

TARTU ÜLIKOOL
MATEMAATIKA-INFORMAATIKATEADUSKOND
Arvutiteaduse Instituut
Tarkvarasüsteemide õppetool
Informaatika eriala

Merle Sibola

ELEKTROONILINE DOKUMENDIHALDUS

Magistritöö

Juhendaja: Anne Villems

Autor: "..." 2002

Juhendaja: "..." 2002

Õppetooli juhataja: "..." 2002

Tartu 2002

Sisukord

SISSEJUHATUS	4
1. INFORMATSIOON, DOKUMENT JA MEEDIUM	6
1.1 MÕISTED	6
1.2 DOKUMENTIDE LIIGITAMINE.....	10
1.3 ERINEVAD DOKUMENDI MEEDIUMID	13
1.3.1 <i>Paber</i>	13
1.3.2 <i>Mikrograafika</i>	14
1.3.3 <i>Elektroonilised formaadid</i>	15
1.4 DOKUMENDI ELUTSÜKKEL.....	17
1.5 MILLAL MILLINE MEEDIUM?.....	20
2. DOKUMENDIHALDUS JA -SÜSTEEMID	22
2.1 SISSEJUHATUS DOKUMENDIHALDUSSE	22
2.1.1 <i>Dokumendihalduse areng</i>	22
2.1.2 <i>Dokumendihalduse vajadus</i>	23
2.1.3 <i>Dokumendihalduse põhimõtted</i>	25
2.2 DOKUMENDIHALDUSSÜSTEEMIDE TEGEVUSED	26
2.2.1 <i>Dokumendi loomine</i>	27
2.2.2 <i>Dokumentide otsing</i>	29
2.2.3 <i>Dokumendi versiooni haldus</i>	29
2.2.4 <i>Dokumentide säilitamine</i>	30
2.2.5 <i>Väljundi haldus</i>	31
2.2.6 <i>Kasutajad</i>	33
2.2.7 <i>Lisavõimalused</i>	33
2.3 DOKUMENDIHALDUSSÜSTEEMI RAKENDAMINE	34
3. DOKUMENDIHALDUSEGA SEOTUD STANDARDID JA TEHNOLOOGIAD	40
3.1. ODMA JA DMA	40
3.1.1 <i>ODMA</i>	40
3.1.2 <i>DMA</i>	42
3.2 XML	44
3.3 DIGITAALALKIRI.....	46
3.4 OCR, ICR JA OMR.....	49
3.4.1 <i>OCR ja ICR</i>	49
3.4.2 <i>OMR</i>	51
3.5 ANDMEBAASID.....	53
3.6 DOKUMENDI GRAAFILINE KUJUTAMINE	55

3.7 VÄLJUNDI HALDUS.....	56
3.8 VEEBI SISU HALDUS	58
KOKKUVÕTE.....	62
ELECTRONIC DOCUMENT MANAGEMENT.....	63
KASUTATUD KIRJANDUS.....	64
LISA 1. PABERIL PÕHINEV JA ELEKTROONILINE DOKUMENDIHALDUS.	67
LISA 2. ELEKTROONILINE DOKUMENDIHALDUSSÜSTEEM: MIS SEE ON JA KAS MA PEAN SELLE OSTMA?.....	69
LISA 3. UUE DOKUMENDIHALDUSSÜSTEEMI NÕUDEID KOGUDA AITAV KÜSIMUSTIK.....	73

Sissejuhatus

Inimesed ja organisatsioonid vajavad oma igapäevaseks toimimiseks infot. Info võib esineda mitmel viisil (nt paberil pastakaga kirjutatud, failis pildina, andmebaasis tabelitena jm) ning erinevate nimetuste all (nt kirje ehk rida andmebaasis, faks, protokoll). Käesolevas töös vaadeldakse infohulki, mida nimetatakse dokumentideks. Nendeks võivad olla paberitel või failides sisalduvad tekstid, pildid, joonised, aga ka audio- ja videoklipid lintidel või digitaliseerituna arvutis.

Dokumendid on seotud mingi kindla omanikuga (inimese või asutusega), kes neid loob ja kasutab. Käesolevas töös käsitletakse omanikuna organisatsiooni, seda piiritlemata mingisse liiki, see võib olla nii haridus-, kaubandus- või ehitusettevõtte vms. Töös viidatakse dokumentide omanikule erinevate sõnadega (asutus, ettevõtte, firma, organisatsioon).

Käesoleva töö eesmärgiks on anda eestikeelne ülevaade elektroonilisest dokumendihaldusest: süsteemidest, mille peamiseks ülesandeks on hallata dokumente ning analüüsida, kuidas on elektrooniline dokumendihaldus seotud teiste infotehnoloogia (IT) distsipliinidega. Elektroonilise dokumendihalduse materjalidega tutvudes leidis autor mitu ingliskeelset mõistet, mis on kuidagi seotud dokumendihaldusega: *Document Management*, *Record Management*, *Content Management*, *Knowledge Management*, *Output Management* jt. Käesoleva tööga püüabki autor lahti seletada nende tähenduse ja seose üksteisega.

Töös on esitatud näiteid Tartu Ülikooli dokumendihalduse praktikast, kus võeti elektrooniline dokumendihaldussüsteem kasutusse 2000 aastal.

Esimeses peatükis tutvustatakse dokumendihaldusega seotud mõisteid. Leitakse vastused küsimustele: mis on dokument, missugused tegevused on seotud dokumendiga vastavalt tema elutsüklile, kuidas neid liigitatakse ning millisel andmekandjail dokumendid asuvad. Andmekandjana vaadeldakse paberit, mikrograafikat ja elektroonilisi formaate.

Teine peatükk on pühendatud dokumendihaldusele ja -haldussüsteemidele. Esitatud põhimõtted kehtivad suvalisel andmekandjal asuvate dokumentide haldamise kohta, lähemalt tutvustakse ainult elektroonilisi dokumendihaldussüsteeme: missugused tegevused on neisse kaasatud ning kuidas uut süsteemi rakendada ja kasutada nii, et see oleks ettevõttele kasulik.

Kolmandas peatükis püütakse esitada dokumendihalduse kohta infotehnoloogias. Kuna infotehnoloogia maailm muutub kiiresti, siis on IT vahendeil rahvusvahelised standardid, mis aitaksid ühtlustada antud meetodi (tehnoloogia) kasutamist. Ka dokumendihalduse jaoks on loodud standardeid, neist mõned on tutvustatud käesolevas töös. Lisaks on kolmandas peatükis käsitletud tehnoloogiaid, mis on seotud dokumendihaldusega: mida kasutab dokumendihaldussüsteem või mis kasutab dokumendihaldust.

Tööl on kolm lisa. Esimeses võrreldakse paberkandjal põhinevat ja elektroonilist dokumendihaldust ning esitatakse mõningad faktid paberdokumentide kasutamise kohta. Teises lisa on toodud küsimustik, mille abil saab ettevõtte kindlaks teha, kas ta vajab elektroonilist dokumendihaldust ning kui vajab, siis kuidas edasi tegutseda. Kolmandas lisa toodud küsimustik aitab paika panna uue dokumendihaldussüsteemi nõuded.

Inglisekeelsete terminite tõlkimisel kasutas autor Arvutikasutaja Sõnastikku ([1]).

Kõik töös viidatud nimed ja nimetused, mille kohta on teada, et nad on kas registreeritud või registreerimata kaubamärgid, kuuluvad nende omanikele.

1. Informatsioon, dokument ja meedium

Iga organisatsioon kasutab oma töös infot. Kvaliteetsete otsuste tegemisel on vajalik aktuaalse info kiire kättesaamine. Aastatega on arenenud uued tehnoloogiad info töötlemiseks. Spetsialiseerudes mingile alale, nimetatakse info ümber, sellest saab kas dokument, teadmus (ik *knowledge*) või lihtsalt andmed. Käesolevas peatükis tutvustatakse, mis on dokument ning vaadatakse selle erinevaid andmekandjaid ehk meediume.

1.1 Mõisted

Kui Euroopa kultuuritaustaga inimesele öelda sõna „dokument“, siis seostuvad sellega tüüpiliselt lepingud, ärikirjad ja aktid. See tähendab, et Euroopas mõeldakse dokumendi all pigem juriidilist teksti, mis asub allkirjastatuna paberil. Seevastu näiteks USAs on dokumendiks suvaline tekstitöötlusprogrammiga kirjutatud tekst, mida näitab ka .doc laiend tekstifailide juures ([2]).

Nii nagu sõnast „dokument“ saadakse erinevates maades erinevalt aru, on ka dokumendiga seotud mõisteid ja definitsioone erinevaid. Inglise keeles on (vähemalt) kaks sõna, mida saab eesti keelde tuua „dokumendina“: „*document*“ ja „*record*“. On allikaid, mis peavad neid sõnu sünonüümideks ja on allikaid, mis eristavad neid. Käesolevas töös käsitletakse edaspidi mõlemat tüüpi dokumente („*record*“ ja „*document*“) üheskoos nimetuse „dokument“ all. Kui on erinevusi nende käsitluses, siis tuuakse see eraldi välja.

Järgnevalt on esitatud mõned dokumendi definitsioonidest.

„Document“:

1. Austraalia valitsus: dokument on salvestatud informatsioon, millel on äratuntav struktuur; mis asub suvalisel meediumil ning on arusaadav ilma edasise töötluseta väljaarvatud ekraanil esitamiseks või paberile trükkimiseks ([3]).

2. Levien, 1989: dokument on salvestatud informatsiooni ühik, mis on struktureeritud inimlikuks tarbimiseks ([4]).
3. Michalski, 1991: dokument on väljavõtte mingist informatsiooni-hulgast, mis ([4]):
 - ühendab mitu kompleksinfo tüüpi;
 - eksisteerib mitmes kohas üle arvutivõrgu;
 - sõltub teiste dokumentide infost;
 - muutub automaatselt (kui alamdokumendid muutuvad);
 - omab keerukat struktuuri või kompleksandmetüüpe nagu heli ja video;
 - on ligipääsetav ja muudetav mitme inimese poolt üheaegselt (kui neil on selleks õigused).
4. Eesti arhiiviseadus: dokument on mis tahes teabekandjale jäädvustatud teave, mis on loodud või saadud asutuse või isiku tegevuse käigus ning mille sisu, vorm ja struktuur on küllaldane faktide või tegevuse tõestamiseks ([5]).

„Record“:

1. Austraalia valitsus: dokument on ettevõtte või individuaalsete funktsioonide, tegevuste ja transaktsioonide salvestatud tõendus. Tõendiks oleval dokumendil peab olema sisu, kontekst ja struktuur ning peab olema seda haldava (säilitava) süsteemi osa ([3]).
2. Minnesota Riigiarhiiv: dokumendid on kaardid, kirjavahetus, plaadid, kaustad, memod, mikrofilmid, paberid, fotod, salvestised, raportid, linnid, kirjutised, optilised kettad, teised andmed, informatsioon või dokumentaalmaterjal; sõltumata füüsilisest vormist või iseloomust, säilitusmeediumist või kasutustingimustest, tehtud või saadud riigiametniku või -agentuuri või maa-, linna-, koolipiirkonna, munitsipaalasutuse või teise agentuuri poolt ([6]).

„Document“ võib olla „record“ ja „record“ võib olla „document“, aga üks-ühest vastavust nende vahel ei ole. Mõlema mõiste esimesest definitsioonist on näha, et „record“, millega tegelemiseks on vaja arvutitöötlust (va ekraanil näitamiseks ja trükkimiseks), ei ole „document“. Näiteks informatsioon, mis on relatsioonilises andmebaasis äritransaktsiooni tõenduseks, on „record“, aga mitte „document“.

Analoogiliselt leidub dokumente, mis ei ole „*record*“. Sõltuvalt organisatsiooni olemusest võib nendeks liigitada umbes 90% või rohkem dokumente. Näiteks puhtalt refereerivad dokumendid ei ole „*record*“, kui nad just pole viimatinimetatu osa. Informatsioon muutub dokumendiks „*record*“-mõttes siis, kui selle loonud äriprotsess (alamprotsess) on tähtis organisatsiooni eesmärkide suhtes.

Toodud dokumendi definitsioonidest saab välja lugeda, et neis peab olema selgelt arusaadavad järgmised kolm komponenti:

1. **sisu** – faktiline informatsioon;
2. **kontekst** – info, mis näitab, kuidas dokument on seotud organisatsiooni äritegevusega;
3. **struktuur** – dokumendi tehniline info (formaad, andmete paigutus lehel, päis, jalus jt).

Dokumendiks olemine ei sõltu sellest, millisel andmekandjal see loodi või säilitati. Dokumendiks võib olla nii E-kiri, töötaja avaldus paberil, veebi lehekülj kui ka fotol asuv kujutis. Käesoleva peatüki alajaotuses 1.3 tutvustatakse lähemalt dokumendi andmekandjaid ehk **meediumeid**.

Iga dokumendiga käivad kaasas **metaandmed** (ik *metadata*) ehk teisisõnu dokumendi atribuudid. Metaandmed kirjeldavad andmeid endid ([7]), nendeks on näiteks dokumendi

- identifikaator;
- loomise aeg;
- muutmise aeg;
- autor;
- pealkiri;
- võtmesõnad;
- adressaadid;
- jõustumise kuupäev;
- liik.

Metaandmete hulk ja liigitus on sõltuvuses organisatsioonist, kes ise määrab, millised metaandmed on vajalikud. Metaandmed on pidevas muutuses, täna kehtivatele

andmetele võib homme lisanduda uued (samuti olemasolevad kaduda). Näiteks soovib firma, et alates järgmisest aastast lisatakse info selle kohta, kuidas dokument firmasse saabus (faksina, E-kirjana või paberil). Atribuutide hulka piirab/laiendab ka dokumentidega tegelemiseks kasutatav tarkvara. Sellest lähemalt järgmises peatükis.

Elektroniline dokument. Minnesota riigiarhiiv ütleb, et elektroniline dokument (mõiste „*record*“) on dokument, mis on loodud, genereeritud, saadetud, saadud või säilitatud elektroniliselt ([6]). See tähendab, et kui näiteks töötaja vormistab avalduse arvutis tekstitöötlusprogrammi kasutades ja esitab selle siis paberkujul, siis võib tekstitöötlusprogrammis asuvat teksti nimetada elektroniliseks dokumendiks „*record*“ mõttes. Kui aga asutus töötleb saadud paberkujul avaldust edasi paberkujul, siis seda ei saa nimetada elektroniliseks dokumendiks „*document*“ mõttes. Jüri Vain ja Juhan-Peep Ernits kirjeldavad elektronilist dokumenti tema sisu ja formaadi kaudu ([8]). Sisu on kriitiline informatsioon, mis pakub soovitud teadmist (suund teadmuse haldus, inglise keeles *Knowledge Management*). Formaat on meetod, mille kaudu informatsioon on säilitatud ja esitatud. Elektronilised dokumendid võivad sisaldada nii tekstifaile, tabelitöötluse faile, esitlusi, skaneeritud kujutisi, fakse, video ja audio infot.

Igas ettevõttes on erinev hulk dokumente. Näiteks Tartu Ülikoolis luuakse iga päev umbes 140 ametlikku dokumenti (lisanduvad veel dokumendid „*record*“ mõttes jm), millega kaasneb vajadus nende haldamiseks. **Dokumentide haldamiseks** nimetatakse administratiivset tegevust, mille läbi organisatsioon süstemaatiliselt loob, organiseerib, säilitab, taastab, kasutab, kaitseb ja hävitab temale kuuluvaid dokumente ([9]). Dokumentide haldamise läbi saab organisatsioon kontrollida tema poolt loodud informatsiooni kvaliteeti ja kvantiteeti ning säilitada informatsiooni kõige ratsionaalsemal kujul. Samuti saab organisatsioon hävitada vananenud või enam mitte tarvitavat informatsiooni ja seeläbi arendada majanduslikku efektiivsust ja tulemuslikkust. Kõige selle tarvis peab organisatsioon olema võimeline kindlustama dokumentide säilitamise, garanteerima nende kättesaadavuse ja hävitama mittevajalikud dokumendid nõuetekohaselt.

Dokumendid on tihedalt seotud organisatsiooni töövooga (ik *workflow*). **Töövoog** on äriprotsesside osaline või täielik automatiseerimine, mille käigus dokumente,

informatsiooni või ülesandeid edastatakse ühelt osaliselt (inimene, masin, protsess) teisele, kusjuures kogu tegevus on kindlalt seotud protseduuriliste reeglitega ([10]). Lühidalt öeldes töövoos haldussüsteemid täidavad äriprotsesse. Tavalised töövoospetsifikatsioonid sisaldavad teostatavaid tegevusi, kontrole ja andmete liikumist nende tegevuste vahel; agente, millel on lubatud teostada tegevusi ning organisatsiooni keskkonda kirjeldavaid poliitikaid. Näiteks Tartu Ülikoolis peab töötaja puhkuseavaldus liikuma esmalt tema vahetule töökorraldajale, kes annab oma nõusoleku; seejärel teaduskonna dekaani (kui tegu on teaduskonna töötajaga) juurde, kes omakorda kinnitab avalduse. Alles pärast kõikide osapoolte nõusolekut liigub avaldus personaliosakonda, kus seda edasi töödeldakse. Töövoos haldussüsteemi ülesandeks oleks avalduse saatmine vajalikele osapooltele õiges järjekorras ning osapoolte tegevuse kontroll (nt teatud aja jooksul nõusoleku andmine).

1.2 Dokumentide liigitamine

Ühes organisatsioonis on enamasti palju dokumente. Need jagunevad ([3]):

- **korporatiivsed dokumendid:** organisatsiooni dokumendid, mis on organisatsiooni poolt kasutatavad ja millel on ajaliselt jätkuv väärtus. Siia kuuluvad nii õigusaktid kui ka revisjonidokumendid; samuti dokumendid väljastpoolt organisatsiooni, kui need on firma dokumendi osaks. Need dokumendid peavad olema ligipääsetavad (sõltuvalt juurdepääsuõigustele) neile, kes neid vajavad.
- **töös olevad dokumendid:** kas korporatiivsete dokumentide koopiad, mida arendatakse või dokumendid, mis sobiva kvaliteeditaseme saavutades registreeritakse korporatiivseteks dokumentideks.
- **isiklikud dokumendid:** individuaalseks kasutamiseks, näiteks märkmed. Neid võib kasutada töös olevate dokumentide arendamisel, aga pole ettenähtud ise saamaks korporatiivseks dokumendiks.

Kuna korporatiivsed dokumendid peavad olema kättesaadavad, siis peavad need ühtlasi olema kergesti leitavad kasutajatele. Otsimise lihtsustamiseks kasutatakse dokumentide liigitamist gruppidesse mingi omaduse järgi.

Dokumentide liigitus peaks olema ([9]):

- **loogiline**: üles ehitatud lihtsal ja kergesti mõistetaval kujul.
- **praktiline**: abistama ametkonna personali dokumentide registreerimise ja otsimise tegevuses ning olema üles ehitatud igapäevasel terminoloogial ja väljenditel.
- **lihtne**: lihtsalt kasutatav ja jätma vähe ruumi interpretatsioonile.
- **funktsionaalne**: lähtuv ametkonna tööülesannetest, sõltumata ametkonna füüsilisest asukohast.
- **mobiilne**: lähtuv säilituskavast ja võimaldama mitteaktiivsete ning lõppenud säilitusperioodiga dokumentide teisaldamist.
- **paindlik**: võimaldama struktuurimuutusi ja andmete mahu kasvu.
- **standardiseeritud**: vastavuses kirjutatud korrale, kus kasutatakse standardset terminoloogiat.

Liigitamine mitte ainult ei lihtsusta otsingut, vaid aitab omakorda asutusel käsitleda dokumendigruppe tervikutena, lähtudes samadest säilitus-, teisaldus- ja hävituseeskirjadest. Igale dokumendiliigile saab eraldi määratleda, kes selle õigsuse eest vastutab, kellele on lubatud juurdepääs, kus ja kui kaua säilitatakse jm.

Näiteks töötajate isikutoimikuid tuleb säilitada 50 aastat, aga personalialast aruandlust vaid 5 aastat. Ootuspäraselt kuuluvad need eraldi kategooriatesse.

Veel ühe näitena on tabelis 1.1 toodud Tartu Ülikooli dokumentide jaotus funktsioonipõhiselt ([11]). Toodud loetelu on aluseks dokumentide täpsemaks liigituseks, mille näiteid on toodud tabelis 1.2 ([11]).

Funktsioon	Allfunktsioon
Juhtimine	
Õppetöö	
	Õppetöö kolledžites ja teaduskondades
Teadus- ja arendustöö	
Personalitöö	
	Töötervishoid- ja tööohutus

Finantsarvestus	
	Finantsplaneerimine
	Raamatupidamine
Haldus- ja majandustegevus	
	Varade haldamine
	Hangete korraldamine
Info haldamine	
	Asjaajamine ja arhiividus
	Avalikkussuhted
	Infohaldus
	Raamatukogu- ja muuseumitöö
Täienduskoolitus	
Välissuhted	

Tabel 1.1: Tartu Ülikooli dokumentide jaotus funktsioonipõhiselt.

Dokumendi liik	Funktsioon	Säilitus- tähtaeg	Juurdepääsu- piirang
Ülikooli nõukogu otsused	Juhtimine	Alatine	
Rektori käskkirjad	Juhtimine	Alatine	
Üliõpilaste, internide, residentide, magistrantide ja doktorantide isikutoimikud	Õppetöö	50 a.	Asutusesisene, sisaldab konfidentsiaalset infot
Kirjavahetus ja dokumendid õppetöö küsimustes	Õppetöö	5 a	
Personalikorraldused	Personalitöö	7 a	

Tabel 1.2: Tartu Ülikooli dokumentide liigituse näited.

Tabelis 1.2 on esitatud vaid üksikud näited dokumentide liigitusest Tartu Ülikoolis, kuna seal on erinevaid dokumendi liike kokku umbes 170. Esitatud liigitus kehtib

Tartu Ülikooli 2002. aastast ja see annab ühtse aluse dokumentide identifitseerimiseks (liik + number) üle ülikooli. Enne seda oli dokumentide identifitseerimine rohkem struktuuripõhine, iga allüksus koostas ise oma dokumentide liigituse alused. Näiteks nii „Matemaatika-informaatikateaduskonna dekaani korralduste“ kui ka „Administratsioonidirektori korralduste“ alla kuuluvad personalikorraldused on nüüd sama liigi alla kuuluvad (varem erinevad).

1.3 Erinevad dokumendi meediumid

Dokumendi definitsioonidest on näha, et dokumendi olemus ei sõltu kasutatavast andmekandjast. Dokument saab olla paberil, elektrooniliselt nii failides kui ka andmebaasides. Millal millist meediumit on otstarbekas kasutada? Selleks, et sellele küsimusele vastata, tutvustatakse käesolevas alajaotuses meediumeid ning järgmises dokumendi elutsükli. Aluseks on võetud materjalid allikast [12].

1.3.1 Paber

Paber on siiani kõige tähtsam meedium dokumendi ajaloos. Kuigi igapäevaselt elektroonilise meediumi kasutamine suureneb, on suurem osa dokumentidest siiski ainult paberil ning kui dokument luuakse elektroonselt, siis edasiseks jagamiseks ja säilitamiseks trükitakse see enamasti siiski paberile.

Paberil on teiste meediumite ees rida eeliseid. Kuna paberil olev info on inimesele kergesti loetav (nii inim-keelne kui ka silmale arusaadav), siis ei ole vaja dokumendi kasutamiseks (viitamiseks) spetsiaalset arvuti tarkvara. Peale selle on paber tuntud ja mugav kasutada; saadaval eri värvides, materjalides ja suurustes. Paberdokumentidele saab paigutada nii tekstilist kui graafilist sisu, sellele saab lisada nii annotatsioone kui allkirju ilma suurema vaevata (tehniliste vahendite sekkumiseta).

Vaatamata kõigile eelpool toodud eelistele on paberil ka omad piirangud:

- paberdokumendid on kergesti kuluvad ja purunevad nende pideval kasutusel;
- suurtes kogustes paberi kasutamine vajab suuri ruume (sobiva temperatuuri, valguse ja niiskusega);

- paberdokumentide indekseerimine, paigutamine ja otsimine on aeganõudev ja töömahukas;
- kui üks dokument on mitme teise dokumendi osa, siis viitamise asemel kasutatakse koopiaid, mis suurendab paberite hulka;
- dokumenti saab kasutada üldjuhul ainult üks inimene korraga, kui just ei kasutata koopiaid.

Paberi kasutamine dokumendi andmekandjana jääb püsima veel pikaks ajaks. Selle väljatõrjumiseks tuleb lahendada mitu inimlikku (nt inimesed on paberiga harjunud) ja tehnoloogilist faktorit (nt kinno on vaja vahendeid mitte paberil oleva pileti müümiseks ja kontrollimiseks).

1.3.2 Mikrograafika

Mikrograafika hõlmab endas mikrofilme, mikrofišše (ik *mikrofišche*), apertuurkaarte (ik *aperture cards*) või teisi mikrovorme, mis on saadud lähtedokumentkaamerate (ik *source document cameras*) või COM (ik *computer-output microfilm*) salvestajate abil. Mikrofilm on dokumendi vähendatud kujutis filmil. Kui see on filmirull, siis on tegemist mikrofilmiga; kui mitu pilti asub indekskaardi suurusel lehel, siis on tegemist mikrofiššiga. Apertuurkaardid on segu vanadest perfokaartidest ja mikrofilmist. Väljalõigatud apertuurala kaardil sisaldab dokumendi 35mm mikrofilmi pilti, tüüpiliselt mingi objekti plaani või teisi tehnilisi jooniseid. Info joonise kohta (pealkiri, versioon, lehekülj jne) on augustatud kaarti (nn Hollerithi koodis) ja on trükitud piki ülaseri. Apertuurkaardi skänner loeb augustatud infot ja skaneerib mikrofilmi ekraanile. Tulemuseks on digitaalne pilt, mis on sarnane paberskänneri tulemusele. Augustatud tekstilist infot kasutatakse skaneeritud kujutise automaatseks indekseerimiseks, lubades tuhandete seosetute kaartide sisestust korraga.

Mikrograafika on dokumendihalduses laialt tunnustatud. Spetsiaalsed mikrograafika tooted on loodud rakendustele, milles asuvaid dokumente on tihti vaja kätte saada (otsimine). Näiteks ise-keerlevad kassetid (ik *self-threading cartridges*) lihtsustavad 16mm mikrofilmi laadimist ja vaatamist; apertuurkaardid hõlbustavad insenerjooniste, kaartide ja teiste suurte dokumentidega tegelemist, reprodutseerimist ja jagamist.

Võrreldes paberil olevate dokumentide väljatrukiga, COM tehnoloogia kiirendab ja lihtsustab nt arvete jt arvuti poolt loodud dokumentide tootmist ja jagamist.

Definitsiooni kohaselt mikrograafika muudab info "miniatuurseks". See pakub dokumentidele kompaktsust, ökonoomilist ja usaldatavat ladustamise võimalust. Võrreldes paberiga, võib mikrograafika vähendada laopinda 95% või rohkemgi.

Mikrograafika kujutised on nagu paberdokumentidki inimesele arusaadavad, need vajavad vaatamiseks või trükkimiseks ainult kujutise suurendamist. Võrreldes elektroonilises formaadis olevate dokumentidega on mikrograafika vähem sõltuv riist- ja tarkvarast.

1.3.3 Elektroonilised formaadid

Elektrooniline formaat hõlmab endas nii digitaliseeritud kujutisi (loodud skännerite abil), tekstipõhiseid (ik *character-coded*) andmeid (loodud tekstitöötlusprogrammide, E-kirja süsteemide kui ka teiste arvutisüsteemide poolt) kui ka tabelitöötluse andmeid jm.

Elektroonilised formaadid on parimad rakendustele, mis vajavad dokumentide kiiret kättesaamist (nii kohalikku kui läbi võrgu) või dokumentide kontrollitud teekonda läbi organisatsiooni töövoos. Elektroonilisi dokumente on lihtne kuvada, printida ning saata edasi teistele kasutajatele. Kuna kasutajad ei saa endale dokumendi ainuversiooni, siis on võimalik dokumenti samaaegselt kasutada mitmel erineval kasutajal, nii et kasutajad ei pea dokumendi pärast võistlema.

Võrreldes paberdokumentide süsteemiga on faili ülalpidamise ja integreerimise probleemid minimaalsed. Kuna elektroonilisi dokumente kasutamisel ei eemaldata hoiukohast, siis elektroonilises süsteemis puuduvad sorteerimise ja tagasipaigutuse ülesanded (nagu paberil). Peale eelneva on elektroonilisi dokumente raskem (peaaegu võimatu) kahjustada kasutamisel või paigutada valesse kohta. Kui individuaalne säilitusmeedium kahjustub või varastatakse, siis on olemas tagavarakoopiaid dokumendi jätkuvaks kasutamiseks. Miinuseks võrreldes pabersüsteemiga kaasneb elektroonilise meediumiga suurem oht andmete vargusele: lisaks ukse

mahamurdmisele on võimalik arvutivõrgus olevatele andmetele ligi pääseda ka arvutivõrku kasutades.

Kui paberi ja mikrofilmi jaoks on defineeritud stabiilsust iseloomustav tunnus, eluiga ja optimaalsed säilitustingimused, siis arvutimeediumil sellised rahvusvahelised standardid puuduvad. *National Archives and Records Administration* on välja pakkunud oma regulatsioonis 36 CFR osa 1234 magnetketaste hoidmise temperatuuriks 62-68 F ja niiskuseks 35-45% ([13]). Püüdes leida elektrooniliste meediumite vastavaid iseloomustusi, kasutatakse kiirendatud eluea teste. **Kiirendatud eluea testid** on loodud reaalse töötingimuste reprodutseerimiseks seadmetele, need testid lubavad analüüsida seadmete halvenemist pika perioodi vältel ilma, et peaks tulemusi ootama mitu aastat ([14]). See saavutatakse eseme kontrollitud kokkupuutumise kahju tekitavate faktoritega: kõrgendatud temperatuur, muutuv niiskus jms. Baseerudes kiirendatud eluea testidele, pakuvad paljud tootjad oma optilisele kettale stabiilset teoreetilist eluiga kuni 100 aastat. Siiski on paljude optiliste ja magnetketaste praktiline eluiga 30 aastat või veelgi vähem. Probleemiks ei ole niivõrd tehniline vananemine kui moraalne (seadmed meediumi kasutamiseks puuduvad, ilmuvad uued seadmed uute meediumite jaoks). See määrab piirangud dokumentide säilitamisele elektroonilisel meediumil.

Kauaaegset säilitamist takistab ka olukord, et elektroonilisi dokumente ja nendega seotud indeksandmeid saab kasutada ainult spetsiaalse rakendusprogrammi abil. Tarkvara uueneb ja muutub veelgi kiiremini kui riistvara. Tarkvara võib muutuda isegi nii palju, et eelnevalt salvestatud materjal pole enam kättesaadav. Seda eriti juhul, kui rakenduse tootjat enam pole (ei tegele tootega, pankrotis vms). Elektroonilise dokumendi kättesaadavust saab pikendada, kui aeg-ajalt salvestada info uuemale meediumile või formaati. Selleks on loodud spetsiaalsed konverteerimissüsteemid. Võib öelda, et täna installeeritud süsteem kas asendatakse või on uuendatud versioonis juba 5-6 aasta pärast. Tuleb aga arvestada, et mida aeg edasi, seda rohkem on infot ja seda rohkem on tõenäoliselt erinevaid formaate, mida konverteerida.

Elektroonilistel formaatidel on ka järgmised puudused ([15]):

- vastuolud erinevate operatsioonisüsteemide vahel;

- erinevate tarkvarafirmade poolt loodud failiformaatide mitteühildumine;
- faili konverteerimisel ühest formaadist teise kannatab dokumendi kujundus;
- erinevad fondid: nende puudumisel dokumendist arusaamine muutub võimatuks;
- eelneva punktiga sarnane probleem: rahvuslike sümbolite puudumine (näiteks „õ“, „ž“).

Elektroonilised dokumendid ei oma sellist füüsilist vormi, mis on paberil. Kui uuenevad riistvara ja tarkvara, siis dokumentide „rännakul“ ühelt platvormilt teise toimub nendes muutus. Samuti tagavarakoopia tegemine ja sellelt taastamine tekitab koopia dokumendist, mis võib erineda esialgsest. Seega kerkib esile küsimus dokumendi autentsusest. Kuidas olla kindel, et saadud dokument on originaaldokumendi õige (täpne) esitus? Lahenduseks oleks versioonikontroll ja logifaili pidamine, milles on kirjas kõik toimunud muutused.

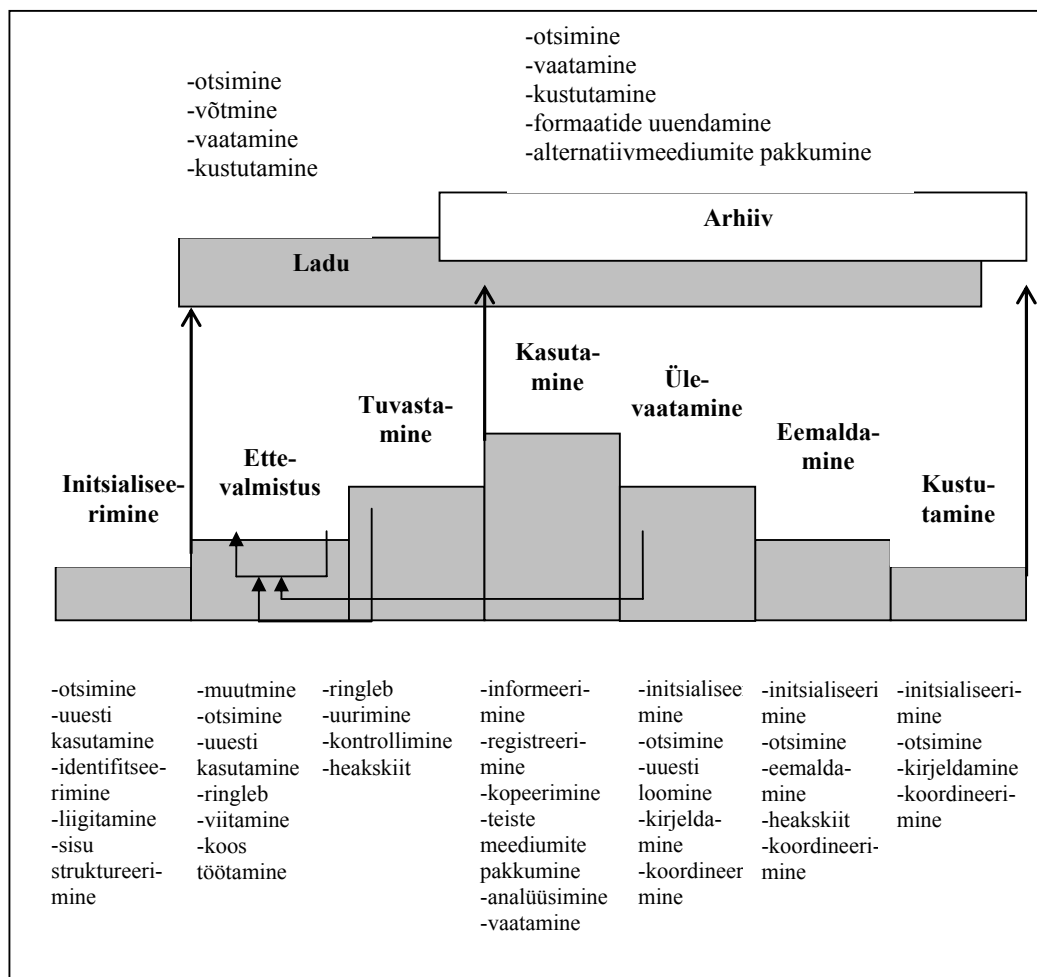
1.4 Dokumendi elutsükkel

Igal dokumendil on oma elutsükkel (ik *life-cycle*), mis väljendub dokumendi loomises, kasutamises, säilitamises ja hävitamises. Üldiselt jaguneb tsükkel järgmiselt:

- **initsialiseerimine** - dokument luuakse esimest korda kas paberil või elektrooniliselt. Nimetatud tegevus toimub ka siis, kui dokument saabub organisatsiooni väljastpoolt.
- **töötlemine** - dokumenti liigub organisatsiooni töövoos, kus seda vajadusel täiendatakse-parandatakse.
- **kasutamine** - dokumenti kasutatakse organisatsiooni töös, sellele viidatakse vajadusel, aga enam ei muudeta.
- **säilitamine** - dokument paigutatakse arhiivi.
- **hävitamine** - dokument hävitatakse jäädavalt.

Joonisel 1.1 on toodud tegevused, millega on seotud dokumendi elutsükli osad ([7]). Initsialiseerimisel saab dokument identiteedi ja ta klassifitseeritakse (liigitatakse).

Ettevalmistamise (nn töötlemise) osas dokumenti arendatakse kuni valmimiseni. Tuvastamise (ik *establishment*) osas dokument vaadatakse üle ning kas antakse nõusolek valmisoleku (sobivuse) kohta või lükatakse tagasi ettevalmistuse etappi. Kasutamise osas dokumenti jagatakse, kopeeritakse, analüüsitakse jne. Ülevaatamise osas dokument nii-öelda avastatakse uuesti ja seda kasutatakse uue dokumendi põhjana. Pärast teatud aja möödudes dokumenti enam ei kasutata ja see eemaldatakse (arhiveeritakse). Viimaseks etapiks on dokumendi kustutamine.



Joonis 1.1. Dokumendi elutsükel.

Dokumendi elutsükel on jagatud dokumendi kasutamise aktiivsuse järgi kahte suurde faasi: aktiivne ja mitteaktiivne faas ([12]).

Aktiivne faas on seotud info õigeaegse kättesaamisega, et toetada organisatsiooni äriprotsesse. Selles faasis peab säilitusmeedium lubama sobivat (mugavat) dokumendi

kättesaamist selle muutmiseks või jagamiseks. Aktiivse faasi pikkuseks saab lugeda tavaliselt päevi, kõige rohkem nädalaid või kuid.

Mitteaktiivne faas on seotud dokumendi säilitamisega. Sel ajal peab säilitusmeedium pakkuma dokumendi odavat ja turvalist hoidmist, tihti pikemaks ajaperioodiks. Mitteaktiivne faas kestab tavaliselt aastaid.

Eelpool toodud kahe faasi vahele saab paigutada **semi-aktiivse faasi**: dokumentidele viidatakse sagedasti, aga mitte iga päev. Sellised dokumendid võidakse paigutada mõnesse odavamasse meediumisse seni, kuni neid enam ei vajata.

Erinevad allikad pakuvad erinevaid definitsioone ka mitteaktiivse faasi jaoks, näiteks ([9]):

- **mitteaktiivne faas**: dokumentide aktiivne ja semi-aktiivne faas on möödunud ja nad ei täida enam seda funktsiooni, milleks nad olid loodud;
- **teisaldamise faas**: dokumendid kas säilitatakse igapäevaseks kasutamiseks, arhiveeritakse või siis hävitatakse.

Dokumendi liikumine tema elutsükli faaside vahel on tihedalt seotud tema liigiga (alajaotus 1.2), millega sisuliselt ettevõtte määrab, millal saab dokumenti lugeda lõplikuks st ei saa enam muuta (va uueks dokumendiks) või millal kuulub dokument hävitamisele. Teiselt poolt on väga raske tõmmata joont aktiivse ja semi-aktiivse faasi vahele. Probleem tekib küsimuses, missuguseid dokumente kasutatakse rohkem, missuguseid vähem ning kui kaua neid vaja hoida nn semi-aktiivse faasi hoidlas? Autori arvates selline faas tavaliselt puudub ning piirduakse ainult aktiivse faasiga ja arhiiviga. Näiteks Tartu Ülikoolis on paberdokumendid toimikutes aktiivses faasis kuni aasta st kalendriaasta lõppedes viiakse need arhiivi (eraldi ruumi). Elektroonilised dokumendid on kõik ühes süsteemis, ainult vanemad dokumendid on paigutatud erinevasse kausta (nimetus “arhiiv”).

1.5 Millal milline meedium?

Uued tehnoloogiad (mikrograafika, elektroonilised formaadid) pole siiani suutnud täielikult asendada vanemaid dokumendihalduse metodoloogiasid (pabersüsteem). Erinevad meediumid täiendavad teineteist, aga ükski neist ei suuda veel pakkuda täielikku parimat lahendust kõikidele dokumendi elutsükli faasidele.

Paber on tuttav, sobiv vahend viitamiseks, samas vajab suurt pinda säilitamiseks ning nende süsteeme on raske organiseerida ja ülal pidada. Elektroonilised formaadid on parimad dokumendi elutsükli aktiivses faasis, kuna arvutil baseeruv indekseerimine lubab teha kiirelt keerulisi päringuid ning on võimalus dokumente integreerida teiste rakendustega; samas on elektroonilised formaadid vähesobivad info säilitamiseks läbi mitmete aastate (dokumendi mitteaktiivne faas). Mikrograafika tehnoloogia pakub oma stabiilsuse ja kompaktsusega head säilitusvahendit ja on sobiv ka aktiivses faasis, aga kasutajad eelistavad siiski täielikult arvutiseeritud süsteemi. Põhjuseks on ligipääsu kiirus ning elektrooniliste rakenduse võimalus pakkuda dokumendi kontrollitud liikumist organisatsiooni töövoos.

Ameerika organisatsioonid *Uniform Photographic Copies of Business and Public Records as Evidence Act* jt on väljendanud seisukohta, et mikrofilm on dokumendi reprodutseerimiseks sobiv meedium ([12]). Mikrofilm on eelistatuim meedium dokumentide kaitsmisel (tagavarakoopia loomiseks), kuna on kergelt ja ökonoomselt dubleeritav.

Suhteliselt levinud on dokumendihaldussüsteemid, kus on kombineeritud nii paberi kui ka mikrograafika kasutamine. Dokumendid on aktiivses faasis paberil ja hiljem säilitamiseks paigutatud mikrofilmile. Analoogselt võib kokku panna elektroonilise meediumi ja mikrograafika. Sellisel juhul tuleb paberdokumentidest vabanemiseks (suure tõenäosusega igas firmas) kasutada skännereid, mikrofilmimist või teisi tehnoloogiaid.

Kindlasti on sobiva meediumi valimisel tähtis ka selle hind. Just mitteaktiivses faasis, kus dokumente tuleb säilitada läbi mitmete aastate, on vaja leida odav, aga siiski

kvaliteetne säilitusmeedium. Dokumendi aktiivses faasis on tähtsaim omadus meediumile juurdepääsu kiirus. Järgnevalt on toodud tabel 1.3, kus on näidatud erinevate meediumite hierarhia, järjestatuna meediumile ligipääsu kiiruse ja hinna võrdluse järgi (aluseks võetud [16]).

Meedium	Ligipääsu kiirus	Hind
Arvuti põhimälu	Väga kiire	Väga kallis
Vahemälu (ik <i>cache</i>)	Väga kiire	Kallis
DASD	Kiire	Keskpärase/mõõdukas
DVD	Mitte aeglane	Kallis
Optiline ketas	Mitte aeglane	Mitte kallis
CD-ROM	Mitte aeglane	Odav
Mikrofišš	Aeglane	Odav
Paber	Väga aeglane	Odav / Mitte kallis

Tabel 1.3 Meediumite võrdlus.

Tabelist 1.3 on näha, et aktiivses faasis on parimad mälu ja DASD, mitteaktiivse faasi jaoks optiline ketas, CD-ROM või mikrofišš. DASD (ik *Direct Access Storage Device*) on meedium, mis võimaldab mäluaadressi (kus info sellel asub) ette antuna kiiresti leida soovitud andmed kõiki teisi samal meediumil asuvaid andmeid üle vaatamata. See oli suur edasimineku võrreldes magnetketaste kasutamisega, kuna magnetkettal tuli soovitud andmeteni jõudes kõik eelnevad andmed üle vaadata. Tabelis toodud kriteeriumite järgi on ka paber üks potentsiaalne säilitusmeedium, aga siis lisandub laopind ja muud tingimused, mis on toodud alajaotuses 1.3.1, mille järgi paber ei mahu parimate säilitusmeediumite hulka.

Lisas 1 on toodud paberil põhineva ja elektroonilise dokumendihalduse kasutamise võrdlus.

2. Dokumendihaldus ja -süsteemid

Dokumendihaldust kasutatakse igal pool, kus dokumente luuakse, töödeldakse, jagatakse, hoitakse. Käesolevas peatükis on toodud dokumendihalduse eesmärgid ja tegevused, millega üks dokumendihaldusprogramm peaks kindlasti tegelema. Dokumendihaldusega seotud standardeid ja uuemaid tehnoloogiaid tutvustatakse järgmises peatükis.

2.1 Sissejuhatus dokumendihaldusse

Järgnevalt tutvustatakse dokumendihalduse arengut, eesmäärke ja põhimõtteid.

2.1.1 Dokumendihalduse areng

Elektrooniline dokumendihaldus sai alguse 1980ndatel aastatel, mil tollal tavalised infotehnoloogilised süsteemid ei toetanud dokumentide graafilist kujutamist, töövoogu ja digitaalset optilist säilitamist ([2]). Esialgu moodustas dokumendihaldus omaette nurgakesse infotehnoloogias, nüüdseks on dokumendihaldus kaasatud mitmesse IT lahendusse (näiteks veebihaldus, projektitöö jt).

Algselt tähendas elektrooniline dokumendihaldus rakendust, mis asendas hierarhilist failide/kataloogide süsteemi ning pakkus lihtsat dokumentide lukustamist, versiooni- ning dokumendi atribuutide haldust. Neid programme nimetatakse nüüd „traditsiooniliseks“ dokumendihaldussüsteemideks. Mida aeg edasi, seda rohkem võimalusi lisati süsteemi, näiteks grupitöö ja töövoogude toetamine ning väljundi haldamine. Nii muutus „traditsiooniline“ dokumendihaldus praeguses mõistes „dokumendihalduseks kitsamas tähenduses“, mis sisaldab:

- traditsioonilist dokumendihaldust;
- dokumentide graafilist kujutamist (ik *imaging*);
- elektroonilist arhiivindust ja „*record*“-dokumentide haldust;
- E-vorme;
- väljundi haldust;
- skaneerimist;
- grupitöö ja töövoogude toetamist;

- Interneti ja Intraneti võimaluste kasutamist;
- jm.

Areng pole loomulikult peatunud, praegu saab dokumendihaldusesse lisada järgmised teemad:

- multimeedia andmebaasid;
- dokumentide andmelaod;
- teadmuse haldus (ik *Knowledge Management*);
- jpm.

Tänapäeval pole enam puhast dokumendihaldust, see tehnoloogia seguneb järjest enam teiste IT tehnoloogiatega. Dokumendihalduse funktsionaalsus integreeritakse operatsioonisüsteemidesse või nende lisateenustesse; andmebaaside juhtimissüsteemid pakuvad järjest enam vahendeid dokumentide haldamiseks ja nende otsimiseks. E-kaubandus toob kaasa kommertstarkvara, milles sisalduvad nii kontori- kui ka grupitöö vahendid.

2.1.2 Dokumendihalduse vajadus

Dokumente leidub igas ettevõttes ja seda mitte ainult paberkujul kuskil kapis tolmu kogumas: avaldused, protokollid, aruanded, videod firma omaniku kõnedest, isegi fotod sünnipäevadest. Mida vanem organisatsioon, seda rohkem dokumente ning vajadus nende haldamiseks kasvab. Kui iga töötaja vastutab ise oma dokumentide ja elektronposti eest, võib juhtuda, et koos inimeste töölt lahkumisega lahkub ka firma „mälu“ - kogemused, sidemed, väljakujunenud struktuur, stiil jms ([17]). Selle vältimiseks tuleb firmas rakendada dokumendihalduse süsteem ning jälgida, et selle reeglitest ka kinni peetaks. Dokumendihaldus võimaldab organisatsioonil ([9]):

- toetada ja põhjendada programmide vastuvõtmist ja otsuste formeerimist;
- olla vastavuses seaduste ja teiste õigusaktidega kehtestatud nõuetele;
- kaitsta enda ja samuti oma töötajate ning klientide huvisid;
- toetada ja arendada organisatsiooni töö tulemuslikkust;

- kaitsta ennast kohtuvaidluste ja uurimiste korral ning teistel sarnastel juhtudel tänu võimalusele pakkuda täpseid andmeid organisatoorsete toimetuste kohta;
- toetada ja pakkuda tõendusmaterjali organisatoorsetes küsimustes ja otsustamisprotsessis vastu võetud otsuste tulenemise kohta;
- dokumenteerida oma toimetusi, otsuseid ja saavutusi;
- juurdepääsu täpsele ja terviklikule informatsioonile ning seeläbi parandada oma majandusliku tegevuse ratsionaalsust ja tulemuslikkust;
- vältida ebaotstarbekaid kulusid dokumentide säilitamisele;
- kindlustada protsesside jätkuvus;
- toetada uurimisprogramme.

Elektroniilise dokumendihalduse suurim väärtus on tema võime laiendada infohalduse tegevusala faktidest, mis on andmebaasides andmete kujul, põhimõtetele ja ideedele, mida üldiselt hoitakse ja säilitatakse dokumentide kujul ([4]). Dokumendihaldus aitab parandada dokumentide efektiivsust ja tõhusust nende rollis kui peamine vahend põhimõtete ja ideede salvestamiseks ning suhtlemiseks organisatsiooni (ja nende gruppide ja indiviidide) siseselt ja –vaheliselt ([4]). See on tähtis firmadele, kus töö käib gruppides või mis on geograafiliselt hajutatud ning kes on rohkem sõltuvad oma oskustest hallata mitmesugust infot. Dokumendihaldusel on ka teisi rolle ja eesmärges ettevõtte elus, need on esitatud tabelis 2.1 ([4]).

Roll	Näited
Salvestada või dokumenteerida lepingud ja kokkulepped	Töölepingud, hoolduslepingud, laenud, liisingud, ostulepingud jne.
Salvestada poliisid, standardid ja protseduurid	Protseduuri käsiraamatud, standardi spetsifikatsioonid, juhendid, täidesaadetavad memod jne.
Esitada reaalsuse vaade teataval ajahetkel	Staatuse aruanded, probleemi analüüsid, strateegilised plaanid, eelarved jne.
Luaa pilt või kujutis	Igaastased aruanded, turu-uuringud, TV ja raadio reklaamid jne.
Luaa tuluartikkel kui toode	Konsultatsioonifirma aruanne, mis müüakse tema kliendile jne.

Käituda kui infovahetuse mehhanism inimeste ja gruppide vahel	Memod, kirjad, esitlused, E-kirjad, teated jne.
Käituda kui vahend organisatsiooni protsesside jaoks	Tellimused, saatelehed, nõusoleku kirjad, enamus ärivorme, jne.
Pakkuda korda põhimõtete ja ideede kinnitamiseks ja liigendamiseks	Peaaegu kõik dokumendid, mis kannavad põhimõtteid ja ideid

Tabel 2.1 Dokumendihalduse rollid ettevõttes.

Küsimusele, kas konkreetne firma vajab dokumendihaldust, aitab vastuse leida küsimustik lisas 2.

2.1.3 Dokumendihalduse põhimõtted

Järgnevad põhimõtted kehtivad dokumentidele suvalisel meediumil ning need baseeruvad informatsiooni haldamise põhimõtetel (arendatud Parer'i ja Parrott'i poolt 1994. a.; [3]):

1. hallatud peab olema kogu dokumendi elutsükkel;
2. dokumendid peavad olema identifitseeritud;
3. dokumentides asuv info kvaliteet peab olema tagatud;
4. dokumendid peavad olema kaitstud;
5. dokumentidele peab olema pakutud sobiv juurdepääs;
6. dokumente peab olema võimalik säilitada.

Esimese ja viimase punkti osas eristatakse dokumendi- ja „*record*“-dokumendi haldust. Nimelt „traditsiooniline“ dokumendihaldus tegeleb ainult valmisdokumentidega st ei hallata kogu dokumendi elutsükli (loomine ja säilitamine puudub). Niipea kui süsteem tegeleb ka dokumendi loomisega (säilitamisega), on süsteem „*record*“-dokumendi haldus (arhiivindus). Aga nagu eespool on öeldud, püüab tänapäeval dokumendihaldus tegeleda kõigega, mis on seotud dokumentidega.

Elektrooniline dokumendihaldus peab täitma kõiki tingimusi, mida oodatakse suvalisel meediumil oleval dokumendihalduselt, nendeks tingimusteks on ([6], [8])

- **usaldatavus.** Dokumentides asuv info peab olema usaldusväärne ja autentne. Võtmesõnaks on seaduslik vastuvõtlikkus, st kas elektroonilisi andmeid aktsepteeritakse kohtus kui tõendeid.
- **Täielikkus.** Dokumentides peab asuma kogu info, mis kindlustaks nende kaua-aegse kasulikkuse. Ärikeskkonnas on dokumendid seotud näiteks transaktsioonidega, koosoleku protokolliga vms, mis tähendab, et dokumendiga kaasneb kontekst. Ka selle konteksti säilitamine on vajalik, võimaluse selleks annab metaandmete kasutamine.
- **Ligipääsetavus.** Kõikidele, kellele see on tarvilik ja vajalik, peab olema tagatud sobiv juurdepääs andmetele. Arvestama peab ka erineva asjaoluga, et kasutajatel on erinevad tarkvara ja riistvara võimalused.
- **Turvalisus.** Ainult autoriseeritud isikutel on juurdepääs dokumentidele. Võimaliku andmekao eest kaitsmiseks peab olema tagavarakoopia tegemise ja sellelt taastamise vahendid.
- **Vastupidavus.** Dokumendid peavad saama hoitud (säilitatud) nende ettenähtud aja vältel. Kasutatav meedium ja formaat peab olema ajakohane ja parima kvaliteedi-hinna suhtega.
- **Integreeritus.** Toetama ettevõtte teisi IT lahendusi.

Elektrooniline dokumendihaldus peab olema ettevõtte töötajatele võimalikult sarnane eelnevalt kasutusel olnud nt paberil põhineva dokumendihaldusega. Kasutades samu põhimõtteid ja töökorraldust (niivõrd kui see on võimalik ja kasulik) on garanteeritud uue süsteemi töötajatepoolne vastuvõtt.

2.2 Dokumendihaldussüsteemide tegevused

Dokumendihaldussüsteemid peavad täitma mitmeid ülesandeid, mille eesmärgiks on dokumendihalduse põhimõtete (alajaotus 2.1.3) tagamine. Järgnevalt on esitatud dokumendihaldussüsteemi põhiülesannete kirjeldus. Lisaks toodud ülesannetele on vajalik garanteerida süsteemi turvalisus ja töökindlus.

2.2.1 Dokumendi loomine

Dokumendi elutsükkel saab alguse dokumendi loomisega, milleks pakuvad dokumendihaldussüsteemid erinevaid võimalusi: dokument luuakse kas süsteemi enda või mõne muu rakenduse vahendeid kasutades. Kui dokument luuakse väljaspool süsteemi, siis tavaliselt lisatakse see süsteemi valmiskujul. See tähendab näiteks, et selline dokument tekib väljaspool süsteemi poolt toetavat töövoogu või see dokument on formaadis, mille loomist süsteem ei toeta (näiteks audio- ja videofailid). Kui süsteem on integreeritud organisatsiooni teiste IT lahendustega, näiteks finantssüsteemiga või elektronkirjade süsteemiga, siis on need programmid ka potentsiaalsed dokumendiloojad (finantsaruanded, eelarve lisad, E-kirjad vms).

Kaasaegsed dokumendihaldussüsteemid toetavad elektroonilisi dokumendi loomise vorme ehk E-vorme. E-vormid pakuvad kasutajale sõbralikku, mugavat ja kiiret vahendit info sisestamiseks dokumendiks. E-vormid on dokumendi põhjad, milles on „augud“, mida kasutaja täidab kas vabatekstiga (näiteks kirjutab inimese nime) või on antud kasutajale valik, mille hulgast ta saab võtta sobiva väärtuse (näiteks kuu on number 1st 12ni). Seejuures süsteem kontrollib, et sisestatud info oleks täielik ja kehtiv. E-vormide koostamiseks kasutatakse näiteks veebi ja XMLi (vt 3.2) vahendeid.

Veel üheks võimaluseks dokumentide loomisel on dokumentide graafiline kujutamine (ik *Document Imaging*), mis on paberdokumentide konverteerimine elektroonilisteks kujutisteks. Dokumendi sissetoomisel dokumentide graafilise kujutamise süsteemi kasutatakse kolme peamist meetodit ([18]):

- skaneerimine paberkujul dokumentide jaoks - luuakse rasterkujutis, mida saab säilitada arvutis. Kasutades vastavaid meetodeid saab neid kujutisi ka elektrooniliselt töödelda.
- konversioon selleks, et luua muutumatuid pilte elektroonilisest dokumendist. Näiteks Windowsi keskkonnas saab *MS Wordi* faile „trükkida“ pildiks, mille formaat on tavaliselt PDF (ik *Portable Document Format*). Tulemuseks ei ole tekstifail, vaid pildifail.
- importimine selleks, et luua elektroonilisest dokumendist selliseid versioone, mida hiljem saab soovi korral muuta. Importimise korral jäetakse dokumendi

formaat muutmata ja kui soovitakse selle sisu vaadata, siis käivitatakse dokumendi loonud rakendus (või kasutatakse süsteemi lisatud faili vaatajat). Näiteks lisatakse süsteemi MS Exceli fail (sisaldab tabeleid) valmiskujul ja kui hiljem on vaja selles sisalduvat tabelit muuta, siis käivitatakse MS Exceli rakendus.

Dokumenti saab süsteemi lisada kahel viisil: kas valmisproduktina või luua süsteemi vahendeid kasutades. Igal juhul täidetakse (kas süsteemi või kasutaja poolt) dokumendi metaandmed so andmed, mis identifitseerivad ja/või kirjeldavad dokumenti. Kindlasti on vaja teada identifikaatorit (so nagu dokumendi nimi), autorit, lisamise kuupäeva jms, mis on tavaliselt dokumendihaldussüsteemis juba kirjeldatud. Identifikaatoriks on kas lisaja poolt pandud pealkiri või süsteemi poolt pakutav kombinatsioon tähtedest/numbritest. Viimane variant on eelistatuim, kuna dokumentide rohkuse korral võib kasutajal nimede väljamõtlemisel raskusi olla. Tihti kasutatakse identifikaatorit ka andmete indekseerimisel, mis lihtsustab dokumentide otsimist. Autoriks või omanikuks on tavaliselt dokumendi lisaja (süsteemi kasutaja).

Dokumentide atribuudid annavad võimaluse luua dokumentide registrit: selline olukord tekib siis, kui dokumendihaldussüsteem ei säilita dokumendi sisu, vaid ainult dokumentide atribuute. Vajalikke metaandmeid võib lasta dokumendihaldussüsteem organisatsioonil endal süsteemi juurde lisada: millal dokument jõustub, kui kaua kehtib, kellele on lubatud juurdepääs jms. Igale uuele atribuudile saab lisada ka tingimusi, näiteks kas need andmed on kohustuslikud kasutajal täita või mitte jms.

Dokumendi lisamisel süsteemi määrab kasutaja ka selle liigi (vt alajaotus 1.2). Kui süsteemis on erinevate liikidega käitumise toetus, siis on kasutajal sellevõrra lihtsam dokumenti käsitleda. Näiteks kui kasutaja otsib koostöölepingut, siis saab ta seda kergesti leida liigi „lepingud“ alt. Erinevad liigid annavad ka süsteemile võimaluse automaatselt dokumenti vastaval ajal arhiveerida (igale liigile saab määrata sobiva aktiivse faasi pikkuse) või kustutada (dokumendi eluea täitumisel).

2.2.2 Dokumentide otsing

Dokumentide leidmine on üks tähtsamaid tegevusi dokumentidega tegelemisel. Öeldakse, et professionaalid kulutavad 5-15% oma ajast info lugemisele, aga kuni 50% ajast selle otsimisele ([19]). Dokumentide kiireks leidmiseks panevad aluse nende lihtne ja loogiline liigitus ning sobivate metaandmete kasutamine. Samuti peab otsing (kriteeriumite määramine ja tulemuse esitamine) tehniliselt olema kasutajale selge ja arusaadav. Dokumendihaldussüsteemi kasutajatel on tavaliselt erinevad tingimused, millele vastavaid dokumente on vaja leida. Näiteks soovib sekretär teada saada, millistele kirjadele on vastamistähtaeg järgmisel nädalal või soovib peaspetsialist teada, millised ülesanded täpsemalt on temale antud mingi probleemi lahendamiseks. Sekretäri soovi saab lahendada, kui otsida metaandmete hulgast vastamise tähtaegu, peaspetsialist aga peab vaatama koosolekute protokollide sisu. Seega peab kasutajale olema antud võimalus otsida dokumente nii atribuutide järgi, sisu järgi (vabateksti otsing) või kõikide nende võimaluste järgi koos. Olenevalt süsteemist võib tihedamini kasutatavaid otsinguid ka eelnevalt defineerida ning kasutaja saab neid siis temale sobival momendil lihtsalt käivitada.

Otsingul kasutatakse tavaliselt mingit keelt, mis lihtsustab otsingu eesmärgi kirjeldamise. Keeles on tavaliselt vahendid, mis vabateksti otsingul lubavad määrata näiteks, mis sõna peab dokumendis leiduma ja mis ei tohi; või missugused sõnad kindlasti peavad koos leiduma. Veel on üpriski tavaline kasutada spetsiaalsümboleid otsingukriteeriumis. Kui näiteks on soov leida dokumente, milles sisaldub sõna „ülikool“ mingiski vormis, siis saab otsingutingimusse kirjutada „ülikool*“. "*" -i asemele paigutab otsimissüsteem 0 kuni mitu suvalist sümbolit ning tulemuseks sobivad dokumendid, mis sisaldavad sõnu „ülikool“, „ülikoolist“, „ülikoolid“, „ülikoolidest“ jms.

Dokumentide leidmine on kasutajatele kõige vajalikum. Dokumendihaldussüsteemi ülesandeks on anda lihtsad võimalused kiireks ja täpseks otsinguks.

2.2.3 Dokumendi versiooni haldus

Dokumente luuakse iteratiivselt, see tähendab, et enne kui dokument on valmis, läbib see mitu muudatuse ringi. Muudatusteks on näiteks sõnastuse ümbertegemine, lõikude

kaotamine või juurdelisamine, kujunduse parandamine vms. Dokumendihaldussüsteemid võivad pakkuda võimalust, mille korral kõik dokumentide muudatused (selle sisu või metaandmete muutumine) säilitatakse mingil kujul. Sellist võimalust nimetatakse versioonihalduseks.

Versioonihaldus sisaldab ([3]):

- kontrolli, mis keelab kasutajatel süsteemis olemasoleva dokumendi muutmise. Selleks et muuta, tuleb lisada uus dokument (versioon).
- Lihtsustatud viisi dokumendi uue versiooni sisestamiseks, st uus versioon saab endale automaatselt eelmise versiooni metaandmed.
- Lihtsat viisi, millega kasutaja teeb vahet, milline on dokumendi uusim versioon.
- Andmeid dokumendi muutuste kohta. See on näiteks info selle kohta, kes ja millal lisas uue versiooni. Sellised andmed peaksid olema nähtavad kasutajale, kellel on õigus vaadata dokumendi versioone.

Versioonikontroll on vajalik mitte ainult valmisdokumentide korral, vaid ka siis, kui süsteem lubab luua uusi dokumente, mille loomisest võtavad osa mitu inimest. Sellist olukorda nimetatakse **grupitööks**. Grupitöö tarkvara koosneb lisaks versioonikontrollist ka aja- ja ressursiplaneerimise ning diskussioonivahenditest jms.

2.2.4 Dokumentide säilitamine

Traditsiooniline dokumendihaldus ise ei tegele dokumentide arhiveerimisega, aga kaasaegsetesse süsteemidesse on kaasatud arhiivihaldus (ik *Archival Management*), mille põhiülesandeks on dokumentide säilitamine teatud aja jooksul, nende leidmise vahendite pakkumine ning dokumentide hävitamine.

Dokumentide säilitamiseks on vaja leida sobiv meedium, mille valimiseks võib tuua järgnevad tingimused ([13], [6]):

- dokumentide eluiga;
- dokumentide säilitamise ja kättesaamise hind;
- dokumentide hulk;
- säilitatud dokumentidele ligipääsuks kuluv aeg;

- meediumi sobivus (so töötab mitme tootja seadmel) ja võime kanda info ühelt meediumilt teisele (näiteks optiliselt kettalt magnetkettale);
- meediumi võimsus (so kui palju dokumente sellele mahub);
- meediumi vastupidavus (so kui kergesti saab meediat kahjustada, kui kiiresti halveneb/mandub);
- meediumi hind.

Tehnoloogia areneb kiiresti, millest tingituna nii tarkvara ja riistvara, mida kasutatakse dokumentide juurdepääsul, kui ka kasutatav meedium vananevad kiiresti. Selline situatsioon nõuab dokumentide „rännakut“ ühelt tehnoloogialt teisele, nii et dokument oleks kättesaadav tänapäevaste vahenditega. Pikaajalise säilitamise üheks põhitingimuseks on info säilimine läbi nende liikumiste.

Dokumentide hävitamise abistamiseks võib dokumendihaldussüsteem pakkuda automaatset süsteemi, mis kontrollib dokumentide „vanust“ ja kustutab need automaatselt või teatab sellest kasutajale, kelle ülesandeks on valida, kas kustutada dokument või säilitada. Teine viis on eelistatum, kuna automaatsuse korral võivad siiski tähtsamad dokumendid kaduma minna (kui näiteks inimliku vea tõttu on dokumendil vale säilitamisaeg). Dokumentide kustutamisel peab arvestama sellega, et kustutamise protsess ei kahjustaks teisi dokumente, mis on samal meediumil. Kui dokumente on vähe ja/või säilitusruumi on palju, siis võib dokumentide kustutamine süsteemist hoopis puududa.

2.2.5 Väljundi haldus

Dokumendihaldussüsteemi kasutajad otsivad ja leiavad süsteemist dokumente. Tulemuseks on tavaliselt dokumendi sisu näitamine ekraanil (sisu enamasti otsitaksegi), lisaks on võimalik leitud dokumente salvestada kasutaja arvutisse või printida. Sellist väljundi osa on aga edasi arendatud ning sellest on saanud eraldi IT tehnoloogia, mida nimetatakse väljundi halduseks (ik *Output Management*). See pole lihtsalt printimine, vaid vahend, mille eesmärgiks on lubada äriprotsessidel ja inimestel kasutada väljundit sõltumata tehnoloogiast (printerid, arvutid, operatsioonisüsteemid, autoriseerimisvahendid jms; [20]). Väljundi haldus sisaldab nii kasutajate võimalust dokumente kergesti printida ja saata suvalisel viisil sobivasse

sihtpunkti kui ka IT tugiisikute võimalust lihtsalt administreerida kõiki printereid, fakse jms.

Väljundi halduse eesmärgiks on võimaldada kasutajal järgmised tegevused ([20]):

- toimetada suvaline dokument suvalisse ettevõtte sihtpunkti, sõltumata rakendusest või platvormist, milles loodi dokument või millesse see suunati;
- haarata transaktsioonipõhised dokumendid automaatselt ja transparentselt, et neid saaks säilitada põhimõttel leht-lehel;
- saata dokumente sadadele ja tuhandetele kasutajatele ühe lihtsa „trüki“ käsuga;
- „külmutada“ dokument kasutaja väljundi keskkonnas ning säilitada dokumente tsentraalses dokumendihoidlas ilma võimalike muutusteta;
- kergesti parandada/suurendada toore või minimaalselt formaaditud andmete välimust vormide, fontide ja graafikaga;
- saada ja vaadata dokumente ilma neid tootnud rakenduse abita;
- indekseerida automaatselt kõik arhiveeritud dokumendid, et neid hiljem oleks lihtne kätte saada;
- kasutada ilma piiranguteta vaatamist, arhiveerimist, printimise haldust ja elektrooniliste dokumentide jaotamise lahendusi;
- hallata ja kontrollida kõiki ettevõtte printereid ja järjekordi ühe arvuti tagant;
- printida sobivasse printerisse vaatamata printeri lehe kirjelduse keelele (ik *page description language* ehk PDL);
- näha, kuhu ja millal dokument trükiti;
- ülal pidada ainult ühte draiverit kogu ettevõtte süsteemi peale (mitme draiveri elimineerimiseks).

Reaalselt on saavutatud enamlevinud platvormidest, operatsioonisüsteemidest ja rakendustest sõltumatus. Kasutajatele on võimaldatud dokumendi printimise lihtsus (süsteem valib sobiva printeri, teatab millal ja kuhu trükiti jms). Firma *Cypress Corporation* on loonud oma dokumendi formaadi, millesse konverteeritakse kõik dokumendid, millega kasutajad tegelevad ([20]). See lihtsustab dokumentide uuesti kasutamist ja hoiab kokku säilituspinda (nimetatud formaadis dokumendid on väiksemad mahult kui nt PDF-dokumendid). Peale ühtse dokumendi formaadi on nad loonud ka ühtse printeri draiveri, mis lihtsustab printeritega tegelemise.

2.2.6 Kasutajad

Dokumendihaldussüsteemi kasutajad saab jaotada kolme suurde gruppi:

- administraatorid, kellel on õigused süsteemi kasutajaid administreerida (lisada, kustutada, anda/võtta õiguseid);
- lisajad, kellel on õigus mingit liiki dokumente lisada, parandada, kustutada jne;
- vaatajad, kellel on õigus teatud dokumente ainult vaadata.

Kasutajate eristamiseks on dokumendihalduses õiguste süsteem. Olenevalt dokumendihaldussüsteemist on võimalik anda õigusi (vaatamine, lisamine, uuendamine, kustutamine jm) kas mingile dokumendi liigile või igale dokumendile eraldi. Õigused võivad olla hierarhilises järjestuses st näiteks lisamisõigusega kaasneb ka lugemisõigus või saab eraldi määratleda õigused (st on antud lisamisõigus, aga lugeda lisatud dokumenti ei saa).

Näiteks Tartu Ülikoolis on dokumentidele määratud õigused nii, et kõikidele kasutajatele on näha dokumentide atribuudid (pealkiri, number, lisamise kuupäev, lühikirjeldus jms), aga sisule on antud juurdepääsu piirangud (näha saavad ainult teatud kasutajad). Näiteks kirjavahetus dokumente näevad ainult selle omanikud (lisajad).

2.2.7 Lisavõimalused

Dokumendihaldussüsteem pole enam lihtsalt dokumendi registreerija ja säilitaja, sellesse on rakendatud mitmed lisavõimalused, mis aitavad kasutajaid nende igapäevatöös. Näiteks kasutajate süsteemisene E-kirja süsteem, millega saab süsteem automaatselt teavitada kasutajat teda huvitavate dokumentide lisandumisest/muutmisest. Uute dokumentide loomisel saab kasutada grupitöö võimalusi, millega kaasneb dokumentide lukustamise süsteem. Nimetatud süsteem paneb dokumendi lukku, kui keegi on selle võtnud muutmiseks. See tagab, et dokumenti saab hetkel muuta ainult üks töötaja. Lukustamise süsteem võib olla alternatiiviks versioonihaldusele st süsteemis on dokumendist ainult üks versioon, mida saab vajadusel muuta ainult üks kasutaja korraga.

Dokumendihaldussüsteem võib pakkuda ka kalendrit, mis ei ole lihtsalt kuupäevade näitamiseks, vaid iga kasutaja saab sinna sisse kirjutada oma kohustused või teiste kasutajate kohustused. Näiteks koosoleku kokkukutsuja kirjutab kalendrisse koosoleku toimumise aja jms. Kasutaja saab määrata, kas kirjutatud tegevus on privaatne (ainult temale) või kõigile nähtav. Kalendersüsteem võib kasutajat siis ka vastava tegevuse eel teavitada.

Projektitöö toetamiseks võib olla süsteemis ka ülesannete nimekiri (ik *Task List*), millega saab hallata projektis osalejate kohustusi ja vaadata, kui kaugemale keegi on jõudnud.

Kõiki toetatavaid võimalusi pole siin võimalik loetleda, kuna need on süsteemi spetsiifilised. Lisavõimaluste kohta saab lugeda üksikasjalikumalt konkreetse rakenduse dokumentatsioonist.

2.3 Dokumendihaldussüsteemi rakendamine

Dokumendihaldussüsteemi planeerimisest ja rakendamisest peaksid osa võtma organisatsiooni kolme osakonna (grupi) töötajad ([21]):

- operatsioonide osakond;
- infosüsteemide osakond;
- protseduuride osakond.

Operatsioonide osakonnaks loetakse osakonda, mis juhib ärirakendusi. See grupp on vastutav info kasutamises, mis saabub paberi, faksi või E-kirja teel; nad teevad otsuseid saabunud info põhjal ning säilitavad infot kasutades üldisi või personaalseid andmebaase. Operatsioonide grupi ülesandeks on määrata, missuguseid tööülesandeid peaks uus süsteem pakkuma, kuidas informatsioon peaks olema esitatud ning kuidas uus süsteem parandab vanu töömeetodeid. Nemad mõistavad rakendust, ütlevad pakkujatele, mida need tegema peavad, miks nad nii peavad tegema ja mis piirangud on hetkel kasutataval süsteemil. Lisaks sellele operatsioonide grupp:

- on vastutav olemasoleva töövoos kaasamises;

- analüüsib sisestavate dokumentide hulka;
- defineerib olemasoleva süsteemi „pudelikaelad“;
- arvestab välja olemasoleva süsteemi hinna.

Infosüsteemide grupp vastutab organisatsiooni varustamise eest infotehnoloogiaga, mis sisaldab arvutivõrke, keskseid arvutisüsteeme, personaalseid arvutisüsteeme ja tarkvara, mis on tarvilik ärieesmärkide saavutamiseks.

Infosüsteemide grupp mõistab, kuidas uus süsteem sobib olemasolevasse arvutite keskkonda ning missugused muudatused tuleb teha, et toetada uut riist- ja tarkvara. Infosüsteemide grupp on vastutav ka süsteemi tehnilise toe eest, kui see on juba installeeritud. Pakkumiste tegemise ajal vaatab infosüsteemide grupp üle pakkujad ja nende tehnoloogiad, et leida nende hulgast sobiv.

Kolmas grupp on **protseduuride grupp**, kes vastutab lepingute sõlmimise ja tasustamise ajakavade väljatöötamise eest. Protseduuride grupp on abiks pakkujate identifitseerimisel, arendab seadmete hinnad ja vastutab selle eest, et potentsiaalsed pakkujad on majanduslikult stabiilsed. Protseduuride grupp võib olla infokanaliks tehnilise tiimi ja pakkujate vahel.

Kõik nimetatud grupid peaksid arendustiimis olema esindatud võrdselt, mis tagab, et kõik projekti planeerimise ja teostamise ajal esilekerkivad erinevat laadi asjaolud oleks kontrolli all ja lahendatud.

Elektroonilisele dokumendihaldusele üleminek peab olema hästi planeeritud ja läbimõeldud. Halb dokumendihaldus toob kaasa muuhulgas ([3]):

- ajaraiskamise dokumentide otsimisele, mis on säilitatud ilma adekvaatse plaanita või mis on hoopis kustutatud;
- ajaraiskamise dokumendi vale versiooni kasutamise korral;
- säilitusruumi raiskamise dokumentide mittevajalike koopiatega;
- infopuuduse, kui dokumendid on halvasti kättesaadavad;
- jm.

Dokumendihaldussüsteemi nagu iga teise tarkvararakenduse kasutuselevõtt sisaldab järgmisi põhiprotseduure ([6]):

- **nõuete analüüs.** Arendustiimis osalejad analüüsivad olemasolevat rakendust ja uue süsteemi tööülesandeid, olemasolevaid ja uusi infotehnoloogilisi nõudeid jms. Abiks nõuete analüüsil on toodud küsimustik lisa 3.
- **Pakkuja valik.** Kui uus süsteem ostetakse väljastpoolt firmat (st ei looda ise), siis valitakse hoolikalt pakkuja, kellelt ostetakse süsteem. Vaatluse alla tulevad mitmed kriteeriumid:
 - pakkuja taust;
 - süsteemi ühilduvus olemasolevate tehnoloogiatega;
 - süsteemi hind;
 - süsteemis pakutavad tegevused;
 - süsteemi poolt toetatavad formaadid, platvormid jms;
 - süsteemi arendamisvõimalused;
 - süsteemi keel (inglise, eesti vm).
- **Rakendamise plaan,** mis peaks sisaldama tehnoloogilist rakendamise plaani, mis näitab, kuidas ja millal süsteemi installeeritakse ja testitakse ning kuidas ja millal koolitatakse kasutajaid.
- **Rakendamise plaani täitmine.** Süsteemi installeerimine ja testimine, kasutajate koolitus.
- **Hooldus.** Uue süsteemi kasutamisel jätkatakse süsteemi haldust ja esilekerkivate probleemide lahendamist.

Kõiki eeltoodud etappe läbib **dokumentatsiooni loomine**. Selles sisalduvad kõikide tegevuste, plaanide jm kirjeldused ning kasutajate juhendid.

Kasutajate koolitusel tuleb tähelepanu pöörata järgmistele asjaoludele ([3]):

- muutused töö tegemisel;
- varasema paberi või elektroonilise süsteemi vahetamisega seotud muutuse aspektid;
- töötamine jagatud informatsiooni keskkonnas;
- info sisestamine ja saamine;
- ärikohustused ja seaduslikud asjaolud;

- tehnoloogia kasutamine;
- kui vaja, siis kuidas on uus süsteem integreeritud teiste elektrooniliste rakendustega, mida töötajad kasutavad.

Lisas 2 on toodud mõned ideed, kuidas mõõta dokumendihaldussüsteemi kasutamise edu. Süsteemi edu üheks võtmesõnaks on kasutajate oskused ja tahtmine uue süsteemiga töötamisel. Kasutajate koolitusele tasub suurt tähelepanu pöörata. Selle kinnitamise näiteks võib tuua Tartu Ülikooli praktika, kus on loid dokumendihaldussüsteemi kasutamine. Üheks põhjuseks on just kasutajate puudulikud oskused. Kuna ülikoolis on üle kahe tuhande töötaja ja üle kümne tuhande tudengi, kes kõik on potentsiaalsed dokumendihaldussüsteemi kasutajad, siis püüti rakendada tuutorite süsteemi. See tähendab, et administratsioon õpetab välja grupi inimesi, kes siis omakorda viivad läbi koolitusi teistele. Esimestele gruppidele oli koolitus loengu vormis. Esineja näitas ühe arvuti ja projektori abil süsteemi võimalusi, osalejatele jagati osalised süsteemi kasutusjuhendid (eestikeelsed) ja tutvustati süsteemi elektroonilist abi, mis on kahjuks inglisekeelne. Aeg näitas, et sellisest koolitusest jäi väheseks. Teiseks etapiks oli kasutajate külastus nende töökohal, kus juhendaja vastas esilekerkinud küsimustele ning tutvustas süsteemi neid võimalusi, mida kasutaja veel ei tundnud. Kahjuks pole jõutud ülikooli suuruse juures kõikide sisestajate (kes õpetavad teisi) juures käia ja tulemuseks on praegu paljud iseõppijad, kes ei oska kasutada kõiki süsteemi võimalusi. Ka dokumendihaldussüsteemi vaatajaid (dokumentide otsijaid) pole suudetud piisavalt koolitada ning siiani kasutatakse sekretäride abi dokumentide leidmisel, olgugi, et kõikidele on vaatamiseks juurdepääs olemas. Sellise tegevuspraktika tulemusena kasutatakse ära vaid väike osa dokumendihaldussüsteemi võimalustest.

Dokumendihaldussüsteemi käigushoidmine vajab pidevalt hoolt, selle eest vastutavad järgmised töötajad ([3]):

- ärijuhid, kes vastutavad ettevõtte äritegevuse halduse eest. Nende ülesandeks on arendada ja parandada dokumendihalduse protseduure.
- dokumentide haldajad, kelle ülesandeks on dokumendihaldusega seotud ülesannete (sisestamine, muutmine, kustutamine jms) täitmise jälgimine;

- infotehnoloogia haldajad, kelle ülesandeks on pakkuda sobivat IT teenust, mis sisaldab dokumendihalduse riist- ja tarkvara, dokumentide ülekandmise protseduure ühelt formaadilt ja platvormilt teisele jms.
- inividid, kelle ülesandeks teada ja õieti kasutada dokumendihalduse vahendeid.

Ärijuht peab tagama, et organisatsioonile vajalikud dokumendid saaksid identifitseeritud ja registreeritud. Sellega kaasneb ka olemasolevate ja uute töötajate koolitus, et nad tunneksid ettevõtte dokumendihalduse protseduure. Ärijuhi ülesandeks on ka süsteemi probleemsete piirkondade identifitseerimine ja sellest vajalike töötajate (dokumendi- või IT haldajate) teavitamine.

Dokumentide haldajate ülesandeks on dokumentide lisamiseks, ülevaatamiseks ja kustutamiseks vajalikud protseduuride väljatöötamine ja nende protseduuride täitmise jälgimine. Näiteks milliseid dokumente millal hävitatakse jms. Dokumentidega tegelemise protseduuride aluseks on sobiv dokumentide liigitus, mille koostab samuti dokumentide haldaja.

Infotehnoloogia haldaja ülesandeks on süsteemi turvalisuse ja töökindluse pakkumine. Selle alla kuulub süsteemile juurdepääsu kontrollimine (paroolid, serveri kaitse, süsteemi kasutajate logi pidamine jms) ning dokumentidest tagavarakoopiate ja sellelt taastamise vahendite pakkumine. Infotehnoloogia haldaja ülesandeks on ka sobivate meediumite valimine, mida läheb vaja dokumentide aktiivsel kasutamisel või säilitamisel. Säilitamisel on vaja tagada kontrollitud (nt tekkivate muudatuste logifailide pidamine) dokumentide ülekandmine ühelt meediumilt (või formaadist) teisele.

Indiviidi ülesandeks on tunda ettevõtte dokumendihalduse protseduure ja põhimõtteid ning osata leida süsteemist dokumente (selles aitab teda dokumentide haldaja). Peale selle on hea, kui indiidid teab, kuidas on süsteem seotud tema tööga. Vajadusel teavitab indiidid ärijuhti probleemidest, mis tekivad süsteemi kasutamisel.

Näiteks Tartu Ülikoolis on dokumendihalduse arendamiseks loodud eraldi töörühm. Selle liikmeteks on rektori abi kui ärijuht, asjaajamise peaspetsialist kui dokumentide haldaja, arvutispetsialist kui infotehnoloogia haldaja ning esindajad teaduskondadest ning haldus- ja tugistruktuurist kui inividid.

Elektroniline dokumendihaldussüsteem aitab ettevõttel jagada infot nii oma töötajate vahel kui ka teistele ettevõtetele. Sobivalt planeeritud dokumentide liigitus ja nendega tegelemise protseduurid kiirendavad töötajatel info loomist, leidmist ja jaotamist. Dokumendihaldussüsteemid võivad lisaks pakkuda kontoritööks vajalikke vahendeid: kalender, ajaplaneerija, sõnumite vahendamine jms. Sobiva süsteemi leidmisel abistab ettevõtteid IT konsultatsiooni pakkuv firma.

3. Dokumendihaldusega seotud standardid ja tehnoloogiad

Dokumendihaldus ei ole enam üksik iseseisev osa infotehnoloogias. Ta ise kasutab ja tema meetodeid ja vahendeid kasutavad mitmed teised tehnoloogiad. Käesolevas peatükis on toodud ülevaade mõningatest dokumendihaldusega seotud standarditest ja tehnoloogiatest.

3.1. ODMA ja DMA

Aastaid tagasi alustati dokumendihalduse standardite loomist kahel rindel. Mõlema standardi peatähelepanu oli leida standardne viis, kuidas kliendi rakendus suhtleks dokumendihaldussüsteemiga (failide võtmiseks ja saamiseks). Nii loodi avatud dokumendihalduse rakendusliides (ik *Open Document Management API* ehk ODMA), mis lubaks kliendirakendustel ja dokumendihaldussüsteemil suhelda omavahel läbi kõrgeimal tasemel rakendusliidese sõltumata platvormist. Teine standard *AIIM Document Management Alliance* (DMA) on tarkvara arhitektuur, mis lubab dokumendihaldussüsteemide ja dokumendiga tegelevate rakenduste ühendamist ühte suurde ettevõtte dokumendihaldussüsteemi, seda sõltumata riistvara ja tarkvara platvormist.

Järgnevalt tutvustatakse neid standardeid, kasutatud on materjale allikatest [22] ja [23].

3.1.1 ODMA

Avatud dokumendihalduse rakendusliidese loojatel olid väga lihtsad tööreeglid:

- kui standard ei lahenda probleemi, siis seda ei hakata kasutama;
- kui standardi loomine võtab kaua aega, siis see ei lahenda probleemi;
- kui standardit on raske rakendada, siis see ei lahenda probleemi;
- standard peab olema tarnijast sõltumatu;
- standard ei pea lahendama kõikide tarnijate probleeme, kuna vastasel korral muutub see keeruliseks ning seda on raske rakendada;
- kui ei leidu lihtsat viisi, et ühendada rakendused, mis loovad dokumente, rakendustega, mis haldavad dokumente, siis on see kliendi kaotus;

- lihtne integratsioon rakenduste ja dokumendihaldussüsteemide vahel kasvatab tööstust ja suurendab müüki kogu turul;

Standardi algatajad soovisid luua lihtsat APIt, mis oleks tarnija ja platvormi sõltumatu ning samas lihtsalt rakendatav. ODMA eesmärgid on

- teha dokumendihaldussüsteemi teenused lauaarvuti kasutajale kättesaadavaks märkamatul viisil, nii et kasutaja arvates oleksid need tema rakenduse osad;
- vähendada rakenduste loojate koormat tegeleda mitmete dokumendihalduse süsteemide loojatega. Kirjutades ODMAs on rakenduse looja potentsiaalselt ühendanud oma toote kõikide toetatavate dokumendihaldussüsteemidega.
- Vähendada dokumendihaldussüsteemide loojate koormat tegeleda mitmete rakendustega. Toetades ODMAd on süsteem potentsiaalselt integreeritud kõikide rakendustega, mis on kirjutatud ODMAs.
- Vähendada pingutusi ja kompleksust, mis kaasneb dokumendihaldussüsteemide installeerimise ja ülalpidamisega.

ODMA koosneb liideste hulgast, mida rakendused saavad kasutada dokumendihaldussüsteemi tegevuste algatamisel. ODMA on platvormist sõltumatu, aga sellega seonduvad andmetüüpide definitsioonid ja siduv info on platvormipõhised. Esimene suurem ODMA spetsifikatsioon tunnustati 1997. aasta septembris, stabiilne tarkvara on saadaval maist 1998. a. Selles oli suurem osa tööst tehtud MS Windowsi operatsioonisüsteemi jaoks. Aja jooksul loodetavasti lisanduvad teisedki platvormid.

ODMAGA on ühilduvad näiteks dokumendihaldussüsteemidest IBM toode *Lotus Domino.Doc*; *Open Text Corporationi LiveLink* ning *80-20 Software* toode *Document Management Extensions for Microsoft Exchange* jt. Rakenduste näiteks on *Coreli* tooted *Quattro Pro* ja *WordPerfect*; *Microsofti* tooted *PowerPoint* ja *Word* ning *Lotuse WordPro* jt. Täpsemad nimekirjad on saadaval allikas [23].

3.1.2 DMA

The Document Management Alliance (DMA) on Rahvusvahelise Info- ja Graafikahalduse Assotsiatsiooni (ik *The Association for Information and Image Management International* ehk AIIM) töörühm, mis loodi realiseerimaks ühtlustatud lähenemist ettevõtte-suuruste dokumendihaldussüsteemide loomisel ja nendega opereerimisel. DMA peamine tulemus on integratsioonimudelite ja liideste objekt-orienteeritud spetsifikatsioon, mille abil saab rakendused ja teenused integreerida ühtsesse dokumendihalduse lahendusesse.

AIIM DMA töörühm loodi 1995. aastal. DMA 1.0 spetsifikatsioon kinnitati DMA Tehnilise Komitee poolt 1997. a. novembris.

Lõppkasutajad suudavad DMA-d toetava süsteemiga:

- kasutada dokumendihaldussüsteeme, mis pakuvad ühtset juurdepääsu ja integratsiooni strateegiliselt ühendatud organisatsioonide teiste süsteemidega;
- leida ja kasutada elektroonilisi dokumente, mis eksisteerivad suvalises kohas (näiteks ärirakenduses) ja suvalises vormis. Teistele organisatsioonidele on antud juurdepääs avalikele dokumentidele oma kohaliku dokumendihaldussüsteemi abil.
- Saada dokument kasutajale sobivas formaadis. DMA pakub mehhanismi, mis teeb kindlaks dokumendi formaadi ning kui võimalik, käivitab vajaliku rakenduse. Kui kasutajal puudub vajalik tarkvarapakett, siis on dokumendihaldussüsteemis olemas dokumendi koopiad alternatiivses formaadis, milles kasutaja saab dokumenti kasutada.
- Säilitada dokumente aja jooksul. Dokumentid jäävad kättesaadavaks ja kasutuskõlblikuks vaatamata tehnoloogia asendamisele, organisatsiooni ja rakenduste kasvule seni, kuni süsteemid on DMA-ga ühilduvad.

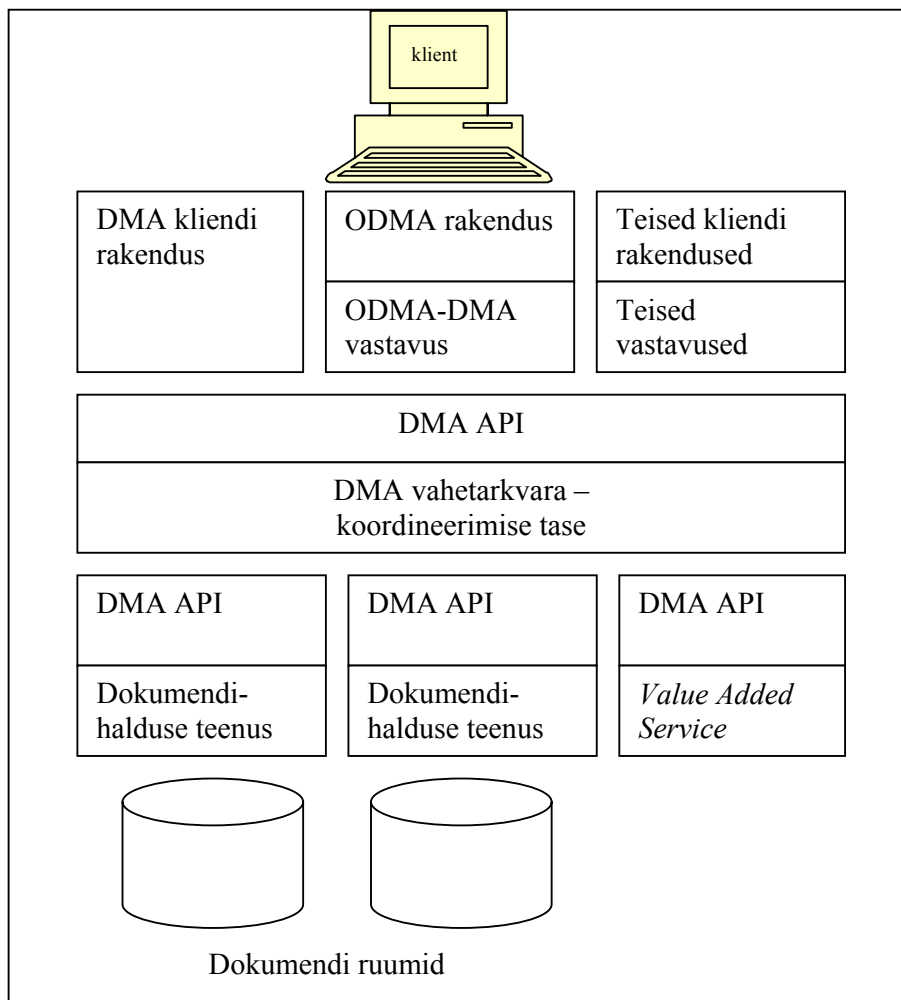
DMA rakenduse arhitektuur koosneb kolmest osast (vt joonis 3.1):

1. DMA kliendirakendused;
2. DMA vahetarkvara;
3. DMA teenuse pakkujad.

DMA kliendirakendused suhtlevad vahetarkvara ja teenuste pakkujatega DMA rakendusliidese ehk API abiga. Kõik kasutaja interaktsioonid ja kasutajaliidesed on kaetud DMA kliendirakendusega. Need rakendused on teostatud mitmetes programmeerimiskeeltes ning neid võib kombineerida teiste dokumendihalduse spetsifikatsioonidega nagu ODMA jt.

DMA vahetarkvaral on kaks peamist funktsiooni: päringute koordineerimine nn dokumendiruumidest (dokumentide andmebaas, andmeladu) ja DMA teenuste pakkujate registreerimine, et need saaksid kättesaadavaks kliendirakendustele.

DMA teenuse pakkujad pakuvad ühe kindla dokumendiruumi teenust.



Joonis 3.1. DMA rakenduse arhitektuur.

DMA standardi väljatöötamisest võtsid osa mitmete firmade esindajad, näiteks firmad *Documentum, Oracle, Xerox, FileNET, Boeing, IBM, USA Justiitsministeerium* jt.

Kokkuvõtteks võib öelda, et DMA on loodud elektrooniliste dokumendihaldussüsteemide omavahelise suhtluse standardiks ja ODMA on kui selle alamosa, mis annab standardi dokumendihaldussüsteemi ja teiste dokumente loovate rakenduste vaheliseks suhtluseks.

3.2 XML

Dokumendihalduses kasutatakse ka XMLi (*Extensible Markup Language*) võimalusi. XML on märgistuskeel, mis loodi 1996 aastal ([15]). Selle keele eesmärkideks võib tuua näiteks keele loomise, mis peab olema lihtsalt kasutatav veebis; keel, mida toetavad mitmed rakendused ning milles loodud dokumendid peavad olema kergesti loetavad ja lihtsalt loodavad nii inimese kui arvuti poolt.

XML dokument on tavaline tekstifail. See koosneb deklaratsioonide osast (*Document Type Declarations* ehk DTD) ja kehast, kus on esitatud dokumendi sisu ühe XML elemendi sees. Elemendiks on järgmine struktuur: `<nimi> tekst </nimi>`. Deklaratsioonide osas kirjeldatakse ära kehas asuva sisu struktuur, see osa võib asuda kehast erinevas failis, millele keha failist viidatakse.

Näiteks on XML dokumendi sisuks järgmine tekst:

```
<leping>
  <esimene_osapool> Keegi Isik </esimene_osapool>
  <teine_osapool> Keegi Firma </teine_osapool>
  <kuupäev> 01.06.2002 </kuupäev>
  <punkt>
    <punkti_pealkiri> Eesmärk </punkti_pealkiri>
    <punkti_sisu> ... </punkti_sisu>
  </punkt>
  <punkt>
    <punkti_pealkiri> Kehtivus </punkti_pealkiri>
    <punkti_sisu> ... </punkti_sisu>
  </punkt>
  ...
</leping>
```

XML kirjeldab dokumendi loogilise ülesehituse. Kui mingi tekst on <esimene_osapool> märgiste vahel, siis on võimalik aru saada, et see tekst sisaldab ühe sõlmitava lepingu osapoole andmed. Peale selle, et XMLis kirjutatud tekst on inimesele arusaadav, on seda ka arvutiprogrammidel kergem töödelda. XMLi eelis on sõltumatus erinevatest riist- ja tarkvaratehnoloogiatest, mis lubab seda keelt kasutada erinevate tehnoloogiate vahelise ühendusena. Näiteks MS Exceli faile avab peale Exceli enda vähe rakendusi, XMLi faili suudavad lugeda mitmed nii kontoritarkvara- kui ka veebirakendused.

Kuna XMLis kirjutatud tekstis on loogiline info eraldatud selle esitlusest (kusagil pole teksti sellel kohta, et näiteks **punkti_pealkiri** asub rea keskel), siis on võimalik seda infot kasutada erineval kujul ilma seda uuesti üle kirjutamata (taaskasutatavus). XML tekstile antakse esitlusvorm teiste vahenditega, näiteks XSL (*Extensible Style Sheet*) abiga. XMLi kasutamisega seotud palju teisigi tehnoloogiaid ja standardeid (enamusel täht X sees), neid käesolevas ei käsitleta. Ingliskeelset materjali nende kohta on võimalik saada näiteks allikast [24].

Kui soovitakse XMLi kasutada dokumendihalduses, siis esimeseks ülesandeks on kirjeldada erinevate dokumendiliikide (leping, määrus, protokoll jms) struktuur DTDna. DTD kirjeldamisel on võimalik ära määrata ka täitmise tingimused: näiteks nimi on kohustuslik, summa sisaldab ainult numbreid, lepingu liikideks on kas ostumüügi leping või tööleping jms. Tingimuste lisamisega peab arvestama ka tulevikuga, näiteks võib summa muutuda reaalarvuks, siis peab lubatud olema ka komakoht. See tähendab, et DTD ei tohi olla liiga jäik, sellesse peab jääma piisavalt ruumi arenguks.

Dokumendihalduses on XML heaks vahendiks dokumentide loomiseks, edastamiseks (eelkõige üle Interneti) ja säilitamiseks, seda muidugi juhul kui esitlusvorm pole tähtis (tähtsam kui sisu). Sellisel juhul aitab näiteks dokumentide graafiline kujutamine (vt alajaotus 3.6). Loomisel on XMList kasu sisestusvormide kasutamisel (E-vorm), DTDsse kirjeldatu abil saab sisestatud teksti kohe kontrollida ja abistada kasutajat teada olevate andmete automaatsel täitmisel.

3.3 Digitaalalkiri

Kuni tänapäevani on õiguslikud dokumendid (lepingud, seadused jms) paber kandjal, seda vaatamata asjaolule, et nende loomisel võidakse kasutada elektroonilisi vahendeid. Probleemiks on dokumendi tõestusväärtuse tagamine, milleks on ([25])

- dokumendi looja (ja ka loomisaeg) peab olema kindlaks tehtav;
- peale dokumendi valmimist pole seda muudetud.

Paber kandjal kinnitab dokumendi loojat ja loomisaega sellel paberil asuv allkiri ja kuupäev, kusjuures allkiri on kindlasti antud käsikirjaliselt. Elektrooniliselt pole selline lähenemine võimalik, kuna sel juhul pole dokument kinnitatud ühele kindlale meediumile (nagu paber, millel asub allkiri), vaid see muutub aja möödudes.

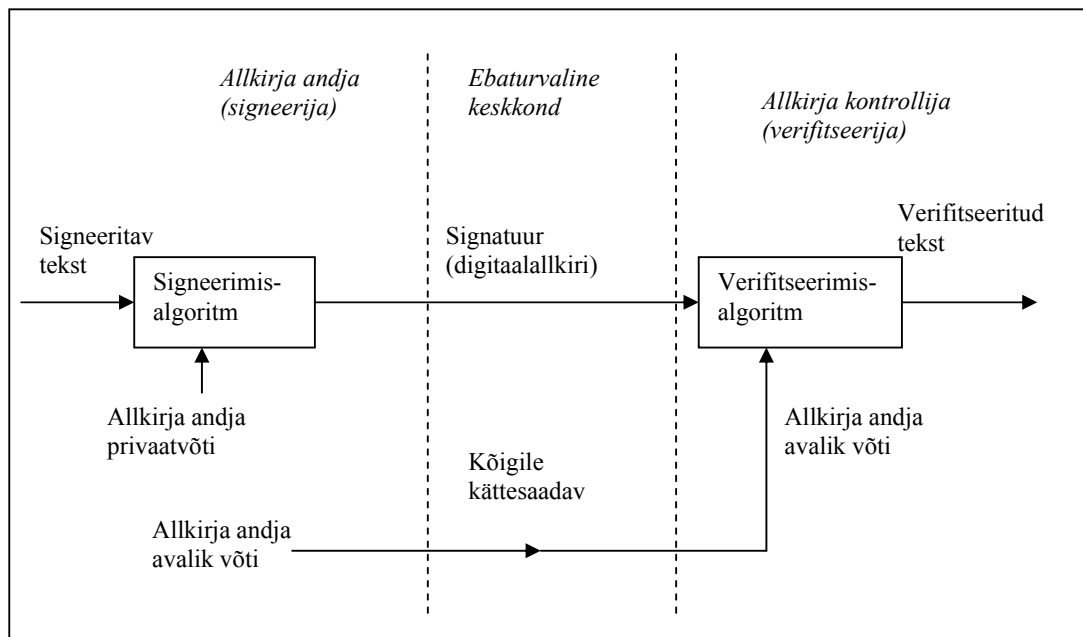
Paber kandjal olevatel dokumentidel on olemas „originaal“, elektroonilisel meediumil selline mõiste puudub. Dokument võib asuda disketil ja temast on tehtud koopia tagavaraks serverisse. Kui diskett hävineb, siis asendatakse sellel olev dokument tagavarakoopiaga, kuna tähtsam asi – sisu – on siis säilitatud. Samuti eristab elektroonilist dokumenti paber kandjal olevast elektroonilise dokumendi vorming. Sama dokument võib erineva väljundprogrammi poolt vaadatuna erineda vormi poolest. Näiteks ühte ja sama HTML-dokument näitavad erinevad veebibrauserid erinevalt.

Eelnevalt toodu põhjal ei saa elektroonilise dokumendi tõestusväärtust siduda dokumendi andmekandjaga, vaid selle sisuga. Elektroonilise dokumendi sisu on esitatud nullide-ühtede jadadena (bitid, baidid) ning nendega on võimalik sooritada matemaatilisi tehteid. Matemaatika annab võimaluse krüpteerimiseks ehk šifreerimiseks, millega viiakse tekst loetamatule kujule. Šifreerimismeetodid kasutavad salajast võtit ning avalikku meetodit (algoritm, mille abil šifreeritakse). Ilma salajase võtmeta ei ole võimalik šifreeritud teksti tagasi algkujule viia (sellist tegevust nimetatakse dešifreerimiseks).

1970ndate aastate lõpus töötati välja avaliku võtmega krüptoalgoritm ([25]). Avaliku võtmega algoritmil on ühe salajase võtme asemel kaks erinevat võtit (nn avalik ja privaatvõti). Esimese võtmega šifreeritud andmeid saab avada vaid teise võtmega ning

teise võtmega šifreeritud andmeid saab avada vaid esimese võtmega. Tekst šifreeritakse privaatvõtmega ja igaüks, kellel on avalik võti, saab siis teksti dešifreerida (pole vaja edastada salajast võtit). Kui avalik võti dešifreerib teksti, siis saab peale teksti enda kindlaks teha, kes on antud teksti loojaks. Seda muidugi eeldusel, et privaatvõti pole liikunud käest-kätte.

Andmete šifreerimist oma privaatvõtmega teabe hilisemaks autorluse tõestamise eesmärgil nimetatakse signeerimiseks või digitaalallkirja andmiseks. Signeeritud andmete dešifreerimist signeerija avaliku võtmega koos kontrollimisega, kas see toimus edukalt, nimetatakse verifitseerimiseks või digitaalallkirja kontrollimiseks (vt joonis 3.2 [25]).



Joonis 3.2. Digitaalallkirja andmine ja kontrollimine.

Peale allkirja lisatakse paberdokumendile veel ka kuupäev, mis on samuti selle õiguslik kinnitus. Digitaalmaailmas on kasutusele võetud ajatempel, milleks on sõnumilühend ja mille annavad välja kindla funktsiooniga ajatemplikeskused ([25]). Sõnumilühendi algoritm võimaldab pikast failist ehk infokogumist kogu olulise sisu paigutada lühikesse bitijadasse (see aitab tõestada dokumendi mittemuutumist). Ajatempel koosnebki eelmisest väljaantud sõnumilühendist (need on nummerdatud) ning dokumendi sõnumilühendist, millele ajatempel võetakse. Digitaalallkirja

varustamisel ajatempliga tuleb võtta kaks ajatemplit (üks enne ja teine pärast allkirja andmist), millega tõestatakse, et digitaalallkiri on välja antud enne (peale) teatud ajamomenti, millest varem (hiljem) ei olnud allkirja olemas. Et kinnitatud saaks ka allkirja andmise ajahetk kellaajaliselt (tavakasutuses oleva ajaga), siis peab lisaks olema veel üks kell, mille näitudele võetakse perioodiliselt, kindla ajavahemiku tagant ajatempleid (nn riigi etalonkell).

Digitaalse allkirjaga on seotud veel mitmedki ohud ja lahendused, aga neid käesolevas töös ei käsitleta. Neist saab lugeda allikast [25].

Digitaalse signatuuri kasutamiseks näiteks XML teksti juures on loodud *World Wide Web Consortium* (W3C) juurde töögrupp, mis tegeleb digitaalallkirja esitamiseks sobiva XML süntaksi ja allkirja verifitseerimise jt protseduuride standardite loomise ja esitamisega.

W3C poolt soovitatud XML digitaalallkiri on esitatud *<Signature>* elemendi sees, mille struktuur on järgmine ([26]):

```
<Signature ID?>
  <SignedInfo>
    <CanonicalizationMethod/>
    <SignatureMethod/>
    (<Reference URI? >
      (<Transforms>)?
      <DigestMethod>
      <DigestValue>
    </Reference>)+
  </SignedInfo>
  <SignatureValue>
  (<KeyInfo>)?
  (<Object ID?>)*
</Signature>
```

SignedInfo on element, mida signeeritakse. *CanonicalizationMethod* on meetod, mida kasutatakse signeeritava elemendi viimisel kanoonilisele kujule. *SignatureMethod* on algoritm, mida kasutatakse kanoonilisel kujul oleva elemendi signeerimiseks märgisesse *SignatureValue*. Täpsemalt digitaalallkirja kohta XMLis saab lugeda allikast [26].

Eestis on digitaalallkiri tihedalt seotud ID-kaardiga, millele on lisatud inimese nii avalik kui ka privaatvõti. ID-kaardi kasutamist toetavad näiteks *E-Maksuamet* ja Ühispanga *U-Netis*. Teiseks näiteks digitaalallkirja reaalse kasutamise kohta saab tuua E-kirja programmi *MS Outlook*, mis lubab saata ja saada digitaalallkirjaga kinnitatud kirju.

3.4 OCR, ICR ja OMR

Optiline sümboli tuvastus (ik *Optical Character Recognition* ehk OCR), intelligentne märgi tuvastus (ik *Intelligent Character Recognition* ehk ICR) ja optiline märgise lugeja (ik *Optical Mark Reader* ehk OMR) on arvutiteaduse osad, mis lubavad vähendada sisestava teksti hulka. Nende tehnoloogiate tutvustamiseks on kasutatud materjale allikatest [27], [28] ja [29].

3.4.1 OCR ja ICR

OCR ja ICR lubavad teksti (masinaga trükitud ja/või käsikirjalist) lugemist paberilt ja selle muundamist sellises vormis elektrooniliseks kujutiseks, millega arvuti oskab manipuleerida (näiteks ASCII koodiks). Need tehnoloogiad võimaldavad saadud kujutise hilisemat elektroonilist säilitamist ja lugemist.

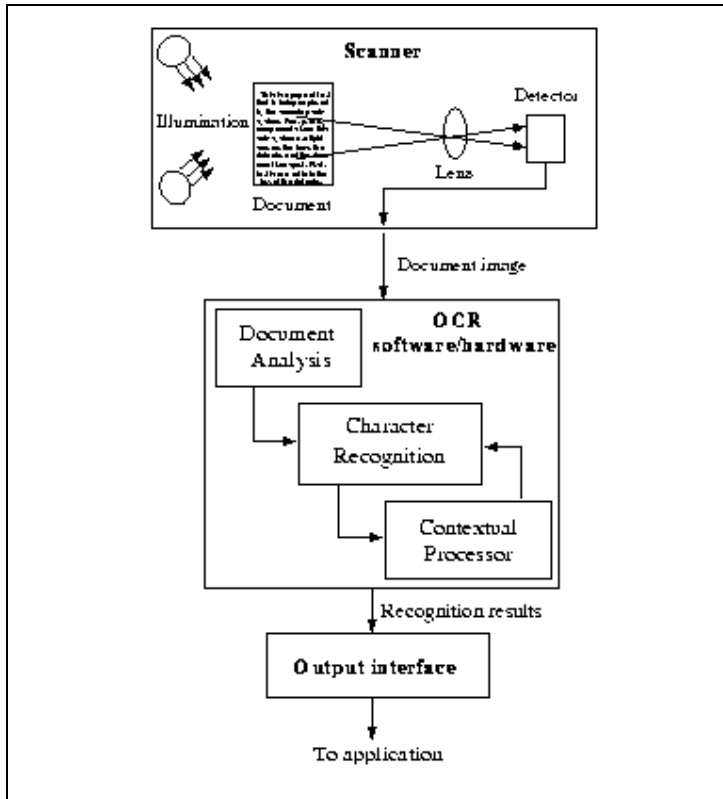
OCRi ja ICRi süsteem sisaldab optilist skännerit teksti lugemiseks, tarkvara graafilise kujutise analüüsimiseks ja väljundi liidest (vt joonis 3.3). Graafilise kujutise analüüs koosneb kolmes protsessist:

- dokumendi analüüs (lahutatakse individuaalsed sümboli pildid);
- nende piltide tuvastamine (kujul põhinev);
- kontekstiline töötlemine.

Dokumendi analüüsi ajal püütakse dokumendist eraldada individuaalsed sümbolid. Analüüsi tulemus on seda usaldusväärsem, mida kvaliteetsem on originaaldokument ja skännerilt saadud dokumendi graafiline kujutis.

Sümboli tuvastamisel on kaks põhikomponenti: iseärasuste lahutaja (ik *feature extractor*) ja klassifitseerimise algoritmid. Iseärasuste analüüs teeb kindlaks kirjeldused, mida kasutatakse kõikide sümbolite kirjeldamiseks. Kui on antud sümboli

graafiline kujutis, siis iseärasuste lahutaja teeb kindlaks kõik selle sümboli eripärad (omadused). Tuletatud iseärasused antakse üle sümboli klassifitseerijale, mille üks põhilisi töövahendeid on mallivõrdlus. Klassifitseerimise ajal võrreldakse sisendsümboli graafilist kujutist iga sümboli hulga prototüüpidega, seejuures kasutatakse igat individuaalset kujutise pikselit. Peale kõikidega võrdlemist antakse sümbolile kõige sarnasema malli identiteet.



Joonis 3.3 Märghsetuvastuse protsess.

Viimane tegevus on kontekstiline töötlemine. Antud kohale paigutatavate sõnade hulka saab piirata, kui teatakse selle koha kõrval olevate sõnade sisu. See tähendab, et kui näiteks tekstis on tuvastatud korrektselt postiindeks, siis tänava nime kohale saab võtta variandi postiindeksile vastavate tänavate hulgast. Pärast kogu tuvastusprotsessi on võimalik vigu parandada ka näiteks õigekirja korrektoriga.

3.4.2 OMR

OMR on tehnoloogia, mille abil saab koguda andmeid optiliste märgiste ("mullid", märkeruudud) ja ribakoodide kohta pabervormidelt. OMR teeb kindlaks märgise olemasolu või puudumise, mitte selle kuju nagu OCR ja ICR. OMRi saab kasutada küsimustike ja uuringute kokkuvõtte tegemisel ning testide hindamisel. Tulemuse saavutamiseks kasutab OMR kas spetsiaalset valgustundlikku tehnikat või vormide protsessi (ik *Forms Processing*).

Valgussensoritega tehnikat kasutades näitab märgise olemasolu paberil märgiselt peegelduv oluliselt väiksem valgus kui seda ümbritsevalt tekstilt. Selleks, et märgist avastada, peab see paiknema paberil korrektselt ja olema tunduvalt tumedam kui ümbritsev tekst. Nimetatud meetod nõuab spetsiaalseid tindipliiatseid või number 2 harilikke pliiatseid.

INDUSTRIAL REHABILITATION INTAKE QUESTIONNAIRE (H) HEALTHSOUTH

(Please answer all applicable questions completely.) (Make heavy black marks that fill the entire oval.) **CORRECT MARK** (To be completed by patient) Patient Name: _____

1) Are you currently working?
 Yes No

2) Which best describes your working status?
 Working without restrictions
 Working with restrictions (modified duty)
 Working different position with restrictions (modified duty)
 Unable to work due to injury
 Other: _____

3) Your Occupation: _____

4) Employer Issues
 Yes No
 My regular job is available when I am able to return
 My employer has a modified duty program
 My job is a union job
 I get along with my supervisors and co-workers

5) Last date worked due to this injury or medical condition.
 Not applicable

MONTH	DAY	YEAR
Jan	01-31	00-99
Feb	01-28	00-99
Mar	01-31	00-99
Apr	01-30	00-99
May	01-31	00-99
June	01-30	00-99
July	01-31	00-99
Aug	01-31	00-99
Sept	01-30	00-99
Oct	01-31	00-99
Nov	01-30	00-99
Dec	01-31	00-99

6) Date returned to work after this injury or medical condition.
 Not applicable

MONTH	DAY	YEAR
Jan	01-31	00-99
Feb	01-28	00-99
Mar	01-31	00-99
Apr	01-30	00-99
May	01-31	00-99
June	01-30	00-99
July	01-31	00-99
Aug	01-31	00-99
Sept	01-30	00-99
Oct	01-31	00-99
Nov	01-30	00-99
Dec	01-31	00-99

7) Is an attorney involved in this case?
 Yes No

8) Date of first doctor visit for this injury or medical condition.
 Not applicable

MONTH	DAY	YEAR
Jan	01-31	00-99
Feb	01-28	00-99
Mar	01-31	00-99
Apr	01-30	00-99
May	01-31	00-99
June	01-30	00-99
July	01-31	00-99
Aug	01-31	00-99
Sept	01-30	00-99
Oct	01-31	00-99
Nov	01-30	00-99
Dec	01-31	00-99

9) Date of surgery due to this injury or medical condition.
 Not applicable

MONTH	DAY	YEAR
Jan	01-31	00-99
Feb	01-28	00-99
Mar	01-31	00-99
Apr	01-30	00-99
May	01-31	00-99
June	01-30	00-99
July	01-31	00-99
Aug	01-31	00-99
Sept	01-30	00-99
Oct	01-31	00-99
Nov	01-30	00-99
Dec	01-31	00-99

10) Darken the appropriate oval.
 No surgeries for this injury?
 One surgery for this injury?
 Multiple surgeries for this injury?
 Took place: In Hospital In Outpatient Surgery Center
 Type of Surgery: _____

11) Please indicate all the following conditions that apply, either presently or in the past.
 High Blood Pressure Emotional/Psychological
 Gout Heart Surgery, Date: _____
 Varicose Veins Arthritis
 Epilepsy / Seizures Diabetes
 Stroke Pacemaker
 Sleeping Problems Metal Implant
 Dizziness or Fainting Currently Pregnant
 Chest Pain, Heart Attack Allergies, Specify: _____
 Cancer Other: _____

12) Have you had other medical or rehabilitation services for this injury?
 (please mark all that apply) Yes No
 Chiropractor Neurologist
 CT Scan Occupational Therapy
 EMG Orthopedist
 General Practitioner Physical Therapy
 Massage Therapy Podiatrist
 MRI Other: _____
 Myelogram

13) Are you currently taking medications?
 (Include Non-Prescriptions) Yes No
 (please mark all that apply)
 Anti-inflammatory Pain medications
 Muscle relaxers Blood pressure medications
 List Medications: _____

14) List any other information that would assist us in your care: _____

15) What are your expectations/goals while in this program?

16) Current Level of Pain
 (0 being no pain, 10 being pain requiring Emergency Room care)
 0 1 2 3 4 5 6 7 8 9 10

17) Physical/functional problems that interfere with your ability to work your regular job (please mark all that apply)
 Decreased lifting ability
 Decreased ability to sit, stand and/or walk
 Decreased ability to bend, and/or stoop
 Decreased ability to squat, kneel and/or crawl
 Decreased ability to climb stairs and/or ladders
 Decreased ability to perform hand and/or arm activities
 Decreased endurance
 Fear of re-injury and/or decreased confidence to do the job
 Pain

611023

Joonis 3.4 OMR vormi näide.

Teiseks meetodiks on vormide võrdlemine. Vormi malli ehk täitmata vormi (nt täitmata test) abil näidatakse OMRi tarkvarale, kus asuvad potentsiaalsed märgised, mida tuleb kontrollida. Kontrollitavad vormid sisestatakse skänneri abil ja tuvastusprotsessi käigus paigutatakse näitevormi mustrit sisendvormile. Joonisel 3.4 on toodud OMR vormi näide.

OMRi tarkvara võimaldab peale märgise tuvastamist ka automaatset tulemuse analüüsi (nt uuringute tulemus, testide automaatne hindamine). Väljundit võib formaatida tavalise tekstina, tabeli kujul (nt MS Excel, Lotus1-2-3), HTML-failina vms. OMRi tehnoloogiat kasutavad näiteks koolid õpilaste tööde hindamisel, firmad turu-uuringute tegemisel ja psühholoogid inimeste hindamisel ja testimisel.

Järgnevalt on toodud tabel 3.1, milles võrreldakse tutvustatud kolme tehnoloogiat.

	OCR, ICR	OMR
Käsitähtkirjalise teksti tuvastamine	Jah - ICR	Ei
Masintähtkirjalise teksti tuvastamine	Jah - OCR	Ei
Mullimärgiste tuvastamine	Jah	Jah
X-de tuvastamine (kas on valitud või ei)	Jah	Jah
Ribakoodi tuvastamine	Jah	Jah
Nõuab vorme	Ei	Jah
Täpsus	Kuni 98%	Kuni 99,9%
Elektrooniline säilitus	Jah	Ei
Kiirus	1200 - 7500 tunnis skänneri kohta	1500-10000 tunnis skänneri kohta

Tabel 3.1. ICR, OCR ja OMR võrdlus.

Tabelist 3.1 on näha, et kui on vaja dokumentide säilitamist, siis tuleb kasutada ICR või OCR tehnikat. OMR aga võimaldab dokumentidest vajalike andmete kiiret ja täpset kogumist.

Märgise tuvastamise tehnikad on vajalikud dokumendihalduses siis, kui on palju paberdokumente, mida on vaja edasi töödelda ning millel puudub elektrooniline vorm (vt ka alajaotus 3.6 Dokumendi graafiline kujutamine). Need tehnikad vähendavad oluliselt teksti sisestamisele kuluvat aega. Samuti kiirendavad ribakoodi lugejad kauba (nt kaupluses müügiprotsessis) ja inventari (ettevõtte vara) käsitlust. Veel üheks näiteks märgise tuvastamise kasutamise kohta võib tuua testide tulemuse automaatse sisestamise.

3.5 Andmebaasid

Elektroonilises vormis olevad dokumendid asuvad failides, mis jagunevad teksti-, audio-, tabelarvutuse- ja teisteks failideks. Dokumendihaldusprogramm tegelebki tavaliselt erinevat tüüpi failidega, aga ta võib pakkuda välja ka oma ühtse viisi dokumentide hoidmiseks (kasutades näiteks XML formaati). Kui dokumendid on erinevates formaatides, siis nendega tegelemisel peab olema kas kõik rakenduste kirjeldused dokumendihaldussüsteemis või peab süsteem kõikide vastavate rakendustega suhtlema. Näiteks MS Wordi faili korral peab kasutajal olema kontoritarkvara rakendus MS Word, kui just dokumendihaldussüsteemis pole kirjeldatud vastav formaat ja kasutaja näeb dokumendi sisu dokumendihaldussüsteemi enda abil. Mida rohkem formaate, seda keerulisemaks läheb dokumentide haldamine ja sellega peavad kõik dokumendihalduse tarkvara tootvad firmad arvestama.

Dokumentide säilitamiseks (kas failidena või muul kujul) kasutab dokumendihaldussüsteem andmebaase. Andmebaaside tarkvara pakuvad firmad aga on võtnud dokumendihalduse põhimõtted ja integreerinud need oma tootesse, pakkudes seega konkurentsi dokumendihalduses.

Näitena dokumendihalduse integreeritusest andmebaasisüsteemi on järgnevalt tutvustatud Oracle iFS võimalusi ([30]).

Erinevad võimalused failidele juurdepääsuks. Failidele saab juurde kasutada MS Windows keskkonda, veebi, FTPd, E-kirja rakendust vms. Kõikidele failidele juurdepääsuks on vaja meeles pidada ainult ühte parooli – andmebaasi oma.

Erinevat tüüpi failid ühes ja samas kohas. Üks kataloog võib sisaldada nii tekstitöötlusfaile, E-kirju kui ka veebifaile. Tavaliselt asuvad näiteks E-kirjad serveris ja tekstitöötlusfailid kasutaja arvutis.

Metaandmed. Andmebaasisüsteemis saab kirjeldada kõik vajalikud metaandmed, millest näiteks loomise kuupäeva ja autori andmed täidab süsteem automaatselt. Kuupäevaks on serveri (kus asub andmebaasisüsteem) kellaeg ja autori andmed saadakse autoriseeritud (parooli sisestanud) kasutaja andmetest.

Versioonihaldus. Kasutaja saab failide salvestamisel valida, kas säilitada faili eelmine versioon või kirjutada see üle.

Failidele juurdepääs projektitööks. Failide sisse-välja registreerimise süsteem (ik *check-in*, *check-out*) lubab lukustada faile, kui keegi sellega tegeleb. Õiguste süsteemiga saab määrata, kellel on õigust faili sisu vaadata ning kellel õigus muuta.

Otsing. Võimalik on kasutada kiiret sisul põhinevat otsingut üle kõikide failitüüpide (nt nii E-kirjade kui veebifailide hulgast).

Tagavarakoopiate tegemine. Andmebaasisüsteem pakub sisseehitatud andmete tagavarakoopia tegemist ja sellelt taastamise võimalusi.

Eelneva põhjal võib öelda, et kui ettevõtte kasutab Oracle andmebaasisüsteemi, siis sellega on kaetud minimaalsed nõuded dokumendihalduseks ja eraldi dokumendihaldussüsteemi pole vaja osta.

3.6 Dokumendi graafiline kujutamine

Dokumentide graafiline kujutamine (ik *Document Imaging*) on paberil ja mikrograafikal asuvate dokumentide konverteerimine elektroonseks graafikafailiks. Dokumentide graafilise kujutamise süsteemi on lisatud peale kujutamise ka dokumentide säilitamine, indekseerimine, otsimine (saamine) ja juurdepääsu haldus.

Graafilise kujutamise süsteeme pakuvad firmad jagunevad kaheks grupiks: on neid, kes tegelevad ainult graafilise kujutistega ja on neid, kes lisavad oma süsteemi ka teisi formaate peale nn pildiformaatide. Viimased süsteemid ei erine eriti dokumendihaldussüsteemidest, ainult nende põhitähelepanu on dokumentide sisestamisel graafiliste kujutamise vahenditega (skaneerimine, OCR jms).

Dokumentide indekseerimisel kasutatakse kolme vahendit ([31]):

- indeks- ja OCR sõnad on dokumendi sees;
- dokumendid salvestatakse erinevatesse kaustadesse ja
- dokumendimallile lisatakse indeksväljad.

Sobiva indekseerimismeetodi kasutamisest sõltub dokumentide leidmise efektiivsus. Kasutajale võimaldatakse otsinguid kaustade järgi, võtmesõnade järgi või täisteksti järgi (OCR meetodi korral).

Graafilise kujutamise tarkvara pakuvad firmad kasutavad tavaliselt TIFF (ik *Tagged Image File Format*) ja/või firma Adobe Systems poolt loodud PDF (ik *Portable Document Format*) formaati ([31],[18], jt). TIFF on mugav vahend dokumentide graafiliseks kujutamiseks, kuna see pole eriti keeruline formaat ja selles on infokadu väiksem kui teistel pildiformaatidel. Tema puuduseks on tulemusfaili liiga suur maht. PDF dokumendis võib peale dokumendi enda olla veel printimisjuhised (nt keelata printimine üldse), märksõnu otsingute tegemiseks ja indekseerimiseks, märkusi jms ([15]). PDF eelistatakse, kui on vaja kindlasti keelata dokumendi edasine muutmine, TIFF dokumenti võib pilditöötlusprogrammide abil muuta.

Säilitusmeediaks pakutakse tavaliselt peale magnetmeediate laserkettaid: CD või DVD. CD on eelistatud seepärast, et selle kasutamine ei nõua keerulisi riist- ja tarkvaratehnoloogiaid. CDle mahub umbes 12 000 lehekülge dokumente ([18]). DVD

kasutamine pole veel niipalju levinud, kuna DVD-ketaste lugemise/kirjutamise seadmed on kallid ning kasutusel on mitu kettale kirjutamise formaati (pole ühtsust nagu CD korral). DVD eeliseks on selle mahutavus: sellele saab salvestada kuni 30 korda rohkem dokumente kui CDle.

Graafilise kujutamise süsteemile on juurdepääs autoriseeritud. Nagu dokumendihaldussüsteemides on ka seal kasutajatel omad õigused (missuguseid dokumente võivad vaadata). Õiguste süsteem on lihtsustatud, kuna graafilise kujutamise süsteemis pole dokumentide muutmist. Ligipääs süsteemile võimaldatakse kasutajale tema lauaarvutist kas omaette rakenduse või veebibrauseri abil.

Dokumendi graafilist kujutamist vajab ettevõtte, kui tema sooviks on kõik ajaloolised dokumendid (kas paberil või mikrograafikal) viia elektroonilisele kujule. Ettevõttel on valida skänneriga dokumendihaldussüsteemi ja graafilise kujutamise süsteemi vahel. Viimase eeliseks on dokumentide konverteerimise ja säilitamise vahendite olemasolu.

3.7 Väljundi haldus

Väljundiks saab nimetada ASCII (ik *American Standard Code for Information Interchange*) koodis teksti ekraanil või laserprinteris olev printeri juhtkeelset (ik *Printer Control Language* ehk PCL) väljundit. Ekslik on arvata, et väljundiks on ainult inimesele arusaadav „tekst“, näiteks on ribakoodid masinloetavad ning OCRi väljund on mõeldud nii inimesele kui ka masinale. Väljundi haldus hõlmab endas kõiki tegevusi, mida läheb vaja väljundi saatmiseks sihtpunkti sobivas formaadis soovitud kiirusega ([32]).

Milleks väljundi haldus? Iga ettevõtte tegevusega kaasneb mingil kujul olev väljund: faksid, E-kirjad, veebi lehed, transaktsioonid vms. Kõige tavalisem väljund on mingi paberile printitud dokument. Kasutajatele on seoses printimisega tüüpiliseks probleemiks printeri järjekord. Enne koosolekut on vaja kiiresti trükkida vajalikud materjalid, aga kaastöötaja prindib 50leheküljelist aruannet. Lahenduseks ei pruugi olla uue printeri ostmine, vaid parem printimistöde haldus ja juhtimine üle olemasolevate printerite. Väljundihalduse rakendus valib kasutajale sobiva printeri

automaatselt ja annab teada, kuhu ja millal dokument trükiti. Kasutaja ülesandeks on ainult „trüki“ käsk anda. Väljundihaldussüsteemid püüavad lahendada mitmeid teisigi kasutajale väljundiga tegelemisega kaasnevaid probleeme. Järgnevalt neist lähemalt, kasutatud on materjale allikast [20]. Toodud on näited printimise ja printerite kohta, sest see on peamine väljundi tehniline vahend.

Tehnoloogiate ja rakenduste erinevus. Nagu juba varem öeldud, on arvutimaailmas palju erinevaid operatsioonisüsteeme, arvutivõrke ja tarkvararakendusi. Tulemuseks on hägus tehnoloogia infrastruktuur, mis pole optimeeritud infovahetuseks kasutajate ja seadmete vahel, rääkimata info haldamisest. Kui kasutaja soovib saabunud faksi edasi saata E-kirjaga, siis skänneri puudumisel tuleb see tal käsitsi sümbolhaaval sisestada ja siis E-kirja programmiga adressaadile saata. Väljundi haldus püüab lahendada analoogsed probleemid, pakkudes kasutajale võimalust ühe trükikäsuga valida, kas väljund printida, faksida või saata E-kirjaga.

Dokumendid on seotud originaalrakendusega. Kuidas printida uuesti dokumenti, mis on kord juba prinditud ja saada täpselt sama tulemus? Printimise tulemus sõltub paljuski rakendusest, millega dokument loodi ja printerist, mille abi trükitakse. Finantssüsteemist trükitud aruande uuesti saamiseks on vaja aruanne samas rakenduses uuesti käivitada. Kirjeldatud probleemi lahenduseks on kogu väljund salvestada mingis ühtses formaadis eraldi andmebaasi, kust seda saab otsida ja uuesti kasutada.

Erinevad printeri draiverid. Printeri draiveri ülesandeks on luua tarkvararakenduse väljundist printimiseks sobiv fail. Kohtvõrgus on vaja tavaliselt igale arvutile eraldi installeerida oma draiverid ja kui on vaja draiverit uuendada, siis tuleb taas kõik arvutid üle käia. Paljud organisatsioonid on loobunud draiverite uuendamisest, mis omakorda suurendab printeritega kaasnevaid probleeme (nt ei trüki õigeid sümboleid, osa printeri võimalustest jäävad kasutamata jms). Lahenduseks on universaalse printeri draiveri loomine, mis kaotab vajaduse draiverite vahetamiseks.

Sobimatu printimistöö staatuse näitamine ja ajalugu. Väga tihti küsivad kasutajad, kas dokument on trükitud, kuhu dokument prinditi jne. Kui printimisel tekib viga, siis on tihti arusaamatu, mis vea põhjustas ja kuidas seda parandada. Kui trükitakse sadu

lehekülgi, siis on kasutajale väga ebamugav teada saada, missugusest leheküljest alates ta peab uuesti printima. Samuti võib mõnikord olla vajalik teada dokumendi trükkimise täpne kuupäev ja kellaaeg. Kui seda ei paku tarkvara, kust dokument trükiti ja kasutaja pole seda ise käsitsi dokumendile peale kirjutanud, siis polegi võimalik seda teada saada. Lahenduseks oleks jällegi eraldi andmebaas, kus on kirjas andmed selle kohta, mis kell, mis kuupäeval, missugusesse printerisse, mitu koopiat trükiti jms.

Erinevad väljundi seadmed. Kui ettevõttes on mitu erinevate võimalustega printerit, siis peab kasutaja tihti meeles pidama, missugusesse printerisse tema saab oma dokumendi trükkida. Probleemiks pole mitte ainult printeri asukoht (nt lukustatud uksega ruumis), vaid ka selle tehnoloogilised võimalused. Paljud rakendused loovad väljundi, mille formaadiks on spetsiaalne lehekülje kirjelduse keel (PDL), mida valitud printer ei toeta. Samuti ei pruugi printer toetada dokumendi lehekülje suurust (A3) jne. Väljundi haldussüsteemid on sellisel juhul kasutajale abiks sobiva printeri valimisel.

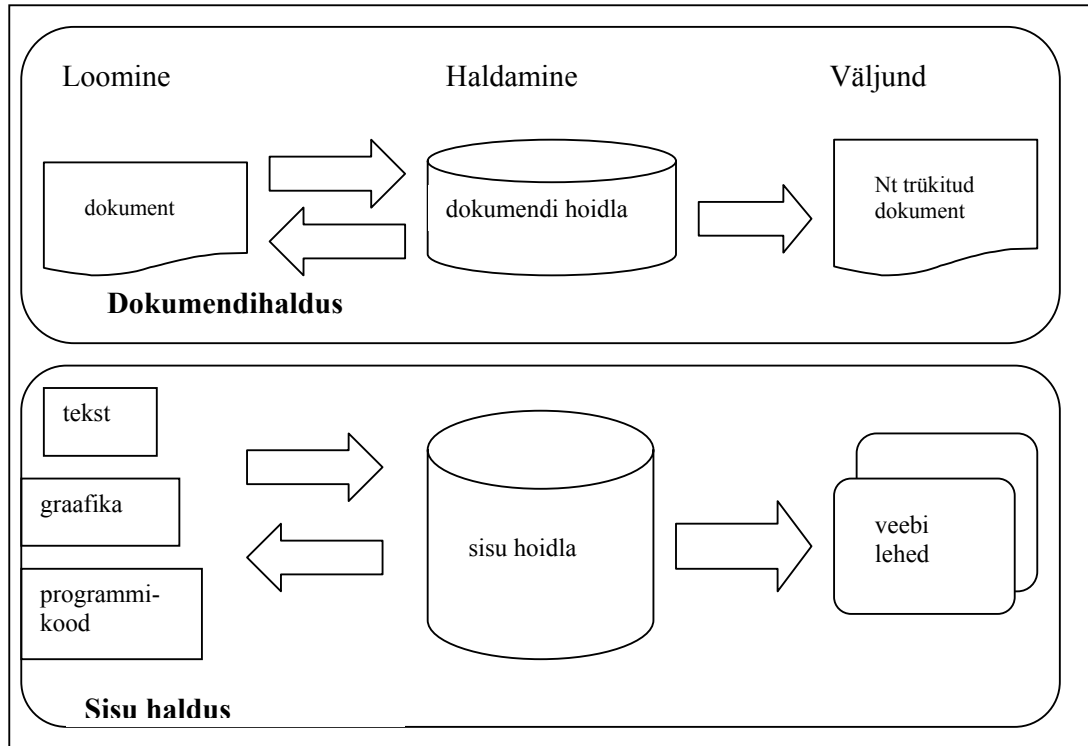
Väljundihalduse peamised komponendid ongi väljundi vaatamine, arhiveerimine, aruannete koostamine ja seadmete haldus. Kõik on loodud selleks, et kasutajatel ja IT tugiisikutel oleks lihtsam väljundiga opereerida.

3.8 Veebi sisu haldus

1990ndate aastate keskel hoogustus veebilehekülgede loomine, mis tõi kaasa nende ülalpidamise probleemi (näiteks ajakohasus). Aluseks võttes dokumendihalduse ja programmi lähteteksti kontrollsüsteemid (ik *source control system*) arenes välja sisu haldus (ik *Content Management*), mis on üheks võimaluseks nimetatud probleemi lahendamisel. Käesolevas alajaotuses on sisu halduse tutvustamiseks kasutatud materjale allikatest [33] ja [34].

Dokumendi - ja sisuhaldus pole samaväärsed tehnoloogiad, nende võrdlus on toodud joonisel 3.6. Suurim erinevus on väljundi ehk veebilehtede dünaamilisus, dokumendid on seevastu enamasti staatilised st ei muutu enam. Teiseks erinevuse on asjaolu, et

dokumendihalduse üks peamisi ülesandeid on dokumentide hoidmine (säilitamine), sisuhalduse peamiseks ülesandeks on sisu avalikustamine, säilitamine on teisejärguline ülesanne.



Joonis 3.5 Dokumentihalduse ja sisuhalduse võrdlus.

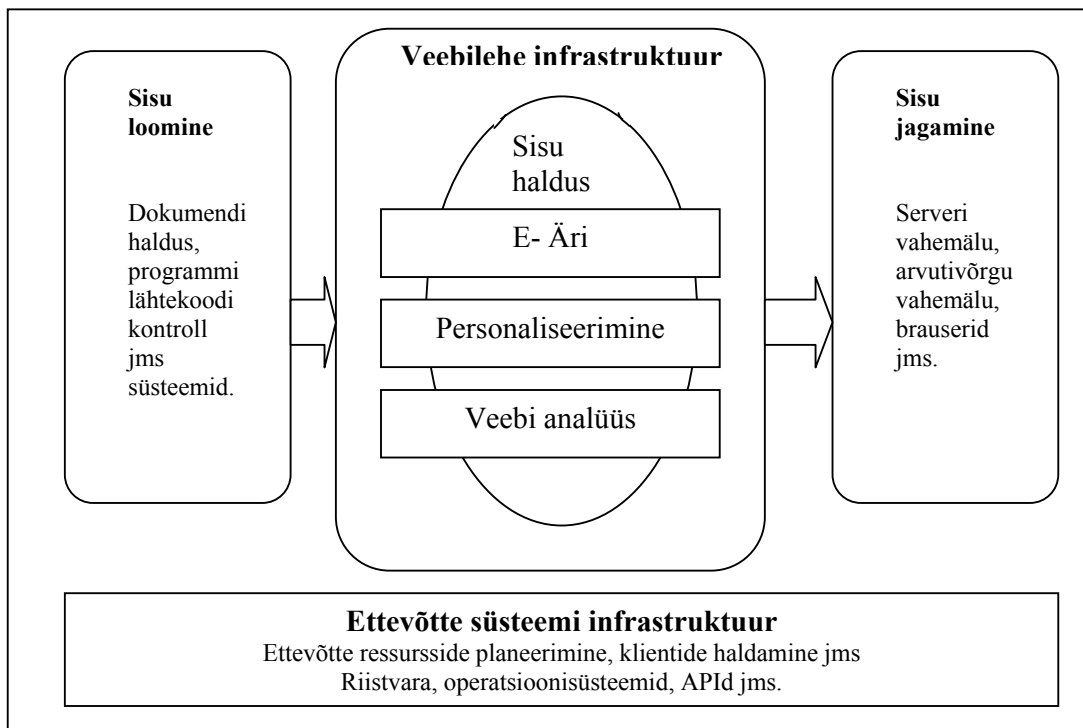
Kui võrrelda sisu haldust failisüsteemiga, mis on samuti kasutatav vahend veebi haldamiseks, siis saab tuua välja sisuhalduse eelised (enamusest võetud dokumentihaldusest):

- järjepidevuse haldus ja versioonihaldus. Failisüsteemis saab üht faili muuta mitu inimest, aga peale jääb viimane salvestaja. Sisuhaldus toetab versioonihaldust, mis säilitab kõik muudatused.
- kooskõlaline vaade üle kogu veebi. Sisuhaldus pakub standardiseeritud vahendeid lehekülgedel navigeerimiseks ning lehekülje kujundamiseks. Näiteks kui soovetakse muuta üht ikooni kogu veebil, siis tuleb seda teha ühes kohas ja sisuhalduse süsteem hoolitseb selle eest, et ikoon saab vahetatud kõikjal, kus see esineb.
- juurdepääsu kontroll. Sisuhalduse süsteem pakub kasutajate haldust ja õiguste süsteemi.

- kiired ja efektiivsed otsingud.

Sisuhalduse süsteem töötab koos mitmete organisatsiooni süsteemide ja – protsessidega, mis jagunevad kolme peamisse kategooriasse (joonis 3.7):

- sisu: sisuhaldussüsteem asub sisu loojate ja sisu edastajate vahel. See võtab sisu teistest hoidlatest (dokumendihaldus, digitaalsed varahaldussüsteemid jms) ja annab selle edasi süsteemidele, mis kiirendavad sisu edastamist, näiteks vahemälule.
- veebileht: E-äri muudab veebi järjest keerulisemaks, mis sisaldab mitmeid veebi rikastavaid ja optimeerivaid süsteeme, näiteks personaliseerimine, analüüsi süsteemid jms. Kõige sellega peab sisuhaldus oskama suhelda.
- ettevõtte: kuna veebileht suhtleb mitmete ettevõtte süsteemidega, peab ka sisuhaldus nendega arhitektuuriliselt sobima, olgu need Microsofti- või Java-põhised vms.



Joonis 3.6 Sisuhaldus ettevõtte infrastruktuuris.

Sisu loojateks on näiteks rakendused, mis loovad puhtalt ainult HTML koodis lehti. Teiselt poolt on ettevõttes palju neid andmeid, mis pole algselt veebi jaoks loodud,

aga soovitakse neid sinna panna. Näiteks info tellimuste laekumise kohta, mis asub vastavas andmebaasis või firma logo, mis asub varahaldussüsteemis vms. Kui sisuhaldussüsteem ei suuda võtta andmeid veebi jaoks mitmetest süsteemidest, siis tuleb ettevõtte töötajatel see info uuesti luua ainult veebi jaoks.

Kui sisu on süsteemis olemas, siis tuleb see transportida sisu tarbijale. Tavaliselt on see inimene, kes kasutab veebi brauserit või hoopis veebitelefoni kasutaja. Seepärast peab sisuhaldussüsteem olema võimeline saatma sisu mitmele erinevale seadmele, tundes nende protokolle ja formaate ning tehnoloogiaid, mis kiirendavad sisu jõudmist tarbijani.

Sisuhalduse süsteemile peaksid eelkõige mõtlema need firmad, kelle veeb on pidevalt muutuv ja sõltub paljuski teistest ettevõtte süsteemidest. Näiteks firmad, kes pakuvad veebikauplusi või lihtsalt tootekatalooge veebis. Sisu haldus on kasulik ka koolitajatele, kes soovivad tutvustada oma kursusi või hoopis neid läbi viia veebi kaudu: kui sisu haldav rakendus on olemas, siis ei pea ostma spetsiaalset veebil põhinevat kaugkoolitussüsteemi.

Sisuhalduse tutvustamisega piirdub käesolevas töös dokumendihaldusega seotud tehnoloogiate tutvustamine.

Kokkuvõte

Käesolevas töös on toodud elektroonilise dokumendihaldusega seotud mõisted, põhimõtted ja tehnoloogiad. Dokumendihaldus ei ole enam eraldiseisev osa infotehnoloogias: kus iganes tegeletakse dokumentidega, seal kasutatakse dokumendihalduse lahendusi. Dokumendihalduse tarkvara ja riistvara areneb pidevalt: luuakse uusi standardeid, uusi meetodeid ja vahendeid. Nii on ka käesolev töö mittetäielik, käsitlemata jäid näiteks töövooga seotud standardid ja tehnoloogiad. Selle kohta on saadaval ingliskeelset materjali allikates [9] ja [35]. Samuti jäi katmata huvitav suund ettevõtte info haldamisel – teadmuse haldus (ik *Knowledge Management*).

Elektroonilise dokumendihalduse ülesandeks on säilitada „elektroonilist mälu“, mis muutub aastatega väärtuseks omaette. Hästi korrastatud dokumentide kogu võimaldab läbi viia täpsemaid ettevõttele vajalikke analüüse. Dokumentide säilitamisel tuleb arvestada mitmete dokumentide formaatide, suurte andmehulkade, turvalisusega, samas lihtsa juurdepääsu nõudega jms. Nii dokumendihalduse lahenduste pakkujatel kui ka selle (tulevastel) kasutajatel on veel palju avastada ja õppida nimetatud eesmärgi saavutamiseks.

Iga asutus toodab dokumente ja haldab neid. Kas see haldus peab olema elektrooniline, sellele saab vastata iga firma ise vaadates oma dokumentide hulka ja vahendeid. Kindlasti suureneb elektroonilise dokumendihalduse osakaal iga päev võrreldes paberil põhineva dokumendihaldusega. AIIM Internationali andmetel kasvab Euroopas dokumendihalduse tehnoloogia sektoris ennustatav aastane kasum umbes 25% aastas, mis tähendab, et aastaks 2003 on see kuni 11 miljonit eurot ([2]).

Electronic Document Management

Merle Sibola

Abstract

Documents are the "memory-keepers" of enterprise, which have to be preserved in some way. Today documents are on any medium: paper, micrographics or electronic media. This Master Degree Paper is mainly dedicated to electronic document and its management. Electronic document can contain almost anything: texts, digitalized speech and videos, spreadsheets, multimedia objects, etc. All that has to be handled and stored by electronic document management systems.

In the beginning, electronic document management technology was an independent discipline, but has now become integral component of different systems for all imaginable application fields. Wherever documents are created, processed, distributed, stored or printed, document management technologies are in use today. Electronic document management is one of the fastest growing submarkets within information and communications field. A study by AIIM International predicted annual growth in profits in the document management technology sector in Europe to be around 25%, which would mean an increase of up to 11 billion euros by 2003 ([2]).

This paper explores some of the standards and technologies, which are connected to document management; for example Open Document Management API (ODMA), Document Management Alliance (DMA), Extensible Markup Language (XML), Digital Signature, Document Imaging and Output Management.

Kasutatud kirjandus

1. Arvutikasutaja sõnastik
<http://ee.www.ee/AKS/>; 30.04.2002.
2. Ulrich Kampffmeyer; Electronic Document Management Market: Technologies and Solutions.
http://www.aiim.org/article_docrep.asp?ID=18302; 30.04.2002.
3. Office of Government Information Technology; Improving Electronic Document Management
<http://www.defence.gov.au/imsc/edmsc/iedmtc.htm>; 30.04.2002.
4. Ralph H. Sprague, Jr.; Electronic Document Management: Challenges and Opportunities for Information Systems Managers
http://www.documentmanagement.org.uk/articles/Challenge_Opport.htm; 30.04.2002.
5. Eesti Vabariigi Arhiiviseadus, Vastu võetud 25. märtsil 1998. a, Riigi Teataja I 1998, 36/37, 552.
6. Minnesota Historical Society, Electronic Records Management Guidelines
<http://www.mnhs.org/preserve/records/electronicrecords/erguidelinestoc.html>; 30.04.2002.
7. IEC SC3B Documentation; IEC 62045 Management data (metadata) for technical documents
<http://www.iec.ch/tc3/txt/3b263.pdf>; 30.04.2002.
8. Jüri Vain, Juhan-Peep Ernits; Electronic Document Management
<http://www.cc.ioc.ee/training/unesco/onlinegov/docsys/detailed/>; 30.04.2002.
9. Riigikantselei; Strateegiliste juhtimisinvesteeringute (*Strategic Management Invest*) ja Infotehnoloogiaalaste (*Information Technology*) tegevusplaanide kokkulangevus (teine versioon)
<http://www.riik.ee/riigikantselei/ahb/is/tolked/halduste.htm>; 30.04.2002.
10. Future Strategies Inc; E-workflow – the workflow portal
<http://www.e-workflow.org/>; 30.04.2002.
11. Tartu Ülikooli dokumentide loetelu
https://dok.ut.ee/liveliink/liveliink/fetch/2000/866945/3208/3437/254243/Dokumentide_loetelu_kinnitamine.html?nodeid=856297&vernum=1; 30.04.2002.

12. William Saffady; The Document Life Cycle: A White Paper,
<http://www.documentconversion.com/news2.htm>; 30.04.2002.
13. National Archives and Records Administration;
Regulations: 36 CFR Part 1234 – Electronic Records Management;
<http://www.nara.gov/nara/cfr/cfr1234.html>; 30.04.2002.
14. Ruscan Technology; Accelerated Aging Test
http://www.ruscan.com/e_products/e_test_accelerated.html; 30.04.2002.
15. Tiina Tamme; Elektrooniline dokumendivahetus – standardid ja probleemid,
magistritöö.
16. W. H. Inmon; Building the Data Warehouse
17. Ivar Hütt; Dokumentide uputus nõuab korralikku haldust
http://www.aripaev.ee/1536/rubr_artiklid_153207.html; 30.04.2002.
18. Laserfiche®; Basics for Document Imaging & Management Systems
<http://www.laserfiche.com/basics/index.html>; 30.04.2002.
19. Document Imaging Solutions, Inc.;
<http://www.disusa.com/>; 30.04.2002.
20. Cypress Corporation White Paper, Output Management: It's More than Just
Managing Printers;
<http://www.cypressdelivers.com/WPOM1.htm>; 31.03.2002.
21. Electronic Records Management (ERM) Project Planning©;
<http://www.erms.com/RFPResources/project.htm>
22. AIIM DMware; The Document Management Alliance;
<http://www.infonuovo.com/dma/>; 30.04.2002.
23. AIIM DMware; The Open Document Management API
<http://www.infonuovo.com/odma/>; 30.04.2002.
24. World Wide Web Consortium;
<http://www.w3.org/>; 30.04.2002.
25. Valdo Praust; Digitaalalkkiri Tee paberivabasse maailma. AS Kirjastus Ilo,
2001.
26. XML-Signature Syntax and Processing - W3C Recommendation;
<http://www.w3.org/TR/xmlsig-core/>; 30.04.2002.
27. Sargur N. Srihari & Stephen W. Lam; Character Recognition;
<http://www.cedar.buffalo.edu/Publications/TechReps/OCR/ocr.html>;
30.04.2002.

28. NCS Pearson™; ICR, OCR, and OMR - A Comparison of Technologies;
<http://www.ncspearson.com/scanners/compare.htm>; 30.04.2002.
29. VisionShape, Inc.; What is OMR?
<http://www.visionshape.com/omr.html>; 30.04.2002.
30. Oracle®; Internet File System User's Guide;
http://download-uk.oracle.com/otndoc/oracle9i/901_doc/ifs.901/a75154/intro.htm; 30.04.2002.
31. Corporate Technology Developers, Inc;
<http://www.docxtodisc.com/>; 30.04.2002.
32. Microsoft TechNet; Print and Output Management Operations Guide;
<http://www.microsoft.com/technet/treeview/default.asp?url=/TechNet/prodtechnol/windows2000serv/maintain/opsguide/pomgmtog.asp>; 30.04.2002.
33. AberdeenGroup; Content Management: At the Center of e-Business. An Executive White Paper. Juuli 2001;
<http://www.aberdeen.com/ab%5Fabstracts/2001/07/07012572.htm>;
30.04.2002.
34. Oracle®; Managing Web Content - From File System to Database; May 2000
<http://otn.oracle.com/products/oracle8i/htdocs/cms/cmsdb.html>; 30.04.2002.
35. Workflow Management Coalition; WfMC Standards
<http://www.wfmc.org/standards/docs.htm>; 30.04.2002.
36. George Parapadakis; Document Management System: What is it and should I buy one?; Versioon 2.1, Oktoober 1996;
<http://www.documentmanagement.org.uk/articles/dmswhat.htm>; 30.04.2002.
37. George Parapadakis; Requirements Gathering Questionnaire; September 1998,
<http://www.documentmanagement.org.uk/articles/URS%20QuestionnaireH.html>;
30.04.2002.

Lisa 1. Paberil põhinev ja elektrooniline dokumendihaldus.

Paberil põhineva dokumendihaldussüsteemi ja elektroonilise süsteemi võrdlus on toodud järgmises tabelis ([19]):

	Paber	Elektrooniline formaat
Dokumentide võtmine	Võtab aega minutitest tundideni, olenevalt sellest, kas dokument oli õigesse kohta paigutatud või mitte.	Paarist sekundist kuni mõne minutini. Oleneb, kas dokument on onlainsüsteemis või mitte.
Kadunud dokumendid	Dokumendid kaovad tihti just valesti paigutamise tõttu.	Peaaegu võimatu kaotada.
Dokumentide paigutamine	Mõnest minutist mitme tunnini.	Peaaegu silmapilkselt õiges kohas
Dokumentide jagamine	Dokumentide jagamine on võimalik ainult koopiaid tehes	Kõik kasutavad ühte ja sama dokumenti.
Dokumentide saatmine	Kirja teel 1-5 päeva, faksiga sekundite-minutitega (koopia).	Printimine, faksimine, E-kirja saatmine. E-kirja saab parooliga kaitsta, nii et lugeda saab ainult see, kellele see on mõeldud.
Kaitse õnnetuste eest	Paber võib kahjustuda tule ja vee läbi.	Saab teha elektroonilisi tagavarakoopiaid, mida hoida väljaspool firmat.
Hoiuruum	Suur ruum koos vajalike kappidega.	Umbes 12 000 lehekülge dokumente on võimalik hoida ühel CD-l. Kasutusel olevad dokumendid asuvad kohalikus serveris.

Huvitavaid fakte paberil põhineva dokumendihalduse kohta on välja toonud *Coopers & Lybrand* ([19]):

- 90% ettevõtte mälust eksisteerib paberil.
- Keskmine dokument saab kopeeritud 19 korda.
- Dokumentide hulgast, mida iga päev kasutatakse, umbes 90% aetakse lihtsalt segamini.
- 7,5% dokumentidest kaob, 3% ülejäänust paigutatakse valesi.
- Professionaalid kulutavad 5-15% oma ajast info lugemiseks, kuni 50% ajast aga selle otsimiseks.
- Firmad kulutavad 20\$ dokumendi loomiseks, 120\$ kadunud dokumendi otsimiseks ja 220\$ kadunud dokumendi reprodutseerimiseks.
- Töötajad kasutavad oma ajast 20-40% dokumentide käsitsi otsimiseks.
- Kuni 90% dokumente kasutatakse ainult korra, selleks, et neid hävitada

Lisa 2. Elektrooniline dokumendihaldussüsteem: Mis see on ja kas ma pean selle ostma?

Kasutatud on materjale allikast [36].

1. Mida sisaldab endas elektrooniline dokumendihaldussüsteem?

Kindlasti on selles järgmised võimalused:

- Elektrooniliste dokumentide säilituskoht.
- Meetod dokumentide lisamiseks säilituskohta.
- Meetod dokumentide leidmiseks ja identifitseerimiseks säilituskohas.
- Meetod dokumentide võtmiseks säilituskohast.

Lisaks võivad süsteemis olla järgmised võimalused:

- **Sisse-välja registreerimissüsteem:** lukustusmehhanism, mis lubab ainult ühte kasutajat korraga muuta dokumenti.
- **Versioonikontroll ja kontrolljalg** (ik *audit trail*): meetod, mis haldab kõiki dokumendi muudatusi.
- **Turvalisus:** kontroll selle üle, missugused kasutajad omavad ligipääsu millistele dokumentidele.
- **Organisatsiooniline struktuur:** meetod, mis jagab loogiliselt seotud dokumendid gruppi, nt kausta või raamatuks.
- **Vaba-tekstiline otsing:** võimalus leida dokumente nende sisu järgi.
- **Dokumendi atribuudid:** dokumendiga seotud info (nt autor, loomise kuupäev jms); seda infot kutsutakse ka metaandmeteks.
- **Töövoog:** võimalused dokumendi saatmiseks ühelt kasutajalt teisele süsteemi poolt kontrollituna.
- **Graafiline kujutamine, OCR, ICR:** meetodid paberdokumentide konverteerimiseks elektrooniliseks.
- **Avalikustamine:** dokumentide sidumine kolleksioonideks ja nende saatmine sihtpublikule.

2. Kuidas ma tean, et vajan elektroonilist dokumendihaldust?

Kui mõnele järgnevale küsimusele on vastuseks "jah" või "võib-olla", siis tuleks mõelda mingisugusele dokumendihaldussüsteemile.

- Kas Te *regulaarselt* loote dokumente?
- Kas Te *regulaarselt* saate dokumente?
- Kas Teil tuleb tihti viidata dokumendile, mis on eelnevalt säilitatud?
- Kas Teil on mitu dokumendi kappi või arhiiviruumi?
- Kas mingi dokumendi loomisega on seotud rohkem kui 3 inimest?
- Kas ühte ja sama dokumenti kasutab samaaegselt mitu inimest erinevatest paikadest (erinevad kontori ruumid jne)?
- Kas Teie dokumente jälgib / auditeerib / reguleerib mingi väline organisatsioon?
- Kas Teil on dokumente, mis kriitilise tähtsusega Teie ärile (lepingud, litsentsid jms)?

3. Kust alustan?

Teil tuleb tuvastada järgmised kolm asja:

1. Kas elektrooniline dokumendihaldus toob endaga kaasa äri kasumi suurenemise?
2. Mis on kasu ulatuspiirkond? Kui see on liiga väike, siis on kasu väike ja see võib isegi pidurdada Teie ülesandeid. Kui see on liiga suur, siis võtab selle paigutamine ettevõtte infrastruktuuri liiga kaua aega. Te vajate lähenemist, mis toob kõige olulisema kasu mõistliku ajahulgaga.
3. Missugust mõju võib sellisel süsteemil olla Teie organisatsiooni töötajatele?

Alustuseks võiks üle vaadata hetkeolukorra:

- Kui palju on dokumente?
- Kas need on elektroonilises vormis?
- Kui pikad/suured need on?
- Kust need pärit on (tekstitöötlustest, skaneeritud jne)?
- Kui palju neid luuakse/saadakse kuus/aastas?
- Kui palju kasutajaid on seotud nende dokumentidega?
- Kus need kasutajad asuvad?

- Kas on olemas selge tee, mida need dokumendid järgivad enne valmimist?
- Kus neid dokumente praegu hoitakse?

Kui olete alustanud nende vastuste kogumisega, siis on kasulik konsulteerida kolmanda osapoolega (tavaliselt sõltumatu IT konsultatsioonifirma, kellel on dokumendihaldusega kogemusi), kes suudab hinnata Teie situatsiooni ja suudab pakkuda parimat viisi edasiminekuks. Nad tavaliselt analüüsivad Teie hetkelist dokumendihalduse protsesse, selgitavad, missugused on alternatiivid ja aitavad Teil ehitada ettevõtte juhtu.

4. Kui palju dokumendihaldussüsteem mulle maksma läheb?

Tihti on inimesed šokeeritud, kui kohtuvad dokumendihaldussüsteemiga seotud summadega. Siiski ei pea see maksma hingehinda. Kui Teil on paigas peamised kasud, mida toob endaga kaasa elektrooniline dokumendihaldussüsteem, süsteemi ulatuspiirkond ja esimene projekt nõuetest, siis tuleb alustada turu-uuringut. Missugused tooted on saadaval, missugused tehnoloogiad ja standardid?

Kui sobiva toote tüüp on identifitseeritud, siis on enam-vähem selge ostu ja/või arendushinnad. Sellele lisanduvad rakendamise, koolituse, halduse ja igapäevase käigus hoidmise kulud.

Jällegi aitab sõltumatu konsultant identifitseerida kõik silmnähtavad ja peidetud kulud.

5. Kuidas ma saan garanteerida edu?

Kuidas mõõta edu? Parimat süsteemi saab tõlgendada kasutuks, kui kasutajad ei kasuta seda efektiivselt. Miski ei saa garanteerida edu. Siiski on mõningad faktorid, mille poole saab püüelda:

- **Määratleda, mis edu on:** kuidas Te hakkate mõõtma süsteemi, et kindlaks teha, et esialgsed kasud on saavutatud?

- **Kasutajate kuulamine:** nemad peavad süsteemiga tegelema päevast päeva. Nemad peavad nägema, missugused on süsteemid kasud *neile* (ei pruugi olla ettevõtte kasu).
- **Piiritleda ulatuspiirkond:** vaadake väga valvsa pilguga üle nõudmised. Missugused toovad kaasa tõelise paranemise ja missugused "hea oleks kui on"? Kui süsteemi vundament on korrektne, siis on võimalik alati lisada funktsionaalsust.
- **Hea plaan:** suvaline projektiplaan peab kindlasti sisaldama kõiki väliseid tegevusi (sh kommunikatsioon, reisimine, töölt eemaldumine jms). Kõige tavalisem probleem selliste süsteemide rakendamisel on vajalike väliste allikate alahindamine.
- **Alustamise faasi peab suunama kõik võimalikud pingutused:** elektroonilise dokumendihalduse rakendamine on harva lihtne, kui süsteem on üles ehitatud korrektselt.
- **"Mõttele suurelt, tegutse väiksel":** kuigi süsteemi algne ulatuspiirkond on piiritletud, püüdke näha suurt pilti. Püüdke vältida otsuseid, mis "panevad Teid kasti", kui on vaja tulevikus süsteemi laiendada.
- **Otsige nõu:** kuulake teiste inimeste kogemusi analoogse süsteemi rakendamisel ja püüdke asetada ennast ja organisatsiooni nende asemele.

6. Järeldus

Enamus ettevõtteid vajavad mingilgi kujul dokumendihaldust. Enamustel juhtudel paberdokumentidega tegelemise hind ja protsessikiiruse parandamine enam kui õigustab elektroonilise dokumendihalduse rakendamise hinna. Teie tegelikud nõudmised määravad süsteemi hinna, aga pole põhjust, miks Teie dokumendihaldussüsteem maksaks rohkem kui keskmine selline süsteem.

Lisa 3. Uue dokumendihaldussüsteemi nõudeid koguda aitav küsimustik

Kasutatud on materjale allikast [37].

Projekti eesmärk

Mis on peamiseks tõukejõuks projekti juures? Mis on praegune või tulevane probleem, mida püütakse lahendada?

Ulatuspiirkond

Defineerige nõuete ulatuspiirkonna mingid kokkuleppelised piirid. Näiteks vaadake ainult mingit ettevõtte osa või kindlat dokumentide hulka.

Piirangud

- Kas on olemas mingi eksisteeriv süsteem, millega käesolev projekt peab ühilduma?
- Kas on mingeid ajalisi piiranguid? Näiteks ei saa alustada enne maid vms.
- Kas on mingeid ressursi piiranguid? Näiteks „Mais ja juunis pole vabu äriressursse seoses tööga hõivatusena“.
- Kas on mingeid Nõusoleku/Reguleerimise piiranguid, mis on seotud käesoleva projektiga.
- Kas on mingeid seonduvaid Infrastruktuuri piiranguid? Näiteks peab toetama Microsoft Exchange.
- Kas on mingeid eelarvelisi piiranguid?
- Kas on mingeid organisatoorseid/poliitilisi piiranguid, mida peab arvestama?
- Kas on mingeid defineeritud sihte, millega süsteem peab tegelema?
- Kas on mingi kindel kliendi/pakkuja suhe, mida peab arvestama?

Eeldused

Kas projektil on mingeid eeldusi? Näiteks „Me ei vaja uut serverit, kuna olemasolevas on piisavalt ressursse“ või „Igal lauaarvutil on Intranetile juurdepääs“.

Projektiga seotud dokumendid

Loetelu eksisteerivatest dokumentidest, mis on seotud käesoleva projektiga.

Lahenduse nõuded

Kui võimalik, siis selgitage, kas allpool esitatud nõuded on vajalikud või valikulised.

Informatsiooni ülevaade

Kirjeldage, millist tüüpi infot hakkab süsteem käsitlema:

- mis info, mis dokumendi liigid, spetsiaalsed indekseerimise nõuded, metaandmed, formaadid jne.
- Kuidas kasutatakse ja hoitakse seda infot praegu?
- Kui palju on paberil ja kui palju on elektroonilisel vormil?
- Kas on olemas mingid dokumenteeritud protsessid, mis kontrollivad selle info loomist ja kasutamist?
- Kuidas on defineeritud dokumendi omandiõigus?
- Milline on turvalise juurdepääsu mudel? See tähendab, kellel on õigused mida teha millise dokumendiga?

Funktsionaalsed nõuded

Loomine ja ülapidamine

- Missugused dokumendid loob väline kolmas pool? Kus kohast? Mis formaadis?
- Missugused dokumendid luuakse sisemiselt? Kas ühe dokumendiga on seotud rohkem kui üks autor? Milline on autoriseerimise protsess?
- Kas on olemas dokumendid, mida peab süsteemi tooma? Mitu protsenti nendest on paberil (elektroonilised)? Kas on vajalik eksisteerivad dokumendid skaneerida süsteemi või neile lihtsalt viidatakse?
- Kas on olemas mingi indekseerimise süsteem, mida on vaja lisada süsteemi (integreerida süsteemiga)?
- Kas on vajalik integreeruda süsteemi paberarhiividega?

Protsessi ülevaade

- Kas süsteemi osaks on dokumendi ülevaatusprotsess?

- Kas on olemas defineeritud protsess nende dokumentide ülevaatamiseks?
- Mitu inimest on seotud ülevaatusega?
- Mitu sammu on tüüpilisel ülevaatusprotsessil?
- Kui püsiv on ülevaatus protsess üle erinevate dokumendi liikide?
- Kas hakatakse kasutama digitaalalkirja?

Dokumentide säilitus

- Kui tihti dokumente vajatakse (juurdepääs)?
- Kui kiiresti on neid vaja?
- Kui kaua dokumente hoitakse?
- Mis määrab, kas dokumenti võib kustutada?
- Missugused dokumendi hoidmise reeglid on paigas?
- Missugused on seaduse nõuded säilitamisele?
- Kas on olemas mingid arhiivid, millega tuleb arvestada?
- Kui dokumente tuleb hoida pikemal ajaperioodil, siis kuidas käsitletakse formaadi uuendamist?

Dokumentide saamine

- Kuidas paigutatakse dokumendid? Milline on peamine dokumentide leidmise kriteerium (hierarhiline organisatsioon, metaandmed/võtmesõnad, sõnaraamat, vaba-teksti indeks)?
- Kuidas kasutatakse saadud dokumente (viitamiseks, uuendamiseks, uue dokumendi aluseks, printimiseks, jagamiseks kolmandatele pooltele)?
- Mitu inimest vajab dokumentidele juurdepääsu?
- Kuidas garanteerida, et alati kasutatakse dokumendi kõige uuemat versiooni, kui töötajad prindivad dokumente?
- Kas ettevõttevälised inimesed vajavad juurdepääsu dokumentidele?

Jaotus

- Mis on süsteemi sihtrahvas?
- Kuidas jaotatakse dokumente praegu?
- Kuidas on kasutajad geograafiliselt jagunenud?

- Kas on vaja dokumente kasutajatele jagada või anda neile lihtsalt juurdepääs (võtavad, millal vaja)?
- Kas on vaja kinnitust, et kasutajad on dokumenti näinud?

Integratsioon

- Kas süsteem on eraldi seisev või on vaja integreerida seda teiste olemasolevate (või tulevaste) süsteemidega?
- Kui nii, siis millised liidesed on võimalikud?

Teised funktsionaalsed nõuded

Kõik teised funktsionaalsed nõuded tulevad siia.

Teostuse nõuded

- Kasutajate hulk:
 - Praeguste kasutajate hulk (numbrites)
 - Kuidas muutub see number tulevikus?
 - Kuidas on kasutajad geograafiliselt jagunenud?
 - Mis on oodatav süsteemi kasutatavus?
- Informatsiooni hulk:
 - Kui palju dokumente?
 - Kui palju see number suureneb kuus/aastas?
 - Kui tihti neid uuendatakse?
 - Mitu liiki on dokumente?
 - Mis on tüüpiline dokumendi suurus (kas A4 lehtedes või kilobaitides)?
- Kui suur on keskmise suurusega dokumendile juurdepääsu aeg?
 - Dokumendi leidmise aeg?
 - Dokumendi avamise aeg?
 - Kui kaua dokumenti kasutatakse (iga kord, kui võetakse)?

Rakendamise nõuded

(Ei pea täitma enne, kui teised nõuete osad on täidetud)

- Kus hakkab süsteem paiknema?
- Kuidas pääsetakse süsteemi?
- Mitmel kasutajal on süsteemi parandamise võimalused?

- Installeerimise plaan (kogu süsteem korraga kõikidele; pilootgrupp, millele järgnevad kõik teised korraga; samm-sammuline).
- Koolitus:
 - Kui palju on vaja?
 - Kui mitmele kasutajale?
 - Missugusel IT-oskuse tasemel?
 - Värskendavad kursused?
- Andmete ülekandmine
 - Kõik eksisteerivad andmed
 - Andmed alates mingist kuupäevast
 - Nõudmisel
 - Mitte midagi

Operatsioonilised nõuded

- Mis ajaks peab süsteem valmis (kättesaadav) olema:
 - Kohalikuks juurdepääsuks?
 - Kaugjuurdepääsuks?
 - Kas tagavaraseadmed on vajalikud (nt CDROM)?
- Rollid ja vastutused
 - Kes on süsteemi omanik (IT poolelt, äri poolelt)?
 - Kes on peamine ärikontakt?
- Tagavarakoopia nõuded
- Toetuse nõuded

Kvaliteedi nõuded

- Turvalisus
 - Mil määral on turvaline juurdepääs tähtis?
 - Kes vajavad turvalist juurdepääsu?
 - Kuidas dokumente turvaliselt transporditakse?
- Missugused on kehtivuse nõuded süsteemile?
- Kas süsteem on GMPga ühilduv?
- Kas süsteem peab toetama mingeid teisi regulatsioone.