuse tellimuse juures registreeritakse tellimuse tegemise aeg, et saaks täita nõuet, mille kohaselt Internetis sõlmitud lepingust on kliendil õigus taganeda 14 päeva jooksul alates teenuse tellimisest.	Jah	Jah
4	Vastavalt Isikuandmete kaitse seadusele (IKS) ei kogu ettevõtte liigseid andmeid klientide kohta.	Jah	Jah
5	Vastavalt Isikuandmete kaitse seadusele (IKS) on kliendil õigus olla ära unustatud.	Jah	Jah

Kõik nõuded on täidetud. Seega töö alustamise hetkel on reeglipärasuse andmekvaliteedi nõuded ettevõtte X poolt täidetud.

Ettevõtte X järgib Eestis kehtestatud seadusi, sh isikuandmete kaitse seadust ja tarbijakaitseseadust. Kõik e-poe kasutajate ja klientide õigused ja kohustused on üles märgitud e-poe ostutingimustes ja privaatsuspoliitikas. Näiteks on ostutingimustes toodud välja peatükk kaupade ja toodete tagastamise kohta, milles on mainitud, et kliendil on võimalik tagastada kaup või toode 14 päeva jooksul, kui ostetud kaup või toode mingil põhjusel ei sobi.

### 6.6.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida vastavust järgmiste meetmetega [1], [12]:

- Juhul, kui organisatsiooni töötaja märkas, et andmed ei vasta reeglipärasuse nõuetele, tuleb need andmed dokumenteerida Trellos ning võimalusel parandada ja märkida Trellos, et probleem lahendati. Juhul, kui töötajal ei ole piisavalt õigusi, siis tuleb Trellos andmed dokumenteerida ja märkida, et antud probleemiga on vaja tegeleda.
- Andmete parandamise vajadus võib tekkida sellest, et töötaja vaatab ise perioodiliselt andmeid üle või märkab nõuetele mittevastavust muu töö käigus.

### 6.6.3 Kriitiliste nõuete põhjendus

Kõik tabelis 22 välja toodud reeglipärasuse andmekvaliteedi nõuded on kriitilised põhjusel, et vastasel juhul ei jälgigi ettevõtte X Eesti Vabariigis kehtestatud seadusi.

## 6.7 Konfidentsiaalsus

Konfidentsiaalsus iseloomustab andmete kättesaadavust vaid selleks volitatud tarbijatele ja selle mittekättesaadavust selleks volitamata tarbijatele [12]. Näiteks ligipääsu veebipoes esitatud tellimustele ja hinnapakumise taotlustele saavad vaid need ettevõtte töötajad, kes tegelevad tellimuste/hinnapakumise käitlemisega. Ettevõtte töötajatel, kes näiteks tegelevad õppematerjalide loomisega, ei ole vajadust saada ligipääsu klientide tellimustele/hinnapakumiste taotlustele.

### 6.7.1 Olemasoleva andmebaasi vastavus

Tabelis 23 tuuakse välja konfidentsiaalsuse nõuded ja hinnang nende täidetusele olemasoleva andmebaasi korral. Tabel on sorteeritud nõuete kriitilisuse ja tähestikulises järjekorras.

Tabel 23. Konfidentsiaalsuse nõuded ja nende täidetud ning kriitilisus ettevõttes.

ID	Nõue	Täidetud	On kriitiline
1	Ettevõtte X veebimajutusele ja kasutuses olevatele ( <i>live</i> ) andmebaasidele saab ligi vaid ettevõtte juhataja.	Jah	Jah
2	Hindamistöölaua prototüüpi on võimalik kasutada vaid neil töötajatel, kellel on seda vaja tööülesannete täitmiseks.	Jah	Jah
3	Kõik registreerinud kliendid omavad staatust „customer“ ja neil puudub ligipääs veebipoe administraatoripaneelile.	Jah	Jah
4	Tehtud varukoopiad on turvaliselt hävitatud.	Jah	Jah



ID	Nõue	Täidetud	On kriitiline
5	Tehtud varukoopiad on turvaliselt hoiustatud.	Jah	Jah
6	Toodete, kaupade ja teenuste andmetega kataloogidele pääsevad ligi vaid selleks volitatud isikud.	Jah	Jah
7	Töötajatel on võimalik kasutada veebisaidi administraatoripaneelil ainult selliseid osi, mis on neile tööülesannete täitmiseks vajalikud.	Jah	Jah
8	Veebipoele saab ligi vaid nendest riikidest, kuhu ettevõtte saab oma kaupu, tooteid ja teenuseid müüa.	Jah	Ei

Kõik nõuded on täidetud. Seega töö alustamise hetkel andmebaasis olevad andmed ja nende töötlemine vastavad konfidentsiaalsuse tunnusele kehtivatele nõuetele.

Näiteks veebipoe administraatoripaneelile pääsevad ligi vaid need töötajad, kes täidavad järgmiseid tööülesandeid: tegelevad veebilehtede loomisega, kaupade/toodete/teenuste sisestamise ja muutmisega, veebisaidi statistika jälgimisega ja veebipoes kasutatavate pistikprogrammide uuendamisega.

### 6.7.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida andmete vastavust järgmiste meetmetega [1], [12]:

- Perioodiliselt kontrollida konfidentsiaalsust. Juhul, kui konkreetset töötajat ei ole õigust omada ligipääsu mingitele andmetele (sh ta lahkub töölt), siis tuleb koheselt need õigused kustutada. Juhul, kui töötajal puudub ligipääs andmetele, millele on tal vastavalt tööülesannetele ligipääsu vaja, siis tuleb antud ligipääs töötajale anda.

### 6.7.3 Kriitiliste nõuete põhjendus

Nõue 1 on kriitiline, sest vastasel juhul võivad andmebaasis hoitavad andmed sattuda valedesse kättesse või olla hävitatud.

Nõuded 2, 3, 6 ja 7 on kriitilised, sest kasutajate üleliigsed andmete kasutamise võimalused loovad turvaauku. Kasutades neid turvaauke võivad kräkkerid veebisaidile/ettevõtte pilve paigutada faile viirustega või muuta näiteks arve numbrit (asetavad selle enda omaga ja kliendid tasuvad mitte ettevõttele, vaid kräkkeritele).

Nõuded 4 ja 5 on kriitilised, sest vastasel juhul võivad varukoopiad ja seetõttu ka varundatud andmed sattuda valedesse kättesse.

## 6.8 Ühekordsus

Andmekvaliteedi tunnus „ühekordsus“ iseloomustab seda, et organisatsioon ei küsi kasutajatelt samu andmeid mitu korda. Näiteks kui klient on registreerunud ja esitab tellimuse, siis ei ole tal vaja uuesti sisestada oma e-posti aadressi või telefoninumbrit [12].

### 6.8.1 Olemasoleva infosüsteemi vastavus

Tabelis 24 tuuakse välja ühekordsuse nõuded ja hinnang nende täidetusele olemasolevas infosüsteemis. Tabel on sorteeritud nõuete kriitilisuse ja tähestikulises järjekorras.

Tabel 24. Ühekordsuse nõuded ja nende täidetud ning kriitilisus ettevõttes.

ID	Nõue	Täidetud	On kriitiline
1	Sisselogitud kasutajalt ei küsita hinnapakkumise taotluse/ostu sooritamise korral kohustuslike kontaktandmeid (nt ees- ja perekonnanimi, e-posti, aadress).	Jah	Jah
2	Sisselogitud kasutajalt ei küsita kontaktvormi täitmisel ees- ja perekonnanime ning e-posti.	Ei	Jah

Nõuetest on täidetud 1 ja täitmata 1. Seega töö alustamise hetkel andmebaasis olevate andmete kasutamine ei vasta ühekordsuse tunnusele kehtivatele nõuetele.

Näiteks praegu, kui kliendil on olemas kasutajakonto, milles on eelnevalt juba sisestatud ees- ja perekonnanimi ning tellimuse kohaletoimetamise aadress, siis hinnapakkumise taotluse esitamisel määratakse automaatselt nende väljadesse väärtused.

### 6.8.2 Vastavuse tagamise meetmed

Organisatsioon saab kontrollida andmete vastavust järgmiste meetmetega [1], [12]:

- Perioodiliselt kontrollida organisatsiooni protsesse. Juhul, kui leidub protsesse, milles küsitakse registreeritud kasutajatelt ühte ja sama informatsiooni, siis tuleb see dokumenteerida Trellos ja märkida sellele tegelemise/tähtsuse staatused.

### 6.8.3 Kriitilised nõuded

Nõue 1 ja 2 on kriitilised, sest vastaselt juhul tuleb kliendil, kes on juba eelnevalt need väljad täitnud, uuesti nende juurde tagasi tulla ja täita. Tegemist on kliendi aja raiskamisega.

## 6.9 Mitteliiasus

Andmekvaliteedi tunnus „mitteliiasus“ iseloomustab seda, kuivõrd organisatsiooni andmekogu andmestruktuurides ja andmetöötluse protsessides on välditud andmete liiasust. Andmete liiasus tähendab, et ühte ja sama fakti on võimalik andmebaasist tuletada rohkem kui ühel viisil. [1], [12]

Andmete liiasust andmestruktuurides (nt tabelites) aitab vältida andmestruktuuride täiendav normaliseerimine ja ortogonaalse disaini printsiibi rakendamine. Täiendav normaliseerimine vähendab andmete liiasust ühe andmestruktuuri piires ja ortogonaalse disaini printsiibi rakendamine üle erinevate andmestruktuuride. Samas, ka kõrge tasemeni normaliseeritud andmestruktuurides võib endiselt olla andmete liiasust. [87] Andmete liiasus tingib andmete muutmise anomaaliate tekkimise. Kõrge tasemeni normaliseeritud tabelites on igal tabelil oma teema, ühes tabelis ei ole koos omavahel nõrgalt seotud andmeid (andmeid erinevat tüüpi olemite kohta) ja seega oleks ka andmebaasi struktuur vaatlejale (sh ettevõtte X töötajad) paremini arusaadav [1], [12].

### 6.9.1 Olemasoleva andmebaasi vastavus

Tabelis 25 tuuakse välja mitteliiasuse nõuded ja hinnang nende täidetusele olemasolevas andmebaasis. Tabel on sorteeritud nõuete kriitilisuse ja tähestikulises järjekorras.

Tabel 25. Mitteliiasuse nõuded ja nende täidetuse ning kriitilisuse ettevõttes.

ID	Nõue	Täidetuse	On kriitiline
1	Andmebaasist ei eksisteeri koopiaid erinevates asukohtades, mida põhikoopia muutmisel ei värskendata.	Jah	Jah
2	Andmebaasi tabelid on kõik viiendal normaalkujul (5NK).	Ei	Ei
3	Andmebaasis kehtib ortogonaalse disaini printsiip.	Ei	Ei
4	Andmebaasis pole muud andmete liiasust.	Ei	Ei

Nõuetest on täidetud 1 ja täitmata 3. Seega töö alustamise hetkel andmebaasis olevad andmed ei vasta mitteliiasuse tunnusele kehtivatele nõuetele.

Ettevõtte andmetes esineb liiasust. Näiteks tellimuse korral salvestatakse käibemaksu suurus kolme tabelisse - *wc\_order\_tax\_lookup*, *wc\_order\_product\_lookup* ja *wc\_order\_stats*.

### **6.9.2 Vastavuse tagamise meetmed**

Organisatsioon saab kontrollida nõuetele vastavust järgmiste meetmetega [1], [12]:

- Perioodiliselt kontrollida salvestatuid varukoopiaid.
- Ettevõtte kasutab valmistoodet (e-poe programm ja selle andmebaas) ja selles muudatusi teha ei ole ettevõttele jõukohane (tegemist ei ole IT-ettevõttega). Seega andmebaasis oleva andmete liiasuse puhul sõltutakse sellest, kui võrd hästi suudab e-poe programm vältida liiasusest tingitud vastuoluliste andmete andmebaasi sattumist. Ettevõtte saaks seda jälgida töölaual olevate kontrollpäringute abil.

### **6.9.3 Kriitilised nõuded**

Nõue 1 on kriitiline, sest vastasel juhul võib koopiast taastamine toimuda vanemast versioonist, milles ei pruugi olla viimaseid muudatusi.

## 7 Protsessid

Selles peatükis käsitletakse andmekvaliteediga tegelevate töötajate rolle, andmete haldamise ja andmekvaliteedi tagamise protsesse. Samuti on selles peatükis käsitletud andmekvaliteedi nõuete muutmise protsess.

### 7.1 Rollid

Käesolevas peatükis on toodud välja ettevõtte töötajate rollid, kes tegelevad andmete haldamise, andmekvaliteedi tagamise ning nõuete muutmise ja uute nõuete lisamise protsessidega.

Ettevõttes on kolm rolli, kelle tööülesanded on seotud andmekvaliteediga.

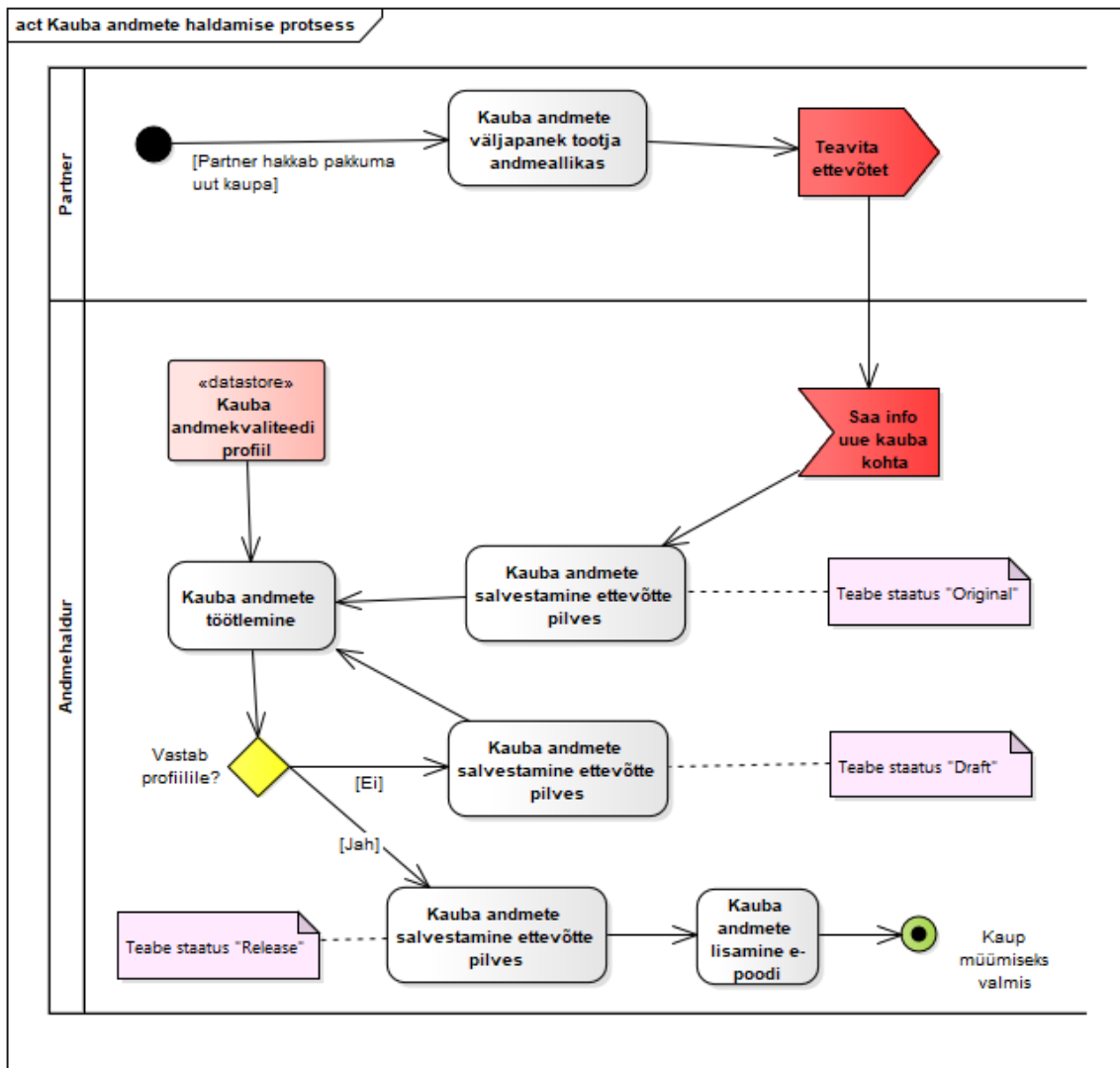
- **Andmehalduse juht.** Andmehalduse juht on töötaja, kes tegeleb andmekvaliteedi nõuete muutmise ja uute andmekvaliteedi nõuete lisamisega (v.t jaotist „7.4 Andmekvaliteedi nõuete lisamise ja muutmise protsessid). Andmehalduse juht võib samuti enda poolt pakkuda luua uusi andmekvaliteedi nõudeid või muuta olemasolevaid. Enne andmekvaliteedi nõuete muutmist või lisamist tuleb tal juhatuselga muudatused kooskõlastada. Andmehalduse juht vastutab andmekvaliteedi nõuete lisamise ja muutmise protsessi eest.
- **Andmehaldur.** Andmehaldur on töötaja, kes tegeleb uute andmete lisamisega (nii ettevõtte pilve, kui ka e-poodi), andmete töötlemisega (kuni nad vastavad andmekvaliteedi nõuetele) ja andmekvaliteedi kontrollimisega. Juhul, kui andmed ei ole kvaliteetsed, tegeleb andmehaldur ka probleemsete andmete töötlemisega, et need hakkaksid vastama andmekvaliteedi nõuetele. Andmehaldur vastutab andmete haldamise protsesside ja andmekvaliteedi tagamise protsessi eest. Andmehaldur võib pakkuda muudatusi andmekvaliteedi nõuete lisamiseks või muutmiseks.
- **Juhatus.** Ettevõtte juhatuselga kooskõlastatakse igasuguseid muutusi andmekvaliteedi nõuetes. Ilma juhatuselga loata, ei või teha muudatusi andmekvaliteedi nõuetes ja hindamistöölaua prototüübis. Ettevõtte juhatus võib samuti pakkuda teha muudatusi nõuetes ja pakkuda kasutuselevõtuks uusi andmekvaliteedi nõudeid.

- Andmete omanik. Andmete omanik teenuse andmete halduse protsessis on ettevõtte töötaja, kes parajasti tegeleb teenuste osutamisega ja kes on välja töötanud uue teenuse. Tema ülesandeks on dokumenteerida teenuse osutamise kohta käivad andmed ettevõtte pilves.

## **7.2 Andmete haldamise protsessid**

Joonisel 2 on esitatud kauba andmete haldamise protsess. Protsessi eest vastutab ettevõtte töötaja, kes parajasti tegeleb uue kauba kohta käiva teabe korjamise ja töötlemisega (andmehaldur). Ühe protsessi eest saavad vastutada ka mitu töötajat (iga protsessi osa eest vastutab üks töötaja).

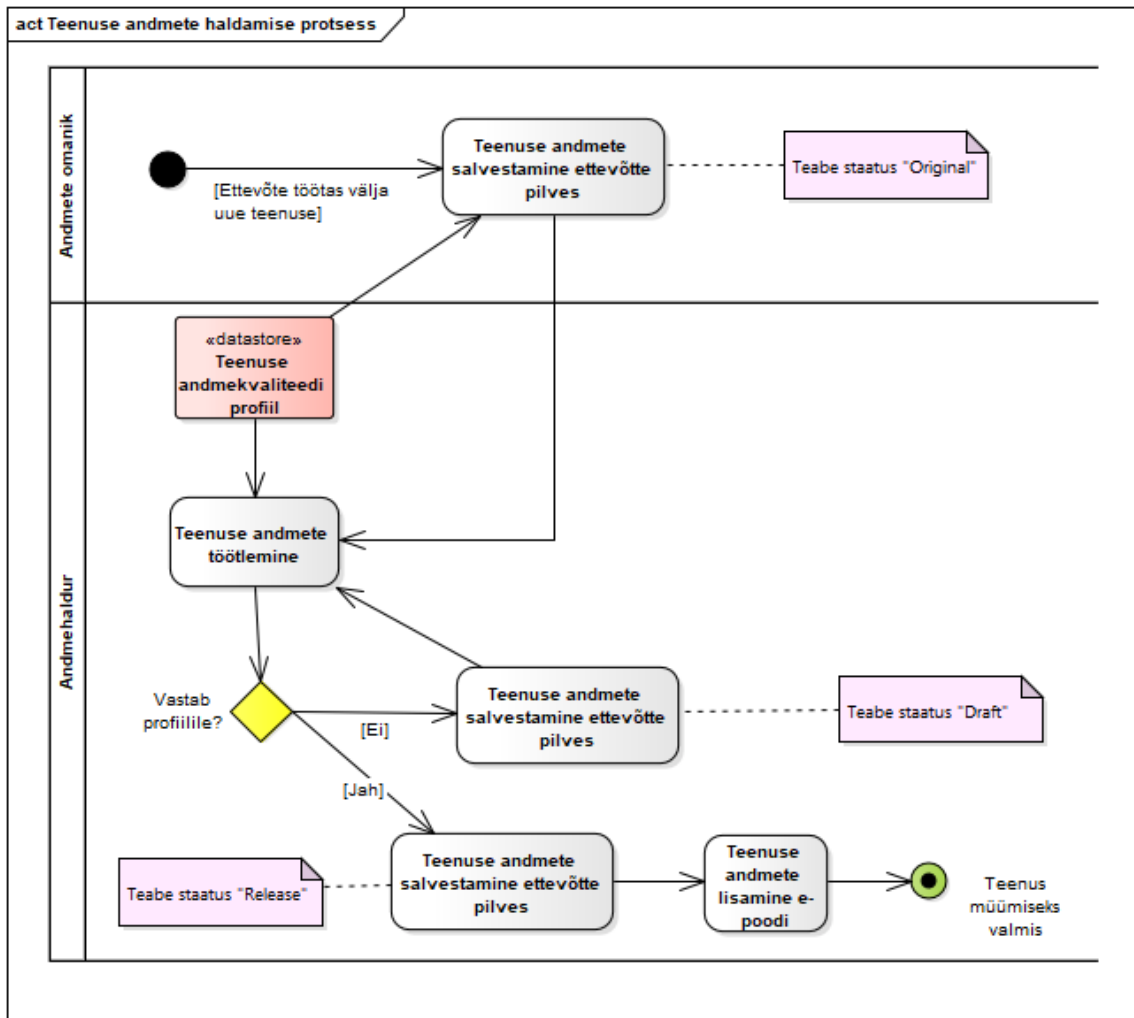
Protsess algab sellest, et ettevõtte partneritel tuleb välja uus kaup, mille informatsioon on esialgu kättesaadav vaid partnerite andmeallikatest. Ettevõtte töötajal, kes tegeleb kaupade tootekaartide loomisega, tuleb siseneda partneri andmeallikasse ja salvestada sealt andmed failina ettevõtte pilve. Pilve salvestamisel peab töötaja olema kursis ettevõtte pilve salvestamise nõuetega (nt salvestada kaup õige tootja alla, kausta „Original“, kus hoitakse veel töötlemata e originaalfaile). Edaspidi toimub vastutava töötaja poolt andmete töötlemine, et andmed vastaksid kauba andmekvaliteedi profiilile (v.t jaotis 8.1.1). Peale andmete töötlemist hindab ettevõtte töötaja, kas andmed vastavad andmekvaliteedi profiilile. Juhul, kui andmed vastavad profiilile, siis salvestatakse need ettevõtte pilve kausta „Release“. Juhul, kui andmed peale töötlemist ei vasta andmekvaliteedi profiilile, siis salvestatakse need pilve kausta „Draft“, milles hoitakse mustandeid ja töötluses pooleli olevaid andmeid ja tulevikus korratakse uuesti andmete töötlemist. Peale seda, kui andmed on töödeldud ja vastavad profiilile, saab neid kasutada kauba andmete lisamiseks e-poodi.



Joonis 2. Kauba andmete haldamise protsess.

Joonisel 3 on esitatud teenuse andmete haldamise protsess, mis sarnaneb joonisel 2 kujutletud kauba andmete haldamise protsessile. Kahe protsessi erinevus on osapoolte ülesannete jaotamises ja teenuse kohta käiva teabe asukoht. Võrreldes kauba andmete haldamise protsessiga loob andmete omanik teenuse kohta käivad andmed kohe ettevõtte pilves (kaustas „Original“). Edaspidi võib andmehaldur või andmete omanik tegeleda andmete töötlemisega, et need vastaksid teenuse andmekvaliteedi profiilile (v.t jaotist 8.1.2). Ideaalis peaks andmete omanik samuti profiili järgima, et töötlemise käigus oleks minimaalselt vaja andmeid korrastada. Juhul, kui andmed vastavad peale töötlemist profiilile, siis salvestatakse need kohe ettevõtte pilve nõuetele vastavate andmete jaoks mõeldud „Release“ kausta, kust neid saab võtta, et lisada teenuse andmed e-poodi. Juhul, kui teenuse andmed ei vasta andmekvaliteedi profiilile, siis salvestatakse need kausta „Draft“ ja need peavad uuesti läbima töötlemise, kuni need vastavad andmekvaliteedi

profiilile. Protsessi eest vastutab ettevõtte töötaja, kes parajasti tegeleb uue teenuse kohta käiva teabe loomisega ja töötlemisega. Ühe protsessi eest saavad vastutada ka mitu töötajat (iga protsessi osa eest vastutab üks töötaja).



Joonis 3. Teenuse andmete haldamise protsess.

Toodete andmete haldamise protsess sarnaneb teenuste andmete haldamise protsessile. Ainukene erinevus on selles, et toodete andmete töötlemisel ja andmekvaliteedi kontrollimisel võetakse aluseks kauba ja toote andmekvaliteedi profiil (v.t. jaotist 8.1.1).

Partnerite andmeallikatest võetud andmete omanikeks on partnerid, ettevõttel X on olemas luba nende andmete kasutamiseks ja töötlemiseks. Andmekvaliteedi profiili alusel e-poodi sisestatud andmed kuuluvad ettevõttele. Ettevõtte X loodud andmete omanikuks on ettevõtte X ise.

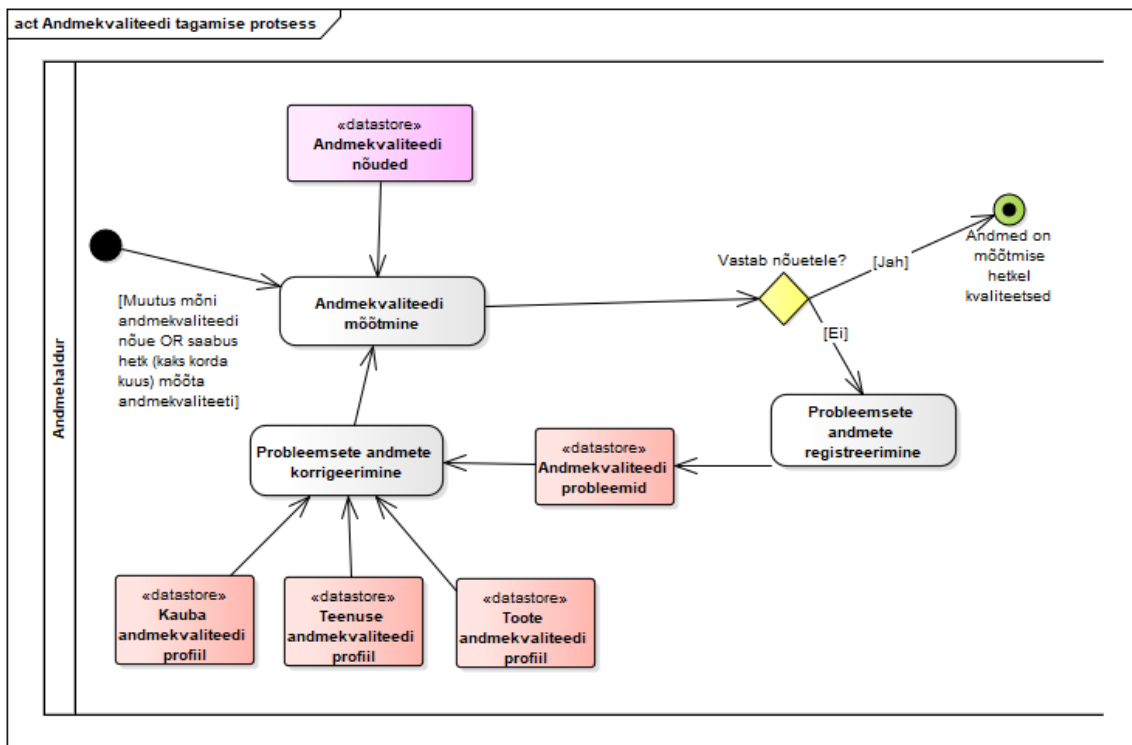


### 7.3 Andmekvaliteedi tagamise protsess

Joonisel 4 on esitatud andmekvaliteedi tagamise üldine protsess. Sellist andmekvaliteedi kontrollimise protsessi tuleb läbida regulaarselt kaks korda kuus, et tagada hea andmekvaliteedi tase ettevõttes. Samuti tuleb seda protsessi läbida peale seda, kui tehti muudatusi andmekvaliteedi nõuetes.

Protsessi sisendiks on ettevõtte andmekvaliteedi nõuded ja profiilid. Protsessi esimeseks sammuks on mõõta andmete kvaliteeti. Andmekvaliteedi mõõtmise käigus võivad ilmnedu andmekvaliteedi nõuetele mittevastavad andmed, mis seega vajavad muudatusi. Selliste andmete esinemine tuleb registreerida, et selle põhjal saaks analüüsida, miks sellised andmed tekivad ja mida tuleb tulevikus ette võtta selliste andmete süsteemi jõudmise vältimiseks.

Järgmise sammuna tuleb ettevõtte töötajatel muuta andmed nõuetele ja andmekvaliteedi profiilidele vastavaks (v.t jaotis 8.1). Peale muudatuste tegemist hinnatakse andmekvaliteeti uuesti ja kui probleemseid andmeid mõõtmise käigus enam ei leita, siis saab protsessi lõpetada ja pidada andmeid hetkeseisuga kvaliteetseks.



Joonis 4. Andmekvaliteedi tagamise protsess.

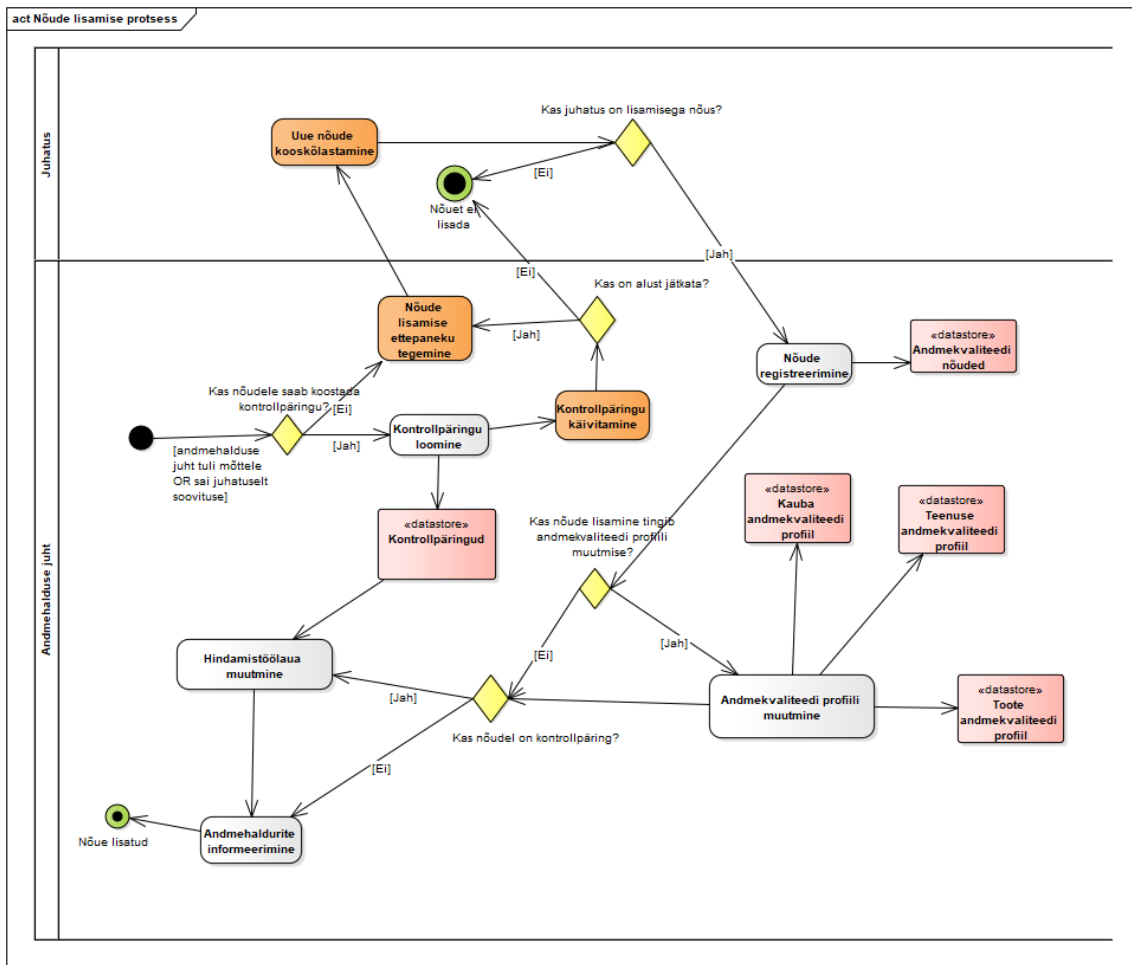
Andmekvaliteedi tagamise protsessi eest vastutab andmehaldur, kes tegeleb andmekvaliteedi mõõtmise läbiviimisega ning parendab andmeid.

#### **7.4 Andmekvaliteedi nõuete lisamise ja muutmise protsessid**

Andmekvaliteedi hindamise käigus saadud tulemused võivad näidata, et korduvate vigade vältimiseks on vaja olemasolevaid andmekvaliteedi nõudeid täiendada või lisada uus nõue.

Joonisel 5 on kujutatud nõuete lisamise protsess, mille sisendiks on andmehalduse juhi mõte või saadud soovitus lisada uus andmekvaliteedi nõue. Esimeseks sammuks on andmekvaliteedi nõude määratlemine e kas nõuet on võimalik kontrollida kasutades kontrollpäringuid või tuleb seda teha käsitsi. Kui tuleb teha käsitsi, siis liigutakse kohe nõude lisamise ettepaneku tegemisele. Juhul, kui nõuet on võimalik kontrollida kontrollpäringuga, siis luuakse see päring ja käivitatakse, et hinnata, kas uue nõude lisamiseks on vajadust või mitte. Vastavalt hindamise tulemustele lõpetatakse andmekvaliteedi nõude lisamise protsess või jätkatakse nõude lisamise ettepaneku tegemisega.

Enne andmekvaliteedi nõude registreerimist ja vastavalt vajadustele hindamistöölaua prototüübi või andmekvaliteedi profiili muutmist, tuleb uue andmekvaliteedi nõude lisamine kooskõlastada juhatusega. Juhul, kui juhatus ei ole nõus pakutud andmekvaliteedi nõuet lisama, siis nõude lisamise protsess lõppeb. Juhul, kui juhatus andis oma nõusoleku, siis registreeritakse andmekvaliteedi nõue ning vastavalt vajadusele muudetakse andmekvaliteedi profiile ja hindamistöölaua. Andmekvaliteedi nõude protsessi lisamise viimaseks sammuks on informeerida tehtud muudatustest andmehaldurit.



Joonis 5. Andmekvaliteedi nõuete lisamise protsess.

Andmekvaliteedi nõuete lisamise protsessi eest vastutab andmehalduse juht, kelle ülesanne on andmekvaliteedi nõuete ja profiilide välja töötamine. Andmehalduse juht võtab andmekvaliteedi nõuete määramisel arvesse andmehaldurite ja andmete omanike tähelepanekuid.

Andmekvaliteedi nõuete muutmise protsess sarnaneb nõuete lisamise protsessile. Erinevused seisnevad vaid selles, et tegemist on olemasoleva andmekvaliteedi nõude muutmisega. Andmekvaliteedi muutmise protsessi eest vastutab samuti andmehalduse juht. Andmekvaliteedi nõuete muutmisel säilitatakse ka info vana nõude versiooni kohta ja selle vana kontrollpäringu versiooni kohta.

## 8 Andmekvaliteedi profiil ja andmete sellele vastavuse hindamise töölaua prototüüp

Selles peatükis on toodud välja andmekvaliteedi profiilid ja informatsioon andmekvaliteedi hindamise töölaua prototüübi kohta – kontrollpäringute näited, tagarakendus ja kasutajaliides.

### 8.1 Andmekvaliteedi profiilid

Antud jaotises on toodud välja kaks andmekvaliteedi profiili – andmekvaliteedi profiil kaupadele ja toodetele (jaotis 8.1.1) ning andmekvaliteedi profiil teenustele (jaotis 8.1.2).

#### 8.1.1 Kauba ja toote andmekvaliteedi profiil

Tabel 26 esitab kauba ja toote andmekvaliteedi profiili, millest ettevõtte töötajad peavad kinni pidama, kui nad lisavad või muudavad kaupu ja tooteid e-poes. Tabel on sorteeritud vastavalt uue kauba või toote väljade täitmise järjekorrale.

Tabel 26. Kauba ja toote andmekvaliteedi profiil.

Tootekaardi väljad	Väärtus
Tootekaardi nimetus ( <i>product name</i> )	[Tootekaardi nimetus] Tootekaardi nimetamisel pea meeles, et nimetus ei või olla liiga pikk ja sisaldada kirjavahemärke (v.a punkti ja sidekriipsu). Nimetus läheb automaatselt ka tootekaardi <i>slugiks</i> ja mida lühem <i>slug</i> välja tuleb – seda parem. Minimaalne nimetuse pikkus võiks olla 9 tähemärki. Maksimaalne nimetuse pikkus võiks olla 60 tähemärki.
SKU	[Tootekood]
Lao staatus ( <i>stock status</i> )	[Staatust] <i>Instock, outofstock, onbackorder</i> – vali neist üks
Kogus laos ( <i>stock quantity</i> )	[Arv]
Tootja ( <i>brand</i> )	[Tootja] Valige õige kauba/toote tootja. Juhul, kui tegemist on komplektiga, siis võib märgitud tootjate arv olla >1.
SEO võtmesõna ( <i>keyword</i> )	[võtmesõna] Tootekaardi nimetus + 2-3 täiendavat märksõna kauba/toote kohta.

Tootekaardi väljad	Väärtus
SEO metakirjeldus	[Metakirjeldus] Kuni 145 tähemärki pikk lühiinformatsioon tootekaardi kohta kasutades võtmesõnu.
Lühikirjeldus ( <i>product short description</i> )	[Kirjeldus] 2-3 lühikest lauset kauba/toote kohta. Veendu, et vähemalt üks võtmesõna on olemas.
Kirjeldus ( <i>product description</i> )	[Kirjeldus] Vähemalt 80 sõna pikk. Kirjelduses kajastatavad teemad: <ul style="list-style-type: none"> <li>▪ Lühidalt kauba/toote kohta</li> <li>▪ Kaupa/toodet saab kasutada ...</li> <li>▪ Eriomadused</li> <li>▪ Spetsifikatsioon</li> <li>▪ Vajalik tarkvara</li> <li>▪ Materjalid/brošüürid/kataloogid (lingid)</li> <li>▪ Komplekti kuulub</li> </ul> Kirjelduses peab võtmesõna korduma kindlasti vähemalt: <ul style="list-style-type: none"> <li>▪ 1 x pealkirjas (<i>heading</i>)</li> <li>▪ 1 x tekstis</li> <li>▪ 1 x lühikirjelduses</li> </ul>
Kategooria ( <i>category</i> )	[Kategooria] Valige vähemalt üks sobiv kategooria. Mitme kategooria märkimisel ära unusta märkida, milline kategooria on peamine ( <i>primary</i> ).
Sildid ( <i>tags</i> )	[Silt] Tootekaardi temaga sobivad sildid. Näiteks pH puhul: aluseline, happeline.
Peapilt ( <i>product image</i> )	[Pilt] Kauba/toote pilt .webp formaadis.
Lisa pildid ( <i>product gallery</i> )	[Pilt] Kui on täiendavaid pilte, siis lisa need siia. Pildi formaat .webp.
Hind ( <i>regular price</i> )	[Hind] Sisesta hind eurodes käibemaksuta. Käibemaks lisandub automaatselt juurde. Sisestatav hind ei tohi olla negatiivne või 0.
Kampaaniahind ( <i>sale price</i> )	[Hind]

Tootekaardi väljad	Väärtus
	Juhul, kui kampaaniahind puudub, siis jäta väli tühjaks. Sisestatav hind ei tohi olla negatiivne või 0. Kampaaniahind ei või olla suurem kui tavahind ( <i>regular price</i> ).
Seotud tooted/kaubad ( <i>linked products cross-sells</i> )	[Tootekaardid] Juhul, kui seda kaupa/toodet on võimalik täiendada mõne teise kauba/tootega või teenusega, siis märgi see siin.

### 8.1.2 Teenuse andmekvaliteedi profiil

Tabel 27 esitab teenuse andmekvaliteedi profiili, millest ettevõtte töötajad peavad kinni pidama, kui nad lisavad või muudavad olemasolevaid teenuseid e-poes. Tabel on sorteeritud vastavalt uue teenuse väljade täitmise järjekorrale.

Tabel 27. Teenuse andmekvaliteedi profiil.

Nimetus	Väärtus
Tootekaardi nimetus ( <i>product name</i> )	[Tootekaardi nimetus] Tootekaardi nimetamisel pea meeles, et nimetus ei või olla liiga pikk ja sisaldada kirjavahemärke (v.a punkti ja sidekriipsu). Nimetus läheb automaatselt ka tootekaardi <i>slugiks</i> ja mida lühem <i>slug</i> välja tuleb, seda parem. Minimaalne nimetuse pikkus võiks olla 9 tähemärki. Maksimaalne nimetuse pikkus võiks olla 60 tähemärki.
SKU	[Tootekood] Teenuse tootekoodi loomiseks kasuta järgmist valemit: DF-K-[teenuse abreviaatuur või lühend]
Lao staatus ( <i>stock status</i> )	[Staatus] <i>Instock, outofstock, onbackorder</i> – vali neist üks.
Kogus laos (stock quantity)	[Arv]
Tootja ( <i>brand</i> )	[Tootja] Valige tootjaks ettevõtte.
SEO võtmesõna	[Võtmesõna] Tootekaardi nimetus + 2-3 võtmesõna teenuse kohta.
SEO metakirjeldus	[Metakirjeldus] Kuni 145 tähemärki pikk lühiinformatsioon teenuse kohta kasutades võtmesõnu.
Lühikirjeldus ( <i>product short description</i> )	[Kirjeldus]

Nimetus	Väärtus
	2-3 lühikest lauset teenuse kohta. Veendu, et vähemalt üks võtmesõna on olemas.
Kirjeldus ( <i>product description</i> )	[Kirjeldus] Vähemalt 80 sõna pikk. Kirjelduses kajastatavad teemad: <ul style="list-style-type: none"> <li>▪ Eesmärk</li> <li>▪ Väljundid</li> <li>▪ Teemad</li> <li>▪ Toimumise kuupäev ja aeg</li> <li>▪ Toimumise viis ja koht</li> <li>▪ Toimumise keel</li> </ul> Kirjelduses peab võtmesõna korduma kindlasti vähemalt: <ul style="list-style-type: none"> <li>▪ 1 x pealkirjas (<i>heading</i>)</li> <li>▪ 1 x tekstis</li> <li>▪ 1 x lühikirjelduses</li> </ul>
Kategooria ( <i>category</i> )	[Kategooria] Valige vähemalt üks sobiv kategooria. Mitme kategooria märkimisel ära unusta märkida, milline kategooria on peamine ( <i>primary</i> ).
Sildid ( <i>tags</i> )	[Sildid] Tootekaardi teemaga sobivad märksõnad.
Peapilt ( <i>product image</i> )	[Pilt] Teenust illustreeriv pilt .webp formaadis.
Lisa pildid ( <i>product gallery</i> )	[Pilt] Kui on täiendavaid pilte, siis lisa need siia. Pildi formaat .webp.
Hind ( <i>regular price</i> )	[Hind] Sisesta hind eurodes käibemaksuta. Käibemaks lisandub automaatselt juurde. Sisestatav hind ei tohi olla negatiivne või 0.
Kampaania hind ( <i>sale price</i> )	[Hind] Juhul, kui kampaaniahind puudub, siis jäta väli tühjaks. Sisestatav hind ei tohi olla negatiivne või 0. Kampaaniahind ei või olla suurem kui tavahind ( <i>regular price</i> ).
Seotud tooted/kaubad ( <i>linked products cross-sells</i> )	[Tootekaardid] Juhul, kui seda teenust on võimalik täiendada mõne teise kauba/toote/teenusega ja või selle teenuse kasutamiseks on vaja soetada mõni teine kaup/toode/teenus, siis märgi see.

## 8.2 Kontrollpäringud

Andmekvaliteedi nõuete täidetuse kontrollpäringud loodi SQL keeles. Töö esitamise hetke seisuga (2023. aasta mai) on loodud 46 SQL-päringut kokku nelja andmekvaliteedi tunnuse komplekti kohta. Iga kontrollpäring vastutab ühe andmekvaliteedi nõude eest. Päringud otsivad probleemseid andmeid, mitte ei anna lihtsalt jah/ei vastust probleemide esinemise kohta.

### 8.2.1 Õigsuse kontrollpäringud

Ettevõtte X on kehtestanud 48 õigsuse andmekvaliteedi nõuet, millest 19-t on võimalik SQL-päringute abil kontrollida ja 29 tuleb kontrollida ettevõtte töötajatel käsitsi, kuna neid ei ole SQL-päringute abil võimalik kontrollida. Näiteks ei ole võimalik SQL-päringu abil kontrollida õigsuse nõuet „Igal kaubal ja tootel ja tootega komplekti pandud teenusel (ühel tootekaardil nii teenus kui toode) on registreeritud õige lao kogus“.

Kontrollpäringu näidis: Joonisel 6 on esitatud õigsuse andmekvaliteedi nõude „Ühegi kauba, toote ja teenuse nimetuse pikkus ei tohiks ületada 60 tähemärki.“ kontrollimiseks loodud SQL-päring. Päring leiab need kaubad, tooted ja teenused, mille nimetuse pikkus on üle 60 tähemärgi. Päring leiab tootekaardi ID-d, SKU-d ja tootekaardi nimetused. Neid kolme väärtust teades on ettevõtte töötajal võimalik kergesti üles otsida kaup, toode või teenus, mille nimetus vajab parandamist.

```
SELECT posts.ID, pml.sku, posts.post_title
FROM wc_product_meta_lookup pml
INNER JOIN wp_posts posts ON pml.product_id = posts.ID
WHERE LENGTH(posts.post_title)>60 AND posts.post_type = 'product';
```

Joonis 6. Õigsuse andmekvaliteedi nõue „Ühegi kauba, toote ja teenuse nimetuse pikkus ei tohiks ületada 60 tähemärki“ kontrollpäring.

### 8.2.2 Täielikkuse kontrollpäringud

Ettevõtte X on kehtestanud 21 andmekvaliteedi nõuet, mis puudutab andmete täielikkust. Kõigi, v.a nõude „Kõik Eestis müüdavad avalikustatud kaubad, tooted ja teenused on kättesaadavad eesti, inglise ja vene keeles.“ ja „Igal teenusel, mis on tüüpi koolitus või kursus, on lahti kirjutatud koolituse/kursuse eesmärk, väljundid, teema, toimumise asukoht ja keel.“, kohta oli võimalik luua SQL-päring, mis toob välja nõudele mittevastavad kaubad, tooted ja teenused.

Kontrollpäringu näidis: Joonisel 7 on toodud välja täielikkuse andmekvaliteedi nõude „Igal kaubal, tootel ja teenusel peab olema kirjeldus.“ kontrollimiseks loodud



SQL-päring. Päring otsib üles kõik kaubad, tooted ja teenused, millel puudub kirjeldus. Päring kuvab nõuetele mittevastavate kaupade, toodete ja teenuste tootekaardi ID, SKU ja tootekaardi nimetuse.

```
SELECT posts.ID, pml.sku, posts.post_title
FROM wp_posts posts
INNER JOIN wc_product_meta_lookup pml ON posts.ID = pml.product_id
WHERE posts.post_type = 'product'
AND (posts.post_content IS NULL OR TRIM(posts.post_content)='');
```

Joonis 7. Täielikkuse andmekvaliteedi nõude „Igal kaubal, tootel ja teenusel peab olema kirjeldus“ kontrollpäring.

### 8.2.3 Kooskõla kontrollpäringud

Ettevõtte X on kehtestanud kuus andmekvaliteedi nõuet, mis puudutab andmete kooskõla. Kõigi nende nõuete kohta oli võimalik luua SQL-päringuid, mis tooksid välja nõuetele mittevastavaid kaupu, tooteid ja teenuseid.

Kontrollpäringu näidis: Joonisel 8 on toodud välja kooskõla nõude „Igal avalikustatud kaubal, tootel ja tootega komplekti pandud teenusel, peavad olema kooskõlas lao staatus ja kogus“ kontrollimiseks loodud SQL-päring. Päring täidab ülesannet otsida üles kõik kaubad, tooted ja teenused, mille lao kogus ja lao staatus ei ole üksteisega kooskõlas. Kasutades seda päringut on võimalik leida kõik kaubad, tooted ja teenused, millel:

- lao staatuseks on „instock“ ja lao kogus puudub üldse või on 0,
- lao staatuseks on „onbackorder“ ja lao kogus puudub üldse või on suurem 0st,
- lao staatuseks on „outofstock“ ja lao kogus puudub üldse või on suurem 0st.

Päring kuvab vaid nõuetele mittevastavate kaupade, toodete ja teenuste tootekaardi ID, SKU, tootekaardi nimetuse, lao koguse ja lao staatuse.

```

SELECT posts.ID, pml.sku, posts.post_title, pml.stock_quantity,
pml.stock_status
FROM wc_product_meta_lookup pml
LEFT JOIN wp_posts posts ON pml.product_id=posts.ID
WHERE posts.post_type='product' AND posts.post_status= 'publish'
AND ((pml.stock_status='instock' AND (pml.stock_quantity<1 OR
pml.stock_quantity IS NULL)) OR (pml.stock_status='onbackorder' AND
(pml.stock_quantity>0 OR pml.stock_quantity IS NULL)) OR
(pml.stock_status='outofstock' AND (pml.stock_quantity>0 OR
pml.stock_quantity IS NULL)));

```

Joonis 8. Kooskõla nõude „Igal avalikustatud kaubal, tootel ja tootega komplekti pandud teenusel, peavad olema kooskõlas lao staatus ja kogus“ kontrollpäring.

### 8.2.4 Konfidentsiaalsuse kontrollpäringud

Ettevõttes X on kehtestatud kaheksa andmekvaliteedi nõuet, mis puudutavad andmete konfidentsiaalsust. Seitsme andmekvaliteedi nõude täidetust kaheksast ei ole võimalik SQL-päringute abil kontrollida, kuna need ei ole otseselt seotud andmebaasiga. Näiteks ei ole SQL lausega võimalik kontrollida nõuet „Tehtud varukoopiad on turvaliselt hoiustatud“.

SQL-päringuga saab kontrollida konfidentsiaalsuse nõuet: “Kõigil veebisaidil registreerinud klientidel on staatus „customer“ ja neil puudub ligipääs veebipoe administraatoripaneelile.” SQL-päring nõude täidetuse kontrollimiseks on esitatud joonisel 9.

```

SELECT usermeta.user_id, usermeta.meta_key, usermeta.meta_value
FROM wp_usermeta usermeta
WHERE usermeta.meta_key='wp_capabilities' AND usermeta.meta_value NOT
LIKE '%customer%' AND usermeta.user_id NOT IN (2, 6, 8, 9);

```

Joonis 9. Konfidentsiaalsuse nõude “Kõigil veebisaidil registreerinud klientidel on staatus „customer“ ja neil puudub ligipääs veebipoe administraatoripaneelile” kontrollpäring.

## 8.3 Tagarakendus

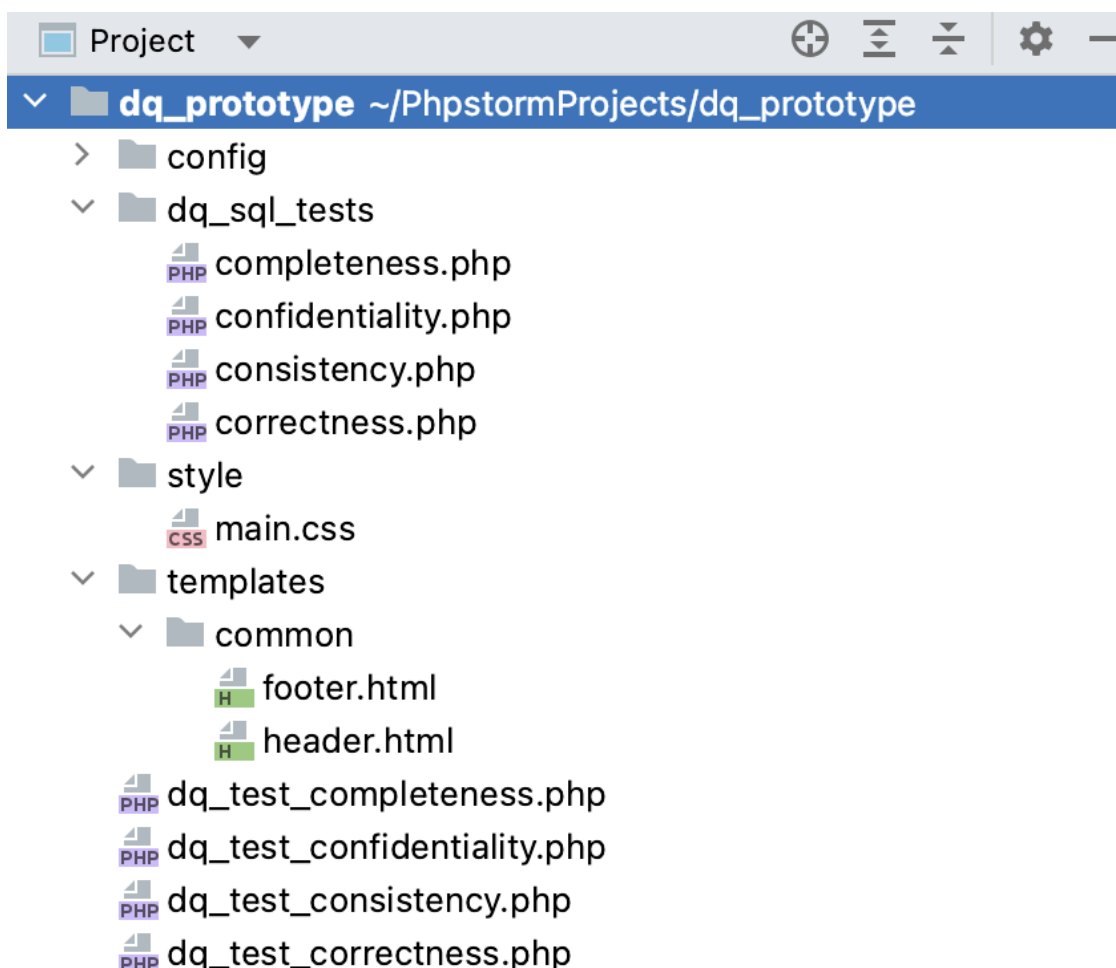
Joonisel 10 on kujutatud hindamistöölaua prototüübi projekti struktuur JetBrains PhpStorm rakenduses.

Kaustas „config“ on *connection.php* fail, mis vastutab selle eest, et hindamistöölaua oleks ühendatud ettevõtte X andmebaasiga. Kausta „dq\_sql\_test“ on koondatud kokku neli .php faili, millest igaüks vastutab ühe andmekvaliteedi tunnuse komplekti eest. Näiteks failis

*completeness.php* on funktsioonid, millest igaüks vastutab ühe SQL-päringu aktiveerimise ja tulemuste kuvamise eest.

Kaustas „style“ on *main.css* fail, mis kirjeldab kuidas tuleb veebilehtede sisu kasutajatele esitada. Main.css failis on näiteks toodud välja navigeerimismenüü värvid, teksti värv, kirjastiil, tabeli stiil. Kausta „templates“ on koondatud .html failid, mida kasutatakse iga lehekülje päises ja jaluses.

Andmekvaliteedi nõuete kontrolli leheküljed (*dq\_test\_completeness.php*, *dq\_test\_confidentiality.php*, *dq\_test\_consistency.php*, *dq\_test\_correctness.php*) asuvad projekti üldises kaustas.



Joonis 10. Hindamistöölaua prototüübi projekti struktuur.

Ettevõtte töötajad, kellel ei ole tööülesandeid ja vastavat luba, et teha muudatusi e-poes, ja ettevõtte asjasse mittepuutuvad isikud ei pääse ligi hindamistöölaua prototüübile ega kontrollpäringutele.

Prototüübi testimise ajal (aprill 2023) oli hindamistöölaua installeeritud ühe ettevõtte töötaja arvutisse. Hindamistöölaua prototüüpi kuskil ei avalikustatud, et ettevõttesse mittepuutuvatel isikutel ei oleks võimalik sellele ligi saada.

## 8.4 Kasutajaliides

Joonisel 11 on kujutatud kasutajaliideses kasutatud värvitoonid. Kirjastiiliks valiti „Palantino“. Nii värvitoonid, kui ka kirjastiil valiti ettevõttes kasutuses olevate stiilide hulgast.



Joonis 11. Hindamistöölaua kasutajaliideses kasutatud värvitoonid.

Andmekvaliteedi prototüübis on kokku neli veebilehte – täielikkuse andmekvaliteedi kontrolli leht, kooskõla andmekvaliteedi kontrolli leht, õigsuse andmekvaliteedi kontrolli leht ja konfidentsiaalsuse andmekvaliteedi kontrolli leht.

Iga veebileht koosneb kolmes osast – päis (*heading*), keha (*body*) ja jalus (*footer*). Kõik eelnimetatud lehed on ühe ja sama päise ning jalusega. Lehe päis koosneb navigeerimisribast (vt joonis 12), millel nimetusele klikkamine viib vastavalt täielikkuse, kooskõla, õigsuse ja konfidentsiaalsuse veebilehtedele.



Joonis 12. Andmekvaliteedi hindamistöölaua prototüübi veebilehe päis.

Veebilehe jalus jaotub neljaks osaks (vt joonis 13).

1. Üldine informatsioon hindamistöölaua prototüübi kohta.
2. Kontaktandmed, kellega võib võtta ühendust hindamistöölaua kohta käivate küsimuste tekkimise korral.

3. Kasulikud lingid, mis viitavad andmekvaliteedi nõuetele, andmekvaliteedi profiilidele ja ettevõtte Trellole.

4. Autoriõigused.

1 **DQ kontroll kohta**  
Andmekvaliteedi hindamise prototüüp.  
Prototüüp on valminud kevadel 2023.

2 **Kontakt andmed**  
irmael@ttu.ee  
irmael@taltech.ee

3 **Kasulikud lingid**  
DQ nõuded  
DQ profiil  
Trello

4 ©2023 Copyright Iris Nael All rights reserved

Joonis 13. Andmekvaliteedi hindamistöölaua prototüübi veebilehe jalus.

Andmekvaliteedi kontrollimise lehed on peaaegu identsed. Erinevus nende lehtede vahel on vaid kontrollpäringute tulemustes ja lehel asuvas tekstis. Andmekvaliteedi nõuete kontrollpäringute tulemused esitatakse kõik tabelite kujul, põhjusel, et sellises esitusviisis on nõuetele mittevastavaid andmeid kerge jälgida. Lisaks on andmete tabelite kujul esitus üks kõige mugavamaid esitusviisi ettevõtte X töötajate jaoks.

Ettevõtte töötajate mugavuse jaoks on iga tabeli kõrvale lisatud kontrollitava andmekvaliteedi nõude sisu. Nii saavad ettevõtte töötajad kiiremini aru, millise nõudega tegemist on ja mida täpsemalt tuleb teha, et parandada tabelis viidata kaupade, toodete ja teenuste andmete kvaliteeti.

Joonisel 14 on kujutatud õigsuse andmekvaliteedi kontrollimise veebilehe fragment.

**DQ kontroll**

Andmete õigsus Andmete täielikkus Andmete kooskõla Andmete konfidentsiaalsus

### ÕIGSUSE ANDMEKVALITEEDI KONTROLL

Kontroll veebilehekülj, mille käivitatakse SQL-päringuid, mis kontrollivad andmete vastavust nõuetele.

**Nõue 1**  
Iga avalikustatud kauba, toote ja teenuse hind peab olema nullist suurem arv.

Tootekood (SKU)	Tootekaardi nimetus	Tavahind	Kampaaniahind
SERETA101	Антимикробные вкладыши для лотков Granel's SortEd, упаковка из 52 шт.	0.0000	0.0000
	LabView DAQ training on-site		
SE-8636A	Glass Stirling Engine	0.0000	0.0000
SE-8636A	Двигатель Стирлинга	0.0000	0.0000
SE-8636A	Glass Stirling Engine	0.0000	0.0000
EX-5542	Vibreerivate joonte eksperiment	0.0000	0.0000
EX-5542	Эксперимент с вибрационными струнами	0.0000	0.0000
PH_AI	Комплект Фотон искусственный интеллект	0.0000	0.0000

Joonis 14. Õigsuse andmekvaliteedi kontrollimise veebilehe fragment.

Joonisel 15 on kujutatud täielikkuse andmekvaliteedi kontrollimise veebilehe fragment.

**DQ kontroll**

Andmete õigsus Andmete täielikkus Andmete kooskõla Andmete konfidentsiaalsus

## TÄIELIKKUSE ANDMEKVALITEEDI KONTROLL

Kontroll veebileheküljel, mille käivitatakse SQL-päringuid, mis kontrollivad andmete vastavust nõuetele.

**Nõue 1**

Igal avalikustatud kaubal, tootel ja teenusel on registreeritud vähemalt üks hind.

**Vastavus nõudele:**

Juhul, kui paremal pool olev tabel ei ole tühi, tähendab see seda, et andmekvaliteedi nõue ei ole täidetud ehk andmed ei vasta andmekvaliteedi nõudele.

Kontrolli, kas kõigil avalikustatud tootekaartidel on hind olemas ja / või 0st suurem.

Tootekood (SKU)	Tootekaardi nimetus
SERETA1L01	Антимикробные вкладыши для лотков Granelis SortEd, упаковка из 52 шт.
	LabView DAQ training on-site
SE-8636A	Glass Stirling Engine
SE-8636A	Двигатель Стирлинга
SE-8636A	Glass Stirling Engine
EX-5542	Vibreerivate joonte eksperiment
EX-5542	Эксперимент с вибрационными струнами
PH_AI	Комплект Фотон искусственный интеллект
roqed_science_training	3D Teaduse kasutamine distantsõppes

Joonis 15. Täielikkuse andmekvaliteedi kontrolli veebilehe fragment.

Joonisel 16 on kujutatud kooskõla andmekvaliteedi kontrollimise veebileht.

**DQ kontroll**

Andmete õigsus Andmete täielikkus Andmete kooskõla Andmete konfidentsiaalsus

## KOOSKÕLA ANDMEKVALITEEDI KONTROLL

Kontroll veebileheküljel, mille käivitatakse SQL-päringuid, mis kontrollivad andmete vastavust nõuetele.

**Nõue 1**

Igal avalikustatud kaubal, tootel ja tootega komplekti pandud teenusel, peavad olema kooskõlas lao staatus ja kogus.

Tootekood (SKU)	Tootekaardi nimetus	Lao kogus	Lao staatus
EX-5534	Эксперимент "Удельное сопротивление"		onbackorder
ST-2997	Kasvuhoone Sense and Control Kit		onbackorder
EP-3580	Лабораторная станция расширения физики		onbackorder
PS-2130	Датчик скорости потока / температуры воды серии PASPORT		onbackorder
PS-2130	PASPORT Voolukiiruse/temperatuuri andur		onbackorder
ST-7830	STEM-комплект PASCO «Умная среда обитания: Звук и свет»		onbackorder
EB-6335	Biology Extension Lab Station		onbackorder

Joonis 16. Kooskõla andmekvaliteedi kontrollimise veebilehe fragment.

Joonisel 17 on kujutatud konfidentsiaalsuse andmekvaliteedi kontrollimise veebilehe fragment. Sellelt on näha, et kui päring ei tagasta mingeid ridu, siis kuvatakse tabeli päis ja tekst, et andmekvaliteedi nõuded on täidetud.

**DQ kontroll**

[Andmete õigsus](#) [Andmete täielikkus](#) [Andmete kooskõla](#) [Andmete konfidentsiaalsus](#)

## KONFIDENTSIAALSUSE ANDMEKVALITEEDI KONTROLL

Kontroll veebilehekül, millel käivitatakse SQL-päringuid, mis kontrollivad andmete vastavust nõuetele.

**Nõue 1**  
Kõik uued veebisaidi registreerinud kliendid omavad staatust „customer“ ja neil puudub ligipääs veebipoe administraator paneelile.

Andmekvaliteedi nõue on täidetud.  
[Kasutaja ID](#) [Rolli kirje DB-s](#) [Rolli väärtus](#)

---

**DQ kontroll kohta**  
Andmekvaliteedi hindamise prototüüp.  
Prototüüp on valminud kevadel 2023.

**Kontakt andmed**  
[irnael@ttu.ee](mailto:irnael@ttu.ee)  
[irnael@taltech.ee](mailto:irnael@taltech.ee)

**Kasulikud lingid**  
[DQ nõuded](#)  
[DQ profiil](#)  
[Trello](#)

©2023 Copyright Iris Nael All rights reserved

Joonis 17. Konfidentsiaalsuse andmekvaliteedi kontrollimise veebilehe fragment.

## 9 Tulemuste analüüs ja järeldused

Käesolevas peatükis analüüsitakse andmebaasi disainivigu, arutatakse andmekvaliteedi tagamise parimate praktikate üle ja analüüsitakse, kuidas neid praktikaid jälgiti antud töö tegemisel ettevõttes X. Peatükis on samuti toodud välja andmekvaliteedi juhtimise raamistiku kohandused väikesele eraettevõttele ning kuidas seda raamistikku ettevõttes rakendati ja millised olid tulemused. Samuti analüüsitakse andmekvaliteedi hindamise töölaua prototüüpi ning tuuakse välja täiendusvõimalused. Peatüki lõpus käsitletakse tehtud töö nõrkuseid ja tehakse ettepanekud edasiseks arenduseks ja uuringuteks.

### 9.1 Andmebaasi disainivigade analüüs

Ettevõtte X andmebaasi skeemi probleemid võivad tingida probleeme ka andmekvaliteedis. Selle väite tõestamiseks või ümberlukkamiseks analüüsiti ettevõtte andmebaasi disaini. Andmebaasi disaini analüüsimiseks kasutati Tallinna Tehnikaülikooli õppeaine „Andmebaasid II“ andmebaaside disaini kontrollpäringuid [34], [88], mida saab käivitada ülikooli õppeserveril.

Disainiprobleemide esinemist otsivad süsteemikataloogi põhised päringud on loodud PostgreSQL andmebaasisüsteemile. Seega eelnevalt teisendati ettevõtte X andmebaasi skeemi loomise laused MySQList PostgreSQLi. Lisaks lausete teisendamisele tuli aega iseloomustavate veergude korral määrata uus vaikeväärtus. Seda tuleb teha põhjusel, et MySQLis määratud vaikeväärtus „0000-00-00 00:00:00“ ei ole PostgreSQLis [89] aktsepteeritav väärtus, sest selle literaali esitatav ajatempel ei kuulu PostgreSQLis TIMESTAMP tüüpi. PostgreSQL dokumentatsioonis rõhutatakse, et ei ole võimalik alustada nullindast sajandist [89]. Seega muutis autor käsitsi kõik ajalised vaikeväärtused väärtuseks „0001-01-01 00:00:00“. Peale vaikeväärtuste muutmist õnnestus luua PostgreSQLis tühjad andmebaasi tabelid, et käivitada nende põhjal andmebaasi disaini kontrollpäringuid. Kontrollitavad tabelid ei sisaldanud andmeid, kuna kontrollpäringute eesmärk on kontrollida andmebaasi skeemi probleemide esinemise suhtes, mitte andmebaasis olevaid andmeid. Järgnevalt on toodud välja disaini probleemid, mida kontrollpäringud leidsid ja mis ei ole seotud vaid PostgreSQLi kasutamisega. Analüüs keskendub andmebaasi kontseptuaalsele skeemile, mille moodustavad baastabelid.



### 9.1.1 Andmebaasiobjektide nimetamine

Andmebaasiobjektide nimetamine mõjutab andmekvaliteeti selle kaudu, et halvasti valitud nimed muudavad skeemi vaatlejale keerulisemaks sellest arusaamise ja selle kaudu ka korrektsete päringute ja andmemuudatuse lausete kirjutamise. Samuti muudavad halvasti valitud nimed keerulisemaks skeemist arusaamise, et seda parandada ja täiendada.

Andmebaasis tuleb pöörata tähelepanu andmebaasiobjektide nimedele. Mida kirjeldavam ja üheselt mõistetavam on nimi, seda paremini saavad kasutajad aru, milliste andmetega on tegemist ja kuidas neid saab kasutada. Robert C. Martin [90] on rõhutanud, et nimi võiks olla kirjeldav, sobival abstraktsioonitasemel, üheselt mõistetav ja kodeeringuteta. Ettevõtte X andmebaasis leidub aga mitmeid tabeleid ja nendes veerge, mille nimed ei vasta eelnevalt nimetatud hea nime tavadele. Probleemide ühisnimetaja on, et paljud nimed on liiga üldised ja ei ava täpsemalt andmete sisu.

Andmebaasi disaini kontrollpäringud [34] abil leiti üles järgmised disaini probleemid:

- Nimed on liiga lühikesed. Ettevõtte andmebaasis esineb 11 tabelit, milles on veerg nimega „id“. See nimi on liiga üldine. Probleemi lahendamiseks tuleks nimesid muuta. Näiteks tabeli *wp\_posts* korral on võimalik veerg nimetada „wp\_post\_id“ ja tabeli *rank\_math\_404\_logs* korral oleks võimalik nimetada veergu „404\_log\_id“. Nii saab muuta veeru nime paremini arusaadavaks. Lisaks on nimede andmisel olnud järjekindlusetu, sest paljudes tabelites ei ole võtmeveeru nimi „id“, vaid nimi on pikem.
- Nimed on liiga üldised ja see muudab nime järgi raskemini mõistetavaks, mis andmed on selles veerus. Selliste nimede näiteks on eelnevalt mainitud „id“ nimelised veerud. Kolm veergu on nimega „type“. Kolm veergu on nimega „timestamp“. Viis veergu on nimega „name“. Leidub ka veerg nimega „position“, veerg nimega „level“ ja veerg nimega „label“. Näiteks sõnad „timestamp“ ja „position“ on SQL standardi kohaselt reserveeritud sõnad, mida pole standardi kohaselt lubatud identifikaatorite e nimedena kasutada [91], [92], [93].
- Nimedes on kasutatud mürasõnu. Kõikide tabelite kõikides veergudes on andmed. Kuid ühe tabeli ühes veeru korral on peetud vajalikuks seda eraldi üle rõhutada, lisades veeru nimesse sõna „data“ – veerg „content\_data“ tabelis *wc\_admin\_notes*. Parem nimi võiks olla „content\_in\_db\_form“.

- Nimed ei ole piisavalt spetsiifilised. 14-s andmebaasi tabelis leitud veerge, mille nimi võrreldes teiste samas tabelites olevate veergude nimedega oli liiga üldine. Teiste sõnadega sisaldus selle veeru nimi mõnes teises sama tabeli veeru nimes. Näiteks tabelis *wp\_comments* on veerg „comment\_author“ ja spetsiifilisema nimega veerg „comment\_author\_email“. Liiga üldise nime probleemi lahendamiseks võiks „comment\_author“ veeru nimetada ümber näiteks „comment\_author\_name“, sest veerus „comment\_author“ hoitakse kommentaari autori nime. Nii saavad ettevõtte töötajad aru, et tegemist on kommentaari autori nimega, mitte kasutajanimega või valitud hüüdnimega.
- Kasutatakse sünonüüme. Ettevõtte X andmebaasi tabelites kasutatakse seisundite kohta infot talletavate veergude nimedes läbisegi sõnu „state“ ja „status“. Parem oleks kasutada sõna „status“, sest „state“ võib tähendada ka riiki.
- Isikunimede veergude nimed „first\_name“ ja „last\_name“ tabelis *wc\_customer\_lookup* ei arvesta sellega, et erinevates kultuurides on nimekomponentide hulk ja järjekord nimes erinev. Parmead veerunimed oleks „given\_name“ ja „surname“ [94]. Sellisel viisil oleks kohandus inimeste järgi, mitte inimeste kohandus süsteemi järgi [95].

### 9.1.2 Pole arvestatud isikuandmete kultuurilise erinevusega

Ettevõtte X andmebaasis on klientide tabelis nii eesnime kui perenime veerud mõlemad kohustuslikud e ei luba NULLe. Mononüümsed isikud [96], kellel puudub ees- või perekonnanimi, peavad ühte nendest väljadest sisestama väljamõeldud andmed. Arvestades andmekvaliteedi tähtsust saaks selle probleemi lahendada võttes kasutusele kahe nime veeru asemel ühe veeru, milles hoitakse lihtsalt isiku nime. Teine võimalus oleks jätta eesnime veerg kohustuslikuks ja muuta perenime veerg mittekohustuslikuks.

### 9.1.3 EAV antimuster

Ettevõtte X andmebaasi tabelites esineb EAV e *Entity-Attribute-Value* antimustrit. Antimuster väljendub selles, et kasutatakse üldist atribuutide tabelit, mis koosneb olemi (*entity*), atribuudi (*attribute*) ja väärtuse (*value*) veergudest. [97] EAV esineb näiteks WordPress tabelist *wp\_postmeta*, milles „entity“ veeruks on unikaalne postituse identifikaator e „post\_id“, atribuudiks on metaväärtus e „meta\_key“ ja väärtuse veeruks on „meta\_value“. „meta\_key“ veerus on 1035 unikaalset atribuuti, näiteks „rank\_math\_seo\_score“, „\_price“, „\_sku“, „\_stock\_status“.

### 9.1.4 Ebajärjekindel tõeväärtuste esitamine

MySQLis saab kasutada tabelite loomisel andmetüübi nime BOOLEAN, kuid vastav veerg on tüüpi TINYINT(1) [98]. Ettevõtte X andmebaasis kasutatakse lisaks tõeväärtustüüpi andmete esitamiseks ka ühte VARCHAR tüüpi veergu. Täiendavad kitsendused (lisaks andmetüübile) kõigil nendel veergudel puuduvad ja tegelikkuses võib nende veergudesse salvestada ka väärtuseid mis ei esita tõeväärtuseid e andmebaasis ei ole tagatud, et nende veergudes on väärtused kindlasti „0/1“ või „Yes/No“.

### 9.1.5 Puuduvad kitsendused

Andmebaasides deklareeritud kitsendused piiravad andmeid, mida saab andmebaasis registreerida [99]. Kõik andmebaasis olevad andmed peavad olema alati kooskõlas kõigi andmebaasis deklareeritud kitsendustega. Andmebaasisüsteem peab tagasi lükkama kõik andmemuudatused, mis üritavad andmebaasi lisada kitsendustega mittekooskõlas olevaid andmeid. Andmebaasis kitsenduste deklareerimine ei taga andmete õigsust. Kuid kitsendustele mittevastavad ja seega süsteemi poolt tagasi lükatavad andmed on kindlasti ebaõiged või mittetäielikud.

Kitsenduste abil on võimalik reguleerida, milliseid andmeid saab (näiteks tootekood peab olema unikaalne) ja millised ei saa (näiteks kauba nimetus ei või olla tühi) andmebaasis registreerida.

Andmebaasi disaini kontrollimise päringute [34] abil tuvastati kitsendustega seotud probleemid, mille ühisnimetaja on see, et kitsendusi on väga (liiga) vähe:

- Pole jõustatud loomulikke võtmeid (*natural keys*). Ettevõtte X andmebaasis on 29 tabelit 52st, mille korral pole andmebaasis jõustatud ühtegi loomulikku e naturaalselt võtit, ehk võtit, kuhu kuuluvad veerud on niigi tabelis juba olemas ja mille väärtused on lõppkasutajate poolt kasutatavad [100], [101]. Teiste sõnadega on sellistes tabelites surrogaatvõti, kuid pole jõustatud ühtegi teist unikaalsuse kitsendust. Ka sellistes tabelis saavad sisuliselt olla korduvad read, mis erinevad üksteisest ainult süsteemi genereeritud unikaalse väärtuse poolest. Jõustamata loomulikke võtmeid leidub ka teistes tabelites. Näiteks tabelis *wc\_customer\_lookup* puudub unikaalsuse kitsendus veerul „username“ ja tabelis *wp\_users* puudub unikaalsuse kitsendus veerul „user\_key“.

- Välisvõtme (*foreign key*) kitsenduste puudumine. Välisvõtme kitsendused aitavad vältida viidete terviklikkus vigu [102]. Ettevõtte X andmebaasi üheski tabelis ei ole välisvõtme kitsendusi. Andmebaasi 52 tabelist 39 põhinevad salvestumootoril MyISAM, mis ei toeta välisvõtmete kitsendusi [102]. Ülejäänule 13-le andmebaasi tabelile oleks võimalik välisvõtme kitsendusi määrata.
- Täielikult puuduvad CHECK kitsendused. Ettevõttes X andmebaasis on 171 tekstitüüpi veergu. Tühi string või ainult tühimärkidest koosnev string nendes veergudes oleks ebaõige väärtus, kuid CHECK kitsendust selle vältimiseks veerul ei ole. Selliste veergude näiteks on *wp\_posts* tabeli veerud „post\_title“ ja „post\_content“. See omakorda võib tähendada ettevõtte jaoks, et näiteks mõne kauba nimetus võib olla tühik (*wp\_posts* tabelis hoitakse samuti ka e-poe tootekaarte). Ühes veerus võib olla kogus ja kahes hind, mis võivad olla negatiivsed väärtused. Neljas veerus võivad olla e-posti aadressid, mis ei sisalda @ märki. Üheksas veerus võib olla URL, kus puudub protokolliga osa. Kõikide selliste andmete registreerimist saaks vältida CHECK kitsenduste abil.
- NOT NULL kitsendused. Andmebaasis on kaks tabelit, kus kõik veerud peale primaarvõtme veeru on mittekohustuslikud veerud e lubavad NULL'e.

### 9.1.6 Sama nimega veergudel on erinevad tüübid või väljapikkused

Sama nimega tabelite veerud on erinevates tabelites erinevate tüüpidega. Näiteks „id“ veerg tabelist *wp\_posts* on INTEGER tüüpi ja tabelis *rank\_math\_404\_logs* on BIGINT tüüpi.

Ettevõtte X andmebaasis leidub samuti ka veerge, mille nimed ja tüübid on samad, aga nende lubatud väljapikkused on erinevad. Nii on näiteks e-posti aadressi veergudega. Tabelis *wp\_comments* on selle lubatud maksimaalseks pikkuseks 100 tähemärki ja tabelis *woocommerce\_downloadable\_product\_permissions* on selle lubatud maksimaalseks pikkuseks 200 tähemärki.

### 9.1.7 Andmebaasi disainiprobleemide mõju andmekvaliteedile

Andmebaasi näol on tegemist populaarse valmistarkvara tabelitest moodustunud andmebaasiga ja seega mõjutavad leitud probleemid kõiki selle tarkvara kasutajaid.

Andmebaasi skeemi probleem, mis võib tuua kaasa andmekvaliteedi probleemide tekkimise on näiteks ees- ja perekonnanime veergude olemasolu ja kohustus registreerida

mõlemad nime komponendid. Kui veebisaidile soovivad registreeruda isikud, kellel ei ole kas ees- või perekonnanime, on nad kohustatud registreerima ühte nendest väljadest suvalise väärtuse. Suvaline tekst andmebaasis ei ole kindlasti olukord, mis iseloomustab kvaliteetseid andmeid.

EAV antimustri esinemine organisatsiooni andmebaasis võib tekitada andmetes liiasust, sest sama olemi sama atribuudi väärtust registreeritakse korduvalt. Andmete liiasus tähendab, et üks ja sama fakt esitatakse mitmekordselt või kui ühte ja sama fakti on võimalik andmebaasist tuletada rohkem kui ühel viisil [1], [12] (v.t jaotist „6.9 Mitteliiasus“). Kontrollimatu liiasus võib tagada olukorra, millel näiteks ühte toote kohta on ühes tabelis regulaarseks hinnaks märgitud 100 eurot ja teises tabelis näiteks 110 eurot. Liiasuse korral on tähtis tagada, et kui andmeid muudeti, siis need muutuvad kõikides tabelites.

Kitsenduste puudumine andmebaasis ei taga andmete kvaliteedi head taset. Andmebaasi veergude unikaalsuse kitsenduste puudumine võib tingida selle, et tabelis on mitmeid ridu, mis võivad olla sisuliselt samad ja erineda vaid süsteemi automaatselt genereeritud ID väärtuse poolest. ID väärtus ei kannu tavaliselt mingisugust väärtust lõppkasutajale [103]. Seega võib andmebaasis sisuliselt esineda korduseid e dubleerimist.

CHECK kitsenduste abil oleks võimalik tagada mõningate andmekvaliteedi nõuete täitmine, sest andmeid kontrollitaks kohe nende lisamisel. Näiteks oleks võimalik lisada kitsendus, et kaupade, toodete ja teenuste hindu ja kogust sisaldavad veerud ei tohiks sisaldada negatiivseid arve. Probleem on selles, et ettevõtte kasutab valmistarkvara ja igasugused muudatused andmebaasi struktuuri tasemel kirjutatakse tarkvara uute versioonide poolt üle.

Andmete kvaliteeti ei taga valede andmetüüpide kasutamine. Näiteks juhul, kui organisatsioonil on vaja hoida veerus väärtusi *true*(1) või *false*(0), siis tuleb selleks võtta veeru tüübina kasutusele BOOLEANi (põhineb MySQLis tüübil TINYINT(1)), mitte VARCHAR. TINYINT(1) või VARCHAR tüüpi ning ilma täiendavate kitsendusteta veergudes olevad väärtused võivad olla midagi muud kui 0 (esitab tõeväärtust FALSE) ja 1 (esitab tõeväärtust TRUE). See on võimalik andmekvaliteedi probleemide tekkimise koht.

Kokkuvõtvalt saab öelda, et probleemid andmebaasi disainis mõjutavad organisatsiooni andmekvaliteeti.

Ettevõtte X paljud kvaliteedi kontrolli päringud otsivad nõuetele mittevastavaid andmeid, mille registreerimist oleks võimalik vältida andmebaasi kitsenduste jõustamisega. Kahjuks ei saa ettevõtte andmebaasis kitsendusi jõustada, põhjusel, et kasutatakse valmistarkvara, mida iga mõne aja tagant uuendatakse. Peale igat uuendust võivad määratud kitsendused andmebaasist kaduda. Samuti tasub märkida, et töö kirjutamise hetkel puudub ettevõttes kompetents selliste muudatuste sisseviimiseks ja tulemuseks oleva lahenduse hooldamiseks.

## **9.2 Andmekvaliteedi tagamise parimad praktikad**

Andmekvaliteedi juhtimise eesmärk on võtta kasutusele abinõusid, mis aitavad ennetada andmekvaliteedi probleeme ja mis aitavad korrastada andmeid, mis on juba andmebaasides kasutusel [104].

Andmekvaliteedi tagamise parimad praktikad, mis kõige enam kordusid läbitöötatud allikates, on järgmised.

- Andmehaldus. Määrata andmekvaliteedi standardid, organisatsiooni töötajate rollid, andmete haldamise protsessid ja andmekvaliteedi tagamise protsessid ning nende eest vastutavad töötajad [104], [105], [106], [107], [108].
- Muuta andmekvaliteet organisatsioonis prioriteetseks teemaks [104], [105], [106], [107], [108]. Kõik organisatsiooni töötajad võiksid olla kursis andmekvaliteedi tähtsusega organisatsioonile ja millised võivad olla andmete halva kvaliteedi tagajärjed organisatsioonile. Organisatsiooni töötajad võiksid läbida andmekvaliteedi koolitusi, mille käigus neile tutvustatakse organisatsioonis kehtivaid andmekvaliteedi nõudeid ja neile näidatakse, kust kohast nad saavad vajadusel vaadata järgi andmekvaliteedi nõudeid ja juhiseid.
- Uurida andmekvaliteedi probleemide tekkimise põhjuseid [104], [105]. Probleemseid andmeid ei tasu koheselt kustutada – neid tuleb samuti analüüsida ja nende põhjal teha järeldusi. Järeldused on hea alus selleks, et tagada kvaliteetsete andmete kogumine. [105]
- Teha regulaarselt andmekvaliteedi kontrolle [104], [109], [106].

- Andmete profileerimine ja kontroll [109]. Allika [109] autor rõhutab, et organisatsiooni sissetulevad andmed peavad läbima andmete kvaliteedi kontrolli, et need andmed vastaksid organisatsioonis sätestatud profiilidele ja nõuetele.
- Määrata organisatsiooni kriitilised valdkonnad, milles tuleb eelkõige kontrollida andmete kvaliteeti ja teha parandusi [107].

Töö kirjutamisel ja andmekvaliteedi juhtimise raamistiku rakendamisel lähtuti samuti nendest andmekvaliteedi tagamise parimatest praktikatest.

### **9.2.1 Andmehaldus**

Andmehaldus on tähtis aspekt, mille parendamise pidi ettevõtte X andmekvaliteedi raamistiku rakendamisel ette võtma. Eesmärgiks seatud andmekvaliteedi küpsustaseme saavutamiseks tehti järgmiseid tegevusi.

- Määratleti ja dokumenteeriti andmekvaliteedi nõuded (sh määratleti kriitilisus). Samuti tehti andmekvaliteedi nõuded ettevõtte töötajatele kättesaadavaks.
- Määratleti ja dokumenteeriti andmekvaliteedi protsessid ja nende omanikud.
- Kirjeldati ja dokumenteeriti andmekvaliteedi profiilid. Profiilid tehti ettevõtte töötajatele kättesaadavaks.
- Tehti täiendusi ametijuhendites, mis puudutab andmekvaliteediga seotud tööülesannete täitmist ja vastutamist andmekvaliteedi tagamise protsesside eest.

### **9.2.2 Andmekvaliteedi koolitused**

Head andmekvaliteeti ja selle tagatavust ei ole võimalik ettevõttes saavutada ilma ettevõtte töötajate protsessi kaasamiseta. Ettevõtte korraldas lühikoolitusi kõikidele töötajatele, et nad oleksid teadlikud, miks on andmekvaliteet tähtis teema ja millised võivad olla halva kvaliteediga andmete tagajärjed. Samuti tutvustati ettevõtte töötajatele wikit, millest nad leiavad kogu ettevõtte kohta käiva tähtsa informatsiooni, sh kehtestatud andmekvaliteedi nõuded, andmekvaliteedi profiilid ja juhendid, mis aitavad töötajatel andmekvaliteedil silma peal hoida ja mittevastavute korral need õigesti dokumenteerida.

Töötajatele, kelle ülesandeks on kaupade, toodete ja teenuste lisamine e-poodi, korraldati täiendavalt lühikoolitusi selle kohta, kuidas tuleks andmeid uutmoodi esitada, et need vastaksid kehtestatud nõuetele. Samuti tutvustati neile uuesti andmekvaliteedi profiile, mida nad saavad kaupade, toodete ja teenuste andmete korjamisel, töötlemisel ja e-poodi lisamisel kasutada justkui malle.

Lisaks korraldati ettevõtte töötajatele lühikoolitus Trello kasutamise kohta. Kõik töötajad on kohustatud registreerima sinna infot nõuetele mittevastavate andmete kohta. Juhul, kui töötaja on volitatud mittevastavust ise korda tegema, siis tuleb tal seda teha ja märkida Trello piletile, et ta sellega tegeleb või et ta on selle probleemi ära lahendanud.

### **9.2.3 Andmekvaliteedi probleemide tekkimise põhjused**

Enamus ettevõtte X kaupadest, toodetest ja teenustest on käsitsi sisestatud ettevõtte töötajate poolt. Nende e-poodi sisestamisega tegelevad mitmed töötajad ja kõigil nendel on enda sisestamise stiil (nt mis järjekorras, mida ja kuidas nad sisestavad) ja siit võivad tekkida ja on ka tekkinud andmekvaliteedi nõuetele mittevastavused.

Samuti on nõuetele mittevastavus tekkinud kaupade, toodete ja teenuste kohta käiva informatsiooni importimisel failist. Juhul kui kaupu, tooteid ja teenuseid tuleb lisada e-poodi faili importimise kaudu, siis tuleb samuti veenduda, et imporditavad andmed vastavad andmekvaliteedi nõuetele.

Kõige suuremaks probleemiks, miks andmed e-poes ei vasta andmekvaliteedi nõuetele on see, et varasemalt ei olnud ettevõttes määratletud ja dokumenteeritud ühtseid andmekvaliteedi nõudeid. Teiste sõnadega, igal töötajal, kes tegeleb või tegeles andmete korjamisega, töötlemisega ja lisamisega e-poodi, olid enda arusaamad sellest kuidas andmeid tuleb esitada ja millised andmed on tähtsad ning millised ei ole.

Kõik need andmekvaliteedi nõuetele mittevastavuse põhjused selgusid peale seda, kui uuriti lähemalt andmeid e-poe andmebaasis ja käivitatud kontrollpäringuid. Samuti korraldati ettevõttes lühikoosolekuid, millel arutati, kuidas varasemalt toimus kaupade, toodete ja teenuste lisamine e-poodi ning kuidas toimus vajaliku teabe kogumine ja töötlemine.

Andmekvaliteedi raamistiku rakendamisel võeti kasutusele Trello keskkond, milles registreeriti kõik andmekvaliteedi nõuetele mittevastavused, mida ettevõtte töötajad avastasid/leidsid või millest kliendid ettevõtte töötajatele teatasid. Trello pileteid ei kustutata peale mittevastavuse likvideerimist, vaid need salvestatakse tahvlile, kus on lahendatud olukorrad. Nii on võimalik iga hetk nende juurde tagasi pöörduda ning analüüsida põhjuseid, miks need probleemid võisid tekkida ja kuidas neid saab lahendada ja tulevikus ära hoida.



#### **9.2.4 Andmekvaliteedi kontrollid**

Peale andmekvaliteedi raamistiku rakendamist on andmekvaliteedi kontrollid tehtavad kahel viisil – käsitsi ja andmekvaliteedi hindamistoolaua prototüübi abil. Kõiki kehtestatuid andmekvaliteedi nõudeid ei ole võimalik katta kontrollpäringutega ja seega tuleb ettevõtte töötajatel käsitsi kontrollida, kas andmed vastavad nõuetele. Näiteks kauba lisamisel, tuleb ettevõtte töötajal kontrollida, et kaubal olev pilt on õige ehk pilt vastab kauba nimetusele ja tootekoodile.

Hetkel tehakse andmekvaliteedi nõuete vastavuse kontrolle kord nädalas ja kohe tehakse parandused mõningatele kaupadele, toodetele ja teenustele. Ideaalne oleks kui töötajad kontrolliksid andmekvaliteeti igapäevaselt ja parandaksid võimalikult palju kaupu, tooteid ja teenuseid, mis muudatusi vajavad. Kuid kuna tegemist on väikeettevõttega ja selle töötajate arv ei ole suur, siis tuleb arvestada ka sellega, et lisaks andmekvaliteedi kontrollile ja paranduste tegemise on töötajatel vaja täita ka teisi tööülesandeid.

#### **9.2.5 Andmete profileerimine ja kontroll**

Kõige kergem viis, kuidas hoida andmekvaliteeti heal tasemel on kontrollida nõuetele vastavust juba andmete korjamise ja töötlemise staadiumis. Ehk andmete korjamisel ja töötlemisel tuleb kohe aluseks võtta ettevõttes kehtestatud profiile ja andmekvaliteedi nõuded. Samuti tasub kasutada ettevõtte wikis olevaid malle ja juhendeid. Ajas muutuvate andmete korral tuleb need andmed kohe parandada, et tulevikus vältida parandamist vajavate andmete kuhjumist.

#### **9.2.6 Kriitiliste valdkondade määramine**

Tehtud andmekvaliteedi kontrollid tõid välja palju kaupu, tooteid ja teenuseid, mis ei vastanud andmekvaliteedi nõuetele. Kõiki korruga ei olnud mõistlik parandama hakata, kuna see võiks ettevõttes mingiks ajaks peatada kogu ülejäänud töö. Selleks, et töö ettevõttes jätkuks ja andmed saaksid parandatud ning viidud andmekvaliteedi profillide ja nõuete abil kehtestatud kujule, määrati ettevõttes kriitilised valdkonnad. Andmekvaliteedi parandamisel tuleb töötajatel eelkõige tegeleda kriitiliste valdkondadega ja peale seda kui sinna kuuluvad andmed on parandatud, saavad töötajad tegeleda vähem kriitilistesse valdkondadesse kuuluvate andmetega. Näiteks on ettevõtte otsustanud, et kauba, toote ja teenuse tootekood (SKU) ja hind on kriitilised valdkonnad. Samuti otsustas ettevõtte, et kriitiliseks valdkonnaks on kaupade, toodete ja teenuste märksõnad ja metakirjeldused.

### **9.3 Andmekvaliteedi juhtimise raamistiku kohandused väikesele eraettevõttele**

Autor tegi RIA andmekvaliteedi küpsusmudelil väikeseid muudatusi, et küpsusmudel oleks sobivam väikeettevõtte vajadustele.

Kategoorias „Juhtimine ja planeerimine“ muudeti kahe väite sõnastust, üks kustutati ja vahetati kahe sama taseme väite järjekorda. Väidete muudatused ei tinginud suurt sisulist muutust. Muudeti kirjastiili, et ettevõtte töötajad saaksid nendest väidetest paremini aru. Esimene väide, mida muudeti oli „Määratletud ja dokumenteeritud on kõik poliitikatele alluvad kriitilised andmed“. Väite uus vorm on „Kriitilised andmed on määratletud ja dokumenteeritud“. Teine väide, mida muudeti oli „Andmekvaliteedi mõõdikud on seotud kõrgemate tasemete strateegiatega või poliitikatega“. Väite uus versioon on „Andmekvaliteedi mõõdikud on seotud kõrgemate tasemete strateegiatega“. Muudatuse põhjus on selles, et väikeettevõttel ei ole poliitikaid.

Kategooriast „Organisatsioon ja vastutused“ kustutati väide „Rühmitusse on kaasatud ka esindajad teistest andmekoguga seotud osapooltest“. Väide kustutati põhjusel, et väikeettevõtte andmekogud ei ole mõeldud kasutamiseks teistele osapooltele kui ettevõtte kliendid, kuid nende esindaja kaasamine oleks sisuliselt raskesti teostatav.

Kategooriast „Töövahendid“ kutsutati väide „On tehtud pingutusi „keskse tõeallika“ andmekoguse loomiseks“. RIA keskse tõeallika all mõtestatakse andmekogu, milles hallatakse riigi põhiandmeid. Väikeettevõttel sarnase andmekogu vajadust ei ole.

### **9.4 Andmekvaliteedi juhtimise raamistiku rakendamise ja selle tulemuste analüüs**

Töö käigus uuriti mitmeid andmekvaliteedi raamistike lootusega leida sobiv raamistik ettevõttele X, et lahendada olemasolevaid probleeme andmete kvaliteediga. Sobiva raamistiku otsimisel pöörati suurt tähelepanu sellele, et seda oleks võimalik rakendada väikesele eraettevõttele, mis tegeleb e-kaubandusega. Uuritud andmekvaliteedi raamistike hulgast tundusid kõige sobivamad järgmised kolm: TBDQ, DQPA ja Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistik. Nende raamistike kohta enamat saab leida jaotistest „3.2 Andmekvaliteedi raamistikud“ ja „3.3 Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistik“.

Ettevõtte X andmekvaliteedi probleemide lahendamiseks otsustati võtta kasutusele Eesti riigi infosüsteemi andmekvaliteedi juhtimise raamistik. Raamistiku valiku otsustavad tegurid olid järgmised.

- Raamistiku saab kohandada väikesele eraettevõttele, mis tegeleb e-kaubandusega.
- Ettevõttele sobib raamistikus rakendatav andmekvaliteedi juhtimise protsess: hinda -> planeeri -> teosta -> kontrolli -> korrigeeri.
- Raamistik oli loodud Eesti riigi infosüsteemide andmekvaliteedi juhtimiseks ja seega kooskõlas Eestis kehtestatud seadustega.
- Raamistiku kohta leidis palju teavet, sh juhend andmekvaliteedi tagamise kohta.
- Raamistiku kirjeldus oli eesti keeles, mis lihtsustab selle tutvustamist erinevatele ettevõttes andmekvaliteedi tagamisega seotud osapooltele.

Andmekvaliteedi raamistiku rakendamine toimus vastavalt Riigi Infosüsteemi Ameti andmekvaliteedi tagamise juhendile [12]. Esialgu tutvuti kogu juhendiga, et paremini mõista, kuidas peaks toimuma andmekvaliteedi raamistiku rakendamine. Edaspidi toimus andmekvaliteedi raamistiku kohandamine väikesele eraettevõtte vajadustele (v.t jaotist „9.3 Andmekvaliteedi juhtimise raamistiku kohandused väikesele eraettevõttele“).

Kohandamise käigus muudetud küpsusmudelit kasutati, et määrata ettevõtte esialgne andmekvaliteedi tase ja panna paika sihttase, milleni soovitakse jõuda. Autor seadis eesmärgiks saavutada töö tulemusena andmekvaliteedi 3. küpsustaseme, sest pidas teostatavaks viia ettevõtte nelja kuuga sellele küpsustasemele ilma muid töökohustusi unarusse jätmata.

Andmekvaliteedi parendamise tegevusplaan 2. ja 3. küpsustaseme saavutamiseks oli võimalik koostada alles peale ettevõtte andmekvaliteedi hindamist (v.t jaotist „5.2.2 Parendamise plaani koostamine“). Edaspidi toimus andmekvaliteedi raamistiku rakendamine juba vastavalt etteseadud parendusplaanidele (v.t jaotis „5.3 Teostamine“).

Suuri raskusi andmekvaliteedi raamistiku rakendamisega ei tekkinud. Siiski esinesid kõrvalekalded andmekvaliteedi nõuete koostamisel. Nimelt andmekvaliteedi nõuete koostamisel ei jõutud koheselt ühisele otsusele, milliseid andmekvaliteedi nõuded on kriitilise tähtsusega ja millised mitte. Hilisemate andmekvaliteedi puudutavate koosolekute käigus toodi välja mitmed uued andmekvaliteedi nõuded, mida andmehaldurid sooviksid lisada õigsuse andmekvaliteedi nõuete hulka. Põhjusel, et

ettepanekud laekusid viimasel hetkel, otsustati esitada pakutud andmekvaliteedi nõuded käesolevas töös ettepanekutena ja võtta need ettevõttes kasutusele järgmises arendustsüklis.

Andmekvaliteedi juhtimise raamistiku rakendamise ajal loodi andmekvaliteedi profiilid, mida ettevõtte töötajad saavad kasutada mallidena uute kaupade, toodete ja teenuste lisamisel. Samuti koostati raamistiku rakendamise ajal ja lisati ettevõtte wikisse mitmed juhised, mida töötajad saavad kasutada ilma, et nad peaksid pöörduma kolleegide poole, et täpsustada, kuidas mõnda lihtsat asja on võimalik teha.

Raamistiku rakendamise käigus võeti kasutusele mitmed uued töövahendid, mis aitavad töötajatel paremini täita oma tööülesandeid (nt kasutades Trello ja andmekvaliteedi hindamistöölaua prototüüpi) ja ettevõtte juhatusel olla paremini kursis andmekvaliteedi seisuga ettevõttes.

Peale iga parendusplaani elluviimise lõpetamist kontrolliti andmekvaliteedi küpsustaset, et näha, kas tegevusplaan aitas viia andmekvaliteedi uuele tasemele või mitte (v.t jaotis „5.4 Kontrollimine“). Kõik muudatused ettevõtte andmekvaliteedis registreeriti wikis, et kõik ettevõtte töötajad saaksid olla kursis ettevõtte andmete kvaliteedi jooksva tasemega. Kokkuvõtvalt võib öelda, et andmekvaliteedi juhtimise raamistiku juurutamine ettevõtte töösse oli edukas. Ettevõtte tegi suure sammu parema kvaliteediga andmete suunas. Andmekvaliteedi tase tõsteti madalamalt võimalikult küpsustasemelt (küpsustase 1) küpsustasemeni 3. Võeti kasutusele uued töövahendid ja loodi ning juurutati andmekvaliteedi profiilid, nõuded ja tagamise protsessid, mis aitavad andmekvaliteedi taset säilitada ja edaspidi tõsta.

## **9.5 Andmekvaliteedi hindamise töölaua prototüübi analüüs ja täiendusvõimalused**

Töö käigus loodi andmekvaliteedi hindamistöölaua prototüüp. Hindamistöölaua loomise peamiseks eesmärgiks oli aidata ettevõtte töötajatel kiiremini leida andmekvaliteedi nõuetele mittevastavaid kaupade, toodete ja teenuste tootekaarte. Loodud prototüüpi saab lisaks nimetatule kasutada järgmistel eesmärkidel.

- Leida üles vigade muustrid ja pakkuda nende alusel välja uus andmekvaliteedi nõue või täiendada olemasolevat.

- Analüüsid andmekvaliteedi nõuetele mittevastavusi ning tulemuste alusel korraldada töötajatele täiendavaid koolitusi andmekvaliteedi kohta ja/või luua täiendavaid juhendeid.

Hindamistöölaua saab lisaks andmehalduritele kasutada ka ettevõtte juhatus, et olla kursis ettevõtte e-poe andmekvaliteediga.

### **9.5.1 Hindamistöölaua prototüübi valideerimine**

Töö Hindamistöölaua valideerimine toimus läbi hindamistöölaua prototüübi A/B testimise ettevõtte X töötajatega. A/B testimine viidi läbi kõigi ettevõtte töötajatega, kes tegelevad andmete lisamisega ja andmete muutmisega ettevõtte e-poes.

A/B testimise eesmärk oli saada teada, kumb andmekvaliteedi nõuetele mittevastavuse leidmise lahendus (käsitsi või töölaua abil) on ettevõtte X töötajatele meeldivam, kumma kasutamine võttis vähem aega ja kummaga leiti ning parandati rohkem vigu.

Testimise käigus püstitati viis hüpoteesi (v.t jaotist „2.3 Tööprotsessi kirjeldus“), et edaspidi nendele toetudes analüüsida tulemusi.

A/B testimise tulemuste analüüs näitas, et ettevõtte töötajatel, kes olid paremini meelde jätanud andmekvaliteedi nõuded ja nende tunnuste komplektidesse kuuluvuse, ei olnud probleeme püstitatud ülesannete täitmisega uue lahenduse abil. Nende puhul said kõik püstitatud hüpoteesid tõestatud (hüpoteese ei lükatud tagasi). Töötajate puhul, kellel oli raskusi andmekvaliteedi nõuete ja nende tunnuste komplektidesse kuuluvuse meeldejätmisega, lükati neli hüpoteesi viiest tagasi, põhjusel et nad klikkasid läbi kõik andmekvaliteedi tunnuse komplektid, et leida üles õige.

Vaatamata sellele, et mõningate töötajate korral lükati enamuse hüpoteese, said nad ülesannetega kiiremini hakkama, kui seda olid teinud töötajad, kes testisid vana andmekvaliteedi kontrollimise lahendust (käsitsi kontroll).

Tagasiside andmisel mainiti, et andmekvaliteedi hindamistöölaua prototüüp oleks poole mugavam kui see võimaldaks tuua ühe tootekaardi kohta välja kõik mittevastavused e kui kauba X andmed ei vasta mitmele andmekvaliteedi nõudele, siis tuuakse need kõik ühes kohas välja. Praeguse lahenduse korral võib ühe tootekaardi juurde olla vaja tagasi pöörduda mitu korda.

Kokkuvõtvalt saab öelda, et hindamistöölaua prototüübi abil on võimalik leida palju kiiremini kaupade, toodete ja teenuste tootekaarte, mis ei vasta andmekvaliteedi nõutele. Tänu probleemsete kohtade kiiremale ülesleidmisele on võimalik neid andmeid kiiremini parandada. Arvestades vigade kiirema avastamise ja parandamise võimalusega, saab eeldada, et hindamistöölaua prototüüpi on parem kasutada, kui seda oli esialgne ettevõtte X lahendus e kõikide kaupade, toodete ja teenuste tootekaartide käsitsi kontrollimine.

### **9.5.2 Prototüübi täiendusvõimalused**

Prototüüpi saab täiendada teiste andmekvaliteedi tunnuste komplektidega. Praeguseks valminud hindamistöölaua prototüübis on toodud välja nelja andmekvaliteedi tunnuste komplekti kontrollpäringud - õigsus, täielikkus, kooskõla ja konfidentsiaalsus. Saab lisada veel kontrollpäringuid reeglipärasuse ja mitteliiasuse kohta. Kõiki tunnuste komplektides olevaid andmekvaliteedi nõudeid ei ole võimalik kontrollpäringute abil kontrollida, kuid leidub neid, mida saab või mille korral saab otsida väärtuseid, mis kindlasti ei vasta nõuetele.

Hindamise töölauda on võimalik lisada uus andmekvaliteedi nõuete mittevastavuse funktsionaalsus, mis kuvab konkreetse tootekaardi nõuetele mittevastavad andmed.

Lisaks on võimalik prototüüpi edasi arendada viisil, et kontrollpäringud registreeritakse andmebaasi. Kontrollpäringute haldamine (lisamine, muutmine ja kustutamine) tähendaks nende andmete haldamist läbi vastava andmebaasirakenduse. Töölauda programm saaks lugeda andmebaasist päringuid, neid järjest käivitada ja raporteerida tulemustest.

## **9.6 Tehtud töö nõrkused**

Järgnevalt on toodud välja tehtud töö nõrkused, millest vabanemiseks tuleks antud teemat edasi uurida ja täiendada töölauda:

- Puuduvad liiasusega seotud kontrollpäringud, mis leiaksid andmete liiasusest tuleneva vastuoluliste faktide registreerimise.
- Üle poole andmekvaliteedi nõudeid ei ole võimalik kontrollida muul viisil, kui käsitsi. Näiteks õigsuse korral oleks võimalik otsida õigsusele kindlasti mitte vastavaid väärtuseid.

## 9.7 Ettepanekud edasiseks arenduseks ja uuringuteks

Käesolevas jaotises on toodud välja ettepanekud, kuidas saaks käesolevat tööd edasi arendada. Ettepanek on jätkata ettevõttes andmekvaliteedi küpsustasemete kasvatamisega, täiendada andmekvaliteedi nõudeid ja profiile ning täiendada andmekvaliteedi hindamistöölaua.

### 9.7.1 Andmekvaliteet küpsustasemele neli ja viis

Ettevõtte X on 2023.a. aprillikuu lõpuks andmekvaliteedi seisukohalt 3. küpsustasemel. Ettepanek on saavutada tulevikus andmekvaliteedi küpsustasemed neli ja viis. Parendamise plaani koostamise aluseks võeti jaotises “5.1 Hindamine” esitatud kategooriate kaupa küpsustaseme hinnang ja RIA küpsustaseme tõstmise juhend [12]. Ühegi väite sõnastust ei peetud vajalikuks muuta. Kolmandalt (standardiseeritud) küpsustasemelt neljandale (juhitud) küpsustasemele liikumiseks on vaja jälgida Tabel 28 esitatud parendamis plaani. Peale neljanda küpsustaseme saavutamist tuleb võtta kasutusele parendamise plaan viienda (optimeeritud) küpsustaseme saavutamiseks (vt Tabel 29).

Tabel 28. Parendustegevused andmekvaliteedi küpsustaseme 4 saavutamiseks.

Kategooria	Parendus-valdkond	Kategoorias taseme 4 saavutamiseks vajalikud parendustegevused
Juhtimine ja planeerimine	Andmekvaliteedi tsoonid ühtlaseks	<ul style="list-style-type: none"><li>Juurutada andmekvaliteedi juhtimine vaadeldavas tsoonis ühtsetel alustel.</li></ul>
Juhtimine ja planeerimine	Regulaarne andmekvaliteedi mõõtmine ja tulemuste analüüs	<ul style="list-style-type: none"><li>Analüüsida perioodiliselt mõõtmistulemusi, mille alusel teha parendusi andmekvaliteedi mõõtmisse ja parendamisse.</li></ul>
Organisatsioon ja vastutused	Andmekvaliteedi vastutajate rühmitus	<ul style="list-style-type: none"><li>Panna kokku andmekvaliteedi eest vastutav juhtivatest ametikohtadest koosnev rühmitus, kuhu kaasata osapooli ka teistest andmekoguga seotud osapooltest.</li></ul>
Protsessid	Andmekvaliteedi parendamine ettevõttes	<ul style="list-style-type: none"><li>Juurutada regulaarne andmekvaliteedi mõõtmine.</li><li>Määratleda ja dokumenteerida andmete kvaliteedi parendamise tegevusplaani.</li></ul>
Protsessid	Andmekvaliteedi mõõtmised kättesaadavaks	<ul style="list-style-type: none"><li>Teha mõõtmistulemused kättesaadavaks ka organisatsiooni juhtkonnale.</li></ul>

<b>Kategooria</b>	<b>Parendus-valdkond</b>	<b>Kategoorias taseme 4 saavutamiseks vajalikud parendustegevused</b>
Teadmus ja kompetentsid	Regulaarsed andmekvaliteedi koolitused	<ul style="list-style-type: none"> <li>▪ Lülitada töötajate koolituskavadesse regulaarsed andmekvaliteedi teemalised koolitused.</li> </ul>
Töövahendid	Andmekvaliteedi kontroll ja raporteerimine	<ul style="list-style-type: none"> <li>▪ Juurutada töövahendid andmekvaliteedi mõõtmistulemuste analüüsiks ja raporteerimiseks.</li> <li>▪ Rakendada vahendites andmekvaliteedi profiilist lähtuvad automaatsed andmeparanduse või –kontrolli protseduurid.</li> </ul>
Töövahendid	Andmekvaliteedi töövahendite eelarve	<ul style="list-style-type: none"> <li>▪ Lülitada andmekvaliteedi protsesse toetavad arendusvajadused arendustegevuste portfelli.</li> </ul>

Tabel 29. Parendustegevused andmekvaliteedi küpsustaseme 5 saavutamiseks.

<b>Kategooria</b>	<b>Parendus-valdkond</b>	<b>Kategoorias taseme 5 saavutamiseks vajalikud parendustegevused</b>
Juhtimine ja planeerimine	Tugevdada andmekvaliteedi juhtimine	<ul style="list-style-type: none"> <li>▪ Juurutada andmekvaliteedi juhtimine poliitikate regulaarse ülevaatus, muutmise ja avaldamise abil.</li> </ul>
Organisatsioon ja vastutused	Andmekvaliteedi vastutussisikud	<ul style="list-style-type: none"> <li>▪ Rakendada regulaarne andmekvaliteedi vastutuste ülevaatus ja täiendamine.</li> </ul>
Protsessid	Regulaarne andmekvaliteedi analüüs ja parendamine	<ul style="list-style-type: none"> <li>▪ Rakendada regulaarne andmekvaliteedi mõõtmistulemuste ülevaatus parendamisvõimaluste leidmise eesmärgil.</li> </ul>
Teadmus ja kompetentsid	Regulaarsed andmekvaliteedi koolitused ja -materjalide ülevaatus	<ul style="list-style-type: none"> <li>▪ Rakendada regulaarne koolituste programmi ja koolitusmaterjalide ülevaatus ja täiendamine.</li> </ul>
Töövahendid	Andmekvaliteedi töövahendite juurutamine	<ul style="list-style-type: none"> <li>▪ Lülitada põhiandmete põhimõtete järgimine andmekogu arendusnõuetesse.</li> <li>▪ Rakendada arendustegevuste portfelli regulaarne ülevaatus ja täiendamine ka andmekvaliteedi aspektidest.</li> </ul>



### 9.7.2 Andmekvaliteedi nõuete täiendamine

Andmekvaliteedi teemat ettevõttes koosolekutel arutades otsustati täiendada andmekvaliteedi nõudeid ja andmekvaliteedi profiile. Üheks selle põhjuseks on ettevõttes müüdavate kaupade, toodete ja teenuste valiku laiendamine. Teiseks märkimisväärseks põhjuseks on soov ettevõtet ning selle kaupu, tooteid ja teenuseid senisest rohkem passiivselt reklaamida. Ettepanek on lisada andmekvaliteedi nõuete hulka järgmised andmekvaliteedi nõuded. Neid nõudeid tuleb lisada enne kui liikuda edasi ettevõtte andmekvaliteedi küpsustasemete tõstmisega (kolmandalt küpsustasemelt neljanda küpsustasemele).

- Igal kaubal, tootel ja teenusel, millega on kaasas kaup või toode, on märgitud vähemalt üks video.
- Igal kaubal, tootel ja teenusel, millega on kaasas kaup või toode, on märgitud vähemalt üks õppematerjal.
- Igal kaubal ja tootel on registreeritud kauba/toote suurus (sentimeetrites).
- Igal kaubal ja tootel on registreeritud kauba/toote kaal (kilogrammides).
- Igal kaubal, tootel ja teenusel on märgitud selle kriteeriumid. Näiteks traadita anduri üheks kriteeriumiks võib olla „Bluetoothiga“.
- Iga kauba, toote ja teenuse pildi nimetus vastab ettevõtte X formaadile „[Tootekood]\_[tootenimetus]\_vaata\_[veebileht]“.
- Iga kauba, toote ja teenuse õppematerjali faili nimetus vastab ettevõtte X formaadile „[Tootekood]\_[nimetus]\_vaata\_[veebileht]“.

Vastavalt andmekvaliteedi nõuete täiendamisele tuleb teha ka vastavad muudatused andmekvaliteedi profiilides ning luua töölauale kontrollpäringud.

### 9.7.3 Andmekvaliteedi hindamistöölaua prototüübi täiendamine

Andmekvaliteedi hindamise prototüübi osas võiks järgnevalt teha selliseid täiendusi.

- Lisada võimalus leida ja vaadata kõiki andmekvaliteedile mittevastavusi ühe konkreetse kauba, toote ja teenuse tootekaardi kohta. Praegu otsitakse ja kuvatakse probleemseid andmeid andmekvaliteedi nõuete kaupa. Ühe kauba, toote või teenuse kõigi andmete analüüsimiseks tuleb vaadata läbi mitmete päringute tulemused.

- Luua roppsõnu või kahtlaseid linke sisaldavate kommentaaride otsimiseks mõeldud päringud.
- Arendada edasi hindamistöölaua prototüüpi (nii eesrakendust kui tagarakendust), et registreerida kontrollpäringud andmebaasis ja võimaldada vastava andmebaasirakenduse kaudu päringuid lisada, muuta ja kustutada. Töölaud loeks andmebaasist päringuid, käivitaks neid ja esitaks tulemused. Praegu tuleb muudatuse tegemiseks muuta prototüübi PHP koodi.
- Viia hindamistöölaua prototüüp üle lokaalsest keskkonnast ettevõtte sisesele serverile.

## 10 Kokkuvõte

Organisatsioonidele on selle kasutatavate andmete kvaliteet tähtis, sest tuginedes andmetele tehakse otsuseid, mis mõjutavad kõiki organisatsiooni huvitatud osapooli. Seega organisatsiooni jaoks on tähtis, et selle käsutuses olevad andmed oleksid kvaliteetsed – see võib neile anda konkurentsieelise ja samuti võimaldab võtta vastu täpsemaid äriotsuseid.

Töö põhiprobleemiks on väikese eraettevõtte X hallatavate andmete ebapiisav kvaliteet, mille tulemuseks on ettevõtte poolt üleliigsete ressursside kulutamine. Ettevõtte peab e-poodi kasutades populaarset sisuhaldussüsteemi WordPress ja selle pistikprogrammi WooCommerce.

Käesoleva töö eesmärk oli püsivalt parandada andmekvaliteeti ettevõttes X.

Eesmärgi saavutamiseks uuriti tänapäeval kasutuses olevaid andmekvaliteedi raamistikke, et leida sobivam andmekvaliteedi hindamise ja parandamise raamistik väikesele eraettevõttele. Töö käigus koostati ja dokumenteeriti ettevõtte wikisse andmebaasi mudel, sh tabelite ning veergude vabatekstilised kirjeldused.

Töös kasutatakse Eesti riigi infosüsteemi andmekogude andmekvaliteedi juhtimise raamistikku, mida on kohandatud väikese eraettevõtte jaoks.

Magistritöö sisaldab muuhulgas ettevõtte andmekvaliteedi analüüsi ja andmebaasi disainiprobleemide analüüsi.

Püstitatud andmekvaliteedi nõuetele mittevastavate andmete leidmiseks loodi ja juurutati veebipõhise andmekvaliteedi hindamistöölaua prototüüp. See programm teeb andmebaasi põhjal päringuid, et leida andmekvaliteedi nõuetele mittevastavaid andmeid.

Hindamistöölaua prototüübi valideerimine toimus ettevõtte X töötajatega, kelle igapäevane tööülesanne on uute kaupade, toodete ja teenuste kohta informatsiooni kogumine, selle töötlemine ja e-poe andmebaasi sisestamine. Valideerimine toimus läbi

hindamistöölaua prototüübi A/B testimise, milles võrreldi ettevõtte esialgset lahendust e käsitsi otsimist ja prototüüpi.

Töö tulemusena viidi ettevõtte andmekvaliteet esimeselt (kõige madalamalt) küpsustasemelt kolmandale küpsustasemele (viietasemelisel skaalal). Teiste sõnadega viidi andmekvaliteet reaktiivselt tasemelt standardiseeritud tasemele. Saavutati olukord, kus ettevõtte andmekvaliteet võiks olla püsivalt sellel küpsustasemel.

Töö kokkuvõttena saadi vastused kõigile püstitatud uurimisküsimustele, leiti ja rakendati sobiv andmekvaliteedi juhtimise raamistik ning loodi andmekvaliteedi hindamistöölaua prototüüp. Andmekvaliteedi püsiva paranemise tagab andmekvaliteedi regulaarne kontrollimine ja parandustegevuste läbiviimine vastavalt kindlaksmääratud juhistele.

Töö võiks pakkuda huvi kõigile WordPressi ja WooCommerce'i kasutajatele, sest töös esitatakse nii nende vahendite andmebaasi mudel kui analüüsitakse ka andmebaasi disainiprobleeme ning osutatakse ebakvaliteetsetele andmetele, mis osalt ka nende disainiprobleemide tõttu on andmebaasi jõudnud. Töö võiks pakkuda huvi väikestele eraettevõtetele, sest demonstreerib samme ja väljakutseid andmekvaliteedi parandamise juures. Töö võiks pakkuda huvi Eesti riigi infosüsteemi andmekogude pidajatele, sest sisuliselt on tegemist Eesti riigi infosüsteemi andmekogude andmekvaliteedi juhtimise raamistiku rakendamise juhtumianalüüsiga – tõsi küll koostajate ettenähtust erinevas kontekstis.

## Kasutatud kirjandus

- [1] J. Linros, M. Lauk, V. Kähari ja J. Tepandi, „Andmekvaliteedi uuringu lõpparuanne,“ Riigi Infosüsteemi Amet, 26.08.2016.
- [2] Riigi Infosüsteemi Amet, „Andmekvaliteedi uuring andmekogudes,“ 26.08.2016.
- [3] AltexSoft, „What is an API: Definition, Types, Specifications, Documentation,“ 18. 06. 2019. [Võrgumaterjal]. Available: <https://www.altexsoft.com/blog/engineering/what-is-api-definition-types-specifications-documentation/>. [Kasutatud 11. 10. 2022].
- [4] Greenwich 2000, „What is Greenwich Mean Time? (GMT),“ [Võrgumaterjal]. Available: <https://greenwichmeantime.com/what-is-gmt/>. [Kasutatud 09. 04. 2023].
- [5] Justin, „Payment Token API in 2.6,“ WooCommerce, 04. 04. 2016. [Võrgumaterjal]. Available: <https://developer.woocommerce.com/2016/04/04/payment-token-api-in-2-6/>. [Kasutatud 09. 04. 2023].
- [6] R. Awati, „MIME (Multipurpose Internet Mail Extensions),“ TechTarget, 14. 10. 2021. [Võrgumaterjal]. Available: <https://www.techtarget.com/whatis/definition/MIME-Multi-Purpose-Internet-Mail-Extensions>. [Kasutatud 09. 04. 2023].
- [7] PayPal, „How PayPal works,“ [Võrgumaterjal]. Available: <https://www.paypal.com/us/digital-wallet/how-paypal-works>. [Kasutatud 09. 04. 2023].
- [8] IBM, „What is a REST API?,“ IBM, [Võrgumaterjal]. Available: <https://www.ibm.com/topics/rest-apis>. [Kasutatud 09. 04. 2023].
- [9] Statistikaamet ja Riigi Infosüsteemi Amet, „Andmekvaliteedi juhised,“ 2020.
- [10] Amazon, „What Is SQL (Structured Query Language)?,“ [Võrgumaterjal]. Available: <https://aws.amazon.com/what-is/sql/>. [Kasutatud 09. 04. 2023].
- [11] D. Treece, „Stripe vs. PayPal,“ Business News Daily, 28. 02. 2023. [Võrgumaterjal]. Available: <https://www.businessnewsdaily.com/16289-stripe-vs-paypal.html>. [Kasutatud 09. 04. 2023].
- [12] Riigi Infosüsteemi Amet, „Andmekvaliteedi tagamise juhend andmekogu omanikele,“ Riigi Infosüsteemi Amet, 14.09.2016.
- [13] Riigi Infosüsteemi Haldussüsteem, „TSD,“ Riigi Infosüsteemi Haldussüsteem, 09. 12. 2019. [Võrgumaterjal]. Available: <https://www.riha.ee/Infosusteemid/Vaata/tsd>. [Kasutatud 09. 04. 2023].
- [14] Riigi Infosüsteemi Haldussüsteem, „TÖR,“ Riigi Infosüsteemi Haldussüsteem, 09. 12. 2019. [Võrgumaterjal]. Available: <https://www.riha.ee/Infosusteemid/Vaata/tor>. [Kasutatud 09. 04. 2023].
- [15] GeeksForGeeks, „Unified Modeling Language (UML) | An Introduction,“ 15. 12. 2022. [Võrgumaterjal]. Available: <https://www.geeksforgeeks.org/unified-modeling-language-uml-introduction/>. [Kasutatud 09. 04. 2023].
- [16] V. Srivastava, „What’s The Difference Between URLs and URIs?,“ NORDIC APIS, 02. 02. 2023. [Võrgumaterjal]. Available: <https://nordicapis.com/whats-the-difference-between-urls-and-uris/>. [Kasutatud 09. 04. 2023].

- [17] M. Miller, „CMS Market Share Monthly: 64.2% Of Sites Use WordPress,“ Search Engine Journal, 30. 06. 2022. [Võrgumaterjal]. Available: <https://www.searchenginejournal.com/cms-market-share-june-2022/455912/#close>. [Kasutatud 19. 11. 2022].
- [18] W3Techs, „Usage statistics of content management systems,“ [Võrgumaterjal]. Available: [https://w3techs.com/technologies/overview/content\\_management](https://w3techs.com/technologies/overview/content_management). [Kasutatud 20. 11. 2022].
- [19] J. Garcia ja R. Brandl, „WooCommerce Market Share Should be Top Ecommerce Plugin, But Is It?,“ ToolTester, [Võrgumaterjal]. Available: <https://www.tooltester.com/en/blog/wocommerce-market-share/>. [Kasutatud 19. 11. 2022].
- [20] Built With, „Plugin / Module Usage Distribution on the Entire Internet,“ [Võrgumaterjal]. Available: <https://trends.builtwith.com/shop/plugin---module/traffic/Entire-Internet>. [Kasutatud 20. 11. 2022].
- [21] WordPress, „Database Description,“ WordPress, [Võrgumaterjal]. Available: [https://codex.wordpress.org/Database\\_Description](https://codex.wordpress.org/Database_Description). [Kasutatud 11. 11. 2022].
- [22] A. Matkur ja L. Saluvere, „Ettevõttesise tarkvara täiendamine - kasutajaliidese ja andmebaasi kavandamine ning sobivuse hindamine,“ [Bakalaureusetöö], Infotehnoloogia teaduskond, TalTech, Tallinn, Eesti, 2022. [Online]. Available: <https://digikogu.taltech.ee/et/Item/d16e8c7d-a5e9-4efc-a741-65f17285f385>. [Kasutatud 20. 10. 2022].
- [23] A. Siitan, „E-poe andmete migreerimine PostgreSQL andmebaasist Microsoft SQL Serveri andmebaasi,“ [Bakalaureusetöö], Infotehnoloogia teaduskond, TalTech, Tallinn, Eesti, 2022. [Online]. Available: <https://digikogu.taltech.ee/en/item/c72040b7-7f99-4cd6-ace2-3578ab7f573c>. [Kasutatud 20. 10. 2022].
- [24] M. E. O. Oja, „EE Tele2 müügiandmete andmetabelite kvaliteedikontroll ja dokumenteerimine,“ [Bakalaureusetöö], Infotehnoloogia teaduskond, TalTech, Tallinn, Eesti, 2022. [Online]. Available: <https://digikogu.taltech.ee/et/Item/dea11d3f-fd23-4249-80ef-d03225b5225b>. [Kasutatud 20. 10. 2022].
- [25] M. Nemberg, „Suur A/B testimise ülevaade – mida testida, kuidas testida, kuidas planeerida,“ Trinidad Wiseman, 02. 12. 2015. [Võrgumaterjal]. Available: <https://blog.twn.ee/et/suur-ab-testimise-ulevaade-mida-testida-kuidas-testida-kuidas-planeerida>. [Kasutatud 19. 11. 2022].
- [26] Microsoft, „Microsoft 365 Onedrive,“ [Võrgumaterjal]. Available: <https://www.microsoft.com/et-ee/microsoft-365?rtc=1>. [Kasutatud 22. 01. 2023].
- [27] Microsoft, „Welcome to Microsoft Teams,“ [Võrgumaterjal]. Available: <https://www.microsoft.com/en-us/microsoft-teams/log-in>. [Kasutatud 22. 01. 2023].
- [28] Slack, „Made for people. Built for productivity.,“ [Võrgumaterjal]. Available: <https://slack.com>. [Kasutatud 22. 01. 2023].
- [29] Docker, „Develop faster. Run anywhere.,“ [Võrgumaterjal]. Available: <https://www.docker.com>. [Kasutatud 22. 01. 2023].
- [30] PhpMyAdmin, „Official phpMyAdmin Docker image,“ [Võrgumaterjal]. Available: <https://hub.docker.com/r/phpmyadmin/phpmyadmin>. [Kasutatud 22. 01. 2023].

- [31] MySQL, „Docker MySQL Official Image,“ [Võrgumaterjal]. Available: [https://hub.docker.com/\\_/mysql](https://hub.docker.com/_/mysql). [Kasutatud 22. 01. 2023].
- [32] JavaTpoint, „phpMyAdmin,“ JavaTpoint, [Võrgumaterjal]. Available: <https://www.javatpoint.com/phpmyadmin>. [Kasutatud 09. 04. 2023].
- [33] Sparx Systems Pty Ltd, „Make Your Vision a Reality,“ [Võrgumaterjal]. Available: <https://sparxsystems.com/>. [Kasutatud 05. 05. 2023].
- [34] E. Eessaar, „Database design queries,“ [Võrgumaterjal]. Available: <https://github.com/erki77/database-design-queries>. [Kasutatud 06. 05. 2023].
- [35] Sequel Ace, „MySQL/MariaDB database management for macOS,“ [Võrgumaterjal]. Available: <https://sequel-ace.com>. [Kasutatud 22. 01. 2023].
- [36] RazorSQL, „RazorSQL - Query, Edit, Browse, and Manage Databases,“ [Võrgumaterjal]. Available: <https://razorsql.com/index.html>. [Kasutatud 22. 01. 2023].
- [37] RazorSQL, „MySQL Database Conversion Tool,“ [Võrgumaterjal]. Available: [https://razorsql.com/features/mysql\\_database\\_conversion.html](https://razorsql.com/features/mysql_database_conversion.html). [Kasutatud 02. 05. 2023].
- [38] JetBrains, „PhpStorm,“ [Võrgumaterjal]. Available: <https://www.jetbrains.com/phpstorm/>. [Kasutatud 22. 01. 2023].
- [39] The PHP Group, „PHP,“ [Võrgumaterjal]. Available: <https://www.php.net/>. [Kasutatud 23. 04. 2023].
- [40] Mozilla, „HTML: HyperText Markup Language,“ [Võrgumaterjal]. Available: <https://developer.mozilla.org/en-US/docs/Web/HTML>. [Kasutatud 23. 04. 2023].
- [41] Mozilla, „CSS: Cascading Style Sheets,“ [Võrgumaterjal]. Available: <https://developer.mozilla.org/en-US/docs/Web/CSS>. [Kasutatud 23. 04. 2023].
- [42] Microsoft, „Microsoft OneNote,“ [Võrgumaterjal]. Available: <https://www.microsoft.com/en-us/microsoft-365/onenote/digital-note-taking-app>. [Kasutatud 22. 01. 2023].
- [43] MediaWiki, „MediaWiki is a collaboration and documentation platform brought to you by a vibrant community.,“ [Võrgumaterjal]. Available: <https://www.mediawiki.org/wiki/MediaWiki>. [Kasutatud 22. 01. 2023].
- [44] Atlassian Trello, „Trello brings all your tasks, teammates, and tools together,“ [Võrgumaterjal]. Available: <https://trello.com/home>. [Kasutatud 22. 01. 2023].
- [45] M. Jusupov, „Põhjalik A/B testimise plaan kodulehe müügile optimeerimiseks,“ 13. 01. 2016. [Võrgumaterjal]. Available: <https://kodulehekoolitused.ee/pohjalik-ab-testimise-plaan-kodulehe-muugile-optimeerimiseks/>. [Kasutatud 04. 2023].
- [46] Majandus- ja kommunikatsiooniministeerium, „a/b testimine,“ [Võrgumaterjal]. Available: <https://digiriik.eesti.ee/tooriist/ab-testimine>. [Kasutatud 30. 04. 2023].
- [47] G. K. Tayi ja D. P. Ballou, „Examining data quality,“ *Communications of the ACM*, kd. 41, nr 2, pp. 54-57, 01. 02. 1998.
- [48] F. G. Alizamini, M. M. Pedram, M. Alishahi ja K. Badie, „Data quality improvement using fuzzy association rules,“ %1 *International Conference on Electronics and Information Engineering*, Kyoto, 2010.
- [49] H. Veregin, „Data quality parameters,“ %1 *Geographical information systems*, 1999, pp. 177-189.
- [50] L. Cai ja Y. Zhu, „The Challenges of Data Quality and Data Quality Assessment in the Big Data Era,“ *Data Science Journal*, kd. 14, nr 2, pp. 1-10, 22. 05. 2015.

- [51] C. Stedman ja J. Vaughan, „Definition data quality,“ TechTarget, [Võrgumaterjal]. Available: <https://www.techtarget.com/searchdatamanagement/definition/data-quality>. [Kasutatud 01. 12. 2022].
- [52] HEAVY.AI, „Data Quality Definition,“ [Võrgumaterjal]. Available: <https://www.heavy.ai/technical-glossary/data-quality>. [Kasutatud 10. 12. 2022].
- [53] Passionned Group, „Data Quality,“ [Võrgumaterjal]. Available: <https://www.passionned.com/bi/data-quality/>. [Kasutatud 15. 12. 2022].
- [54] T. N. Herzog, F. J. Scheuren ja W. E. Winkler, „What is Data Quality and Why Should We Care?,“ %1 *Data Quality and Record Linkage Techniques*, Springer, New York, NY, 2007.
- [55] *Data Quality model*, ISO/IEC 25012. [Online] Available: <https://iso25000.com/index.php/en/iso-25000-standards/iso-25012>.
- [56] C. Cichy ja S. Rass, „An Overview of Data Quality Frameworks,“ *IEEE Access*, kd. 7, pp. 24634 - 24648, 15. 02. 2019.
- [57] L. L. Pipino, Y. W. Lee ja R. Y. Wang, „Data Quality Assessment,“ *Communications of the ACM*, kd. 45, nr 4, pp. 211-218, 04. 2002.
- [58] A. Gupta, „The 6 dimensions of data quality,“ 29. 08. 2022. [Võrgumaterjal]. Available: <https://www.collibra.com/us/en/blog/the-6-dimensions-of-data-quality>. [Kasutatud 03. 2023].
- [59] R. L. Sarfin, „5 Characteristics of Data Quality,“ Precisely, 02. 11. 2022. [Võrgumaterjal]. Available: <https://www.precisely.com/blog/data-quality/5-characteristics-of-data-quality>. [Kasutatud 03. 2023].
- [60] Office Of Elementary & Secondary Education, „Data Quality Component 4: Data Coherence,“ 27. 08. 2020. [Võrgumaterjal]. Available: <https://oese.ed.gov/resources/oese-technical-assistance-centers/state-support-network/resources/data-quality-component-4-data-coherence/>. [Kasutatud 03. 2023].
- [61] National Water Quality Monitoring Council, „The Value of Data Comparability,“ [Võrgumaterjal]. Available: [https://acwi.gov/methods/pubs/over\\_pubs/valcomp\\_fs.pdf](https://acwi.gov/methods/pubs/over_pubs/valcomp_fs.pdf). [Kasutatud 03. 2023].
- [62] R. Vaziri, M. Mohsenzadeh ja J. Habibi, „TBDQ: A Pragmatic Task-Based Method to Data Quality Assessment and Improvement,“ *PLOS ONE*, 2016.
- [63] Y. W. Lee, D. M. Strong, B. K. Kahn ja R. Y. Wang, „AIMQ: a methodology for information quality assessment,“ *Information & Management*, kd. 40, nr 2, pp. 133-146, 05. 05. 2002.
- [64] C. Batini, C. Cappiello, C. Francalanci ja A. Maurino, „Methodologies for Data Quality Assessment and Improvement,“ *ACM Computing Surveys*, kd. 41, nr 3, pp. 1-52, 30. 07. 2019.
- [65] M. Mrkaic, „Information Content of DQAF Indicators—Empirical Entropy Analysis,“ 01. 09. 2010. [Võrgumaterjal]. Available: <https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Information-Content-of-DQAF-Indicators-Empirical-Entropy-Analysis-24186>.
- [66] C. Batini, D. Barone, F. Cabitza ja S. Grega, „A Data Quality Methodology for Heterogeneous Data,“ *International Journal of Database Management Systems (IJDMS)*, kd. 3, nr 1, pp. 60-79, 01. 02. 2011.



- [67] C. Cappiello, P. Ficiaro ja B. Pernici, HIQM: A Methodology for Information Quality Monitoring, Measurement, and Improvement. In: , et al. Advances in Conceptual Modeling - Theory and Practice. ER 2006. Lecture Notes in Computer Science, kd 4231., Springer, Berlin, Heidelberg, 2006.
- [68] Riigi Infosüsteemi Haldussüsteem, „Ülevaade riigi infosüsteemist,“ [Võrgumaterjal]. Available: <https://www.riha.ee/Avaleht>. [Kasutatud 10. 04. 2023].
- [69] Riigi Infosüsteemi Haldussüsteem, „Infosüsteemid,“ [Võrgumaterjal]. Available: [https://www.riha.ee/Infosüsteemid?systemStatus=IN\\_USE&searchText=Maksu-%20ja%20tolliamet&sort=name&dir=DESC](https://www.riha.ee/Infosüsteemid?systemStatus=IN_USE&searchText=Maksu-%20ja%20tolliamet&sort=name&dir=DESC). [Kasutatud 10. 04. 2023].
- [70] Math Rank, „Score 100/100 With Rank Math Post Tests,“ [Võrgumaterjal]. Available: <https://rankmath.com/kb/score-100-in-tests/>. [Kasutatud 02. 2023].
- [71] T. Shellberg, „Getting started with the REST API,“ WooCommerce, 19. 08. 2019. [Võrgumaterjal]. Available: <https://github.com/woocommerce/woocommerce/wiki/Getting-started-with-the-REST-API>. [Kasutatud 10. 04. 2023].
- [72] WooCommerce, „WooCommerce REST API,“ [Võrgumaterjal]. Available: <https://woocommerce.com/document/woocommerce-rest-api/>. [Kasutatud 10. 04. 2023].
- [73] WP Staging, „The WordPress Database Structure,“ 2020. [Võrgumaterjal]. Available: [https://wp-staging.com/docs/the-wordpress-database-structure/#wp\\_users\\_wp\\_usermeta](https://wp-staging.com/docs/the-wordpress-database-structure/#wp_users_wp_usermeta). [Kasutatud 01. 2023].
- [74] Rank Math, „Making the Most of the Analytics in Rank Math,“ [Võrgumaterjal]. Available: <https://rankmath.com/kb/analytics/>. [Kasutatud 01. 2023].
- [75] M. Jolley, „Experiment: Add Rate Limits to Store API,“ 04. 11. 2022. [Võrgumaterjal]. Available: <https://github.com/woocommerce/woocommerce-blocks/pull/5962>. [Kasutatud 03. 05. 2023].
- [76] WooCommerce, „WooCommerce Store API,“ [Võrgumaterjal]. Available: <https://github.com/woocommerce/woocommerce-blocks/tree/trunk/src/StoreApi>. [Kasutatud 03. 05. 2023].
- [77] WooCommerce, „Rate Limiting for Store API endpoints,“ [Võrgumaterjal]. Available: <https://github.com/woocommerce/woocommerce-blocks/blob/trunk/src/StoreApi/docs/rate-limiting.md#rate-limiting-options-filter>. [Kasutatud 03. 05. 2023].
- [78] WooCommerce, „Webhooks,“ [Võrgumaterjal]. Available: <https://woocommerce.com/document/webhooks/>. [Kasutatud 01. 2023].
- [79] D. Leydon, „How Do I Use WooCommerce Webhooks?,“ [Võrgumaterjal]. Available: <https://www.websitebuilderinsider.com/how-do-i-use-woocommerce-webhooks/>. [Kasutatud 01. 2023].
- [80] Kinsta, „How to Clean up Your wp\_options Table and Autoloaded Data,“ 06. 07. 2022. [Võrgumaterjal]. Available: <https://kinsta.com/knowledgebase/wp-options-autoloaded-data/>. [Kasutatud 01. 2023].
- [81] WooCommerce, „Database Description,“ [Võrgumaterjal]. Available: <https://github.com/woocommerce/woocommerce/wiki/Database-Description>. [Kasutatud 01. 2023].
- [82] Rank Math, „How to Monitor 404 Errors?,“ [Võrgumaterjal]. Available: <https://rankmath.com/kb/monitor-404-errors/>. [Kasutatud 01. 2023].

- [83] WooCommerce, „Adding and Managing Products,“ [Võrgumaterjal]. Available: <https://woocommerce.com/document/managing-products/>. [Kasutatud 01. 2023].
- [84] Unnicommerce, „What is an Order, Order item and Shipment? What are their statuses during the life cycle of an order?,“ [Võrgumaterjal]. Available: <https://support.unicommerce.com/index.php/knowledge-base/what-is-an-order-order-item-and-shipment-what-are-their-statuses-during-the-life-cycle-of-an-order/>. [Kasutatud 01. 2023].
- [85] *Eesti Vabariigi Avaliku teabe seadus*. Vastu võetud 01.05.2023. RT I, 07.03.2023, 11.
- [86] *Eesti Vabariigi Tarbijakaitse seadus*. Vastu võetud 28.05.2022. RT I, 01.04.2022, 7.
- [87] E. Eessaar, „Teema 9. Andmebaasi loogilise disaini tulemuse parandamine ja headuse kontrollimine,“ 2023. [Võrgumaterjal]. Available: <https://maurus.ttu.ee/download.php?aine=346&document=32230&tyyp=do>. [Kasutatud 17. 04. 2023].
- [88] E. Eessaar. (2020). Automating Detection of Occurrences of PostgreSQL Database Design Problems. In: Robal, T., Haav, HM., Penjam, J., Matulevičius, R. (eds) Databases and Information Systems. DB&IS 2020. Communications in Computer and Information Science, vol 1243. Springer, Cham. Available: [https://doi.org/10.1007/978-3-030-57672-1\\_14](https://doi.org/10.1007/978-3-030-57672-1_14). [Kasutatud 05. 05. 2023].
- [89] PostgreSQL, „9.9. Date/Time Functions and Operators,“ [Võrgumaterjal]. Available: <https://www.postgresql.org/docs/9.4/functions-datetime.html>. [Kasutatud 02. 05. 2023].
- [90] R. C. Martin, „Chapter 17: Smells And Heuristics,“ %1 *Clean Code: A Handbook of Agile Software Craftsmanship*, Pearson, 08.2008.
- [91] E. Eessaar, „Using in some way reserved (in PostgreSQL) SQL keywords as the names of a database object (detailed view)“. Available: [https://maurus.ttu.ee/design\\_queries/query.php?query\\_id=327](https://maurus.ttu.ee/design_queries/query.php?query_id=327). [Kasutatud 05. 05. 2023].
- [92] PostgreSQL, „Appendix C. SQL Key Words,“ [Võrgumaterjal]. Available: <https://www.postgresql.org/docs/current/sql-keywords-appendix.html>. [Kasutatud 04. 05. 2023].
- [93] MySQL, „9.3 Keywords and Reserved Words,“ [Võrgumaterjal]. Available: <https://dev.mysql.com/doc/refman/8.0/en/keywords.html#keywords-8-0-detailed-T>. [Kasutatud 05. 05. 2023].
- [94] E. Eessaar, „Names of columns that hold personal names but do not take into account cultural diversity“. Available: [https://maurus.ttu.ee/design\\_queries/query.php?query\\_id=310](https://maurus.ttu.ee/design_queries/query.php?query_id=310). [Kasutatud 03. 05. 2023].
- [95] M. Jõgi, „Mõned disainilahendused isikunimede hoidmiseks SQL-andmebaasides,“ [Bakalaureusetöö], Infotehnoloogia teaduskond, TalTech, Tallinn, Eesti, 2016. [Online]. Available: <https://digikogu.taltech.ee/et/item/ab7e8f7e-ce28-4cc1-bb78-40c624d19bf6>. [Kasutatud 10. 05. 2022].
- [96] Wikipedia, „Mononym,“ [Võrgumaterjal]. Available: <https://en.wikipedia.org/wiki/Mononym>. [Kasutatud 05. 2023].

- [97] B. Karwin, „Chapter 6: Entity-Attribute-Value,“ %1 *SQL Antipatterns*, Pragmatic Bookshelf, 06.2010.
- [98] „Boolean vs tinyint(1) for boolean values in MySQL,“ [Võrgumaterjal]. Available: <https://stackoverflow.com/questions/3751853/boolean-vs-tinyint1-for-boolean-values-in-mysql>. [Kasutatud 05. 2023].
- [99] A. Klaos, S. T. Laisar, P. Luik, T. Palts ja E. Ääremaa, „Lisamoodulid: 6 Vaikeväärtused ja kitsendused,“ [Võrgumaterjal]. Available: <https://web.htk.tlu.ee/digitaru/lisamoodulid/chapter/vaikevaartused-ja-kitsendused/>. [Kasutatud 05. 05. 2023].
- [100] E. Eessaar, „Base tables that have a surrogate key and do not have any unique constraints“. Available: [https://maurus.ttu.ee/design\\_queries/query.php?query\\_id=15](https://maurus.ttu.ee/design_queries/query.php?query_id=15). [Kasutatud 05. 05. 2023].
- [101] B. Snaidero, „Surrogate Key vs Natural Key Differences and When to Use in SQL Server,“ [Võrgumaterjal]. Available: <https://www.mssqltips.com/sqlservertip/5431/surrogate-key-vs-natural-key-differences-and-when-to-use-in-sql-server/>. [Kasutatud 05. 05. 2023].
- [102] B. Karwin, „Chapter 5: Keyless Entry,“ %1 *SQL Antipatterns*, Pragmatic Bookshelf, 06.2010.
- [103] B. Karwin, „Chapter 4: ID Required,“ %1 *SQL Antipatterns*, Pragmatic Bookshelf, 06.2010.
- [104] I. Looremata, „Data Quality Management Best Practices,“ 25. 08. 2022. [Võrgumaterjal]. Available: <https://www.linkedin.com/pulse/data-quality-management-best-practices-isaac-looremata?trk=pulse-article>. [Kasutatud 14. 04. 2023].
- [105] B. Calzon, „The Ultimate Guide to Modern Data Quality Management (DQM) For An Effective Data Quality Control Driven by The Right Metrics,“ 30. 09. 2022. [Võrgumaterjal]. Available: <https://www.datapine.com/blog/data-quality-management-and-metrics/#best-practices>. [Kasutatud 14. 04. 2023].
- [106] C. Tozzi, „Data Quality Best Practices,“ 11. 06. 2022. [Võrgumaterjal]. Available: <https://www.precisely.com/blog/data-quality/data-quality-best-practices>. [Kasutatud 14. 04. 2023].
- [107] K. D. Foote, „Data Quality Best Practices,“ 25. 10. 2022. [Võrgumaterjal]. Available: <https://www.dataversity.net/data-quality-best-practices/#>. [Kasutatud 14. 04. 2023].
- [108] I. Mikhailouskaya (Makaranka), „Your Guide to Data Quality Management,“ 13. 12. 2018. [Võrgumaterjal]. Available: <https://www.scnsoft.com/blog/guide-to-data-quality-management#practices>. [Kasutatud 14. 04. 2023].
- [109] S. Shen, „7 Steps to Ensure and Sustain Data Quality,“ 29. 07. 2019. [Võrgumaterjal]. Available: <https://towardsdatascience.com/7-steps-to-ensure-and-sustain-data-quality-3c0040591366>. [Kasutatud 14. 04. 2023].
- [110] WordPress, „Post Status,“ [Võrgumaterjal]. Available: <https://wordpress.org/documentation/article/post-status/>. [Kasutatud 06. 05. 2023].
- [111] raincoaster, „What is a ping?,“ [Võrgumaterjal]. Available: <https://wordpress.com/forums/topic/what-is-a-ping/>. [Kasutatud 06. 05. 2023].

- [112] raincoaster, „What are pings for?“, [Võrgumaterjal]. Available: <https://wordpress.com/forums/topic/what-are-pings-for/>. [Kasutatud 06. 05. 2023].
- [113] WordPress Development, „What is the use of to\_ping and pinged column?“, [Võrgumaterjal]. Available: <https://wordpress.stackexchange.com/questions/100556/what-is-the-use-of-to-ping-and-pinged-column>. [Kasutatud 06. 05. 2023].
- [114] WordPress development, „When is the 'post\_content\_filtered' column in database cleared by WordPress?“, [Võrgumaterjal]. Available: <https://wordpress.stackexchange.com/questions/113387/when-is-the-post-content-filtered-column-in-database-cleared-by-wordpress>. [Kasutatud 06. 05. 2023].
- [115] WordPress experts "Editorial Staff", „WordPress Post Revisions Made Simple: A Step by Step Guide“, [Võrgumaterjal]. Available: <https://www.wpbeginner.com/beginners-guide/complete-guide-to-wordpress-post-revisions/>. [Kasutatud 06. 05. 2023].
- [116] WordPress, „Developer Resources: Post Types“, [Võrgumaterjal]. Available: <https://developer.wordpress.org/themes/basics/post-types/#post>. [Kasutatud 06. 05. 2023].
- [117] WP engine, „MIME Types in WordPress“, [Võrgumaterjal]. Available: <https://wpengine.com/support/mime-types-wordpress/>. [Kasutatud 06. 05. 2023].
- [118] MySQL, „15.1 Introduction to InnoDB“, [Võrgumaterjal]. Available: <https://dev.mysql.com/doc/refman/8.0/en/innodb-introduction.html>. [Kasutatud 05. 05. 2023].
- [119] G. Hamzaj, Z. Dika ja G. Armenski, „An Overview on How to Choose the Data Quality Assessment and Improvement Frameworks“, %1 *ICT Innovations 2021*, 2021.

## **Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks**

Mina, Iris Nael

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „E-kaubandusettevõtte X andmekvaliteedi analüüs ja parendamine“, mille juhendaja on Erki Eessaar.
  - 1.1.reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
  - 1.2.üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

06.05.2023

## Lisa 2 – Ettevõtte X wiki

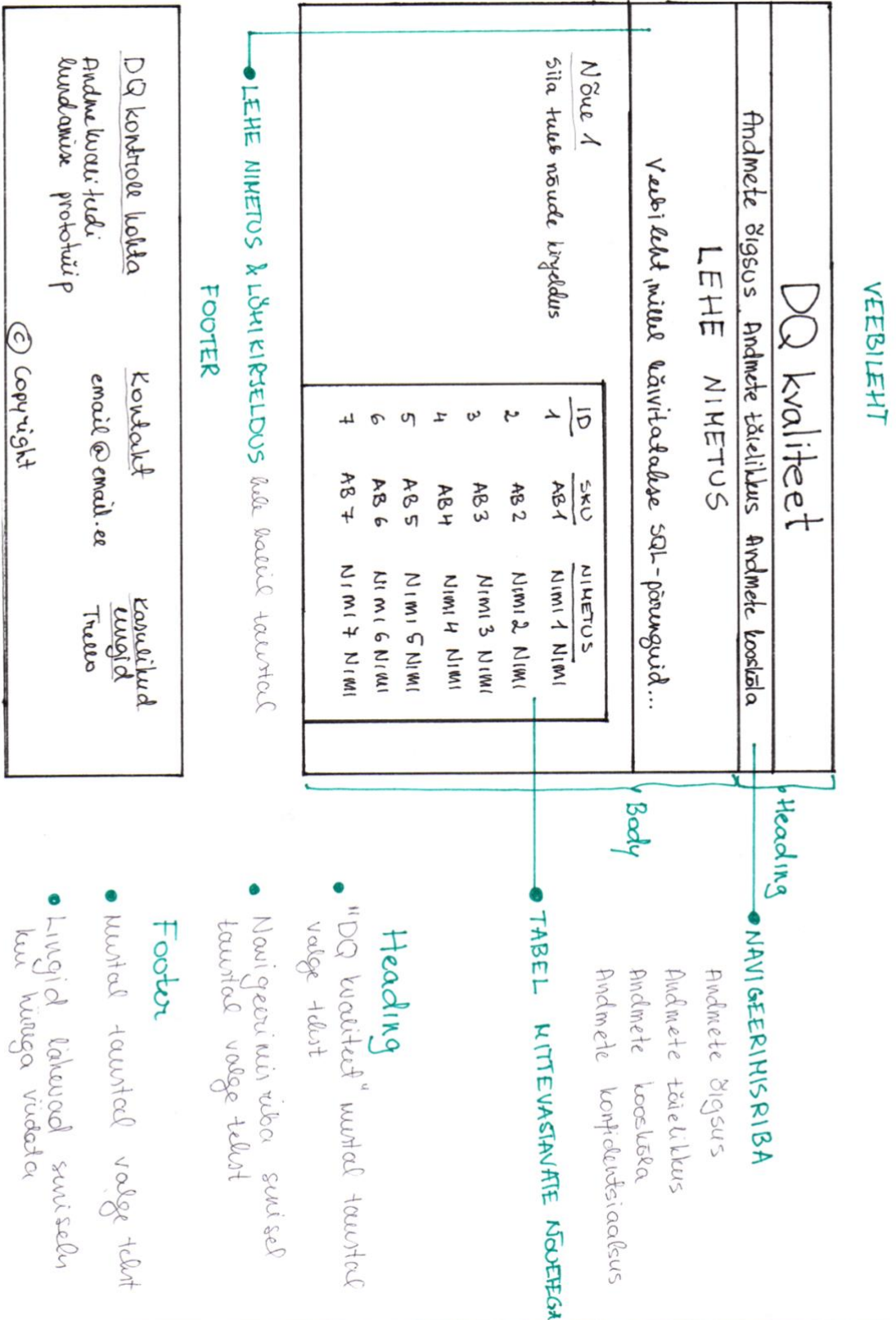
Ettevõtte üldine wiki struktuur.

▼ Ettevõtte Teadmisteait	
> 1. Ettevõtte üldine informatsioon	
> 2. Ettevõtte projektid	
> 3. Ettevõtte kaubad, teenused, kurs...	
> 4. Ettevõtte materjalid	
> 5. Ettevõtte marketing	
> 6. Kasutusjuhendid ja mallid	
> 7. Ettevõtte veebisaidid	
> 8. Enesearenguks	

Andmekvaliteedi raamistik ja andmekvaliteedi nõuded ettevõtte wikis.

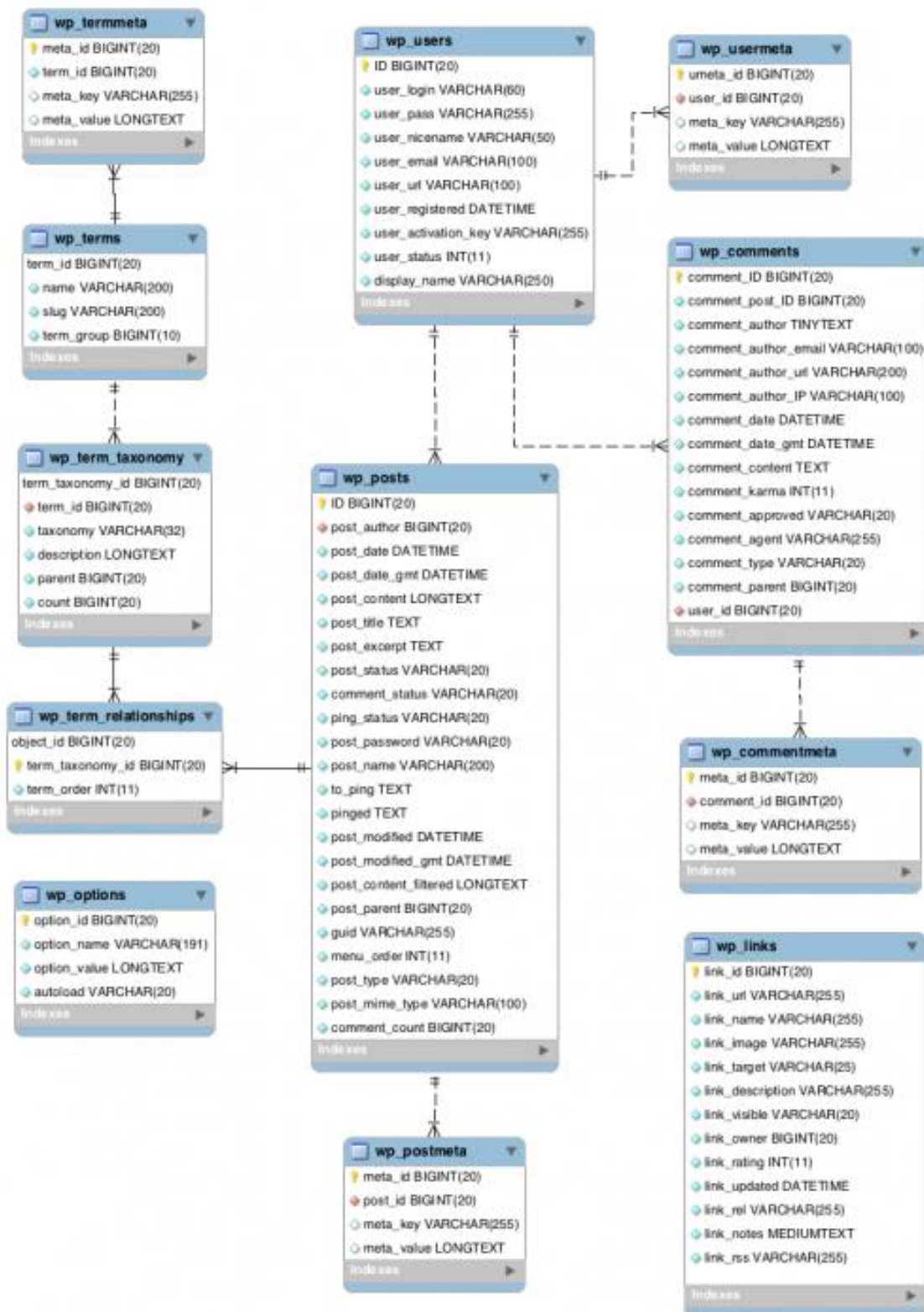
▼ Ettevõtte Teadmisteait	Andmekvaliteedi juhtimise raamistik
▼ 1. Ettevõtte üldine informatsioon	1. Õigsus
> 1.1. Ettevõtte töötajate kontakt ja...	2. Täielikkus
▼ 1.2. Ettevõtte reeglid ja nõuded	3. Kooskõla
Kampaania keskkonna reeglid	4. Usaldusväärsus
Andmekvaliteedi nõuded	5. Ajakohasus
▼ 1.3. Ettevõtte bränd-book	6. Reeglipärasus
> 1.4. Ettevõtte kalender	7. Konfidentsiaalsus
> 1.5. Informatsioon uutele töötajatele	8. Ühekordsus
> 2. Ettevõtte projektid	9. Mitteliiasus
> 3. Ettevõtte kaubad, teenused, kurs...	
> 4. Ettevõtte materjalid	
> 5. Ettevõtte marketing	
> 6. Kasutusjuhendid ja mallid	
> 7. Ettevõtte veebisaidid	
> 8. Enesearenguks	

# Lisa 3 – Andmekvaliteedi hindamistöölaua prototüübi esialgne kasutajaliidese kavand



## Lisa 4 - WordPressi andmebaasi disaini diagramm

WordPressi dokumentatsioonist pärit andmebaasi diagramm, millel on esitatud WordPressi tabelid ja nende vahelises seosed [21].





## Lisa 5 – Müügiobjektide register

WordPressi andmebaasi tabel *wp\_posts*.

Veeru nimi	Veeru tüüp	Kohustus	Veeru sõnaline kirjeldus
ID	BIGINT	Jah	Postituse unikaalne identifikaator.
post_author	BIGINT	Jah	Postituse autori identifikaator.
post_date	DATETIME	Jah	Postituse loomise kuupäev ja kellaeg sekundite täpsusega. Näiteks „2018-12-21 15:00:10“.
post_date_gmt	DATETIME	Jah	Postituse loomise kuupäev ja kellaeg sekundite täpsusega, universaalajas. Näiteks „2018-12-21 13:00:10“.
post_content	LONGTEXT	Jah	Postituse sisu. Näiteks kauba korral on see kauba kirjeldus.
post_title	TEXT	Jah	Postituse pealkiri.
post_excerpt	TEXT	Jah	<p>Postituse sisu kärbitud versioon. Postituse autor võib ise kirjutada postituse lühendatud versiooni või jätta selle tegemata ning kirjelduse lühendatud versioon genereeritakse automaatselt.</p> <p>Kaupade, toodete ja teenuste lisamise puhul on kärbitud versioonis lühikirjeldus ning kui autor jätab selle kirjutamata, siis automaatselt sinna teksti kirjeldusest ei genereerita.</p>
post_status	VARCHAR(20)	Jah	<p>Postituse avalikustamise staatus. WordPressil on olemas kaheksa postituse staatust [110]:</p> <ul style="list-style-type: none"> <li>▪ <i>Publish</i> – postitus on avalikustatud ja kõigile nähtav.</li> <li>▪ <i>Future</i> – postituse avalikustamine on etteplaneeritud kindlale kuupäevale ja ajale.</li> <li>▪ <i>Draft</i> – postitus on mustandina üleväl ja kättesaadav vaid vastava õigustega WordPressi kasutajatele.</li> </ul>

Veeru nimi	Veeru tüüp	Kohustus	Veeru sõnaline kirjeldus
			<ul style="list-style-type: none"> <li>▪ <i>Pending</i> – postitus ootab toimetajalt avaldamisluba.</li> <li>▪ <i>Private</i> – postitus on avalikustatud, kuid nähtav vaid WordPressi administraatorkasutajatele.</li> <li>▪ <i>Trash</i> – postitus on prügikastis.</li> <li>▪ <i>Auto-draft</i> – autori tehtud muudatused on postituses automaatselt salvestati.</li> <li>▪ <i>Inherit</i> – postituses olevad meediafailid omavad saatust „pärib“.</li> </ul>
comment_status	VARCHAR()	Jah	Postituse kommenteerimise võimalus. Kui postituse kommenteerimise staatuseks on „opened“, siis saab seda igäüks kommenteerida. Juhul, kui postituse kommenteerimise staatus on „closed“, siis ei saa seda keegi kommenteerida.
ping_status	VARCHAR()	Jah	Vaikeväärtus on „open“. Peale seda kui ping saab tehtud muutub väärtus automaatselt „closed“ [111]. Pingimine on automaatselt suletud vaid juhul, kui postituse staatuseks on „private“ [111]. WordPressi ping teavitab otsingumootoreid ja teavituste tellijaid sellest, et veebilehel on uus postitus [112].
post_password	VARCHAR(255)	Jah	Postitusele määratud parool. Kõigile avalikustatud postitustele saab määrata parooli, et postitust näeksid vaid need, kellel on parool olemas.
post_name	VARCHAR(200)	Jah	Postituse nimetus <i>slug</i> kujul. Näiteks kui postituse pealkiri on „Tarkvara kasutamise konsultatsioon“, siis selle nimetus <i>slug</i> kujul on „tarkvara-kasutamise-konsultatsioon“. Üldiselt määrab WordPress selle automaatselt, kuid postituse autor saab seda käsitsi muuta.

Veeru nimi	Veeru tüüp	Kohustus	Veeru sõnaline kirjeldus
to_ping	TEXT	Jah	URLide list, millele WordPress peab <i>pingbacke</i> saatma [113].
pinged	TEXT	Jah	URLide list, millele WordPress on <i>pingbacke</i> tagasi saanud [113].
post_modified	DATETIME	Jah	Postituse viimase muutmise aeg kuupäeva ja aeg sekundite täpsusega.
post_modified_gmt	DATETIME	Jah	Postituse viimase muutmise aeg kuupäeva ja aeg sekundite täpsusega, universaalajas.
post_content_filtered	LONGTEXT	Jah	Kasutatakse ajutisteks salvestamiseks. Selle sisu kaob kohe, kui postitust uuendatakse. [114]. Näiteks siia salvestatakse järgmine informatsioon [114]. <ul style="list-style-type: none"> <li>▪ Postituse publitseerimine ajastati automaatselt.</li> <li>▪ Postitust muudeti kasutades „Quick Edit“.</li> </ul>
post_parent	BIGINT	Jah	Vanempostituse unikaalne identifikaator. Tavaliselt on antud veeru väärtus „0“, mis tähendab, et postitusel ei ole vanem-postitust, millelt ta midagi saaks pärida. Nullist („0“) erinevad väärtused antud veerus on üldiselt meediafailide puhul, sest need pärivad automaatselt vanem-postituse identifikaatorit.
guid	VARCHAR(255)	Jah	Antud veerus hoitakse WordPressi sisest linki postitusele. Näide lingist, kus <i>p</i> tähendab postitust ja xxxxxx postituse unikaalset identifikaatorit: <a href="https://domeen.ee/?p=xxxxxx">https://domeen.ee/?p=xxxxxx</a>
menu_order	INT	Jah	Postituste kuvamise järjestus. Vaikeväärtusena määratakse kõigile postitustele järjekorraks „0“. Kõik nullist erinevad väärtused tähistavad, et antud postitust on kasutatud veebilehe menüüs.
post_type	VARCHAR(20)	Jah	Postituse tüübid. Wordpressil on seitse üldist postituse tüüpi [115], [116]:

Veeru nimi	Veeru tüüp	Kohustus	Veeru sõnaline kirjeldus
			<ul style="list-style-type: none"> <li>▪ <i>Post</i> – üldiselt on tegemist ajaveebi postitusega.</li> <li>▪ <i>Page</i> – veebileht. Näiteks koduleht.</li> <li>▪ <i>Attachment</i> – manus. Näiteks pilt, video või dokumendifail.</li> <li>▪ <i>Revision</i> – WordPressi sisseehitatud funktsioon, mis annab võimaluse postituses tehtuid muudatusi tagasi võtta.</li> <li>▪ <i>Navigation menu</i> – WordPressi veebilehe menüü.</li> <li>▪ <i>Block templates</i> – WordPressi postitus koosneb plokkidest, mida on võimalik salvestada ja edaspidi teistes postitustes kasutada. Näiteks võib luua ploki „Kontaktandmed küsimuste korral“ ja luues uut postitust, mitte uuesti käsitsi sisestada kontaktandmeid küsimuste korral, vaid valida eelnevalt salvestatud plokkide seas sobiv.</li> <li>▪ <i>Template parts</i> ehk mall – plokkide mustrid, mida saab kasutada WordPressi postituste loomisel.</li> </ul>
post_mime_type	VARCHAR(100)	Jah	MIME tüübid, mida veebilehitsejad jm programmid kasutavad, et määrata postitusega seotud sisu tüüp [117]. Näiteks .webp ja .png failide MIME tüüp on pilt, .mp4 ja .mov tüüp on video.
comment_count	BIGINT	Jah	Postitusel olevate kommentaaride kogus.

## Lisa 6 – Andmekvaliteedi küpsusmudel

Andmekvaliteedi küpsusmudel Riigi Infosüsteemi Ameti projektist „Andmekvaliteedi tagamise juhend“. Mudelis on tehtud väikseid muudatusi, et see sobiks rohkem väikeettevõtte vajadustele. Väidete sõnastust ei peetud vajalikuks muuta.

<b>Kategooria: Juhtimine ja planeerimine</b>	
<b>Tase</b>	<b>Väide</b>
1	Andmekogu andmekvaliteedi nõuded ei ole määratletud ega dokumenteeritud.
2	Andmekogu andmekvaliteedi nõuded on määratletud ja dokumenteeritud.
2	Kriitilised andmed on määratletud ja dokumenteeritud.
3	Andmekvaliteet on nõuetele vastavuse osas kontrollitud.
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.
5	Andmekvaliteedi protsesside osas toimub järjepidev parendamine.
5	Andmekvaliteeti juhitakse poliitikate regulaarse ülevaatus, muutmise ja avaldamisega.
<b>Kategooria: Organisatsioon ja vastutused</b>	
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.
5	Rühmituses toimub regulaarne andmekvaliteedi juhtimise vastutuste ülevaatus ja täiendamine.
<b>Kategooria: Protsessid</b>	

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.
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.
<b>Kategooria: Teadmus ja kompetentsid</b>	
1	Andmekvaliteedi teadlikkuse tõstmiseks ei ole läbi viidud andmekvaliteedi teemalisi koolitusi.
1	Võtmeisikud ei ole tuvastatud.
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.

3	Jagatud teadmusbasisina dokumenteeritakse parimad praktikad andmekvaliteedi tagamisel.
4	Andmekvaliteedi teemalised koolitused toimuvad regulaarselt.
5	Andmekvaliteedi teemaliste koolituste sisu regulaarselt vaadatakse üle ja täiendatakse vastavalt vajadusele.
<b>Kategooria: Töövahendid</b>	
1	Andmekvaliteedi standardid infotehnoloogiliste vahendite jaoks ei ole välja töötatud ja ei ole dokumenteeritud.
1	Infotehnoloogilised vahendid andmekvaliteedi mõõtmiseks ei ole juurutatud.
2	Andmekvaliteedi standardid infotehnoloogiliste vahendite jaoks on välja töötatud ja dokumenteeritud.
2	Töövahendid andmekvaliteedi mõõtmiseks ja parandamiseks on osaliselt juurutatud.
3	Juurutatud on töövahendid andmekvaliteedi mõõtmiseks ja parandamiseks.
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 andmeparanduse või - kontrolli protseduurid.
5	Andmekogude arendamisel on rakendatud põhiandmete põhimõtteid.
5	Andmekvaliteedi juhtimist toetavate infosüsteemide arendustegevuste sihtportfell vaadatakse regulaarselt üle ja täiendatakse vastavalt vajadusele.