

TALLINNA TEHNIKAÜLIKOOL  
Infotehnoloogia teaduskond

Rain Hallikas

**EUROOPA JA RAHVUSVAHELISTEST XML  
FORMAADIS STANDARDITEST EESTI  
STANDARDITE KOOSTAMISE  
LAHENDUSE VÄLJATÖÖTAMINE EESTI  
STANDARDIKESKUSELE**

Bakalaureusetöö

Juhendaja: Gunnar Piho  
dotsent

Tallinn 2019

## **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: Rain Hallikas

18.05.2019

## **Annotatsioon**

Käesoleva bakalaureusetöö eesmärgiks oli luua Eesti Standardikeskusele lahendus Eesti standardi dokumentide koostamiseks XML baasil, mis vahetaks välja Eesti standardite koostamise PDF baasil ja pakuks rohkem funktsionaalsust.

Töö tulemuseks on XML-HTML-PDF teisendaja, mis suudab Rahvusvahelise Standardimisorganisatsiooni (ISO), Euroopa Standardikomitee (CEN) ja Euroopa Elektrotehnika Standardikomitee (CENELEC) poolt välja antud XML formaadis standarditest kokku panna Eesti standardeid ja nende eelvaateid nii PDF kui ka HTML formaadis.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 35 leheküljel, 5 peatükki, 18 joonist.

## **Abstract**

### **Developing a solution for the Estonian Centre for Standardisation for generating Estonian standards from European and international XML standards**

The aim of this Bachelor's thesis was to create a solution for the Estonian Centre for Standardisation to compose Estonian standard documents from XML, which would replace the Estonian standard compilation from PDF and would offer more functionality.

The result of this thesis is an XML-HTML-PDF converter that can compile Estonian standards and their previews in both PDF and HTML format from XML standards issued by the International Organization for Standardization (ISO), European Committee for Standardization (CEN) and European Committee for Electrotechnical Standardization (CENELEC).

The thesis is in Estonian and contains 35 pages of text, 5 chapters, 18 figures.

## Lühendite ja mõistete sõnastik

EVS	Eesti Standardikeskus [1]
XML	<i>eXtensible Markup Language</i> , informatsiooni tekstiformaat [2]
HTML	<i>HyperText Markup Language</i> [3]
PDF	<i>Portable Document Format</i> [4]
CSS	<i>Cascading Style Sheets</i> [5]
ISO	Rahvusvaheline Standardimisorganisatsioon [6]
IEC	Rahvusvaheline Elektrotehnika Komisjon [7]
CEN	Euroopa Standardimiskomitee [8]
CENELEC / CLC	Euroopa Elektrotehnika Standardimiskomitee [8]
DOC / Word	Microsoft Word Document [9]
STS4i	Standards Tag Suite for interoperability [10]
XSL	<i>eXtensible Stylesheet Language</i> [11]
StaDIS	Eesti Standardikeskuse standardite andmebaasi veebirakendus

## Sisukord

1	Sissejuhatus .....	8
1.1	Standardite üle võtmine .....	8
1.2	Probleem PDF-PDF teisendajaga .....	9
1.3	Töö eesmärk .....	10
1.4	Töö struktuur .....	11
2	Metoodika.....	12
2.1	Nõuded.....	12
2.2	Tööriistad.....	16
2.3	Töö protsess.....	19
3	Realisatsioon.....	20
3.1	Arhitektuur ja tehnilised andmed .....	20
3.2	XML-HTML-PDF teisendaja .....	25
3.3	Kood .....	28
4	Analüüs.....	34
4.1	Funktsionaalsus .....	34
4.2	Parendused.....	38
4.3	Autori hinnang.....	41
5	Kokkuvõte .....	42
	Kasutatud kirjandus .....	43

## Jooniste loetelu

Joonis 1. Dokumendi neljas lehekülg on valesti nummerdatud .....	14
Joonis 2. Kahel järjestikusel lehel on parempoolsed veerised.....	15
Joonis 3 Lahenduse arhitektuur .....	20
Joonis 4 Lahenduse objektid .....	21
Joonis 5 Rakenduse loogika .....	21
Joonis 6 Failide konfigureeritavad asukohad .....	22
Joonis 7. Rakenduse tööks vajalikud failid .....	22
Joonis 8. Rakenduse struktuur .....	23
Joonis 9 EVS XML paketi näide .....	26
Joonis 10 FillEnTitlepage meetod .....	28
Joonis 11 MoveArrayOutOfImage meetod .....	29
Joonis 12 Lahenduse Prince-CSS kood .....	29
Joonis 13 Rakenduse tavaline CSS kood.....	30
Joonis 14 EVS eessõna mall.....	31
Joonis 15 STS4i tööriistade XSL kood .....	32
Joonis 16 ResizeImage testid.....	33
Joonis 17 Pikale tabelile eelnev murdumine .....	37
Joonis 18 MoveHeadingsToTable meetod .....	39

# 1 Sissejuhatus

Käesoleva lõputöö eesmärk on luua Eesti Standardikeskusele (EVS) lahendus standardite üle võtmiseks XML baasil, mis vahetaks välja PDF-PDF teisendusel põhineva standardite koostamise. Uue lahenduse välja töötamine oli aktuaalne, sest ISO, CEN ja CENELEC hakkasid standardeid välja andma lisaks Word ja PDF formaatidele ka XML formaadis. PDF töötamise välja vahetamine oli vajalik, sest PDF formaadis standarditele EVS lehekülgede lisamine tekitas probleeme lehekülje numbrite, peegelveriste, linkide, järjehoidjate, päiste ja muude sarnaste elementide korrektsel viisil kuvamisega.

## 1.1 Standardite üle võtmine

„Tulenevalt rahvusvaheliste ja Euroopa standardimisorganisatsioonide liikmesusest on Eesti Standardikeskusel õigus nende organisatsioonide standardeid (edaspidi rahvusvahelised ja Euroopa standardid) üle võtta Eesti standarditeks“ [12, p. 5].

Eesti Standardikeskus võtab ühe tegevusena üle standardeid Euroopa (CEN ja CENELEC) ja rahvusvahelistelt (ISO ja IEC) standardimisorganisatsioonidelt. EVS andmebaasirakendus (StaDIS) impordib rahvusvahelised ja Euroopa standardid PDF kujul ja seob need StaDIS-e standardiprojektidega. Standardi üle võtmiseks on vajalik standardi tekstist kokku panna dokument, millele on lisatud Eesti standardile omaseid meta-andmeid. Standardimiskoordinaatorid lisavad projektidele vajalikult EVS meta-andmed ja kui projekt on vastuvõtmise etapis, siis kasutatakse Eesti standardi dokumendi genereerimiseks EVS PDF-PDF teisendajat.

EVS PDF-PDF teisendaja on Eesti Standardikeskusele standardite üle võtmise automatiseerimiseks välja töötatud ja StaDIS-esse integreeritud rakendus. Kui standardi projekt on jõudnud vastuvõtmise etappi, siis annab StaDIS PDF-PDF teisendajale sisendiks projektiga seotud PDF faili(d) ja standardi projekti meta-andmed. Kui faile on mitu, siis paneb teisendaja need omavahel kokku. Seejärel lisab teisendaja üksikule või kokkupandud failile EVS tiitellehe, EVS eessõna, tühjad leheküljed (korrektseks printimiseks) ja EVS tagakaane. Standardi sisu puhul asendab teisendaja iga lehekülje



ülaosas olevad originaaltähised EVS tähisega. Teisendaja genereerib lõplikult PDF formaadis Eesti standardist ka PDF formaadis eelvaate, mida kasutatakse klientidele standardi sisu tutvustamiseks enne standardi ostmist. Eesti standardid ja nende eelvaated on PDF formaadis saadaval EVS veebipoes.

Eesti Standardikeskus kasutab teisendajat Euroopa standardite (CEN ja CENELEC) ja rahvusvaheliste alusdokumentidega Euroopa standardite (CEN + ISO ja CENELEC + IEC) üle võtmiseks. Puhtalt rahvusvaheliste standardite üle võtmine vormistatakse üldjuhul algupärase standardina ja pannakse seeläbi kokku käsitsi. Samuti on standardilaadseid dokumente ja muid erandeid, mille üle võtmisel teisendajat ei kasutata. Teisendajat kasutatakse ka vaid Eesti standardiks ülevõtul mitte-eestikeelsena. See tähendab, et standard sisaldab Euroopa ja rahvusvaheliste standardite tekste, kuid standardile lisatakse täiendavaid lehekülgi ja Eesti standardile omast informatsiooni.

Kui PDF-PDF teisendaja on Eesti standardi dokumendi genereerinud, siis läbib see ülevaatuse ja toimetamise. Teade standardi kinnitamise kohta avaldatakse Eesti Standardikeskuse ametlikus väljaandes EVS Teataja [12, p. 9]. Eesti standard tehakse seejärel EVS veebipoes kättesaadavaks ja standard on seeläbi Eesti standardiks üle võetud.

## **1.2 Probleem PDF-PDF teisendajaga**

PDF-PDF teisendaja arendati välja ajal, mil Euroopa ja rahvusvahelised standardimisorganisatsioonid standardeid XML formaadis ei avaldanud. Saadaval olid peamiselt vaid PDF ja Word formaadis standardid. Sellest tulenevalt arendati PDF töötlust teostav PDF-PDF teisendaja, mis oli küll limiteeritud funktsionaalsusega, kuid võimaldas standardite üle võtmise protsessi automatiseerida.

PDF-PDF teisendamisel toimub PDF dokumendi automaatne töötlus. Teisendaja lisab dokumendile lehekülgi ja asendab standardi originaaltähised EVS tähisega, mille käigus toimuvad standardi ülesehituses mõningad ebavajalikud muutused. Näiteks nihkuvad lehekülje numbrid ja peegelveerised valeks, lingid lakkavad töötamast, järjehoidjad kaovad ära ning päises olevad tähised positsioneeritakse vahel valesti. Enne standardi avaldamist kontrollitakse ja parandatakse vaid peegelveeriseid ning tähiste korrektsust. Lehekülgede nummerdamise, linkide ja järjehoidjate parandamine on liialt ajamahukas,

mistõttu neid probleeme käsitsi ei parandata. PDF-PDF teisendaja väljastab ka standardi eelvaate PDF formaadis. Standardi eelvaates kuvatakse tihti rohkem informatsiooni kui vaja, sest teisendaja ei ole alati võimeline tuvastama õiget kohta, kust standardi eelvaateväline sisu algab.

Eesti Standardikeskuses kasutusel olev PDF-PDF teisendaja ei ole praeguste võimaluste puhul parim lahendus Eesti standardite kokku panemiseks. Teisendaja ei suuda käsitsi parandamise abita piisavalt kvaliteetseid Eesti standardeid koostada. Samuti ei ole piisavalt lihtsalt võimalik PDF kujul standardi sisu kombineerida EVS veebilehel oleva otsingumootoriga, sest genereeritud PDF failil pole piisavalt organiseeritust. Lisaks sellele on vaid PDF faili olemasolust tingitult EVS kommenteerimisportaalil kasutuses veebilehel mitte-eelistatud Adobe Acrobat Readeril põhinev tööriist, mida kasutatakse PDF failide lugemiseks ja kommenteerimiseks.

Vajadus genereeritud Eesti standardeid käsitsi parandada ja klientide rahulolematus seoses mittetöötavate linkide ning järjehoidjatega on loonud olukorra, kus olemasolevat lahendust oleks vaja parandada või asendada. Seoses sellega, et standardeid on hakatud avaldama ka XML kujul, on loodud võimalusterohked eeldused uue lahenduse välja töötamiseks.

### **1.3 Töö eesmärk**

Lõputöö eesmärk on välja töötada XML-põhine teisendaja, mis vahetaks välja Eesti standardite üle võtmisel kasutatava EVS PDF-PDF teisendaja. XML-põhine teisendaja peab esialgu suutma genereerida Eesti standardeid CEN, CENELEC ja ISO standarditest. Kui IEC hakkab samuti standardeid XML formaadis avaldama, siis on vajalik ka IEC standardites Eesti standardite genereerimine.

PDF formaadis Eesti standardi puhul peab XML-põhine teisendaja suutma koostada standardit sama efektiivselt kui PDF-PDF teisendaja, kuid parandama kõik sellega seotud vead. Lehekülgede numbrite, peegelveeriste, linkide ja järjehoidjate vigase teisendamise parandamine on vajalik, et vähendada genereeritud Eesti standardi käsitsi parandamise vajadust ja tõsta klientide rahulolu.

Arendatava teisendaja XML-põhisus võimaldab Eesti standardeid genereerida ka teistes formaatides, mistõttu peab teisendaja genereerima Eesti standardi ka XML ja HTML

formaadis. XML formaadis Eesti standard eeldab, et lisaks XML failidele on kaasas ka EVS meta-andmed. HTML formaadis Eesti standard peab olema tulevikus kasutatav EVS veebilehel. XML ja HTML formaadis Eesti standardi genereerimine lisab Eesti Standardikeskusele formaate, milles standardit on võimalik avaldada.

## **1.4 Töö struktuur**

Töö on jaotatud viite peatükki: Sissejuhatus, Metoodika, Realisatsioon, Analüüs, Kokkuvõte. Peatükid kirjeldavad XML failidest Eesti standardi koostamiseks välja töötatud lahenduse arendamise protsessi ja moodustavad ühtse terviku.

Metoodika peatükk kirjeldab lahenduse nõudeid, tööriistade valikut (koos põhjendustega) ja eeldatavat arendusprotsessi. Spetsiifilised nõuded töötas välja autor ise, valis välja sobilikud tööriistad ja pani paika üldise arendusprotsessi.

Realisatsiooni peatükk kirjeldab valminud lahenduse tehnilisi andmeid ja arhitektuuri, lahenduse standardi koostamise töökulgu ning tähtsamaid probleeme koos lahendustega. Peatükki võib käsitleda kui rakenduse dokumentatsiooni.

Analüüsi peatükk kirjeldab valminud lahenduse vastavust nõuetele, kirjeldab uusi mõtte- ja murekohti ning annab ettekujutuse rakenduse edasisest suunast. Peatükk lõpeb autoripoolse hinnanguga tööle.

## 2 Metoodika

Töö realiseerimiseks kasutatavate tööriistade ja arendusmetoodika valimiseks andis Eesti Standardikeskus autorile vabad käed. Uus lahendus peab parandama PDF-PDF teisendaja vigu ja võimaldama Eesti standardeid genereerida lisaks PDF formaadile ka XML ja HTML formaatides. Autor peab analüüsima PDF-PDF teisendaja vigu, panema paika täpsed nõuded, leidma sobilikud tööriistad ja töötama välja rakenduse, mille abil saaks Euroopa ja rahvusvahelistest XML failidest stabiilselt nõuetele vastavaid XML, HTML ja PDF formaadis Eesti standardeid genereerida.

### 2.1 Nõuded

Uus lahendus peab suutma ISO, CEN ja CENELEC poolt välja antud XML formaadis standarditest ja EVS meta-andmetest kokku panna XML, HTML ja PDF formaadis EVS standardeid. IEC ei avalda lahenduse välja töötamise hetkel standardeid XML formaadis ja jääb seetõttu tähelepanuta. Lahendus peab suutma sarnaselt PDF-PDF teisendajale genereerida Eesti standardi põhjal ka PDF formaadis standardi eelvaate. Uue lahenduse arendamise eesmärgiks on parandada hetkel kasutuses oleva lahenduse vigu ja pakkuda EVS standardeid rohkemates formaatides HTML ja EVS XML kujul.

Eesti standardil on vastavalt alusfailide algatajatele ja tüüpidele kindel struktuur. Kokkuvõtvalt on Eesti standardi struktuur on järgnev:

- 1) Eesti standardi tiitelleht
- 2) Eesti standardi eessõna
- 3) Euroopa tiitelleht
- 4) Euroopa eessõna
- 5) Standardi põhitekst

- 6) Muudatus(ed)<sup>1</sup>, mis kordab/kordavad punktide 3-5 struktuuri
- 7) Tühjad leheküljed (0-3)
- 8) Eesti standardi tagakaas

Standardi põhitekst varieerub vastavalt sellele, kas tegemist on puhta Euroopa (CEN/CENELEC) või rahvusvaheliste alusdokumentidega Euroopa (CEN ISO) standardiga. CEN ISO puhul kasutatakse põhitekstina rahvusvahelise standardi põhiteksti, kuid puhta CEN või CENELEC puhul Euroopa standardi põhiteksti. Standardil võib olla ka muudatusi, mille struktuur on sama, mis eelneva struktuuri kirjelduse punktides 3-5. Standard võib olla ka muudatus ise.

Täpsed EVS standardi vormistuslikud nõuded paneb paika EVS Juhend 4 [13]. Samuti on Eesti standardi genereerimise käigus vajalik kokku panna XML fail, mis sisaldaks EVS meta-andmeid, et kaugemas tulevikus võimaldada standardite publitseerimine XML kujul. XML-PDF genereerimise vaheetapina on vajalik genereerida Eesti standard ka HTML kujul, et EVS veebilehel standardi eelvaate kuvamist kvaliteetsemaks ja paindlikumaks muuta.

EVS lehekülgede lisamisel nihkuvad leheküljenumbriid paigast ära. Joonisel 1 on kuvatud Euroopa tiitellehe lõpp ja sisukorra esimene leht, mille järjekorranumbriks on „2“. Samas on dokumendile lisatud EVS tiitelleht ja eessõna (joonisel pole näha), nii et järjekorranumber peaks olema „4“. Arendatav rakendus peab lugema ja nummerdama lehti korrektselt alates esimesest leheküljest.

---

<sup>1</sup> Standardi muudatus - tavalise standardi struktuuriga dokument, mis kirjeldab mõne teise standardi vigu ja/või muutusi

EVS-EN 71-3:2013+A3:2018

Contents	Page
European foreword.....	5
Introduction .....	7
1 Scope .....	7
2 Normative references .....	7
3 Terms and definitions.....	8
4 Requirements .....	9
4.1 Toy material categories (see H.4).....	9
4.2 Specific requirements.....	9
5 Principle.....	10
6 Reagents and apparatus.....	11
6.1 Reagents.....	11
6.2 Apparatus.....	11
7 Sampling and sample preparation .....	12
7.1 Selection of test portions .....	12
7.2 Standards preparation .....	12
7.3 Sample preparation .....	12
7.3.1 General.....	12
7.3.2 Category I: Dry, brittle, powder like or pliable and Category II: Liquid or sticky .....	12
7.3.3 Category III: Scraped-off .....	13
7.4 Migration procedure .....	14
7.4.1 General.....	14
7.4.2 Category I: Dry, brittle, powder like or pliable and Category II: Liquid or sticky .....	14
7.4.3 Category III: Scraped-off .....	15
8 Methods of analysis.....	17
9 Calculation of results.....	17
9.1 General elements.....	17
9.1.1 Calibration curve .....	17
9.1.2 Calculation of migration.....	18
9.1.3 Interpretation of results .....	18
9.2 Chromium (III) and Chromium (VI).....	18
9.2.1 Calibration curve .....	18
9.2.2 Calculation of migration.....	18
9.2.3 Interpretation of results .....	19

2

Joonis 1. Dokumendi neljas lehekül on valesti nummerdatud

EVS lehekülgede lisamisel muutuvad peegelveerised ebakorrapärasteks. Joonisel 2 on kujutatud standardi Euroopa tiitellehe ette lisatud EVS eessõna. Praegune lahendus ei oska tuvastada, millal peab olema lehekülg joondatud paremale ja millal vasakule. Arendatav lahendus peab oskama seda tuvastada ja õigesti realiseerida. Samuti peab rakendus peegelveeriseid üleüldiselt ühtlustama, sest nende rahvusvaheliste standardite äärekaugus erineb Eesti standardite äärekaugusest.

Tagasisidet standardi sisu kohta on võimalik edastada, kasutades EVS-i veebilehel asuvat tagasiside vormi või saates e-kirja meiliaadressile [standardiosakond@evs.ee](mailto:standardiosakond@evs.ee).

ICS 97.200.50

**Standardite reprodutseerimise ja levitamise õigus kuulub Eesti Standardikeskusele**  
Andmete paljundamine, taastekitamine, kopeerimine, salvestamine elektroonsesse süsteemi või edastamine ükskõik millises vormis või millisel teel ilma Eesti Standardikeskuse kirjaliku loata on keelatud.  
Kui Teil on küsimusi standardite autorikaitse kohta, võtke palun ühendust Eesti Standardikeskusega:  
Koduleht [www.evs.ee](http://www.evs.ee); telefon 605 5050; e-post [info@evs.ee](mailto:info@evs.ee)

**The right to reproduce and distribute standards belongs to the Estonian Centre for Standardisation**  
No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, without a written permission from the Estonian Centre for Standardisation.  
If you have any questions about copyright, please contact Estonian Centre for Standardisation:  
Homepage [www.evs.ee](http://www.evs.ee); phone +372 605 5050; e-mail [info@evs.ee](mailto:info@evs.ee)

---

EUROPEAN STANDARD **EN 71-3:2013+A3**  
NORME EUROPÉENNE  
EUROPÄISCHE NORM June 2018

---

ICS 97.200.50 Supersedes EN 71-3:2013+A2:2017

English Version

**Safety of toys - Part 3: Migration of certain elements**

Sécurité des jouets - Partie 3: Migration de certains éléments Sicherheit von Spielzeug - Teil 3: Migration bestimmter Elemente

This European Standard was approved by CEN on 29 Mai 2013 and includes Amendment 1 approved by CEN on 18 August 2014, Amendment 2 approved by CEN on 18 April 2017 and Amendment 3 approved by CEN on 15 March 2018.

CEN members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration. Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the CEN-CENELEC Management Centre or to any CEN member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CEN member into its own language and notified to the CEN-CENELEC Management Centre has the same status as the official versions.

Joonis 2. Kahel järjestikusel lehel on parempoolsed veerised

EVS tähiste lisamisel lakkavad lehekülgedele lingid töötamast. Praegune lahendus käib iga PDF lehekülje elemente läbi, kuni leitakse (tavaliselt) lehe ülemises osas olev tähise element ja muudetakse seda vastavalt EVS tähisele. Selle protsessi käigus muutuvad kõik lingid tavatekstideks. Arendatav lahendus peab XML struktuuris olevad dokumendisisesed lingid realiseerima töötavatena.

EVS tähiste lisamisel kaob PDF järjehoidja, mis pannakse kokku pealkirjade põhjal. Arendatav lahendus peab suutma genereerida pealkirjadele vastava järjehoidja.

Vahel esineb olukordi, kui praegune lahendus ei leia EVS tähise lisamise kohta või lisab EVS tähise dokumendi keskele, mistõttu peab vead parandama käsitsi PDF tööriistadega. Käsitsi dokumendi parandamist oleks tarvilik võimalikult palju ära hoida. Arendatav lahendus peab lisama dokumendi lehekülgede päistesse EVS tähiseid.

Lisaks nimekirjas olevatele nõuetele peab uus lahendus olema pidevalt orienteeritud EVS Juhend 4 vormistusnõuetest ja jätma töökorda ka praeguse lahenduse juures probleemivabad aspektid (näiteks teksti suurus, tabelite murdumine, taanded, piltide ja tabelite suurused jpm). Peamine suunis on EVS Juhend 4-le ja standardi algele PDF formaadi stiilinõuetele vastamine.

Lisaks praeguse lahenduse parandamisele on vajalik ka standardi HTML formaati teisendamine, et see oleks veebilehel kasutamiseks paindlik. Genereerida tuleb ka EVS XML failid, mis kujutavad endast sissetulevaid XML faile koos EVS meta-andmetega, et Eesti Standardikeskusel oleks tulevikus võimalus avaldada standardeid ka XML formaadis.

## 2.2 Tööriistad

Autor leidis tööriistad, millega on optimaalne nõuetele vastavat lahendust realiseerida. Peamine tööriist on Prince, mille ümber autor lahenduse struktureeris ja ka realiseeris. Ülejäänud tööriistad ja keeled on valitud vastavalt Prince vajadustest. Töös kasutatavaid tööriistu ja eraldiseisvaid komponente ühendab .NET raamistikus realiseeritud rakendus.

**Prince** on HTML, XHTML ja XML formaadis dokumentide PDF formaati teisendamise tööriist. Prince suudab toime tulla kõige uuema CSS funktsionaalsusega ja suudab teisendamisel realiseerida päiseid ja jaluseid, leheküljenumbreid, mitmeleheküljelisi



tabeleid, nimekirju, allmärkuseid, viiteid ja muid spetsiifilisi nõudeid. Prince on .NET raamistikuga lihtsasti integreeritav. [13]

Autor valis Prince'i seoses sellega, et tasuta lahendused (Wkhtmltopdf, HiQPdf, EvoPdf jpt) ei pakkunud autori testimisel piisavat funktsionaalsust ja soovitud kvaliteeti. Prince on tasuta, kuid suutis oma hinda õigustada laiaulatusliku ja dokumendispetiifilise funktsionaalsusega. Samuti oli Eesti Standardikeskuse nõuetest lähtuvalt standardi dokumentide avaldamise puhul kvaliteet ja kliendi rahulolu olulisem, kui dokumentide publitseerimiseks kuluv hind.

**.NET raamistik** on keskkond, mis võimaldab rakenduste arendamist ühendades omavahel Common Language Runtime'i<sup>1</sup> ja .NET Class Library<sup>2</sup>. Kõige üldkasutatavam programmeerimiskeskkond .NET rakenduste arendamiseks on **MS Visual Studio**, mida ka autor töös kasutab. [14]

.NET raamistik ja MS Visual Studio on valitud põhjusel, et autor on nendega varasemalt tuttav. Samuti on Eesti Standardikeskuse rakendused valdavas osas .NET veebirakendused.

Olulisemad ebatraditsioonilised rakenduses kasutatavad teegid on:

1. HTML Agility Pack – HTML struktuuri vormindamiseks,
2. Saxon – XSL abil XML-HTML teisendamiseks,
3. PdfSharp – PDF lehekülgede kokku lugemiseks.

**HTML** (*HyperText Markup Language*) on Veebi<sup>3</sup> põhiline märgistuskeel veebilehtede sisu loomiseks. [3]

---

<sup>1</sup> Common Language Runtime – virtuaalmasin, mis tõlgib valitud programmeerimiskeele käsklused masinale loetavaks ja võimaldab seeläbi rakendust käivitada

<sup>2</sup> .NET Framework Class Library – tarkvara komponentide (klasside) kogum, mida arendajad saavad enda rakenduste arendamisel kasutada

<sup>3</sup> Veeb – Internetis toimiv ja hüpertekstil põhinev elektrooniliste dokumentide süsteem [16]

Autor realiseeris XML-PDF teisendamise HTML vaheetapiga XML-HTML-PDF rakendusena põhjusel, et Eesti Standardikeskusel on plaanis standardeid HTML kujul oma veebilehel kuvada. Töö kirjutamise hetkel toimub standardite kuvamine veebilehel PDF kujul. HTML vaheetapp võimaldab lisaks veebilehe-sõbralikule versioonile ka ulatuslikku kujunduslihtsust läbi stiililehtede.

**CSS** (*Cascading Style Sheets*) on veebilehtedele stiili (näiteks kirja, värvide, vahekauguste) lisamiseks kasutatav märgistuskeel. [5]

CSS kasutamine hoiab HTML koodi puhtana ja võimaldab kujundada HTML dokumendi oodatud PDF dokumendi sarnaseks, lihtsustades seeläbi HTML-PDF teisendust. Samuti töötab Prince funktsionaalsus suuresti CSS põhjal.

**JavaScript** on peamiselt veebilehtede skriptimiseks kasutatud programmeerimiskeel. JavaScripti ja selle funktsioonide abil on võimalik vastavalt sündmuse esinemisele veebilehe disaini muuta. [15]

Autor kasutab rakenduses JavaScripti paari üksiku Prince funktsionaalsuse puudumise tõttu. JavaScripti võib ka hiljem rakenduse probleemide lahendamisel vaja minna, kui peaks tekkima mõni eriti ebatraditsiooniline ja Prince funktsionaalsusele mitte alluv probleem. JavaScript võimaldab kontrollida ka näiteks tabelite mahtuvust lehekülje piiridesse, sest tabeli mõõtmed selguvad alles pärast HTML lehe laadimist.

**STS4i tööriistad** on NISO<sup>1</sup> poolt välja töötatud Standards Tag Suite<sup>2</sup> struktuurireeglitele vastavad abivahendid, mis aitavad standardimisorganisatsioonidel teisendada XML formaadis väljastatud standardeid algelisse HTML formaati. [10]

STS4i tööriistade kasutamine lihtsustab XML-HTML teisendamist oluliselt, sest selle puudumisel oleks vaja see mahukas teisendusreeglite kogum niikuinii välja töötada. STS4i ei ole siiski perfektne, sest mõningad erandlikumad teisendusreeglid on jäänud realiseerimata ja mõned üksikud reeglid on veidi valesti realiseeritud või puudulikud.

---

<sup>1</sup> NISO – National Information Standards Organization

<sup>2</sup> Standards Tag Suite – NISO poolt loodud reeglistik, mis kirjeldab XML formaadis oleva standardi struktuurinõudeid [18]

## 2.3 Töö protsess

Autor otsustas arendada .NET lahenduse, mis kasutab Prince HTML-PDF teisendajat, STS4i tööriistaid, HTML malle ning Prince-spetsiifilist CSS ja JavaScript faili. Tulemus on eraldiseisev parameetritega välja kutsutav rakendus, mis on integreeritav Eesti Standardikeskuse standardihalduse veebirakendusse StaDIS.

Autor parandab STS4i tööriistade teisendusreegleid. Muudatused peaksid jääma minimaalseks, et lihtsustada STS4i tööriistade uuenduse korral kasutuses olevate tööriistade välja vahetamist. Samuti püsivad STS4i tööriistad seeläbi rakenduses võimalikult eraldiseisvana.

Autor töötab välja HTML mallid, mis oleksid rakendusele kasutatavad, et lisada standardile EVS lehekülgi. Kui rakendus teostab STS4i tööriistade abil XML-HTML teisenduse, siis kombineeritakse saadud HTML osast ja mallidest kokku ühine HTML fail. Mallide kasutamine lihtsustab dokumentidele korduvate lehtede lisamist.

Autor töötab välja CSS stiililehe, mis vastaks EVS Juhend 4 vormistusreeglitele ja Prince teisendusreeglitele. Autor lähtub standardi põhiosa puhul ISO, CEN ja CENELEC PDF formaadis standardite visuaalsest vaatlusest, et tagada võimalikult sarnane standardi põhiosa stiil.

Kasutades HTML Agility Pack teeki, peab rakendus töötleva kokkupandud HTML faili, et parandada selle struktuuri STS4i teisendusprotsessi järgselt. Töötlus ei ole seotud STS4i tööriistadega ja rakenduse osad on seeläbi paremini eraldatud. Töötluse lõpul väljastatakse HTML fail, millest on vaja genereerida ka HTML eelvaade.

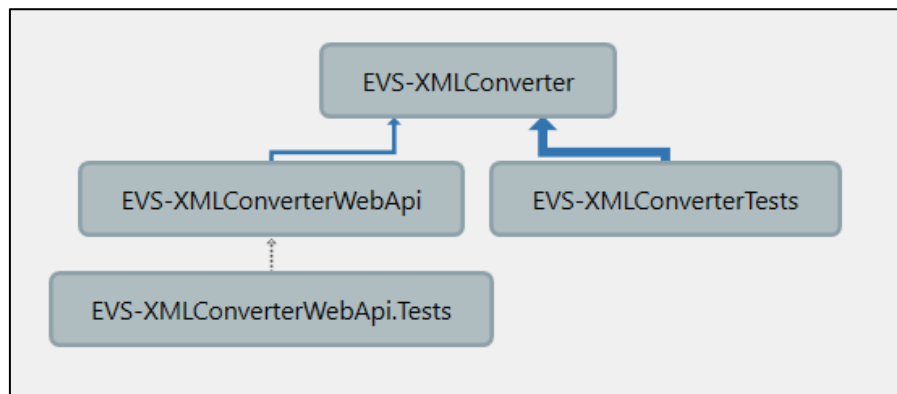
Prince kutsutakse välja .NET rakendusest ja sellel on konfigureeritav asukoht. Sisenditeks on töödeldud HTML failid, CSS stiilileht ja JavaScript skript. Prince peab väljastama standardi ja selle eelvaate PDF dokumendi. Standardi PDF failile peab lisama tühjad leheküljed ja eelvaate PDF failile vesimärgid.

Autor hindab, et lahenduse realiseerimise kõige mahukam osa on HTML töötlus ja CSS stiililehe arendamine, sest nende etappide puhul toimub pidevalt visuaalne võrdlus genereeritava standardi ja originaalselt välja lastud PDF formaadis standardi vahel. Kõik nõuetes kirjeldatud probleemid on lahendatavad ja rakendus on realiseeritav.

## 3 Realisatsioon

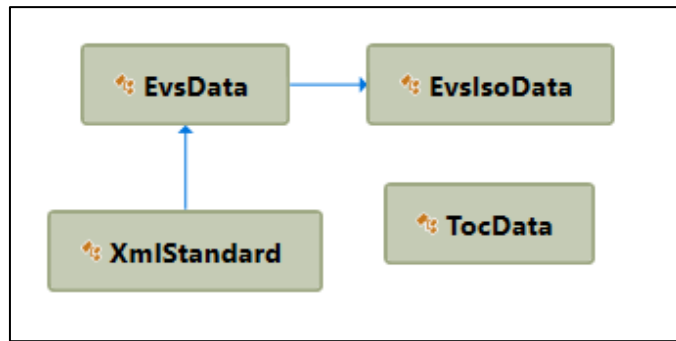
### 3.1 Arhitektuur ja tehnilised andmed

Autor realiseeris lahenduse .NET raamistikus. Lahendus moodustavad rakendus ja rakendusliides (API). Rakendusliidese eesmärk on pakkuda EVS andmebaasirakendusele (StaDIS) võimalust kutsuda rakendust välja ja hoida lahendus StaDIS-est eraldatuna. Joonis 3 kirjeldab lahenduse üldist arhitektuuri ehk projekte, millest lahendus koosneb ja nende omavahelisi suhteid.



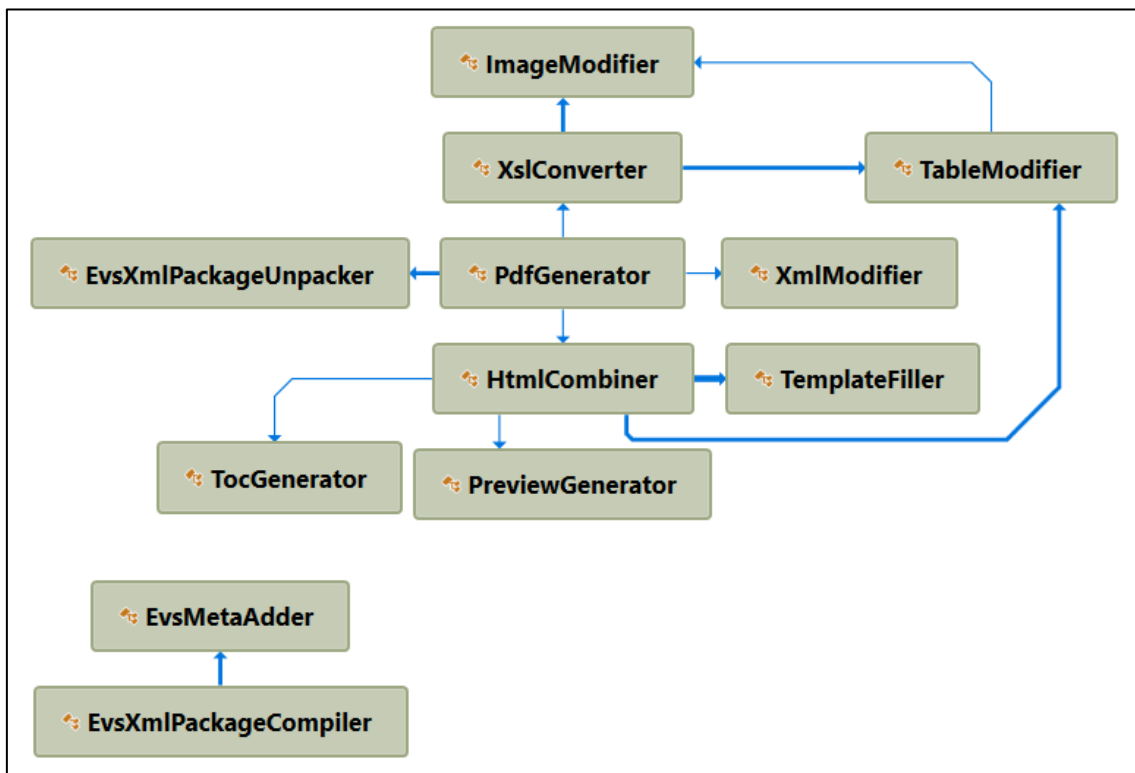
Joonis 3 Lahenduse arhitektuur

Lahendus teostab väga palju XML ja HTML töötlust. XML ja HTML on olemuselt juba struktureeritud, mistõttu ei pea rakendus olukorra haldamiseks palju uusi objekte looma. Rakenduses on vaid neli põhilist objekti: EvsData (EVS meta-andmed), EvsIsoData (lisaandmed ISO standardi üle võtmiseks), TocData (sisukorra koostamiseks) ja XmlStandard (XML failide haldamiseks). Joonis 4 kirjeldab lahenduse objekte ja nendevahelisi seoseid.



Joonis 4 Lahenduse objektid

Rakenduses on 27 klassi ja 156 meetodit (130 testitud). Paljud klassid kasutavad ühiseid abimeetodeid ja objekte, et kood oleks võimalikult puhas. Joonis 5 kirjeldab rakenduse loogikakihi meetodeid, mis tegelevad EVS XML paketi koostamise ja sellest HTML ja PDF failide genereerimisega.



Joonis 5 Rakenduse loogika

Rakenduses kasutatavate failide nimed ja asukohad on rakenduse põhikoodist eraldatud ja realiseeritud konfigureeritavalt App.config failis. See tagab failide nimede ja asukohtade mugava muutmise ning lihtsustab seeläbi integratsiooni StaDIS-ega. Joonis 6 annab ülevaate konfigureeritavate asukohtade struktuurist.

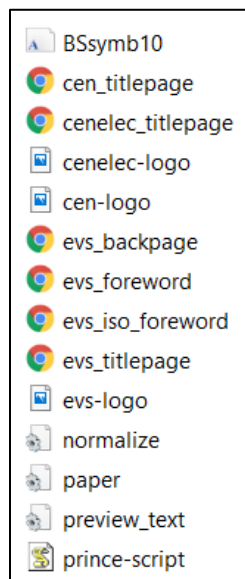
```

<configuration>
  <appSettings>
    <add key="Prince" value="C:\Program Files (x86)\Prince\engine\bin\prince.exe" />
    <add key="StsTools" value="C:\Users\Rain\Documents\Source\Repos\EVS-XMLConverter\EVS-XMLC
    <add key="NisoStsXsl" value="nisoSts2html\nisoSts2html.xsl" />
    <add key="IsoStsDtd" value="schema\isosts\dtd\ISOSTS.dtd" />
    <add key="Templates" value="C:\Users\Rain\Documents\Source\Repos\EVS-XMLConverter\EVS-XMLC
    <add key="Paper" value="paper.css" />
    <add key="Normalize" value="normalize.css" />
    <add key="PreviewText" value="preview_text.css" />
    <add key="PrinceScript" value="prince-script.js" />
    <add key="EvsForeword" value="evs_foreword.html" />
    <add key="EvsIsoForeword" value="evs_iso_foreword.html" />
    <add key="EvsTitlepage" value="evs_titlepage.html" />
    <add key="CenTitlepage" value="cen_titlepage.html" />
    <add key="CenelecTitlepage" value="cenelec_titlepage.html" />
    <add key="EvsBackcover" value="evs_backpage.html" />
    <add key="EvsXmlPackages" value="C:\Users\Rain\Documents\Source\Repos\EVS-XMLConverter\EVS
    <add key="HtmlPreview" value="_preview" />
    <add key="PdfPreview" value="_preview" />
    <add key="PreEvsPds" value="_evsPdfDelete" />
    <add key="Temporary" value="C:\Users\Rain\Documents\Source\Repos\EVS-XMLConverter\EVS-XMLC
    <add key="Html" value="\Html" />
  </appSettings>

```

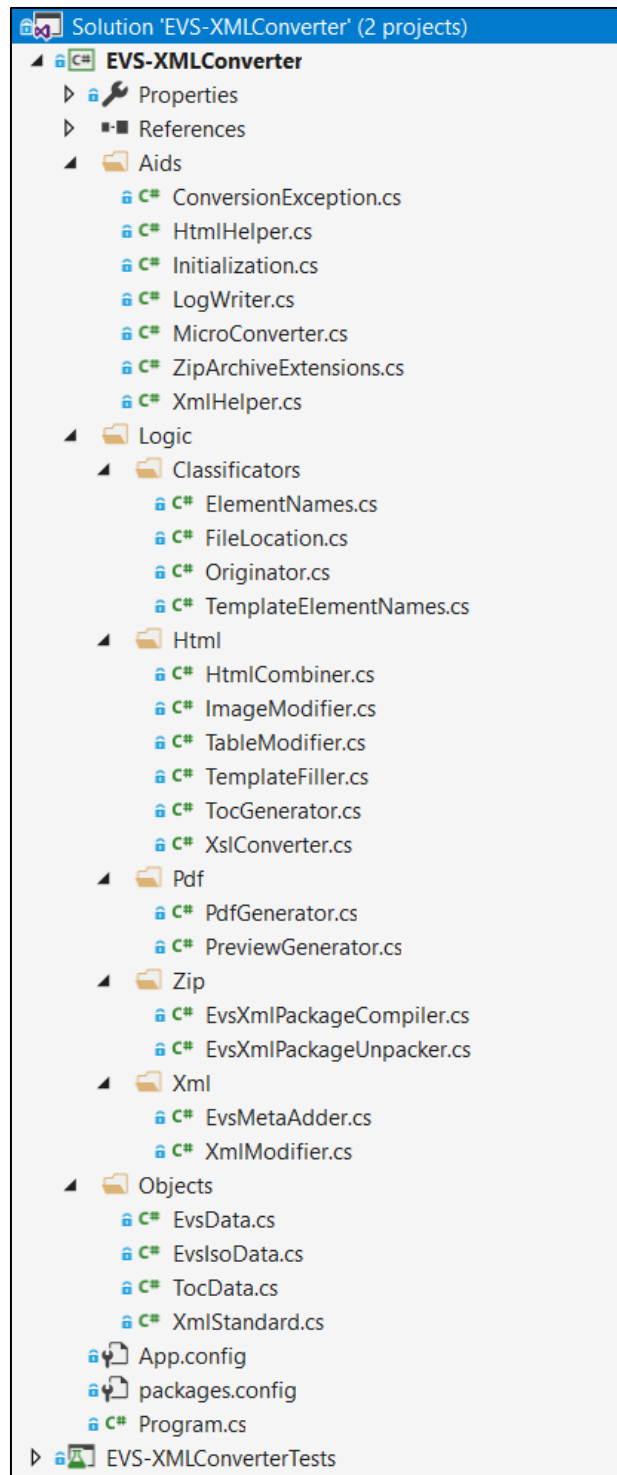
Joonis 6 Failide konfigureeritavad asukohad

Autor parandas STS4i tööriistade vigu, et XML-HTML teisendus toimuks korrektselt. Autor töötas välja EVS tiitellehe, EVS eessõna (Euroopa ja rahvusvaheline) ja EVS tagakaane HTML mallid. Autor töötas välja Prince reeglitele vastava CSS stiililehe, mida kasutavad nii mallid kui ka XML-HTML teisendusest tulevad HTML failid. Joonis 7 annab ülevaate failidest, millest rakendus sõltub.



Joonis 7. Rakenduse tööks vajalikud failid

Rakenduse kood on jaotatud eraldi kaustadesse, et Eesti standardi genereerimise erinevaid etappe paremini hallata. Samuti on see realiseeritud autori oskustele vastavatele puhta koodi tavadele, et tagada võimalikult mugav lahendusest arusaamine, vigade parandamine ja lahenduse edasiarendamine. Joonis 8 kirjeldab rakenduse struktuuri.



Joonis 8. Rakenduse struktuur

**Aids** kaust sisaldab endas abimeetodeid, mis võivad olla seotud rakenduse erinevate osadega. Selles on klassid, mis tegelevad veateadete, logi ja väiksemate teisendustega. Samuti on seal ka XML, HTML ja ZIP abimeetodid.

**Classifiers** kaust sisaldab rakenduse põhiosale kasutatavaid NISO STS XML elementide, HTML mallide elementide ja standardimisorganisatsioonide nimesid, et need oleks vajadusel ühest kohast muudetavad.

**Html** kaust sisaldab endas rakenduse kõige kogukamat kihti. Selles toimub HTML failide muutmise lõpliku PDF faili sarnaseks, et Prince saaks need edukalt PDF failideks teisendada. Enamik meetoditest oleks saanud realiseerida ka STS4i tööriistade parandamisega, kuid autor otsustas STS4i tööriistad ja EVS spetsiifilised muutused eraldi hoida.

**Pdf** kaust sisaldab endas valmis HTML failide PDF formaati teisendamise klasse. Need on seotud Prince'iga ja genereerivad Eesti standardi PDF dokumendi ja selle eelvaate.

**Zip** kaust sisaldab endas EVS XML paketi kokkupanemise ja ka selle lahtipakkimise funktsionaalsust, et alustada XML-HTML teisendust.

**Xml** kaust sisaldab XML failide struktuuri ja vigu parandavat funktsionaalsust, mis on oluline, et võtta XML failidest vaid Eesti standardi genereerimiseks oluline informatsioon. Lisaks on kaustas ka klass, mis paneb sisenditest tulevad EVS meta-andmed XML formaati, et see EVS XML paketti lisada.

**Objects** kaust sisaldab endas rakenduse töö lihtsustamiseks koostatud objekte nagu EvsData (EVS meta-andmed Euroopa standardi üle võtmiseks), EvsIsoData (EVS meta-andmed ISO standardi üle võtmiseks), TocData (sisukorra koostamiseks) ja XmlStandard (XML failiga seotud informatsioon).

Enamik rakenduse meetodeid on paljude võimalike stsenaariumite korral testitud (*TestCase*'idega). Euroopa ja rahvusvaheliste standardite struktuuris on palju variatsiooni, mis teeb meetodite põhjaliku testimise raskeks. Kõige kindlam lõpptulemuse testimine peab toimuma läbi visuaalse võrdluse.



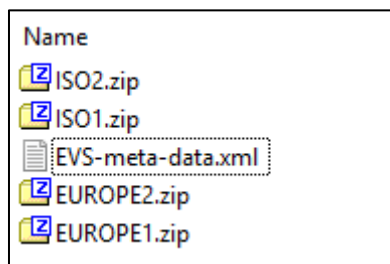
### 3.2 XML-HTML-PDF teisendaja

Lahenduse põhiline komponent XML-HTML-PDF teisendaja. See on .NET rakendus, mille tööks on vajalikud HTML mallid, Prince'i reeglitele vastav CSS stiilileht ja JavaScript skript. Rakendus on välja kutsutav API liidese abil, mida hakkab kasutama StaDIS. Rakendus töötab koostöös STS4i tööriistadega, mis on välja töötatud riiklikele standardimisorganisatsioonidele XML-põhise standardite üle võtmise protsessi lihtsustamiseks. STS4i sisaldab XSL transformaatorit, mis suudab NISO STS-ile vastavaid XML faile teisendada algelisse HTML formaati.

Rakenduse esimene tähtis funktsionaalsus on EVS XML paketi koostamine, mis väljastab Eesti standardi XML kujul. Rakendus loeb sisse XML paketid (XML failid ja pildid) ja StaDIS-est tulevad standardi meta-andmed ning paneb nende põhjal kokku EVS XML paketi. Pakett sisaldab XML pakette ja EVS meta-andmetest kokku pandud NISO STS struktuurile vastavat XML faili. Paketi kokku panemiseks saadetakse API kaudu rakendusele:

1. ZipFileDirectories – XML kujul standardite pakettide asukohad serverikettal;
2. FolderName – standardiga seotud failid kausta nimetus, kuhu genereeritud failid lõpus salvestatakse;
3. FileName – standardi faili lõplik nimetus, mille genereerimist teostab StaDIS;
4. EvsTitle – standardi eestikeelne pealkiri;
5. OriginalTitle – standardi originaalne pealkiri;
6. EvsReference – standardi EVS tähis;
7. IcsGroups – standardiga seotud ICS grupid;
8. BulletinIssue – standardiga seotud EVS Teataja avaldamise kuupäev;
9. Scope – standardiga seotud üldine sissejuhatav tekst;
10. Amendments – standardi muudatustega seotud sissejuhatav tekst;
11. Committee – standardi koostamise algatanud komitee.

BulletinIssue, Scope, Amendments ja Committee saadetakse vaid puhta ISO standardi puhul. Ülejäänud informatsioon on dokumendi korrektseks genereerimiseks vajalik. Rakendus salvestab EVS XML paketi konfigureeritavasse kausta. See pakett on Eesti standard XML kujul. Joonisel 9 on kujutatud näide neljast XML paketist kokku pandud EVS XML paketist. Pakett sisaldab endas Euroopa ja rahvusvaheliste standardite XML failide pakette (XML failid + pildid) ja EVS meta-andmetega XML faili.



Joonis 9 EVS XML paketi näide

Rakenduse teine tähtis funktsionaalsus on XML-HTML-PDF teisendus. Kui rakendus saab sisendiks EVS XML paketi, siis pakib rakendus selle ajutisse kausta lahti ning alustab XML-HTML-PDF teisendamist. Rakendus alustab XML failide meta-andmete põhjal standardi täpse struktuuri tuvastamist. Rakendus loob igale XML failile XmlStandard tüüpi objekti. Igal objektil väärtustatakse Xml, XDoc, XmlDirectory, IsPureIso, IsEuropean, IsEnIso, Originator, SpecificOriginator, MetaStructure, IsAmendment, OriginalTitle ja ReleaseDate. Objektid järjestatakse avaldamise kuupäeva järgi ja kõige pealmisel objektil väärtustatakse ka EvsData (ja EvsIsoData). Struktuuri määramisel võetakse objektid ühe- või kahekaupa ette. Kahekaupa juhul, kui on tegemist alusdokumendiga standardiga, et tuvastada Euroopa XML fail ja selle rahvusvaheline alusdokument. Terviklik standard võib olla struktuurilt:

1. Euroopa standard (CEN või CENELEC)
2. Rahvusvaheline standard (ISO või IEC)
3. Euroopa standard rahvusvahelise alusdokumendiga (CEN ISO või CENELEC IEC)

Pärast XmlStandard objektide atribuutide väärtustamist, standardi struktuuri tuvastamist ja objektide õigesse järjekorda panemist on standard valmis STS4i tööriistade abil XSL

transformatsiooniks. Rakendus on valmis XML faile järjekorras ette võtma ja neid STS4i tööriistade abil HTML formaati teisendama. Enne seda loetakse struktuurist tulenevalt sisse õiged mallid, väärtustatakse nende vajalikud elemendid ja moodustatakse seeläbi standardile baas.

Rakendus loeb sisse EVS tiitellehe HTML malli ja pealmise XML faili EvsData (ja EvsIsoData) väärtuste abil väärtustatakse seal vajalikud väljad. Seejärel tehakse valik euroopapõhise ja rahvusvahelise eessõna HTML malli vahel ning väärtustatakse selle elemendid kasutades kõigi objektide andmeid ja XML faile. Euroopa struktuuri puhul lisatakse ja väärtustatakse veel Euroopa standardi tiitellehe HTML mall.

Kui eesmised lehed on paigas, siis teisendatakse XML failidest vajalik sisu STS4i tööriista abil HTML formaati. Pärast seda viiakse läbi põhjalik töötlus, et HTML oleks hiljem Prince programmiga ühilduv. Töötluse käigus toimub autoriõiguse lisamine, liigsete tühikute ja tühjade elementide eemaldamine, piltide linkimine, tabelite ja piltide asendamine suuruste kohandamine, spetsiifiliste elementide klassifitseerimine (sissejuhatus, bibliograafia, allmärkused, tabelid, teistsuguse orientatsiooniga elemendid, eessõnad), korduvate elementide eemaldamine, klasside täpsustamine ja nende sisu parandamine või liigutamine. Samuti selle kõige käigus ka muid väiksemaid, kuid vajalikke ümbertegemisi. See HTML töötlus on koos CSS stiililehega kõige koodirohkem osa lahenduses ja see annab standardile suuresti lõppdokumendi väljanägemise.

Töötluse läbinud sisu lisatakse mallidest kokku pandud osale otsa ja seejärel pannakse lõppu EVS tagakaas. Selles faasis on dokument juba üsna terviklik. Rakendus genereerib pealkirjadest tulenevalt sisukorra, mis sisestatakse eessõna või Euroopa tiitellehe järele. Samuti toimub veel väiksem ümbertegemine seoses suuremate tabelite vale käsitlesega Prince teisendusel.

Lõpuks salvestatakse HTML fail konfigureeritud kausta. Eelvaate saamiseks tehakse failist koopia, mis läbib ka tükeldamise ja vesimärgistamise. HTML faasi läbimisel on tulemuseks kaks HTML faili, mis on valmis HTML-PDF teisenduseks ja veebilehel kasutamiseks.

Rakendus kutsub välja Prince rakenduse ja annab sellele valminud HTML täisteksti faili koos spetsiaalse CSS stiililehega. Prince määrab lehekülgede murdekohad, lisab lehekülge numbrid, tähised, autoriõiguse kirjed, sisukorra lehekülgede viited ja väljastab PDF

dokumendi. Enamik operatsioone toimub vastavalt CSS stiililehele. Vastavalt saadud lehekülgede arvule lisatakse vajadusel HTML failile tühjade lehekülgede elemendid, et dokument oleks korrektselt printitav ja teisendatakse lõplik fail uuesti. Valminud EVS standardi täisteksti dokument salvestatakse konfigureeritud kausta.

HTML-PDF teisenduse läbib ka eelvaate HTML fail, millele tühjasid lehtesid ei lisata. See-eest antakse Prince rakendusele kaasa Javascript fail sisukorra puuduvate lehtede parandamiseks. Prince väljastab ja salvestab EVS standardi eelvaate dokumendi konfigureeritud kausta.

### 3.3 Kood

Autor üritas realiseerida koodi vastavalt puhta koodi tavadele. Autor proovis hoida meetodite ja muutujate nimed loetavana ning vältida pikkasid ebaloetavaid koodirodusid. Autor toob välja joonistena välja mõned tüüpilisemad näited lahenduse koodist.

Joonis 10 on näide meetodist, mis kujutab endast Euroopa tiitellehe malli täitmist XML faili meta-andmete ja EVS meta-andmetega. See meetod sisaldab endas kogu funktsionaalsust alam-meetodite näol. Autor jagas meetodi väiksematesse alam-meetoditesse, et tagada koodi muutmise mugavus ja üldine loetavus.

```
public static HtmlNode FillEnTitlepage(HtmlDocument htmlDoc, XmlStandard xmlStandard) {
    if (htmlDoc.DocumentNode.SelectSingleNode("//body/div") == null)
        throw new ConversionException("Template on vigane.");
    ReplaceReferenceWithoutYear(htmlDoc, xmlStandard.XDoc);
    ReplaceDateYearAndMonth(htmlDoc, xmlStandard.XDoc, xmlStandard.ReleaseDate);
    ReplaceIcsGroups(htmlDoc, xmlStandard.XDoc);
    ReplaceSupersedes(htmlDoc, xmlStandard.XDoc);
    ReplaceEuropeanTitlepageTitles(htmlDoc, xmlStandard.XDoc);
    ReplaceSections(htmlDoc, xmlStandard.XDoc);
    ReplaceCopyright(htmlDoc, xmlStandard);
    ReplaceReferenceNumber(htmlDoc, xmlStandard.XDoc);
    var htmlBody = htmlDoc.DocumentNode.SelectSingleNode("//body//div");
    return htmlBody;
}
```

Joonis 10 FillEnTitlepage meetod

Joonis 11 on näide ühest lihtsamast HTML elementi muutvast meetodist. Sarnase eesmärgiga meetodeid on rakenduses palju ja need muudavad HTML formaati teisendatud standardi struktuuri, et see oleks võimalikult lähedane Euroopa ja

rahvusvaheliste standardite PDF kujul standardite visuaalsele välimusele. Need meetodid sõltuvad nii NISO STS reeglistikust kui ka autori poolt välja töötatud CSS failist.

```
public static HtmlNode MoveArrayOutOfImage(HtmlNode imageNode) {
    if (imageNode.ParentNode.Attributes["class"]?.Value.Contains(StsFig) == false)
        return imageNode;
    var arrayNode = imageNode.ParentNode.ChildNodes
        .FirstOrDefault(c => c.Attributes["class"]?.Value.Contains(StsArray) == true);
    if (arrayNode == null) return imageNode;
    var parent = arrayNode.ParentNode;
    var duplicate = arrayNode;
    arrayNode.Remove();
    parent.ParentNode.InsertAfter(duplicate, parent);
    return imageNode;
}
```

Joonis 11 MoveArrayOutOfImage meetod

Joonis 12 sisaldab väikest väljavõtet CSS failist, mis on vajalik HTML-PDF etapis Prince abiga teisendamisel. Tegemist on mittetraditsioonilise CSS osaga, mis tagab lehekülgede õiged mõõtmed, peegelveeriste korrektsuse, õigetes kohtades lehekülgede murdumise, lehekülgede nummerdamise, päiste mõõtmed erinevatel lehekülgedel jpm.

```
div.sts-array { page-break-inside: avoid; }
div.sts-fig-index { page-break-inside: avoid; }
div.sts-fig { page-break-inside: avoid; }
div.sts-table-wrap { page-break-inside: avoid; }
div.sts-table-wrap table tr { page-break-inside: avoid; }
div.sts-table-wrap .sts-caption .sts-caption-label { page-break-after: avoid; }
div.sts-tbx-label { page-break-after: avoid; }
div.sts-tbx-term { page-break-after: avoid; }
span[class*="bsFont-char"] { page-break-after: avoid; }
div.sts-ref-list li { page-break-inside: avoid; }
div.sts-footnote { page-break-inside: avoid; }
div.sts-section-bibl { page-break-inside: avoid; }

/*Always break after*/
div.sts-empty-page { page-break-after: always; }
}
@media screen { body { background: #e0e0e0 }
.sts-standard { background: white; box-shadow: 0 .5mm 2mm rgba(0,0,0,.3); margin: 5mm
.sts-landscape-page { background: white; box-shadow: 0 .5mm 2mm rgba(0,0,0,.3); margin
#copyright-notice { visibility: hidden; display: block; width: 0; height: 0; string-se

/*Pages*/
@page { size: A4; margin: 1.2cm 1cm 1cm 1cm; font-family: 'Cambria'; font-size: 11pt;
@page:right { margin: 1.7cm 1.5cm 1.6cm 2.0cm;
    @top { content: string(avsReference); white-space: nowrap; text-align: right; marg
    @bottom-left { content: string(copyright); font-size: 9pt; text-align: left; verti
    @bottom-right { content: counter(page); text-align: right; margin-right: -72mm; ve
}
@page:left { margin: 1.7cm 2.0cm 1.6cm 1.5cm;
    @top { content: string(avsReference); white-space: nowrap; text-align: left; verti
    @bottom-right { content: string(copyright); font-size: 9pt; text-align: right; mar
    @bottom-left { content: counter(page); text-align: left; vertical-align: top; }
}
```

Joonis 12 Lahenduse Prince-CSS kood

Joonis 13 sisaldab väljavõtet tavalisest rakenduse poolt kasutatavast CSS koodist. See kood kirjeldab nii autori poolt välja töötatud malli elementide välimust kui ka XML failist genereeritud HTML faili välimust. See osa annab standardi sisule Euroopa ja rahvusvahelise standardi PDF dokumendile sarnase väljanägemise ja standardi EVS lehekülgedele EVS nõuetele vastava väljanägemise. CSS kood on suures osas kirjutatud vastavalt genereeritud dokumendi visuaalsele võrdlusele Euroopa ja rahvusvaheliste PDF dokumentide ning EVS lehekülgedega.

```
hr { border: 0; background-color: #333; color: #333; height: 0.137em; border-color: #333; }
hr.titlepage-line-top { margin-top: 5pt; }
hr.titlepage-line-bottom { margin: 10pt 0pt 10pt 0pt; }
.sts-footnotes hr { height: 0.868em; margin-bottom: 8pt; text-align: justify; }
.align-center { text-align: center; }

/*BS font*/
@font-face { font-family: 'bsFont'; src: url(BSsyms10.ttf); }
span[class*='bsFont-char'] { font-family: 'bsFont'; margin-bottom: -10pt; font-weight: normal; }
span.bsFont-right { display: flex; margin-bottom: 6pt; justify-content: flex-end; }

/*EVS titlepage template*/
#evs-titlepage-grad { vertical-align: middle; height: 25pt; background-color: #0357a6; background-weight: bold; }
#evs-titlepage-header-left { line-height: 25pt; font-size: 19pt; float: left; vertical-align: middle; }
#evs-titlepage-header-right { position: relative; margin-right: 3pt; line-height: 25pt; font-size: 19pt; middle; display: inline-block; string-set: evsReference content(); }
#evs-title-et { font-size: 19pt; margin-top: 220pt; margin-bottom: 0; text-transform: uppercase; }
#evs-title-en { font-size: 19pt; margin-top: 28pt; margin-bottom: 0; font-weight: bold; }
#evs-titlepage-footer { position: absolute; right: 0; bottom: 0.8cm; }
#evs-logo { vertical-align: bottom; }

/*EVS foreword template START*/
.evs-foreword-table { border-collapse: collapse; position: relative; width: 45.1em; text-align: justify; }
#evs-foreword-table-upper tr th { font-size: 13pt; text-transform: uppercase; text-align: left; border: 1px solid white !important; border-left: 0.068em solid white; border-right: 0.068em solid white; }
.evs-foreword-table tr td { border-right: 0.068em solid black; font-size: 11pt; padding: 4pt 10pt 4pt 10pt; }
```

Joonis 13 Rakenduse tavaline CSS kood

Joonis 14 illustreerib EVS eessõna lehekülje HTML malli. Autor realiseeris EVS eessõna lehekülje HTML kujul ja struktureeris selle muudetavad elemendid rakendusele lihtsasti loetavale kujule identifikaatorite ja klasside abil. Lahendus loeb malle sisse HTML-PDF teisendamisel. Mallide välja töötamine tähendas, et lahendus ei pea EVS lehekülgi iga kord uuesti koostama. Samuti on mallid lahendusest eraldatud ja neid on vajaduse korral lihtne muuta ja täiendada.

<b>EESTI STANDARDI EESSÕNA</b>	<b>NATIONAL FOREWORD</b>
<p>See Eesti standard [EVS tähis] sisaldab Euroopa standardi [EN tähis] ingliskeelset teksti.</p> <p>Standard on jõustunud sellekohase teate avaldamisega EVS Teatajas.</p> <p>Euroopa standardimisorganisatsioonid on teinud Euroopa standardi rahvuslikele liikmetele kättesaadavaks [Date of Availability].</p> <p>Standard on kättesaadav Eesti Standardikeskusest.</p>	<p>This Estonian Standard [EVS tähis] consist of the English text of the European Standard [EN tähis].</p> <p>This standard has been endorsed with a notification published in the official bulletin of the Estonian Centre of Standardisation.</p> <p>Date of Availability of the European Standard is [Date of Availability].</p> <p>The standard is available from the Estonian Centre for Standardisation.</p>
<p>Tagasisidet standardi sisu kohta on võimalik edastada, kasutades EVS-i veebilehel asuvat tagasiside vormi või saates e-kirja meiliaadressile <a href="mailto:standardiosakond@evs.ee">standardiosakond@evs.ee</a>.</p>	
<p><b>Standardite reprodutseerimise ja levitamise õigus kuulub Eesti Standardikeskusele</b></p> <p>Andmete paljundamine, taastekitamine, kopeerimine, salvestamine elektroonsesse süsteemi või edastamine ükskõik millises vormis või millisel teel ilma Eesti Standardikeskuse kirjaliku loata on keelatud.</p> <p>Kui Teil on küsimusi standardite autorikaitse kohta, võtke palun ühendust Eesti Standardikeskusega: Koduleht <a href="http://www.evs.ee">www.evs.ee</a>; telefon 605 5050; e-post <a href="mailto:info@evs.ee">info@evs.ee</a></p> <p><b>The right to reproduce and distribute standards belongs to the Estonian Centre for Standardisation</b></p> <p>No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, without a written permission from the Estonian Centre for Standardisation.</p> <p>If you have any questions about copyright, please contact the Estonian Centre for Standardisation: Homepage <a href="http://www.evs.ee">www.evs.ee</a>; phone +372 605 5050; e-mail <a href="mailto:info@evs.ee">info@evs.ee</a></p>	

Joonis 14 EVS eessõna mall

Joonis 15 illustreerib ühte autori poolt tehtud muudatust STS4i tööriistade XSL koodis. Autor oli sunnitud tegema mõningaid muudatusi ja parandusi ISO poolt välja töötatud tööriistades, sest nendes oli puudujääke. Autor lisas igale muudatusele kommentaari, et see oleks lihtsasti leitav ja originaalkoodist eraldatud. Mõningaid parandusi oleks saanud teha ka .NET rakenduse-siseselt, kuid kui autor leidis, et tegemist on pigem veaga, siis oli parem parandada XSL koodi.

```

<div class="{sts-prefix}tr--title-dot-1">
  <xsl:apply-templates select="title-wrap[@xml:lang = $doc-lang]/intro,
                                title-wrap[@xml:lang = $doc-lang]/main,
                                title-wrap[@xml:lang = $doc-lang]/compl,
                                custom-meta-group/custom-meta[@xml:lang = $do
                                "
                                mode="#current">
    <xsl:with-param name="is-main-meta" select="$is-main-meta"/>
  </xsl:apply-templates>
<!-- EVS Rain Hallikas 2018
Fixed a problem where the <full> title was not always rendered -->
</div>
<xsl:choose>
  <xsl:when test="title-wrap[@xml:lang = $doc-lang]/main = ''">
    <xsl:apply-templates select="title-wrap/full" mode="#current"/>
  </xsl:when>
  <xsl:otherwise>
    <xsl:apply-templates select="title-wrap[@xml:lang != $doc-lang]/full"
    </xsl:otherwise>
  </xsl:choose>
</div>
</xsl:template>

```

Joonis 15 STS4i tööriistade XSL kood

Joonis 16 kirjeldab ResizeImage meetodi testimist. Autor üritas näha võimalikult palju erinevaid olukordi, mis teisendamisel esinevad võivad ja koostas igale olukorrale TestCase-põhise testi. Autor ei testinud kõiki meetodeid ja kõiki võimalike variatsioone, sest see oleks olnud ajamahukas, kuid põhilised stsenaariumid said kaetud. ResizeImage meetodit on samuti võimalik testida veel põhjalikumalt näiteks teises suuruses piltide.



```

[TestMethod]
[DataRow(NameInfo.IsoCenRegularImageId, ElementNames.Cen, 2009, "width: 668px;")]
[DataRow(NameInfo.IsoCenRegularImageId, ElementNames.Cen, 2014, "width: 133.65px;")]
[DataRow(NameInfo.IsoCenRegularImageId, ElementNames.Cenelec, 2009, "width: 668px;")]
[DataRow(NameInfo.IsoCenRegularImageId, ElementNames.Cenelec, 2014, "width: 133.65px;")]
[DataRow(NameInfo.IsoCenRegularImageId, ElementNames.Iso, 2014, "width: 285.12px;")]
[DataRow(NameInfo.IsoCenRegularImageId, ElementNames.Iso, 2009, "width: 285.12px;")]
0 references | Rain Hallikas, 2 days ago | 1 author, 2 changes
public void ResizeImage_RegularImage_WidthDecreasedBasedOnOriginatorAndYear(
    string imageId, string originator, int year, string expected) {
    var image = IsoCenHtml.GetElementById(imageId);
    Assert.IsNull(image.Attributes["style"]);

    var xmlStandard = new XmlStandard {
        Originator = Originator.GenerateOriginator(originator, false),
        ReleaseDate = new System.DateTime(year, 1, 1)
    };
    image = ImageModifier.ResizeImage(image, xmlStandard, NameInfo.IsoCenRegularImageDirectory, false);
    Assert.AreEqual(expected, image.Attributes["style"]?.Value);
}

```

#### Joonis 16 ResizeImage testid

Autor lähtus koodi kirjutamisel komponentide ja funktsionaalsuse piisavast eraldatusest, kuid siiski ka sellest, et kõik lahenduse osad moodustaksid ühtse terviku. Samuti üritas autor hoida koodi loetava ja arusaadavana. Autor üritas koodi kirjutamisel vältida rohket kommentaaride kasutamist, et hoida kood võimalikult müravabana. Autor kasutas kommentaare vaid keerulisemates ja ebatraditsioonilisemates koodilõikudes.

## 4 Analüüs

### 4.1 Funktsionaalsus

EVS XML-HTML-PDF teisendaja põhiline arendusfaas on lõppenud ja ootab suuremat testimist. Autor parandas eelmise teisendaja vigu, kuid XML failidest standardi koostamisel esineb ka uusi vigu. Siiski on teisendaja võimekam ja Eesti standardikeskusele parem lahendus, kui PDF-PDF teisendamisega jätkamine.

PDF-PDF teisendaja ei suuda leheküljenumbreid otsast-otsani õigesti lugeda, sest Eesti standard kleebiti kokku juba nummerdatud PDF osast ja uutest EVS lehekülgedest. See tekitas olukorra, kus eesmised EVS leheküljed nummerdati rooma numbritega. Sisu oli juba nummerdatud ja kleebiti EVS lehekülgedele otsa. Uus teisendaja alustab nõuetele vastavalt leheküljenumbrite lugemist alates tiitellehest. XML failis ei ole leheküljenumbreid, kuid see sisaldab standardi täielikku sisu. Pärast EVS lehekülgede genereerimist ja nende nummerdamist saab teisendaja sellele sisu lisada. Prince murrab lehekülgesid vastavalt autori seatud EVS juhendist tulenevatele nõuetele ja lisab igale leheküljele vastava numbri.

PDF-PDF teisendaja jätab Eesti standardile ebakorrapärased peegelveerised. See on tingitud sellest, et standardi sisu võetakse standardi PDF dokumendist ja sellel on veerised seeläbi juba paigas. Euroopa ja rahvusvaheliste standardite peegelveeriste mõõtmed pole EVS nõuetega kooskõlas ja Eesti standardi dokumenti jäävad seetõttu ebahühtlased veerised. XML-HTML-PDF teisendaja genereerib EVS leheküljed ja alustab seejärel XML failist sisu lugemist. HTML kujul on genereeritud Eesti standard terviklik ja HTML-PDF etapis paigutatakse see kindla suurusega lehekülgedesse, mille vasak- ja parempoolsed veerised on alati kordamööda.

PDF-PDF teisendaja genereeritud Eesti standardi lingid ei tööta ja PDF järjehoidjat ei ole. Linkide töötamine on seotud PDF dokumendi tähistega muutmisega. Lehekülje elemente käiakse ükshaaval läbi ja muudetakse kindlatel koordinaatidel asetsevaid tekstiblokke. Pärast seda protsessi algsed lingid enam ei tööta ja PDF järjehoidjat kaob samuti. XML

failides on lingid ja pealkirjad NISO STS struktuurile vastavalt märgitud. Rakendus klassifitseerib XML-HTML etapis kõik lingid ja pealkirjad nii, et HTML-PDF teisendamisel suudaks Prince lingid ja järjehoidja realiseerida. Sellest tulenevalt realiseerib Prince kõik dokumendisisesed lingid töötavatena. Samuti koostab Prince korrektselt määratud pealkirjade hierarhia põhjal järjehoidja automaatselt.

PDF-PDF teisendajal esineb mõnikord probleem EVS tähise lisamisega, mis nõudis käsitsi parandamist. Tähise lisamine toimus läbi PDF dokumendi elementide töötuse koordinaatide põhised, mistõttu tekkis vahel probleem, kui leheküljed olid keeratud või struktuur oli ebatavaline. Prince võimaldab PDF lehekülgedele lisada päise. Autor defineeris nõuetele vastava päised CSS failis, et Prince saaks päiseid lisada igale leheküljele samamoodi.

PDF-PDF teisendaja on võimeline standardi sisust eelvaate genereerimiseks tuvastama vajaliku peatüki, kuid ei oska selle sisu eelvaate viimaselt leheküljelt eemaldada. Seetõttu jääb eelvaatesse rohkem sisu, kui vaja oleks. XML-HTML-PDF teisendaja suudab samuti leida eelvaate lõikepunktiks oleva peatüki ja sellele järgneva sisu edukalt leheküljelt välja jätta. Uus teisendaja kuvab seeläbi genereeritud eelvaates täpselt õige informatsiooni.

XML failidest Eesti standardi koostamisel on siiski ka murekohad. XML formaadis standardid genereeritakse Word formaadis standarditest, mis võivad sisaldada inimsilmale nähtamatuid vormistusvigu. Seetõttu jõuavad need vead ka XML formaadis standardisse ja XML-HTML-PDF teisendaja ei suuda neid vigu tuvastada. See tähendab, et Eesti Standardikeskus peab genereeritud Eesti standardeid põhjalikumalt üle vaatama. Vigade leidmisel on võimalus saata standardi avaldanud organisatsioonile kiri või parandada viga XML failis ise. Sõltuvalt vigaste XML failide hulgast ja vigade parandamise keerukusest peab Eesti Standardikeskus langetama otsuse, kuidas seda probleemi lahendada.

NISO poolt välja töötatud XML spetsifikatsioon on üsna hästi välja töötatud, kuid standardimisorganisatsioonidel on siiski palju vabadust, kuidas seda täpselt kasutada. See tähendab, et XML faile antakse tihti välja erineva vormistuse või puuduliku täpsustava informatsioonita. Samuti toimub ISO, CEN ja CENELEC XML failide kokkupanemine Word faili põhised, mis tähendab, et vormistuslikud vead väljenduvad ka XML failis. Lisaks sellele on funktsionaalsuse puudulikkust ka Prince tarkvaral.

NISO spetsifikatsioonis on tabeli teksti suuruse atribuut küll määratav, kuid Euroopa ja rahvusvahelised standardimisorganisatsioonid jätavad XML failides olevatel tabelitel selle väärtustamata. Mõningatel juhtudel on kaasa pandud kommentaar, kuid selle töötlemine ei ole eelistatav. Tabelitel on kaasa pandud laius, mille järgi saab neid dokumendi jaoks õigesse suurusesse kohandada, kuid kirja suurus pole alati sama.

Autor valis läbi katse-eksitus meetodi probleemi lahenduseks jätta kuni teatud laiuseni tabelite vaikimisi kirja suuruse muutmata. Kui tabelid on äärest-ääreni, siis on aga tõenäosus, et tabeli tekst võib olla väiksemaks tehtud. Rakendus käsitleb kõiki äärest-ääreni tabeleid väiksema tekstiga tabelitena. Meetodi puuduseks on see, et vahel kuvatakse tavapärase kirja suurusega mõeldud tabelid väiksema kirja suurusega. Samas välditakse probleemi, et kui tabeli tekst on tavapärase kirja suurusega, kuid mõeldud väiksemana, siis ei lähe tabel dokumendi piiridest välja. Teine võimalus tabeli piiridest väljumise vältimiseks oli kõigile tabelitele määrata vaba sõnade murdumine, kuid sõnade grammatiline ebaõigsus oli häirivam kui tabeli teksti veidi väiksem suurus.

NISO spetsifikatsioonis on tabeli orientatsiooni atribuut määratav, kuid enamik juhtudel seda ei kasutata. See tähendab, et kunagi ei saa täpselt teada, millal on tegemist keeratud tabeliga. Samas on tabelite keeramine vajalik, sest need lähevad dokumendi piiridest välja.

Autor leidis erinevate XML standarditega katsetamisel tabeli laiuse, millest alates on üldjuhul tegemist keeratud tabeliga. Kuigi mõnel üksikul juhul on tegemist äärest-ääreni tabeliga, on kindlam ja vähem riskantne alati tabelit keerata. Rakendus keerab orientatsiooni atribuudi väärtuse puudumisel ja laiuse määratud väärtuse ületamisel tabelit alati. Autor realiseeris ka tabeli orientatsiooni atribuudile vastava tabelite keeramise, sest uuematel juhtudel on atribuut vahel siiski määratud.

Kuigi Prince on üks parimaid HTML-PDF teisendamise tööriistu, siis on sellel ka teatud puudused. Nimelt on probleem tabelielementidega, mille sisu on pikem kui üks lehekülge. Kui Prince tuvastab, et tegemist on mitme lehelise tabeliga, siis alustatakse selle kuvamist uuel lehel. Joonis 10 illustreerib olukorda, kus tabeli pealkirjad jäävad leheküljele üksi.

## D.4.1 Datasheet for HVAC equipment — DX cooling coils

EVS tähis

D.4.1	HVAC EQUIPMENT DX COOLING COILS DATASHEET		Page 1 of 2
Package No.	Doc. No.		Rev.
Tag No.		Location/Module	
System No.		Number required	
Size and type		Enquiry No.	
Supplier		Quote No.	
Manufacturer		Purchase order No.	
Model		Job No.	
		Serial No.	

Joonis 17 Pikale tabelile eelnev murdumine

Autor realiseeris olukorra lahendamiseks tingimuse, et kui tabelile eelnevad pealkirjad, siis tõstetakse need tabeli elemendi sisse päiseks. Murdumise parandamine on realiseeritud vaid pealkirjadega, sest need on kõige kindlamad ja sagedamini esinevad juhud. Realiseerida sai elementide tõstmist tabelisse olenemata nende olemusest, kuid nendel juhtudel oli Prince'i teisendamine veelgi probleemsem.

Autor kaldus ka nõuetele vastamisest kõrvale seoses Eesti Standardikeskuse eelistuste kujunemisele töö arendamisprotsessi käigus. Otsused on langetatud koos Eesti Standardikeskuse töötajatega, et standard oleks loetavam, lugejasõbralikum või loogilisema ülesehitusega. Kõrvalepõiked nõuetest on originaalsisukorra ja sisupealkirjade realiseerimine vastavalt puhta Eesti standardi vormistusreeglitele.

Standardi originaalsisukord on Euroopa ja rahvusvahelistel standarditel eri struktuuriga, kuid autor realiseeris sisukorra hoopis puhta Eesti standardi nõuete kohaselt. Tavaliselt on standardi üle võtmisel jäänud standardi sisu välimus puutumata, kuid XML failidest standardi genereerimine võimaldas muutust. Sisukorras on alati kuvatud kirjed kuni teise alampealkirjani ja esimese kihi pealkirjad on suurte tähtedega. Standardi sisu tekstisisesed esimese kihi pealkirjad on samuti realiseeritud suurte tähtedega pealkirjadena.

## **4.2 Parendused**

EVS XML-HTML-PDF teisendaja põhiline arendusfaas on lõppenud ja teisendaja ootab suuremat testimist. Autor leiab, et lahenduse puhul on aspekte, mida saab paremini teha ja edasi arendada. Autor viis analüüsi, arendamist ja testimist läbi üksi, mistõttu valmis lahendus vaid autori võimete kohaselt ja arendamise murekohad said selgeks alles lahenduse valmimisel.

Lahenduse arendamisel ja testimisel kasutas autor ligi 40 erinevat XML standardipaketti. Paketid on erineva struktuuriga (CEN, CENELEC, CEN ISO, ISO) ja sisaldavad mitmekesiseid elementide kombinatsioone. Siiski leiab autor, et lahendus pole lõplik, sest XML standardid võivad olla ootamatult ebatavalisel kujul. Kindlasti on struktuurilisi variatsioone, mille korrektsele teisendamisele lahendus veel ei oska tähelepanu pöörata. Neid variatsioone aitab leida suurem testimine, mis kujutab endast suure hulga dokumentide genereerimist ja teiste EVS töötajate kaasamist dokumentide visuaalsesse

võrdlusprotsessi. See aitab leida lahenduse murekohti ja testida, kas lahendus on valmis PDF-PDF teisendaja väljavahetamiseks.

Autor leiab, et kuigi rakendus sai struktureeritud ja arendatud suuresti vastavalt puhta koodi tavadele, on siiski ka puudujääke. Autor tegeles arendamise käigus pidevalt visuaalse testimise ja vea parandamisega, mis koosnes originaaldokumendi võrdlemisest genereeritud failidega. Pärast võrdlemist asus autor tähele pandud nõudeid realiseerima ja vigu parandama. Tihti koostades igale veale ja nõudele eraldi meetodi. Mõni viga nõudis pikema ja põhjalikuma meetodi kirjutamist, mistõttu tuli mõni meetod pikem või konarlikum. Vigade ja vormistuslike nõuete rohkusest tingituna üritas autor arendada tempokalt, jättes kohati realiseerimata abimeetodeid ja meetodeid tükeldamata. Seetõttu leidub meetodeid, mida saaks vajadusel analüüsida ja paremini kirjutada (Joonis 11).

```
public static HtmlDocument MoveHeadingsToTable(HtmlDocument htmlDoc) {
    var tableContainers = htmlDoc.DocumentNode.Descendants()
        .Where(d => (d.Attributes["class"]?.Value.Contains(StsTableWrap) == true
            || d.Attributes["class"]?.Value.Contains(StsArray) == true
            || d.Attributes["class"]?.Value.Contains(StsFig) == true)
            && d.Attributes["class"]?.Value.Contains(StsTableWrapFoot) == false
            && d.Attributes["class"]?.Value.Contains(SideturnedTable) == false).ToList();
    foreach (var tableContainer in tableContainers) {
        var mainNode = tableContainer;
        if (tableContainer.Attributes["class"].Value.Contains(StsLandscapePage)) continue;
        while (mainNode.ParentNode.Attributes["class"]?.Value.Contains(StsMain) == false) {
            var previousNode = mainNode.PreviousSibling;
            var counter = CountPreviousNodes(previousNode);
            if (counter < 2) MoveNodesIntoTableContainer(tableContainer, mainNode);
            if (counter > 1) break;
            MoveCoParentNodesIntoTableContainer(tableContainer, mainNode);
            break;
        }
    }
    return htmlDoc;
}
```

Joonis 18 MoveHeadingsToTable meetod

Autor toob Joonis 11 näitel välja mõned aspektid, mida saab rakenduses parandada. Joonisel on spetsiifiliste tabelite leidmiseks kasutusel pikk välistamine, millele saaks kirjutada abimeetodi, mis vähendaks selle pikkust. Teine võimalus oleks tabelite leidmine kirjutada eraldi meetodisse FindRegularTableContainers, et meetod oleks lühem ja loetavam. Tabelite käsitlemisel teeb meetod mitut asja, mida saaks omakorda eraldada paari väiksemasse meetodisse. Hetkel on olukord, kus mõne spetsiifilisema murekoha leidmisel peab autor enda kirjutatud meetodisse süvenema, et viga leida. Samas tasub mainida, et liigsel tükeldamisel võib tekkida ka olukord, kus kood on liiga lai ja ei jäta terviklikku muljet. Sellest tulenevalt jättis autor hetkel mõned meetodid hilisemaks

analüüsimiseks ja parendamiseks. Kõige olulisem oli lahendus suureks testimiseks valmis saada, et lahenduse integratsioon Eesti standardi genereerimise protsessi saaks alata.

Lahenduse meetodid ei ole saajaprotsendiliselt testitud. Seoses sellega, et dokumendi elementide vormistuse variatsioon on ääretult palju, kasutab autor testimisel peamiselt ühte suurt CEN ISO standardit. Standard sisaldab peamisi võimalikke olukordi ja autor annab testidele sisendiks elemente sellest standardist. Kõige rohkem on HTML-põhiseid meetodeid, mis ootavad sisendiks erinevaid HTML struktuure. Kõigi võimalike HTML olukordade loomine on pikem protsess ja vajaks eraldi tähelepanu. Seetõttu realiseeris autor vaid enim esinevate olukordade testid. Samuti jäid mõned meetodid testimata. Nende testimine jääb autorile järgmiseks arendustsükliks.

Lahendus ei realiseeri Eesti standardile dokumendiväliseid linke. Dokumendivälised lingid sisaldavad XML struktuuris tavaliselt identifikaatoripõhist viidet standardile. Potentsiaalne edasiarendus oleks leida lingi identifikaatori kaudu EVS, ISO, CEN või CENELEC veebilehtedelt vastava standardi viide ja siduda see lingiga. Dokumendivälised lingid jäid skoobist välja, sest nende arendamine oli liialt aeganõudev ja pole niivõrd oluline, kui lahenduse ülejäänud funktsionaalsus.

Lahendus ei ole võimeline genereerima Eesti standardeid IEC XML failidest, sest IEC avalda veel standardeid XML formaadis. Kui IEC hakkab standardeid XML formaadis avaldama, siis on lahendust vastavalt edasi arendada. Hetkel ei ole arendus võimalik, sest ei ole täpselt teada, kuidas erinevad IEC XML failid teiste standardimisorganisatsioonide omadest. IEC XML failidest Eesti standardi genereerimise arendus jääb ootele, kuniks IEC hakkab väljastama standardeid XML formaadis.

Kui lahendus läbib suurema testimise, tähtsamad probleemid saavad lahendatud ja integratsioon StaDIS-ega on edukas, siis saab autor tegeleda tähelepanuta jäänud aspektidega. Plaanis on veelgi põhjalikum koodi meetodite tükeldamine ja loetavamaks muutmine vastavalt puhta koodi tavadele, ühiktestide põhjalikum kirjutamine ja dokumendiväliste linkide realiseerimine. Lahenduse funktsionaalsuse muutmine on võrreldes PDF-PDF teisendajaga üsna paindlik ja seeläbi on ka uute muudatuste läbiviimine lihtsustatud. Autor on valmis ja võimeline parandama lahendusega seotud vigu ja realiseerima uut funktsionaalsust.



### 4.3 Autori hinnang

Algselt oli autori ülesanne koostada uuele teisendajale nõuded ja koostada lähteülesanne, et Eesti Standardikeskuse välised arendajad saaksid lahenduse välja töötada. Autor leidis, et lahenduse välja töötamine oleks seeläbi nõudnud väga palju edasi-tagasi suhtlust. Autor otsustas seeläbi uurida võimalusi lahenduse realiseerimiseks ja leidis, et see on jõukohane projekt üksi arendamiseks.

Lahenduse välja töötamiseks kulus autoril ligikaudu kaks kuni kolm kuud intensiivsemat arendamist töö käigus. Pärast seda tegeles autor lahenduse väiksemate probleemide lahendamise ja testimisega rahulikumalt. Hetkel on lahendus suuresti valmis ja ootab suuremat visuaalset testimist kaastöötajate abiga. Lahendus on vaja integreerida StaDIS test-keskkonnaga, et oleks ligipääs suurele hulgale imporditud Euroopa ja rahvusvaheliste standardite XML failidele. Autor leiab, et lahendus on selle jaoks valmis ja eeldab, et testimise käigus esineb uusi olukordi, millega lahendus veel kokku puutunud ei ole.

Lahenduse välja töötamine oli autori jaoks väljakutsuv, sest arendus oli mitmekesine. Autor oli küll kokku puutunud .NET raamistikku, C# keele, HTML-i, CSS-i, JavaScripti ja XSL-i ga, kuid polnud veel sellist XML-HTML-PDF teisendaja sarnast rakendust arendanud. Samuti õppis autor kasutama ka Prince tarkvara. Autor leiab, et lahenduse välja töötamine on olnud hea kogemus, sest see pani proovile autori olemasolevaid oskusi, kinnistas neid ja õpetas ka uut. Autor on nüüd võimeline arendama sarnast rakendust kiiremini ja puhtamalt, kui enne.

Autor leiab, et lahendus tuli hästi välja. Lahenduse kõige suurem tugevus on võimalus seda mugavalt täiendada. Samuti on lahendus paindlikum ja funktsionaalsem kui PDF-PDF teisendaja. Lahendus jääb esialgu kasutusse koos PDF-PDF teisendajaga, seoses sellega, et XML failid võivad olla vigased. Samuti pole uus lahendus olnud veel kasutuses ja see pole usaldusväärne. Kui uus lahendus tagastab vea, siis pöörduetakse vana PDF-PDF teisendaja poole. Autor on arendusega rahul.

## 5 Kokkuvõte

Lõputöö eesmärk oli vahetada välja Eesti Standardikeskuse standardite üle võtmisel kasutatav PDF-PDF teisendaja XML baasil põhineva lahenduse vastu. PDF-PDF teisendaja välja vahetamine oli vajalik, sest genereeritud Eesti standarditel oli probleeme lehekülje numbrite, peegelveeriste, linkide, järjehoidjate ja päiste korrektsel viisil kuvamisega. Lisaks pakub XML baasil põhinev lahendus Eesti Standardikeskusele rohkem võimalusi standardite üle võtmisel ja avaldamisel.

Autor töötas välja täpsemad nõuded ja leidis lahenduse realiseerimiseks sobilikud tööriistad. Autor otsustas arendada .NET lahenduse, mis kasutab Prince HTML-PDF teisendajat, STS4i tööriistasid, HTML malle ja Prince-spetsiifilist CSS ja JavaScript faili. Tulemuseks oli eraldiseisev parameetritega välja kutsutav XML-HTML-PDF teisendaja, mis pärast suuremat testimist ja vigade parandamist integreeritakse Eesti Standardikeskuse standardihalduse veebirakendusse StaDIS. Lahendus suutis parandada kõik PDF-PDF teisendaja vead ning pakkuda uusi Eesti standardi formaate XML ja HTML näol.

Lahendus on valmis suuremaks testimiseks ja piisavalt paindlik, et parandusi või uut funktsionaalsust kiiresti realiseerida. Lahenduse kood on üldiselt puhas ja korralik, kuid autor leiab, et seda on võimalik parandada ning rohkem ühiktestidega katta. Seoses vigaste XML failidega võivad teisendusel tekkida standardi sisus vead, mida autor edasise testimise käigus võimalikult suurel hulgal likvideerida üritab. Seoses võimalike sisuliste vigadega ei kao standardite avaldamiseelne visuaalne ülevaatus. Lahendus vastab siiski nõuetele ja võimaldab Eesti Standardikeskusel mugavamalt ja kvaliteetsemalt standardite üle võtmist. Autor leiab, et XML-HTML-PDF teisendaja arendus oli edukas.

## Kasutatud kirjandus

- [1] „EVS,“ [Võrgumaterjal]. Available: <https://www.evs.ee/>. [Kasutatud 26 2 2019].
- [2] „XML,“ [Võrgumaterjal]. Available: <https://www.w3.org/XML/>. [Kasutatud 26 2 2019].
- [3] „HTML,“ [Võrgumaterjal]. Available: <https://www.w3.org/html>. [Kasutatud 25 2 2019].
- [4] „PDF,“ [Võrgumaterjal]. Available: <https://acrobat.adobe.com/us/en/acrobat/about-adobe-pdf.html>. [Kasutatud 26 2 2019].
- [5] „CSS,“ [Võrgumaterjal]. Available: <https://www.w3.org/Style/CSS/>. [Kasutatud 25 2 2019].
- [6] „ISO,“ [Võrgumaterjal]. Available: <https://www.iso.org/home.html>. [Kasutatud 26 2 2019].
- [7] „IEC,“ [Võrgumaterjal]. Available: <https://www.iec.ch/>. [Kasutatud 26 2 2019].
- [8] „CEN & CENELEC,“ [Võrgumaterjal]. Available: <https://www.cencenelec.eu/Pages/default.aspx>. [Kasutatud 26 2 2019].
- [9] „WORD,“ [Võrgumaterjal]. Available: <https://products.office.com/en-us/word?SilentAuth=1&wa=wsignin1.0>. [Kasutatud 26 2 2019].
- [10] „STS4i Tools,“ [Võrgumaterjal]. Available: <https://github.com/sts4i/sts4i-tools>. [Kasutatud 14 9 2018].
- [11] „XSL,“ [Võrgumaterjal]. Available: <https://www.w3.org/Style/XSL/WhatIsXSL.html>. [Kasutatud 26 2 2019].
- [12] „EVS Juhend 5,“ [Võrgumaterjal]. Available: <https://www.evs.ee/LinkClick.aspx?fileticket=sYbJ665UfpI%3d&tabid=207>. [Kasutatud 09 04 2019].

- [13] „EVS Juhend 4,“ [Võrgumaterjal]. Available: <https://www.evs.ee/LinkClick.aspx?fileticket=XOhgFmsuYd8%3d&tabid=207>. [Kasutatud 12 9 2018].
- [14] „Prince XML,“ [Võrgumaterjal]. Available: <https://www.princexml.com>. [Kasutatud 13 9 2018].
- [15] „NET Framework,“ [Võrgumaterjal]. Available: <https://docs.microsoft.com/en-us/dotnet/framework/get-started/index>. [Kasutatud 25 2 2019].
- [16] „Javascript,“ [Võrgumaterjal]. Available: [https://developer.mozilla.org/en-US/docs/Web/JavaScript/About\\_JavaScript](https://developer.mozilla.org/en-US/docs/Web/JavaScript/About_JavaScript). [Kasutatud 25 2 2019].
- [17] „Veeb,“ [Võrgumaterjal]. Available: <https://et.wikipedia.org/wiki/Veeb>. [Kasutatud 25 2 2019].
- [18] „NISO STS,“ [Võrgumaterjal]. Available: <https://www.niso-sts.org/>. [Kasutatud 15 01 2019].