

ARKISTOINTI DIGITAALISELLA AIKAKAUDELLA

**Mitä muutoksia ja ongelmia elektroniset tallenteet
tuovat pitkäaikaiseen arkistointiin?**

Tietojärjestelmätieteen
pro gradututkielma
Sari Koivuaho
30.5.2002

Jyväskylän yliopisto
Tietojenkäsittelytieteiden laitos
Digitaalinen media

TIIVISTELMÄ

Koivuaho, Sari Katriina

Arkistointi digitaalisella aikakaudella. Mitä muutoksia ja ongelmia elektroniset tallenteet tuovat pitkäaikaiseen arkistointiin? / Sari Koivuaho

Jyväskylä: Jyväskylän yliopisto, 2002.

83 s.

Pro gradututkielma

Pitkäaikaisessa arkistoinnissa on pitkät perinteet paperitallenteiden käsittelyssä. Vähitellen digitaaliselle aikakaudelle siirryttäessä on yhä enemmän tallenteista tuotettu sähköisesti tietokoneen avulla. Tietoverkkojen kehittyessä yhä enemmän tietoa on saatavilla ja säilytettynä sähköisessä muodossa. Tässä tutkielmassa selvitetään pitkäaikaisen arkistoinnin periaatteita, käsitteitä ja käytäntöjä valtionhallinnossa ja selvitetään, kuinka ne soveltuvat elektronisten tallenteiden pitkäaikaiseen arkistointiin. Elektronisten tallenteiden pitkäaikainen arkistointi on huomattavasti mutkikkaampaa kuin paperitallenteiden arkistointi, joten tässä tutkielmassa selvitetään keinoja elektronisten tallenteiden pitkäaikaisen arkistoinnin ongelmien ratkaisemiseksi.

Tutkielma on sekä kirjallisuuskatsaus että kyselytutkimus. Työssä on perehdytty tutkimuskirjallisuuteen ja arkistolakiin ja haastateltu arkistolaitoksessa elektronisten tallenteiden parissa työskenteleviä henkilöitä.

Tutkielmassa on esitetty sekä perinteisten että elektronisten tallenteiden pitkäaikaisen arkistoinnin pääperiaatteet tutkimuskirjallisuuden ja arkistolaitoksen näkökulmasta. Työ on laadittu siten, että sitä myös muu kuin pitkäaikaisen arkistoinnin ja elektronisten tallenteiden asiantuntija pystyy ymmärtämään. Tutkielman keskeinen tulos on selvitys elektronisten tallenteiden pitkäaikaisen arkistoinnin ongelmiin ja ratkaisukeinoihin liittyvästä tutkimuksesta ja tämänhetkisestä tilanteesta arkistolaitoksessa elektronisten tallenteiden käsittelyssä.

AVAINSANAT: pitkäaikainen arkistointi, elektroniset tallenteet ja arkistolaitos.

ABSTRACT

Koivuaho, Sari Katriina

Archiving in digital age. What kind of influences electronic records have on longterm archiving? / Sari Koivuaho

Jyväskylä: Jyväskylä University, 2002.

83 s.

Thesis

Longterm archiving has long traditions in dealing with paper records. Along the digital age more and more records have gradually been created electronically with computers. After the development of web an increasing accessible number of records is widely in electronic form. The longterm archiving of electronic records is much more complicated than archiving of paper records. The thesis introduces the strategies for solving the problems of archiving of electronic records and describes the longterm archiving methods and practices in Finnish public and how usefull they are in longterm archiving of electronic records.

The data collection is based on literature review and interviews. The main literal sources consist of research literature, the archives act and instructions published by the Finnish Archives.

In this thesis the main principles of archiving of paper and electronic records have been presented from the perspectives of research literature and Finnish Archives. This report has also been written so that it does not require special knowledge about longterm archiving and electronic records. The main result of the thesis is the description of the possibly strategies for solving the problems of archiving of electronic records and the description of the situation in dealing with electronic records in Finnish Archives.

KEYWORDS: thesis, longterm archiving, electronic records, Finnish Archives.

1	JOHDANTO.....	5
1.1	TAUSTAA.....	5
1.2	TUTKIMUSKYSYMYKSIÄ.....	6
1.3	TUTKIMUSMENETELMÄ JA TAVOITTEET.....	7
1.4	TUTKIMUKSEN KESKEISET KÄSITTEET.....	7
2	ARKISTOINTIKÄYTÄNNÖT VALTIONHALLINNOSSA.....	11
2.1	ARKISTOINNISTA VASTAAVAT ORGANISAATIOIDEN.....	11
2.2	ASIAKIRJAHALLINTOON KOSKEVAT SÄÄNNÖKSET.....	14
2.3	PITKÄAIKAINEN ARKISTOINTI.....	15
2.3.1	PITKÄAIKAISEN ARKISTOINNIN PERIAATTEET ARKISTOTIETEESSÄ.....	15
2.3.2	PITKÄAIKAINEN ARKISTOINTI SUOMESSA.....	17
3	SIIRTYMINEN ELEKTRONISEEN ARKISTOINTIIN.....	20
3.1	DIGITAALISEEN AIKAAN SIIRTYMINEN.....	20
3.2	DIGITAALISUUDEN HAASTEET TALLENTEITA TUOTTAVILLE JA ARKISTOIVILLE ORGANISAATIOILLE.....	23
3.2.1	ELINIÄN LYHENEMISEN SYNNYTTÄMÄT ONGELMAT.....	23
3.2.2	NÄKEMYKSIÄ ARKISTOKÄYTÄNTÖJEN MUUTTAMISTARPEISTA.....	24
4	KÄYTETTÄVYYS, ALKUPERÄISYYS JA MUUTTUMATTOMUUS ELEKTRONISESSA ARKISTOINNISSA.....	27
4.1	ARKISTOTIETEESSÄ JA ARKISTOLAITOKSESSA TALLENTEILLE ASETETUT VAATIMUKSET.....	27
4.2	ELEKTRONISTEN TALLENTEIDEN ARKISTOINNILLE AIHEUTTAMAT ONGELMAT.....	29
4.3	ELEKTRONISESSA ARKISTOINNISSA YLEISIMMIN KÄYTETTÄVÄT TOIMENPITEET.....	31
4.3.1	YLEISTÄ.....	31
4.3.2	TALLENTEIDEN UUDISTUS.....	32
4.3.3	TALLENTEIDEN KOPIOINTI.....	33
4.3.4	MUUNNOS.....	35
4.3.5	ANALOGISTEN TALLENTEIDEN KOPIOIMINEN DIGITAALISILLE MEDIOLLE.....	38
4.3.6	ELEKTRONISTEN TALLENTEIDEN ASIANMUKAINEN SÄILYTYS JA TALLENTEIDEN KUNNON TARKKAILU.....	38
4.4	TUTKIMUKSESSA SUOSITELTAVAT ELEKTRONISEN ARKISTOINNIN MENETELMÄT.....	41
4.4.1	STANDARDIEN KEHITTÄMINEN ARKISTOINTIA VARTEN.....	41
4.4.2	TIETOJÄRJESTELMIEN SUUNNITTELU ARKISTOINTIA VARTEN.....	42
4.4.3	ELEKTRONISTEN TALLENTEIDEN ARVIOIMINEN SÄILYTYSTÄ VARTEN.....	44
4.4.4	METATIEDON KÄYTTÖ.....	47
4.4.5	TALLENTEIDEN LUOTETTAVUUDEN, AITOUUDEN VARMISTAMINEN.....	51
4.4.6	TALLENTEIDEN MIGRAATIO.....	54
4.4.7	TALLENTEIDEN EMULAATIO.....	62
5	ELEKTRONISEN ARKISTOINNIN KÄYTÄNNÖN RATKAISUT ARKISTOLAITOKSESSA.....	65
5.1	ELEKTRONISTEN TALLENTEIDEN KÄSITTELYOHJEIDEN SUUNTA-AVIIVOJA.....	65
5.2	SÄHKÖISTEN TALLENTEIDEN KÄSITTELYSTÄ ARKISTOLAITOKSESSA.....	67
5.2.1	ARKISTOLAITOKSEN HAASTATTELU.....	67
5.2.2	HAASTATTELUUN VASTAUKSET JA NIIDEN ANALYSOINTI.....	68
6	YHTEENVETO.....	74
	LÄHTEET.....	79
	LIITTEET.....	85
	LIITE 1: PITKÄAIKAISEN ARKISTOINNIN KÄSITTEITÄ.....	85

1 JOHDANTO

1.1 Taustaa

Rakkaassa maailmassamme on tapahtunut aikojen saatossa monia mullistuksia milloin mistäkin syystä. Suurin osa niistä on juontanut juurensa poliittisista ja taloudellisista syistä ja epäkohdista. Mutta myös tekniikan kehitys on aikanaan aiheuttanut teollisen vallankumouksen. Teollinen vallankumous oli mullistava koko maailman historiassa. Voisiko kenties informaatioteknologian huima kehitys saada mullistusten tapaisia vaikutuksia aikaan? Informaatioteknologian nopean kehityksen ja tiedon määrän valtavan kasvun myötä olemme jo tälläkin hetkellä keskellä suuria muutoksia. Historiallisesta perspektiivistä katsottuna tieto ei ole kauaa ollut sähköisessä muodossa, ja sen jälkeen kun sähköisiä tallenteita alkoi olla, on tapahtunut paljon. Tällä hetkellä yhä enemmän tietoa on pelkästään sähköisessä muodossa tallennettuna digitaalisilla laitteilla ja medioilla. Tietoa on digitaalisessa muodossa niin luottokorttitietoina, lääketieteellisinä tallenteina, lentoliikenteen tietokantoina ja verotusjärjestelmien tiedostoina. Yhä suurempi osa ihmisistä käyttää sähköpostia ja katsoo digitaalisia filmejä tai kuuntelee digitaalista ääntä. Viime vuosina on tapahtunut paljon muutoksia informaation alalla ja muutokset vain jatkuvat. Tulevaisuutta informaation tulvan ja informaatioteknologian kehityksen keskellä ei tällä hetkellä pystytä varmasti ennustamaan. Kukaan ei tarkalleen tiedä, kuinka kauan nyt käytössä olevat käyttöjärjestelmät, ohjelmat ja niillä luodut sähköiset tallenteet ja asiakirjat ovat käyttökelpoisia, toimivia ja säilyviä sekä kuinka olisi parasta turvata tietojärjestelmissä liikkuvan tiedon aitous, luotettavuus, muuttumattomuus ja autenttisuus eli alkuperäisyys.

Perinteisten asiakirjojen pitkäaikainen arkistointi ja säilytys on tarkoittanut esimerkiksi monien historiallisten aineistojen, kuten esimerkiksi 1600-luvun tuomiokirjojen kohdalla, satojen vuosien pitkäaikaista säilytystä. 1600-luvulta peräisin olevat alkuperäiset tuomiokirjat ovat vielä tänäkin päivänä luettavassa kunnossa. Eri organisaatiot ovat viime vuosisadan ja viime vuosikymmenten aikana siirtäneet pitkäaikaiseen säilytykseen tarkoitettujen asiakirjansa arkistolaitosten huomaan. Suurin osa näistä asiakirjoista on paperimuodossa. Arkistolaitoksella on omat arkistolakiin perustuvat määräyksensä asiakirjojen arkistoinnista, ja se on hoitanut tehtäviään näihin tukeutuen jo vuosikymmeniä. Tiedon määrän valtava kasvu on myös lisännyt sähköisten tallenteiden määrää.

Sähköisten tallenteiden ja asiakirjojen määrän kasvun myötä arkistolaitoskin joutuu miettimään, kuinka näitä sähköisiä tallenteita tulisi käsitellä ja arkistoida. Perinteisen arkistotoimen kannalta sähköisiin tallenteisiin ja asiakirjoihin liittyy monenlaisia ongelmia. Tietojärjestelmissä liikkuvan tiedon integriteetti (luotettavuus ja aitous) on keskeinen ongelma. Säilyykö tieto elinkaarensa kaikissa vaiheissa luotettavana ja aitona? Voidaanko olla todella varmoja siitä, että aineisto on varmasti peräisin siitä lähteestä, mistä sen väitetään tulevan ja ettei aineistoa ole asiattomasti manipuloitu? Pystytäänkö sähköiset tallenteet säilyttämään niin, että ne ovat käyttökelpoisia vielä pitkien aikojenkin kuluttua? Jos tulevaisuutta varten ei ole valmistauduttu kunnolla, sähköisten tallenteiden pitkäaikaisessa säilyttämisessä ja arkistoinnissa syntyvät ongelmat voivat pahimmassa tapauksessa aiheuttaa kallisarvoisen kulttuuriperintömme katoamisen.

1.2 Tutkimuskysymys

Ensisijainen tutkimuskysymys näkyy pro gradu -työni otsikossa: Arkistointi digitaalisella aikakaudella. Mitä muutoksia ja ongelmia elektroniset asiakirjat tuovat pitkäaikaiseen arkistointiin?

Tutkimukseni ensimmäinen kysymys koskee eri tutkijoiden ja asiantuntijoiden näkemyksiä elektronisten tallenteiden ja asiakirjojen pitkäaikaisesta arkistoinnista ja sen ongelmista. Sopivatko esimerkiksi vanhat arkistointiperiaatteet elektronisten dokumenttien käsittelyyn vai aiheuttavatko elektroniset tallenteet huomattavia muutoksia tähän asti noudatettuihin menettelytapoihin? Minkälaiset strategiat olisivat tarkoituksenmukaisia?

Tutkimukseni toisen kysymyksen muodostaa sen asian selvittäminen, miten Suomen arkistolaitoksessa on ajateltu meneteltävän sähköisten tallenteiden kanssa ja minkälaisia kokemuksia siellä on sähköisistä tallenteista.

1.3 Tutkimusmenetelmä ja tavoitteet

Tutkimukseni on sekä kirjallisuuskatsaus että kyselytutkimus, joka tarkoittaa keräämäni ja hankkimani tutkimus- ja haastatteluaineiston kokoamista, käsittelyä ja analysointia. Osan tästä aineistosta olen kerännyt haastattelemalla arkistolaitoksen henkilökuntaa. Loppu aineistosta on tutkimuskirjallisuutta ja muuta aihepiiriin liittyvää kirjallisuutta. Jonkinlaisena teoriapohjana käytän arkistolakiin (23.9.1994/831) perustuvia arkistolaitoksen määräyksiä pitkäaikaisesta arkistoinnista, tallenteiden ja asiakirjojen käsittelystä mm. aitouden, luotettavuuden ja säilyvyyden takaamiseksi. Vuonna 2001 juuri sopivasti työtäni ajatellen arkistolaitoksessa saatiin valmiiksi ohjeistus sähköisten tietoaineistojen ja tallenteiden käsittelystä. Tutkielmani tavoitteena on selvittää elektronisten tallenteiden ja asiakirjojen pitkäaikaisen arkistoinnin ongelmia sekä mahdollisia ongelmien ratkaisukeinoja tutkijoiden esittämien ajatusten ja esimerkkitapausten pohjalta.

1.4 Tutkimuksen keskeiset käsitteet

Keskeisimpiä käsitteitä tutkimuksessani ovat asiakirja, arkisto, arkistonmuodostaja ja tallenne. Muut käsitteet on selostettu liitteessä 1.

Asiakirja:

Arkistolaissa (Arkistolaki 23.9.1994/831, 6§) määritellään asiakirja (document) seuraavasti: ”Asiakirjalla tarkoitetaan tässä laissa kirjallista tai kuvallista esitystä taikka sellaista sähköisesti tai muulla vastaavalla tavalla aikaansaatu esitystä, joka on luettavissa, kuunneltavissa tai muutoin ymmärrettävissä teknisin apuvälinein.” Arkistolain määritelmässä otetaan huomioon sähköisesti aikaansaadut tuotokset. Asiakirjaksi määritellään myös sähköisesti aikaansaatu tuotos, mikä on itse asiassa välttämätöntä, koska nykyään suurin osa asiakirjoista tehdään sähköisillä, teknisillä laitteilla.

Australialaisen arkistoalan tutkijan Greg O`Shean (O`Shea, G., 1996) mukaan asiakirja on median ja informaation yhdistelmä, joka on tallennettu medialla tai mediaan. Sitä voidaan käyttää todisteena tai neuvottelua varten. Asiakirja on yksittäinen arkistointi-, tallenne- tai käsikirjoituskappale. Informaatioteknologiassa Shean mukaan asiakirja määritellään taas nimetyksi, rakenteiseksi tekstiyksiköksi, joka voidaan yksittäisenä yksikkönä varastoida, hakea ja vaihtaa järjestelmien ja käyttäjien kesken. Asiakirja on tietoa ja media, jolla se on tallennettu. Asiakirja on pysyvä ja se on sekä ihmisten että

koneiden tulkittavissa. Informaatioteknologian sekä arkistoinnin ja tallenteiden hallinnan määritelmät asiakirjalle ovat lähempänä toisiaan kuin tallenteen määritelmät.

Kanadalaisen tutkijan Charles Dollarin (Dollar, C., 1998, 10) mukaan asiakirja on tallennettua tietoa, jonka sisältö on välitetty tai esitetty aakkosnumeerisella tekstillä, vektoritietona, digitaalisena bittikarttakuvana, taulukkolaskentaohjelmana, tietokantoina, liikkuvina kuvina sekä äänidatana. Tällä tallennetulla tiedolla on loogiset ja fysikaaliset rajat, kuten esimerkiksi alku ja loppu, tai joku muu määritelty raja.

Arkisto ja arkistonmuodostaja:

Lain (Arkistolaki, 6§) mukaan arkistoon kuuluvat asiakirjat, jotka ovat saapuneet arkistonmuodostajalle sen tehtävien johdosta tai syntyneet arkistonmuodostajan toiminnan yhteydessä. Arkistonmuodostajalla tarkoitetaan yhteisöä tai henkilöä, jonka tehtävien hoitamisesta tai toiminnasta syntyy tai on syntynyt yksi tai useampia arkistoja (ks. liite 1). Arkistolla tarkoitetaan yhteisön tehtävien hoitamisesta tai henkilön toiminnasta kertyneiden asiakirjojen kokonaisuutta (ks. liite 1).

Greg O`Shean (1996) mukaan arkisto ja arkistointi määritellään informaatioteknologiassa hieman toisin kuin arkistoalalla. Arkisto on yhden tai useamman tiedoston kopio tai kopio tietokannasta, joka on tallennettu tulevaisuutta varten tai sen varalta, että alkupe-
räinen tieto vahingoittuu tai katoaa. Arkistointi on tiedostojen ja julkaisujen varastointia ennalta määritellyksi ajaksi. Arkistoinnin ja tallenteiden hallinnan ammattipiirissä arkistolla nähdään olevan kaksi merkitystä. Ensinnäkin arkisto on instituutio, joka on vastuussa keräämisestä, hankinnasta, säilyttämisestä ja arkistojen välisestä tiedonkulusta. Toisekseen arkistossa on tallenteiden luomisesta vastuussa olevien henkilöiden tai heidän seuraajiensa säilyttämät arkistointiarvoiset ei-nykyaikaiset tallenteet (määräajaksi tai pysyvästi säilytettävä hallinnollinen, taloudellinen, laillinen, todistusvoimainen ja informoiva tieto).

Tallenne:

Tallenne (record) määritellään Australiassa säädetyin Archives Act 1983 mukaan asiakirjaksi, joka sisältää mitä tahansa kirjoitettua tai tulostettua aineistoa. Tallenne voidaan käsittää myös objektiksi, joka sisältää äänitteen, koodatun muistilaitteen, magneettinauhan tai -levyn, mikrofilmin, valokuvan, filmin, kartan, piirroksen, mallin, maalauksen

tai muun graafisen tuotoksen. Se on tallennettu sisältämänsä, johonkin tapahtumaan, henkilöön, olosuhteisiin tai asiaan liittyvän informaation vuoksi. Tämän määritelmän mukaisesti tallenne on fyysinen objekti, kuten paperi, nauha, levy jne. Sähköisten tallenteiden kohdalla määritelmästä aiheutuu ongelmia. Ensiksi: sähköiset tallenteet eivät ole pelkkiä asiakirjoja eivätkä objekteja. Toiseksi: ne ovat riippumattomia välineestä, jolla ne kulloinkin sijaitsevat. Kolmanneksi: yksittäiset tallenteet ja niiden osat eivät fyysisesti sijaitse loogisessa järjestyksessä eivätkä usein samassa tilassakaan. Neljänneksi: määritelmä ei tarjoa mitään aitouteen liittyvää testausta. (O'Shea 1996)

Tallenne määritellään Australian kansallisarkistossa seuraavasti: "Kun tietoa luodaan, käytetään ja viestitetään viraston toiminnan ja liiketoimien yhteydessä, siitä muodostuu hyödyllistä todistusaineistoa eli tallenne viraston toiminnasta ja toimenpiteistä" (National Archives of Australia, Electronic Records). O'Shea (1996) mukaan "tallenne on sellainen, joka luodaan ja pidetään todisteena viraston tai yksilön toiminnasta, toimenpiteistä ja vuorovaikutuksesta. Jotta tallenne olisi todistusvoimainen, sillä täytyy olla sisältö, rakenne ja konteksti eli asiasisältö ja sen pitää olla osa tallennetta ylläpitävää järjestelmää".

Charles Dollarin (Dollar, C., 1992, 45) mukaan tallenne määritellään informaatioksi, joka liittyy fyysiseen entiteettiin eli kohteeseen. Siten tallenne on ensikäden todistus jostakin toimenpiteestä. Dollarin (1998, 10) mukaan tallenne syntyy joko liiketoimen tai fyysisen henkilön aikaansaamana ja se on tuotettu juridisesti tunnistettavassa ja hyväksyttävässä muodossa. Se pitää sisällään myös todisteen jostain tapahtumasta tai toimenpiteestä. Tallenteella tulisi olla yhteys muihin tallenteisiin, jotka on joko luotu tai jotka ovat syntyneet kanssakäymisen tuloksena. Täydellisessä tallenteessa on Dollarin mukaan seuraavia selvästi tunnistettavia elementtejä:

- Tekijän tai luomisesta vastuussa olevan henkilön tunnistus
- Päätöstoimenpide tai päätöksestä informoiminen
- Väline, joka tekee tallenteesta pysyvän ja siten tunnistettavan
- Muoto, johon sisältyvät elementteinä päiväys, luomisajankohta, teksti ja oikeaksi todistaminen (allekirjoitus)

Edellä selostetut tallenteen määritelmät sisältyvät arkistotieteeseen. Informaatioteknologiassa tallenne määritellään hieman toisin. Greg O`Shean (1996) mukaan tallenne on erilaisia ominaisuuksia sisältävistä tieto-objekteista koostuva kokonaisuus. Näillä tieto-objekteilla on tavallisesti identifioivat tunnukset. Tallenne on joukko tietoa, jota käsitellään yksikkönä tai yhden tai useamman tietoalkion ryhmänä.

Tallenteen ja asiakirjan keskeisimmät eroavaisuudet voidaan kiteyttää seuraavalla tavalla: Tallenne on todiste jostakin tapahtumasta tai toimenpiteestä. Tallenteet eivät ole siten pelkkää tietoa tai asiakirjoja, vaan ne luodaan ja säilytetään toimenpiteiden ja toiminnan suorituksena. Asiakirja on hyödyllinen todistusvoimainen tallenne, jolla pitää olla lisäksi sisältö, rakenne, konteksti eli asiasisältö (dokumenttien suhteet tekijöiden kanssa sekä suhteet dokumenttien välillä). Asiakirjan tulee myös olla osa tallennetta ylläpitävää järjestelmää.

2 ARKISTOINTIKÄYTÄNNÖT VALTIONHALLINNOSSA

Kysymykset, kuten 1) mitkä organisaatiot vastaavat arkistoinnista, 2) mitä säännöksiä liittyy asiakirjahallintoon ja 3) mitä pitkäaikainen arkistointi on arkistotieteen näkökulmasta ja miten sitä on hoidettu Suomessa, ovat keskeisiä tämän tutkimuksen lähtökohdissa. Tässä luvussa esitetään vastauksia näihin kysymyksiin.

2.1 Arkistoinnista vastaavat organisaatiot

Suomessa arkistointia säädetään 1.10.1994 voimaan tulleessa arkistolaissa (23.9.1994/831). Laki koskee erikseen määriteltyjä arkistonmuodostajia, joita ovat seuraavat (1§):

- valtion virastot, laitokset, tuomioistuimet ja muut lainkäyttöelimet sekä muut valtion viranomaiset
- kunnalliset viranomaiset ja toimielimet
- Suomen Pankki, Helsingin yliopisto, kansaneläkelaitos ja muut itsenäiset julkisoikeudelliset laitokset
- valtion ja kunnan liikelaitokset
- ortodoksinen kirkkokunta ja sen seurakunnat
- muut yhteisöt, toimielimet ja henkilöt niiden suorittaessa julkista tehtävää lain tai asetuksen taikka lain tai asetuksen nojalla annetun säännöksen tai määräyksen perusteella siltä osin kuin niille tämän tehtävän johdosta kertyy viranomaisten toiminnan julkisuudesta annetussa laissa (621/1999) tarkoitettuja asiakirjoja. (21.5.1999/689)

Julkishallinnossa arkistoinnista vastaavat viranomaisten arkistot. Muusta kuin julkishallinnon arkistoinnista vastaavat yksityisluontoiset arkistot. Näiden lisäksi myös kirjastoissa tehdään arkistointia, kun käsitellään ja säilytetään tallenteita. Suomessa viranomaisten arkistoja ovat: arkistolaitos (kansallisarkisto ja maakunta-arkistot), hallinnolliset keskusarkistot, muut valtion viranomaisten arkistot, kunnalliset arkistot ja kirkolliset arkistot. (Tampereen yliopiston historian laitos, Arkistotietokannat ja -luettelot, kuvatie-

tokannat, 1999). Seuraavassa on tarkemmin selostettu Suomessa toimivat a) viranomais-ten arkistot ja niiden tehtävät sekä b) yksityisluontoiset arkistot.

a) Viranomaisten arkistot:

- Arkistolaitos: Kansallisarkisto ja sen alaisina piirihallintoviranomaisina maakunta-arkistot
 - arkistolaitoksen tehtävänä on varmistaa kansalliseen kulttuuriperintöön kuuluvi-en asiakirjojen säilyminen ja niiden käytettävyys, edistää tutkimusta sekä ohjata, kehittää ja tutkia arkistointia (Arkistolaki, 4§).
 - valtionhallinnon pysyvästi säilytettävät aineistot, lukuun ottamatta hallinnollis-ten keskusarkistojen ja eduskunnan arkistoja (Tampereen yliopiston historian laitos 1999)
 - arkistolaitokseen voidaan ottaa säilytettäväksi muidenkin yhteisöjen ja yksityis-henkilöiden arkistoja (Arkistolaki, 17§).
- Hallinnolliset keskusarkistot
 - säilyttävät oman alansa aineistot
 - sota-arkisto
 - ulkoasianministeriön arkisto
- Muut valtion viranomaisten arkistot
 - Eduskunnan arkisto, ei siirretä Kansallisarkistoon
 - kaikilla valtion viranomaisilla on oma arkistonsa, jossa uusin materiaali säilyte-tään
- Kunnalliset arkistot
 - kaupunkien ja kuntien arkistot
 - ei yleensä siirretä Kansallisarkistoon tai maakunta-arkistoon (voidaan siirtää)
- Kirkolliset arkistot
 - seurakuntien arkistot

- ei yleensä siirretä Kansallisarkistoon tai maakunta-arkistoon (voidaan siirtää, vanhempia aineistoja on paljon siirretty)
- väestökirjanpitoon liittyvät vanhimmat aineistot mikrofilmattu (Tampereen yliopiston historian laitos 1999)

b) Yksityisluontoiset arkistot

- Järjestöt, ammattiyhdistykset, yritykset, talot, suvut, yksityishenkilöt ym. (Tampereen yliopiston historian laitos 1999)
- Keskeisiä ovat yksityiset keskusarkistot, joihin kerätään vanhaa aineistoa. Yksityisiä keskusarkistoja ovat:
 - Ammattiyhdistysarkisto, Kansan Arkisto, Keskustan ja Maaseudun Arkisto, Porvarillisen Työn Arkisto, Suomalaisen Kirjallisuuden Seuran Kansanrunousarkisto, Suomalaisen Kirjallisuuden Seuran Kirjallisuusarkisto, Suomen Elinkeinoelämän Keskusarkisto, Suomen Urheiluarkisto, Svenska centralarkivet, Svenska litteratursällskapet i Finland (Folkkultursarkivet, Språkarkivet, Historiska och litteraturhistoriska arkivet, Österbottens traditionsarkiv), Toimihenkilöarkisto, Työväen Arkisto ja Yksityiset keskusarkistot ry (Agricola – Suomen historiaverkko).
- yksityinen arkisto voidaan siirtää myös Kansallisarkistoon tai maakunta-arkistoon (Arkistolaki, 17§)
 - uusin materiaali aina yhteisöllä tai henkilöllä itsellään (Tampereen yliopiston historian laitos 1999).

Yksityinen arkisto tai siihen kuuluvia asiakirjoja voidaan arkistolain (17-19§) mukaan siirtää tai siirättää arkistolaitokseen tai jonkun muun laissa tarkoitetun arkiston säilytettäväksi ja hoidettavaksi. Jos yksityisen hallussa oleva tutkimuksen tai muun syyn vuoksi merkityksellinen arkisto tai yksittäinen asiakirja on vaarassa tuhoutua tai hävitä, arkistolaitoksella on oikeus lunastaa ko. arkisto tai asiakirja tai ottaa siitä jäljennös. Arkistolaitos voi myös määrätä, että yksityinen arkisto, kokoelma tai asiakirja on välittömästi siirrettävä arkistolaitokseen tai muuhun viranomaisen varmaan säilytyspaikkaan siihen asti, kunnes kysymys ko. kohteen lunastamisesta tai jäljentämisestä on saatu ratkaistua (20§).

Arkistojen lisäksi tallenteiden säilyttämisestä ja talletuksesta tuleville sukupolville huolehtivat myös kirjastot, lähinnä kansalliskirjastot, joissa kansallista kulttuuriperintöä säilytetään.

2.2 Asiakirjahallintoa koskevat säännökset

Asiakirjahallinto koskee asiakirjoja ja niiden käsittelyä. Asiakirjahallinnolla tarkoitetaan toimenpiteitä, joiden avulla kehitetään ja ohjataan yhteisön toimintaa koskevien asiakirjojen vastaanottamista, tuottamista, käsittelyä, tallentamista, säilytysarvon määrittelyä ja käyttöön saattamista. Valtionhallinnossa suuri osa paperitallenteista on asiakirjoja. Valtionhallinnossa arkistointi on osa asiakirjahallintoa. Asiakirjaa koskevat arkistointiperiaatteet siitä lähtien, kun se on saapunut arkistotoimen piiriin tai kun asiakirja on laadittu. Arkistolaissa määritellään asiakirjahallintoon liittyvät toimenpiteet ja arkistotoimen tehtävät. Suomessa arkistolain mukaan virasto tai laitos vastaa asiakirjahallintonsa suunnittelusta ja käytännön järjestämisestä. (Arkistolaitoksen suositus arkistonmuodostussuunnitelmista ; Documenta-asiakirjahallinnon opas).

Muissa maissa paitsi Suomessa ja Pohjoismaissa asiakirjahallinto ja arkistotoimi erotetaan toisistaan ja molempia hoitavat eri ammattiryhmät. Vasta siinä vaiheessa, kun asiakirjoja ei enää säännöllisesti tarvita, niistä valitaan arkistoon sellaiset, joilla katsotaan olevan pysyvää tutkimusarvoa. Suomessa ja Pohjoismaissa on ajateltu asiakirjahallinnon ja arkistotoimen liittyvän niin läheisesti yhteen, ettei niitä voida erottaa erillisiksi toiminnoksi (Arkistolaitoksen suositus arkistonmuodostussuunnitelmista ; Documenta-asiakirjahallinnon opas).

Asiakirjahallintoa hoidetaan valtionhallinnossa arkistolain sekä arkistolaitoksen ohjeiden mukaan. Arkistotoimen järjestäminen kunnassa kuuluu kunnanhallitukselle, jonka tulee määrätä arkistoinnista vastuussa oleva viranhaltija tai toimihenkilö. Kyseinen henkilö johtaa kunnan arkistointia ja arkistonmuodostusta sekä huolehtii kunnan pysyvästi säilytettävistä asiakirjoista. (Arkistolaki, 9§) Arkistotoimen vaatimukset on otettava huomioon arkistonmuodostajan tieto- ja asiakirjahallinnossa (7§).

Kaikkien arkistolaisissa määrättyjen arkistonmuodostajien, kuten valtionhallinnollisten virastojen ja laitosten, on tehtävä arkistonmuodostussuunnitelma (Arkistolaki, 8§). Valtionhallinnon yksiköt tekevät toiminnastaan arkistointiehdotuksen, jonka kansallisarkisto hyväksyy (Lagus, Antti J., Elektroninen arkistointi vaatii huolellisuutta. Kauppalehti Extra, Informaatiotekniikka, huhtikuu 2002). Arkistonmuodostussuunnitelma sisältää viraston tai laitoksen tehtävät ja asiakirjat, asiakirjojen tallennustavan, arkistointijärjestyksen, kirjaamis- ja rekisteröintitoimenpiteet, säilytyspaikan, säilytysajan työpisteessä, kokonaissäilytysajan, asiakirjojen käytettävyyden sekä arkistointi- ja säilytysvastuun (Arkistolaitoksen suositus arkistonmuodostussuunnitelmista ; Documenta-asiakirjahallinnon opas).

2.3 Pitkäaikainen arkistointi

2.3.1 Pitkäaikaisen arkistoinnin periaatteet arkistotieteessä

Arkistotiede on luonut perustan arkistoinnin käytännöille ja arkistojen toiminnalle. Asiakirjojen ja tallenteiden pitkäaikainen säilytys on ollut pääasiassa arkistolaitosten hoidossa. Keskusarkistot (Suomessa kansallisarkisto, maakunta-arkistot, hallinnonalan keskusarkistot ja virastojen keskusarkistot) ovat olleet perinteisten tallenteiden pääasiallisia säilytyspaikkoja. Tähän on ollut monia hyviä syitä. Ensiksi: keskusarkistot, joilla on ollut laillinen ja fyysinen kontrolli tallenteisiin, ovat pystyneet takaamaan ja varmistamaan tallenteiden yhtenäisyyden. Siihen on vaikuttanut arkistoiden sitoutuminen säilyttämään informaatiota sellaisena kuin se on ollut alun perin organisaatioissa, joista se on arkistoon tuotu. Toiseksi: on ollut halvempaa varastoida ei-aktiiviset perinteiset tallenteet keskusarkistoon kuin ylläpitää niitä aktiivisessa toimistoympäristössä. Kolmanneksi: käyttäjien on ollut paljon helpompaa ja halvempaa asioida keskusarkistoissa kuin käydä etsimässä asiakirjoja ja tallenteita eri organisaatioista, joissa ei-aktiiviset tallenteet olisivat saattaneet olla sekoittuneina aktiivisten tallenteiden kanssa. (Dollar 1992, 48-49)

Arkistotieteen perusta on alkuperän käsitteessä. Charles Dollarin (1992, 53) mukaan alkuperä tarkoittaa sitä, että ylläpidetään dokumenttien luomiseen liittyvää kontekstia (dokumenttien suhteet tekijöiden kanssa sekä suhteet dokumenttien välillä). Kontekstiin

sisältyvät tiedot dokumentin luomispaikasta ja dokumenttiin liittyvästä prosessista. Prosessiin sisältyvät tekijät, kuten 1) mihin lopputulokseen dokumentin kanssa päädyttiin, 2) kenelle se oli tarkoitettu, 3) milloin ja kuinka se päättyi vastaanottajalle ja 4) kuinka se tuli arkistojien käsiin. Jotta tallenteiden alkuperä voidaan ymmärtää, tulee ymmärtää toimintoja ja tapahtumia, jotka liittyvät tallenteiden luomistapahtumaan liittyvään informaatiojärjestelmään. Järjestelmä, jossa tallenteet luotiin, on osa tallenteen kontekstia. Arkistoihmiset ja arkistotieteen asiantuntijat kautta maailman ovat yhtä mieltä siitä, että tallenteen luomiseen liittyvän kontekstin säilyttäminen on tärkeää. Konteksti kertoo tallenteen alkuperästä. Tallenteen alkuperään vaikuttavat myös taidot, joiden avulla tallenteita luodaan. Kysymys on tallenteiden luojaan pätevyydestä ja osaamisesta. Pätevä henkilö pystyy huolehtimaan siitä, että tallenne luodaan ja säilytetään luotettavissa oloissa. Organisaatio ylläpitää tallenteitaan arkistointisuunnitelmiansa mukaan. Arkistointisuunnitelmat sisältävät organisaation vastuulle ottamia tehtäviä tai toimintoja. Organisaatio on toiminnassaan vastuussa siitä, että sen luomat ja ylläpitämät tallenteet ovat aitoja ja luotettavia. Sitten kun tallenteet siirretään arkistolaitokseen, vastuu niiden säilyttämisestä siirtyy arkistolaitokselle.

Paperitallenteiden alkuperä on yleensä ollut varsin helposti tunnistettavissa, koska paperiasiakirjakokoelmien alkuperään liittynyt informaatio ilmaistaan useissa paikoissa: organisaation taulukoissa ja manuaaleissa, menettelytavoissa, käytännöissä, fyysisessä tallenteiden järjestyksessä, vesileimoissa, kirjeiden otsikoissa, sisältö- ja kontekstielementeissä kuten kirjoittajan ja vastaanottajan nimissä, päivämäärissä, viittauksissa toisiin dokumentteihin tai arkistokokoelmiin jne. Lisäksi yksittäiset asiakirjat saattavat sisältää reunahuomautuksia, nimikirjaimia ja muuta sellaista kontekstuaalista informaatiota, josta ilmenee tallenteiden luomisen, käytön ja ylläpidon informaatioympäristö. (Dollar 1992, 48-49)

Tieto tallenteiden alkuperästä on tärkeä, sillä jos yksittäisen tallenteen alkuperästä ei ole varmuutta, ei sitä voida pitää luotettavana lähteenä tutkimuksellisessa eikä muussakaan käytössä. Alkuperä ja luotettavuus ovat tärkeitä. Arkistot ovat olleet vastuussa tallenteiden säilymisestä lukukelpoisina, eli niitä tulee säilyttää oikeanlaisissa olosuhteissa.

2.3.2 Pitkäaikainen arkistointi Suomessa

Arkistolaissa ja arkistolaitoksen säännöksissä määrätään, mitkä organisaatiot vastaavat valtionhallinnossa syntyvien asiakirjojen ja tallenteiden säilytyksestä ja arkistoinnista. Pitkäaikaisesta arkistoinnista vastaa tavallisesti arkistolaitos. Arkistonmuodostajien pysyvästi säilytettävät asiakirjat on arkistolain mukaan siirrettävä kansallisarkistoon, maakunta-arkistoon tai muuhun arkistoon siten kuin arkistolaitos erikseen määrää (Arkistolaki, 14§). Yleiset suuntaviivat arkistointiin on täten luotu arkistolaitoksessa, joka hoitaa tehtäviään arkistolain pohjalta. Arkistolain mukaan arkistotoimen tehtävänä on (7§):

- varmistaa asiakirjojen käytettävyys ja säilyminen
- huolehtia asiakirjoihin liittyvästä tietopalvelusta
- määrittää asiakirjojen säilytysarvo
- hävittää tarpeeton aineisto.

Muita arkistotoimeen kuuluvia tehtäviä ovat siirtäminen päätearkistoon, määräajan säilytettävien asiakirjojen hävittäminen, järjestäminen, kuvailu ja luettelointi sekä pysyvästi säilytettävien asiakirjojen luovuttaminen arkistolaitokseen tai muuhun hallinnon keskusarkistoon. (Arkistolaitoksen suositus arkistonmuodostussuunnitelmista ; Documenta-asiakirjahallinnon opas)

Virastoissa ja laitoksissa määritellään myös niille kertyvien asiakirjojen määräaikaisen säilytyksen pituus. Virasto ja laitos saa päättää siten omien asiakirjojen määräaikaisesta säilyttämisestä. (Documenta-asiakirjahallinnon opas) Sen sijaan asiakirjojen ja niiden sisältämien tietojen pysyvästä säilyttämisestä saa päättää vain arkistolaitos (Arkistolaki, 8§). Asiakirjan kohtaloon vaikuttaa se, kuinka tärkeäksi se arvioidaan säilyttämisen kannalta. Pysyvään säilytykseen määrätty asiakirjat tulee laatia pitkäaikaista säilytystä kestäviä materiaaleja ja säilyvyyden turvaavia menetelmiä käyttäen. Samoin tiedot tulee tallentaa edellä mainittuja ohjeita noudattaen. Asiakirjat tulee säilyttää siten, että ne eivät ole vaarassa tuhoutua, vahingoittua tai joutua asiattomaan käyttöön. Ne asiakirjat, jotka on määrätty säilytettäväksi tietyn ajan, tulee hävittää määräajan kuluttua umpeen siten, että niiden tietosuoja on varmistettu. (11-13§)

Asiakirjoilla ja asiakirjoihin sisältyvillä tiedoilla on elinkaari, joka alkaa asiakirjan laatimisesta tai vastaanottamisesta. Elinkaaren viimeiseen vaiheeseen päädytään eri käsitteilyvaiheiden kautta. Viimeisessä vaiheessa asiakirja joko laitetaan pysyvään säilytykseen

arkistossa tai hävitetään tarpeettomana. Arkistolaitoksen (Documenta-asiakirjahallinnon opas) mukaan asiakirjojen elinkaareen kuuluu kolme vaihetta: aktiivi-, passiivi- ja historiallinen vaihe:

- Aktiivivaiheessa asiakirjoja laaditaan, käsitellään, muokataan ja siirretään. Aktiivivaiheessa asiakirjat säilytetään niitä käsittelevässä työyksikössä tai viraston ja laitoksen kirjaamossa. Asiakirjan aktiivivaihe kestää muutamasta päivästä useisiin vuosiin. Aktiivivaiheessa tärkeitä tekijöitä ovat asiakirjojen käytettävyys, tuottamis- ja käsittelyprosessin tehokkuus sekä asiakirjojen hallinta. Aktiivivaiheen lopussa tarpeeton osa asiakirja-aineistosta hävitetään viraston tai laitoksen tekemän arkistonmuodostussuunnitelman mukaisesti.
- Passiivivaiheessa aktiivikäytöstä poistuneet asiakirjat siirretään lähi- tai päätearkistoon. Passiivivaiheessa asiakirjoja saatetaan vielä tarvita silloin tällöin. Passiivivaihe päättyy yleensä viimeistään 10-20 vuoden kuluttua asiakirjan laatimisesta. Passiivivaiheen lopussa suuri osa aineistosta hävitetään.
- Historiallisessa vaiheessa asiakirjat säilytetään tulevaisuudessa tapahtuvaa käyttöä, kuten tutkimusta ja kulttuuriin liittyviä toimenpiteitä varten. Asiakirjat säilytetään arkistolaitoksessa, hallinnonalan keskusarkistossa tai viraston keskusarkistossa. Passiivi- ja historiallisessa vaiheessa on tärkeää huomioida asiakirjojen esitys- ja säilytysmuotojen yleisyys, asiakirjojen tietosisällön muuttumattomuus sekä tiedon helppo löydettävyys ja käytettävyys. Arkistolaitoksen mukaan alkuperäisestä asiakirja-aineistosta pysyvästi säilytettäviä asiakirjoja on noin 5-15%.

Koska valtionhallinnolliset asiakirjat ja niiden sisältämät tiedot muodostavat erilaisia ryhmiä, on selvitettävä mihin asiakirjoihin kuhunkin valtionhallinnolliseen tehtävään liittyvät tiedot sisältyvät. Tämän jälkeen on mahdollista määritellä yksittäisen asiakirjaryhmän säilytysaikatarve. Asiakirjaryhmistä on muodostettava sarjoja, joiden koostumus tulee päättää. Kunkin sarjan sisälle tulevien asiakirjojen järjestys on myös päätettävä. (Arkistolaitoksen suositus arkistonmuodostussuunnitelmista) Monissa maissa paperidokumenttien arviointi tapahtuu vasta elinkaaren loppuvaiheessa. Näin ei ole kuitenkaan Suomen arkistolaitoksen piirissä, jossa on jo pitkään käytännössä yritetty suunnitella dokumenttien arviointia etukäteen. Näin arkistoja luotaessa tai niiden saapuessa

tiettyyn organisaatioon on jo tiedossa kunkin dokumentin oma säilyttämisaika. Kun asiakirjojen arviointi suunnitellaan jo etukäteen, on arkistoidessa mahdollista laittaa tietty dokumentti alusta lähtien sellaiseen paikkaan, jossa ovat kaikki saman pituiseksi määräajaksi säilytettäväksi tarkoitetut asiakirjat. (Pohjola 1999)

Pitkäaikaiseen arkistointiin liittyy seulonnan käsite, joka tarkoittaa pitkäaikaiseen säilytykseen tarkoitettujen ja määräajaksi säilytettäväksi tarkoitettujen asiakirjojen erottamista toisistaan. Ennen seulontaa asiakirjat pitää arvioida, eli määrätä niiden arvon ja tärkeyden mukainen säilytysajan pituus sekä päättää niiden säilytys- ja arkistointitavoista. Jotta seulonta olisi mahdollisimman yksinkertaista toteuttaa, tulisi pysyvästi säilytettävät asiakirjat pitää erillään määräajan säilytettävistä asiakirjoista. Seulontaan liittyy myös salassa pidettävien asiakirjojen erottaminen julkisista asiakirjoista. (Arkistolaitoksen suositus arkistonmuodostussuunnitelmista)

Osa asiakirjoista on muita tärkeämpiä. Arkistolaitoksen mukaan erityistä suojelua tarvitsevia asiakirjoja on valtionhallinnossa kahdenlaisia:

- toiminnan jatkumisen kannalta välttämättömät asiakirjat
- tutkimuksellisista, juridisista tai taloudellisista syistä suojeltavat asiakirjat (Arkistolaitos, Asiakirjojen säilyttäminen poikkeuksellisissa oloissa).

Säilytysarvon kannalta ehdottoman välttämättömistä ja tärkeistä asiakirjoista tulee ottaa varmuuskopioita mahdollisen alkuperäisen aineiston tuhoutumisen varalta. Varmuuskopioita valmistetaan esimerkiksi mikrofilmaamalla. Varmuuskopiot tulee säilyttää eri paikkakunnilla mahdollisen tuhon varalta. (Arkistolaitos, Asiakirjojen säilyttäminen poikkeuksellisissa oloissa)

3 SIIRTYMINEN ELEKTRONISEEN ARKISTOINTIIN

1980-luvun loppupuolelle asti kirjoituskone oli pääasiallinen asiakirjojen luomisessa käytetty apuväline. Vasta tietokoneiden muuttuessa henkilökohtaisemmiksi työvälineiksi, niitä alettiin käyttää työvälineinä. Elektroniseen arkistointiin siirtyminen on tapahtunut vähitellen tietokoneiden, digitaalisten välineiden ja tallenteiden määrän kasvun ja tietoverkkojen kehittymisen myötä. Tässä luvussa käsitellään aluksi tietokoneiden kehitystä ja luodaan siten pohja koko digitaalisen aikakauden ja sen tuotosten ymmärtämiselle. Tässä luvussa selvitetään myös digitaalisten tallenteiden tuomia haasteita ja ongelmia niitä tuottaville ja arkistoiville organisaatioille.

3.1 Digitaaliseen aikaan siirtyminen

Tietokoneiden kehitysvaiheet

Morrisin (Morris, R.J., 1998, 32) mukaan elektronisesti aikaansaatu tieto on luotu ja luodaan hyvin erilaisissa organisaatio- ja teknologiaympäristöissä. 1950-luvulta lähtien on ollut monenlaisia vaiheita elektronisen tiedon luomiskäytännöissä. 1950-luvulla oli käytössä suurtietokoneet.

Oeschin (1993, 43-44, 52, 54, 58, 60) mukaan tietokone oli kehityksensä alkuvaiheessa tarkoitettu lähinnä tehokkaaksi laskimeksi sekä käsittelemään alfanumeerista ja tekstitietoa, mutta jo siinä vaiheessa sillä oli tärkein tietokoneen ominaisuus eli kaikenlaisen tiedon digitalisoiminen. Digitalisoiminen tarkoittaa tiedon purkamista alkeisyksiköihin, bitteihin. Bitti merkitsee elektronisen komponentin sähköistä ei- tai kyllä-tilaa eli on- tai off-tilaa.

Morrisin (1998, 33) mukaan 1970-luvun lopulla käytössä oli jo paljon PC-tietokoneita ja sanojen prosessointia tehtiin. Voimakas kasvu PC-tietokoneiden laajentumisessa tapahtui 1980-luvulla, jolloin tietokoneista tuli henkilökohtaisia työvälineitä. Tuonaikaisissa PC-tietokoneissa oli alhainen kapasiteetti ja alkeelliset viestintätoiminnot, mutta vähitellen niiden tehokkuus lisääntyi. Tuo aika oli myös ns. lerppujen (silloiset tallenuslevykkeet) aikaa. Mikroprosessoriteknikan kehittymisen myötä tietokoneet tulivat yhä pienemmän kokoisiksi (Oesch 1993, 44). 1990-luvulla käyttöön tulivat ns. korppu-

tallennuslevykkeet, mutta kaikista tärkeintä kehityksen kannalta oli maailmanlaajuisen tietoverkon Internetin ja World Wide Webin tuleminen.

Oeschin (1993, 56-57) mukaan mikroprosessori- ja komponenttitekniologian miniatyrisoituminen on ollut hyvin voimakasta. Mikrosiruista on tullut kuluttajaelektroniikan perusosia. Digitaalitekniologian perustana ovat signaalinprosessointi, videon pakkaus ja nestekidenäytöt. Digitaalista signaalinprosessointia on käytetty aluksi tutkissa ja sotilaallisissa kuvankäsittelysovelluksissa. Digitaaliset signaaliprosessorit eli DSP-sirut pystyvät prosessoimaan ja muuntamaan nopeasti analogisia datavirtoja digitaaliseen muotoon. Tämän vuoksi ne soveltuvat hyvin erilaisiin multimediasovelluksiin, joita tulee koko ajan lisää tekniikan kehityksen myötä. Digitaalisen signaaliprosessointitekniikan avulla CD-soittimien, videokameroiden ja nauhureiden kokoa on voitu pienentää. Oeschin mukaan voidaankin 1980-lukua kutsua henkilökohtaisten tietokoneiden vuosikymmeneksi ja 1990-lukua tietoliikenteen vuosikymmeneksi. 1990-luvulla tiedonsiirto ja tiedonkäsittely tehostuivat valtavasti. Laajakaistaverkot tulivat ja kaistat laajenevat jatkuvasti mahdollistaen tehokkaamman ja suuremman tietomäärän tiedonsiirron samanaikaisesti. Maailmamme käy koko ajan digitaalisemmaksi. Työpaikoille ja koteihin tulee lisää digitaalisia laitteita ja välineitä. Televisio ja tietokone ovat yhdistymässä yhdeksi ja samaksi laitteeksi.

Mediat ja tallenteet

Tietoa on säilytettyä erilaisissa fyysisissä välineissä, medioissa ja formaateissa. Säännöllisesti käytettyjä välineitä Michael Leskin (Lesk, M., 1996 ; Lesk, M., 1992) mukaan ovat olleet ja ovat seuraavat:

Vanhat ja poisjäämässä olevat tallenteet

- Analogiset tallenteet
 - nykyiset analogiset tekniikat: C-kasetit, valokuvat, elokuvafilmit ja VHS-videonauha
 - vanhentuneita tai puoli- tai vanhentuneita analogisia tekniikoita: gramofonilevyt (33, 45 ja 78 rpm (kierrosnopeus)), Beta videonauha, 8-raittaiset äänikasetit ja kelaääninauhat
 - perinteisten äänikasettien sovitimet

- Digitaaliset tallenteet
 - 8 -ja 5.25 -floppy ja 3.5-tuumainen ”korppu”
 - digitaalinen lineaarisesti nauhoittava magneettinauha (1/2 tuumaa, 200, 556, 800, 1600, and 6250 bpi:n (bits per inch) tiheyksiset)
 - reikäkortti (rectangular hole ja round-hole) ja reikänauha
 - vanhentuneet suuremmat tietokonelaitteistot: reikäkortit, erilaiset siirrettävät levypakat, erilaiset magneettinauhat, kuten Dectape, 7-raitainen puolituumainen nauha, 9-raitainen nauha alle 1600 bpi:n tiheyksillä, punched paper tape (reikäpaperinauha) ja monet muut laitteet, kuten laitteet, joissa on yhä kelattava teräsnauha, jota esim. Univac I käyttää

Uudet tallenteet

- Digitaaliset tallenteet
 - erilaiset lineaariset nauhakasetit, kuten 3480, QIC ja DLT (Digital Linear Tape)
 - magneettis-optiset nauhakasetit (enimmäkseen 5.25 tuumaa, mutta jotkut myös 8.5 -ja 3.5 tuumaa)
 - eripituiset ja -tiheyksiset QIC-tyyppiset nauhakasetit, DAT –nauhakasetit (magneettinauha)(digital audio tape)
 - suuremmat tietokonelaitteet: Exabyte 8mm videonauhakasetit (käytetään digitaalisen tiedon tallentamiseen), 9-raitaiset korkeatiheyksiset nauhat (joista 1600 ja 6250 bpi yleisimmät), magneettisoptiset (jatkuvasti monipuolistuvat koot ja tiheydet) ja WORM-levyt (monipuolistuvat koot ja tiheydet)
 - WORM optinen varasto (12 tuumaa, 5.25 tuumaa, CD-ROM-levyt) (Lesk 1996 ; Lesk 1992)
 - MiniDisc (MD)-levyt ja Digital Compact Cassette (DCC) (Oesch 1993, 69)

3.2 Digitaalisuuden haasteet tallenteita tuottaville ja arkistoiville organisaatioille

3.2.1 Eliniän lyhenemisen synnyttämät ongelmat

Informaatioteknologian nopea kehitys ei ole pelkästään hyvä asia, koska se aiheuttaa käytetyn teknologian nopeaa vanhentumista. Suuria muutoksia tapahtuu teknisissä ratkaisuissa, ja ne syrjäyttävät tai korvaavat aikaisemmin luodut tekniset ratkaisut. Mackenzie Owenin ja Wallen (1996, 7-8) mukaan elektronisen informaation elinikä on merkittävästi lyhyempi kuin paperille tulostetun informaation elinikä. Nyky-yhteiskunnassa tiedon katsotaan menettävänsä arvonsa yhä lyhyemmässä ajassa. Esimerkiksi elektronisten julkaisujen sisältämä tieto vanhentuu monilla aloilla jo 1-2 vuodessa. Tiedosta tulee tavallaan kertakäyttötavaraa, joka heitetään pois käytön jälkeen. Uudemmat ja paremmat mediat (esim. 6250 bpi:n magneettinauhat) korvaavat vanhemmat (esim. 1600 bpi:n magneettinauhat). Tällä hetkellä käytettävien varastointimedioiden arvioitu elinikä on korkeintaan 10 vuotta. Teknologian vanhentumista tapahtuu myös käyttöjärjestelmissä ja ohjelmistosovelluksissa, kun valmistajat tuovat markkinoille uusia tuotteita uusilla toiminnoilla. Tällä hetkellä käytössä olevat laitteistot, ohjelmistot ja dataformaatit tulevat vanhentuneiksi todennäköisesti 10-25 vuodessa. Mackenzie Owen ja Walle puhuvat teoksessaan jopa 25 vuodesta, jonka kuluessa laitteistot, ohjelmistot ja dataformaatit vanhentuvat. Se saattoi vaikuttaa siltä vielä 1990-luvun puolessa välissä, jolloin he ovat teoksensa kirjoittaneet. Tällä hetkellä tuo 25 vuotta vaikuttaa monin kerroin liian pitkältä aikaväliltä, varsinkin kun ohjelmistot saattavat joutua poistetuiksi käytöstä jopa muutaman vuoden väliajoin. Yksisy syy tähän on maailmanlaajuisen monopoliaseman tuotteillaan hankkinut Microsoft, joka tuo markkinoille lähes joka vuosi uuden käyttöjärjestelmän, uudet ohjelmistot jne. Valmistajat, kuten Microsoft on pyrkinyt tarjoamaan yhteensopivuuden myös vanhempien laitteiden ja ohjelmien kanssa useiden ohjelmistosukupolvien yli. Tämä ei ole kuitenkaan mikään pysyvä ratkaisu, koska ei ole mahdollista laajentaa tätä toiminnallisuutta kovin pitkälle tulevaisuuteen. Dollarin (1998, 18-19) mukaan teknologian vanhentumista ei voida katsoa tapahtuneeksi niin kauan, kun nykyiset laitteet jatkavat toimintaansa kunnolla ja kun varaosia ja järjestelmien ylläpitoa on saatavilla. Joka tapauksessa kaikki varastointimedit, joiden tulee varmistaa aitojen sähköisten tallenteiden saatavuus kautta aikojen, vanhenevat teknologisesti jossain vaiheessa.

Suhtautuminen elektronisten tallenteiden tuomiin haasteisiin ja ongelmiin vaikuttaa siihen, kuinka hyvin ja kuinka käyttökelpoisina tallenteet pystytään säilyttämään tuleville sukupolville. Digitaalisuuden aiheuttamat ongelmat ovat hyvin monitahoisia. Amerikkalaisen arkistoteoretikon David Bearmanin (Arkistoviesti 2/1998) mukaan odotettavissa olevat vaikeudet, jopa mahdolliset sabotaasit, kannustavat organisaatioita sähköisen asiakirjahallinnon kehittämiseen niin, että aineistojen luotettavuus ja aitous ja niiden turvaamisen käytettävät keinot otetaan huomioon. Koska tähän Bearmanin mukaan menee vielä aikaa, saattaa pahimmassa tapauksessa suuri osa tällä hetkellä tietojärjestelmissä pyörivistä merkittävistä asiakirjoista tuhoutua.

Ihmisten ammatilliset näkemykset vaikuttavat siihen, kuinka he suhtautuvat elektronisten tallenteiden pitkäaikaisen saatavuuden ja käytettävyyden turvaamisen ongelmiin. Charles Dollar (1998, 6) on havainnut työskennellessään tutkimuspiireissä ja ollessaan tekemisissä eri alojen tutkijoiden kanssa, että on olemassa kaksi koulukuntaa suhtautumisessa elektronisten tallenteiden pitkäaikaisen käytettävyyden ja saatavuuden ongelmiin. Toiseen koulukuntaan kuuluu tietokone- ja informaatioteknologian asiantuntijoita, jotka ajattelevat elektronisten tallenteiden pitkäaikaisen saatavuuden ja käytettävyyden olevan ensisijassa teknologinen asia. He ihmiset kiinnittävät vähän huomiota tallenteiden aitouden ja luotettavuuden varmistamiseen. Toiseen koulukuntaan kuuluu taas kirjasto-, arkisto- ja tallenteiden hallinnan alalla työskentelevät ihmiset. He puolestaan ajattelevat elektronisten tallenteiden pitkäaikaisen saatavuuden ja käytettävyyden olevan yhtä paljon ideologinen kuin teknologinenkin asia. Tämä jälkimmäinen näkemys on yleinen myös useissa viime aikoina valmistuneissa tai meneillään olevissa tutkimusprojekteissa, joiden perusajatuksena on sähköisten tallenteiden saatavuus.

3.2.2 Näkemyksiä arkistokäytäntöjen muuttamistarpeista

Perinteisten tallenteiden pitkäaikaisessa arkistoinnissa keskusarkistot ovat olleet keskeisessä roolissa. Digitaalisella aikakaudella perinteisiin arkistointikäytäntöihin vaikuttaneet tekijät eivät ole kuitenkaan enää yhtä määrääviä (Dollar 1992, 53). Informaatioteknologian kehitys tulee muuttamaan yleisiä arkistointikäytäntöjä ja arkistolaitoksen piirissä työskentelevien ihmisten työnkuvaa ja työtehtäviä. Vaasan maakunta-arkiston johtajan Elias Orrmannin (Orrman, Elias, 1999, Structural Complexity of Electronic

Records as a Factor Guiding Decisions on Permanent Retention. DLM-Forum) mukaan keskusarkistot, jotka ovat tähän asti hoitaneet pääosin paperisten tallenteiden pitkäaikais säilytyksen, eivät pysty teknisesti hoitamaan esimerkiksi sähköisten tallenteiden migraatiota eli vanhoista järjestelmistä uusiin järjestelmiin siirtämistä. Elektroniset tallenteet saattavat olla hyvinkin monimutkaisia ja niiden käytettävyyden turvaamisessa pitkäksi aikaa vaaditaan teknistä tietämystä ja osaamista käyttöjärjestelmistä ja ohjelmistoista. Arkistoilta puuttuu teknistä osaamista ja resursseja. Ongelma on sekä taloudellinen että tekninen. Dollarin (1992, 53-54) mukaan tulevaisuudessa elektronisten dokumenttien käsittely-, arkistointi-, muuttamis- ja siirtämistoimenpiteisiin tulee kuluun huomattavan paljon rahaa materiaalien määrän koko ajan kasvaessa, sillä nämä kustannukset kasvavat jopa kumulatiivisesti. Mitä monimutkaisempia tallenteet ovat, sitä kalliimmaksi tulee niiden käytettävyyden ylläpito. Taloudellisiin näkökohtiin ei Dollarin mielestä ole juurikaan kiinnitetty huomiota. Esillä ovat olleet vain tekniset kysymykset, jotka nekin ovat luonnollisesti tärkeitä.

Dollarin (1992, 65-66, 79-80) mukaan arkistolaitoksen tehtäviä ja vastuita tulisi miettiä uudelleen. Hän esitti jo vuonna 1992 näkemyksensä, että sähköisiä dokumentteja luovien organisaatioiden tulisi ottaa itse hoitaakseen omien asiakirjojensa säilytys, osittain edellä mainitun kustannusten kiihtyvän kasvun vuoksi. Kun tallenteita luovat organisaatiot itse vastaisivat pitkäaikaisesta säilytyksestä, olisivat kustannukset pienemmät, koska ne huolehtivat tallenteita ylläpitävistä järjestelmistä.

Australialainen David Roberts (Reed, Barbara, 2000, *Electronic Records Management in Australia*) on Dollarin kanssa samoilla linjoilla siinä, että osan tallenteiden ylläpidosta ja arkistoinnista hoitaisivat niitä luovat organisaatiot ja osan taas arkistot. Hän jakaa selvästi organisaatioiden ja arkistoiden tehtävät omiin kategorioihin. Tallenteita tuottavat organisaatiot ylläpitävät omia tallenteitaan tietoteknisessä ympäristössä ja siirtävät ne tarpeen vaatiessa uusiin laitteistoihin ja ohjelmistoihin. Arkistojen vastuu tallenteiden suhteen on säännellä standardeita ja tarkkailemalla niiden käyttöä, koska näillä seikoilla on keskeinen asema elektronisten tallenteiden ja niiden käytettävyyden ylläpito menetelmissä. Näin arkistot huolehtisivat tallenteiden yhdenmukaisista muodoista, formateista ja standardeista, kun taas tallenteita tuottavat organisaatiot huolehtisivat tallenteista ja niiden käytettävyydestä teknisessä ympäristössä ja sen muutoksissa. Higgsin (Higgs, E., 1998, 145) mukaan tulevaisuudessa arkistotyö kuuluu mahdollisesti osana

informaation hallintaan liittyvään työhön. Varsinaiset informaation hallintaan liittyvät toimenpiteet edeltävät tallenteiden luomiseen liittyviä toimia. Arkistoinnin asiantuntijoiden tulisi näin ollen laajentaa tietämystään ja osaamistaan opettelemalla tulevaisuuden arkistoinnissa tarvittavia taitoja. Higgsin (1998, 189-190) mukaan arkistojat eivät mahdollisesti enää tulevaisuudessa hallitse arkistoissa olevia ns. ”kuolleita asiakirjoja” eli historiallisessa vaiheessa olevia asiakirjoja, vaan he työskentelevät asiakirjoja (ja arkistoja) luovissa organisaatioissa käyttämällä hyväksi arkistointiperiaatteita ohjelmistojärjestelmien kehityksessä.

Australiassa niin kuin myös Suomessa ja muuallakin maailmassa valtionhallinnollisten tallenteiden arkistoinnista ovat huolehtineet arkistot. Tallenteiden luojina ja ylläpitäjinä ovat olleet eri valtionhallinnolliset organisaatiot, virastot ja laitokset. Digitaalisella aikakaudella perinteiset menettelytavat eivät enää sellaisenaan riitä, koska elektronisten tallenteiden luominen, ylläpito ja arkistointi on paljon mutkikkaampaa kuin paperisten tallenteiden ja mikrofilmien kohdalla. Kuten edellä tuli ilmi, Charles Dollar arkistoinnin asiantuntijana on esittänyt kritiikkiä siitä, että teknologia ja sen asiantuntijat ovat liikaa sanelleet kehityksen kulkua ja vallalla olevaa keskustelua. Jotta tallenteet onnistuttaisiin säilyttämään tulevaisuuden tarpeita varten, on tärkeää tarkastella ongelmia mahdollisimman monelta kannalta niin teknisistä, taloudellisista kuin arkistoinnin näkökulmista.

Tässä kappaleessa on käsitelty varsin yleisellä tasolla digitaalisuuden tuomia uusia haasteita ja ongelmia, joita elektroniset tallenteet aiheuttavat niitä luoville ja niiden säilytyksestä ja arkistoinnista vastuussa oleville organisaatioille. Seuraavassa luvussa käsitellään tarkemmin ja syvällisemmin digitaalisuuden aiheuttamien muutosten tuomia ongelmia sekä selvitetään mahdollisia ongelmien ratkaisukeinoja.

4 KÄYTETTÄVYYS, ALKUPERÄISYYS JA MUUTTUMATTO- MUUS ELEKTRONISESSA ARKISTOINNISSA

Kuinka voidaan varmistaa, että elektroniset tallenteet ovat käytettäviä muutaman vuoden tai vuosikymmenen kuluttua? Minkälaisia ovat alkuperäiset ja muuttumattomat tallenteet arkistotieteen näkökulmasta ja millaisilla toimenpiteillä näiden ominaisuuksien säilyminen voidaan varmistaa eri tutkijoiden mielestä? Nämä ovat kysymyksiä, joihin tässä luvussa etsitään vastauksia.

4.1 *Arkistotieteessä ja arkistolaitoksessa tallenteille asetetut vaatimukset*

Arkistotiede asettaa sellaisia vaatimuksia sähköisten tallenteiden pitkäaikaissäilytykselle, joita ei välttämättä puhtaasti teknologian puolelta katsottuna osattaisi ottaa ollenkaan huomioon. Arkistosäilytyksen kannalta sähköisten tallenteiden pitkäaikaissäilytyksessä on tärkeää turvata ja varmistaa tallenteiden käyttökelpoisuus, autenttisuus (alkuperäisyys), muuttumattomuus, luotettavuus ja aitous. Luotettavuus ja aitous voidaan katsoa liittyvän autenttisuuteen eli alkuperäisyyteen ja eheyteen eli muuttumattomuuteen. (Moisio, R., 2002) Suomen arkistolaisissa arkistolaitoksen yhdeksi tehtäväksi määrätään varmistaa kansalliseen kulttuuriperintöön kuuluvien asiakirjojen säilyminen ja niiden käytettävyys (Arkistolaki, 4§). Lisäksi on tärkeää varmistaa tallenteiden luettavuus, selvyys ja ymmärrettävyys, tunnistettavuus, saatavuus, kapselointi yhdessä viitetietojen kanssa sekä palautettavuus.

- *Aitous ja luotettavuus* tarkoittaa sitä, että sähköisillä tallenteilla on arvovaltaa ja uskottavuutta. Ne ovat todisteena ja muistona toiminnasta, johon ne sisältyvät luonnollisena sivutuotteena. Niiden sisältö perustuu luotettaviin tosiasioihin, mikä on voitu todeta jo aiemmin niiden luomiskontekstista ja pitkäaikaisesta käytöstä. Aidot tallenteet ovat myös sellaisia, ettei niitä ole muutettu, vaihdettu eikä muutenkaan vahingoitettu. (Dollar 1998, 12, 25)
- *Luettavuus* tarkoittaa, että sähköisten tallenteiden pohjalla oleva bittivirta voidaan prosessoida tallenteet luoneessa, varastoivassa sekä tulevaisuudessa digitaalista tie-

toa varastoivassa tietokonejärjestelmässä tai – laitteessa. (Dollar 1998, 21)

- *Selvyys* tarkoittaa niin selkeää binääristä bittivirtaa, että tietokone ymmärtää ja pysyy tulkitsemaan, mitä ykköset ja nollat esittävät, jolloin sovellusohjelmassa osataan tehdä oikeat toimenpiteet niihin perusteella. (Dollar 1998, 22)
- *Tunnistettavuus* tarkoittaa tallenteiden järjestämistä ja kuvailua sellaisella tavalla, että käyttäjien tai järjestelmän on mahdollista erotella tieto-objekteja toisistaan niiden yksilökohtaisen ominaisuuden kuten tallenteen nimen, tallenteen ID-numeron tai materiaalin muodon perusteella. (Dollar 1998, 23)
- *Kapseloidut* sähköiset tallenteet ovat itse viittaavia kaikessa tiedossa, kuten metatietossa, sisällössä ja multimedialinkeissä, jotka liittyvät loogisena tai fyysisenä yksikönä olevaan tiettyyn tallenteeseen tai moninkertaisiin tallenteisiin. Kun kaikki sähköisen tallenteen loogiset komponentit yhdistetään yksittäiseksi fyysiseksi tallenteeksi, on kyseessä yksi kapseloinnin muoto, joka voi auttaa vähentämään multimediatyyppisten tallenteiden monimutkaisuutta. Kapselointi voi myös lievittää ja helpottaa sovellusohjelma- tai käyttöjärjestelmäversioiden yhteensopivuusongelmia ja vaikutuksia. Riippuvuus tietyistä versioista on yksi pääasiallinen este sähköisten tallenteiden pitkäaikaiselle saatavuudelle. Kapseloinnissa voidaan käyttää Standard Generalized Markup Language:a (SGML). Dokumenttityypimäärittelyä (DTD) voidaan käyttää siten, että voidaan tunnistaa kapseloitujen tallenteiden looginen rakenne. Dokumenttityypimäärittely voidaan suunnitella niin, että voidaan ottaa huomioon hierarkiset suhteet eri osasten välillä tai multimediatallenteet. (Dollar 1998, 23)
- *Saatavuus* tarkoittaa sitä, että erilliset tieto-objektit tai niiden osat voidaan tuoda esiin. Tieto-objektien saatavuus on suoraan yhteydessä ohjelmistosovelluksen ja käyttöjärjestelmän toimintaan. Tietokoneympäristössä saatavuus vaatii avaimia tai pointtereita, jotka linkittävät yhteen tieto-objektien loogisen rakenteen (tietokentät tai tekstijonot) ja fyysiset muistipaikat. Pointterit sallivat erottelun tieto-objektien välillä ja keskuudessa. (Dollar 1998, 23-24)

- *Ymmärrettävyys* tarkoittaa sitä, että sähköisten tallenteiden tulee olla selviä sekä tietokoneelle että ihmisille. Ihmisille selvät sähköiset tallenteet välittävät tietoa. Kun varmistetaan sähköisten tallenteiden ymmärrettävyys, ollaan tekemisissä hieman eri asioiden kanssa kuin paperitallenteiden yhteydessä. Paperitallenteissa on mukana niiden luomis- ja käyttökonteksti niiden fyysisen olemuksen yhteydessä. Sen sijaan sähköisten tallenteiden luomis- ja käyttökonteksti liittyy tallenteeseen tavallisesti loogisesti eikä fyysisesti, jolloin käyttäjä ei sitä fyysisen olemuksen yhteydessä näe. Metatiedolla on keskeinen osa sähköisten tallenteiden ymmärrettävyydessä, koska silloin jos sitä ei ole tallenteiden mukana, eivät tallenteet ole ymmärrettäviä. (Dollar 1998, 24)
- *Palautettavuus* tarkoittaa sähköisten tallenteiden esittämisen- tai tulostamismahdollisuutta niin, että niillä on sama looginen ja fyysinen rakenne sekä sisältö kuin niillä oli luomis- tai vastaanottohetkellä. Tässä mielessä täydellinen palautettavuus liittyy alkuperäistä tallennetta vastaavaan kopioon (imitative copy)(ks. kpl 4.3.3, s.22), jossa ovat nämä tallenteiden luotettavuuden ja aitouden säilyttämiseen vaadittavat ominaisuudet. (Dollar 1998, 24)

4.2 Elektronisten tallenteiden arkistoinnille aiheuttamat ongelmat

Amerikkalaisen arkistoteoreetikon David Bearmanin (Arkistoviesti 2/1998) mukaan odotettavissa olevat vaikeudet, jopa mahdolliset sabotaasit, kannustavat organisaatioita sähköisen asiakirjahallinnon kehittämiseen niin, että aineistojen luotettavuus, aitous ja niiden turvaamiseen käytettävät keinot otetaan huomioon. Nimittäin jos asiakirjahallintoa ei saada tarpeeksi nopeasti kehitettyä, saattaa pahimmassa tapauksessa tietojärjestelmissä pyörivistä merkittävistä asiakirjoista suuri osa tuhoutua.

Informaatioteknologian kehityksen myötä elektroniset asiakirjat eivät enää välttämättä ole yhdenmukaisia paperisten vastineidensa kanssa, kuten aikaisemmin oli tilanne esimerkiksi master- nauhojen kohdalla. Nykyiset elektroniset asiakirjat voivat sisältää multimediaa (ääntä, videoita, grafiikkaa, animaatioita), integroitua relaatiotietokantoja ja useamman organisaation yhdistäviä monimutkaisia tietokantoja (Arkistolaitoksen suosittus arkistonmuodostussuunnitelmista ; Zweig, R.W., 1998, 54). On huomattavan tärke-

ää, että Suomen arkistolaisissa huomioidaan myös sähköisessä muodossa oleva asiakirja (kts. kappale 1.4), mikä on tärkeää jo senkin vuoksi, että arkistolaki on perustana arkistosäännöksille ja –käytännöille Suomessa. Kuitenkaan se seikka ei enää yksin sellaisenaan riitä, kun ollaan tekemisissä multimedialla ja tietokantoja sisältävien asiakirjojen ja tallenteiden kanssa. Ongelman ydin elektronisten tallenteiden pitkäaikaisessa arkistoinnissa ei ole asiakirjan määritelmässä. Se on pieni tekijä verrattuna siihen, kuinka elektronisten tallenteiden luotettavuus, alkuperäisyys ja eheys saadaan turvattua.

Yksi tärkeä ero paperitallenteiden ja elektronisten tallenteiden välillä on tallenteen ja sen kontekstuaalisen informaation suhde. Paperitallenteissa ja yleensä perinteisissä tallenteissa fyysiset, muodolliset ja kontekstuaaliset osat, komponentit, ovat erottamattomasti yhdessä. Tämän ansiosta yksittäisen tallenteen fyysinen sijainti suhteessa toiseen tallenteeseen ilmaisee sen, kuinka tallenteet luotiin ja kuinka niitä käytettiin määrättyinä ajankohtana. Kun perinteisten tallenteiden alkuperäinen järjestys säilytetään, varmistetaan, että tallenteiden välisiä ja keskinäisiä suhteita pidetään yllä. Huonona puolena tässä on edellä mainittujen tallenteiden välisten ja keskinäisten suhteiden tuleminen vaikeammin tunnistettaviksi, kun jostain syystä tallenteiden alkuperäinen järjestys muuttuu. Hyvä puoli perinteisten tallenteiden kohdalla on niiden fyysisten, muodollisten ja kontekstuaalisten osasten oleminen erottamattomasti yhdessä. Näin ei ole asian laita elektronisten tallenteiden kohdalla. Elektronisten tallenteiden kontekstin tunnistaminen ja ylläpito aiheuttaa suuria haasteita arkistojille, koska elektroniset tallenteet eivät ole olemassa erillisinä fyysisinä entiteetteinä. Lisäksi suuri osa elektronisia tallenteita koskevaa kontekstuaalista informaatiota joko ei ole käyttäjille näkyvässä tai sitä ei automaattisesti poimita. Tallenteen muodostavien elektronisten signaalien järjestelyllä on harvoin mitään suhdetta siihen kuvaan, joka tallenteesta muodostuu monitorin näytölle tai tulostuspaperille. Elektroniselle informaatiolle jaettavan varastointitilan luomisesta huolehtii tietokoneen bittejä ja elektronisten tallenteiden osasia varastoiva toiminnallinen järjestelmä. Siten bittien ja elektronisten tallenteiden osasten fyysinen sijainti ei tarvitse eikä pidä sisällään mitään kontekstuaalista informaatiota. Elektronisten tallenteiden fyysinen ja looginen puoli ovat erillään toisistaan. Ohjelmisto kontrolloi suurta osaa elektronisiin tallenteisiin liittyvää kontekstuaalista informaatiota, jolloin se ei ole käyttäjille silmin nähtävissä. (Dollar 1992, 48-50) Toisin sanoen tallennetta vastaavat nollista ja ykkösistä koostuvat bittijonot, tallenteen luomaa järjestelmää koskeva informaatio ja varsinainen tallenteen sisältö teksteineen ja kuvineen ovat erillään toisistaan.

Tällä hetkellä käytössä olevat ohjelmistotyökalut eivät automaattisesti poimi mukaansa elektronisten tallenteiden kontekstuaalista informaatiota (Dollar 1992, 48).

Pohjolan (1999) mukaan elektroniseen arkistointiin liittyvät keinot, välineet ja tekniikat ovat yhä vielä melko kehittymättömiä, koska esimerkiksi tallenteet täytyy tietyn ajan välein kopioida ja siirtää paikasta toiseen, mikä merkitsee myös lisää kustannuksia jatkuvalla aikavälillä. Säilytettävän informaation määrän kasvaessa koko materiaalin hallussapito tulee ongelmallisemmaksi ja laitteille aiheutuu ylimääräistä rasitusta suuren informaation määrän vuoksi. Tietyn asian ja tiettyjen materiaalien paikantaminen ja sijoittaminen suureen massaan vaikeutuu. Samoin tiedon hakeminen monimutkaistuu, ellei saada kehitetyksi tehokkaita luettelointeja yms. Tämän vuoksi on erittäin tärkeää erottaa toisistaan materiaalit, joilla on eri pituiset säilytysajat. Informaatioteknologian nopean kehityksen vuoksi käytössä olevien laitteiden ja ohjelmistojen käyttöikä jää lyhyeksi, koska ne muuttuvat varsin nopeasti ns. vanhaksi teknologiaksi.

4.3 Elektronisessa arkistoinnissa yleisimmät käytettävät toimenpiteet

4.3.1 Yleistä

Yksi tapa varmistaa tiedon säilyminen on siirtää tieto säännöllisesti uuteen formaattiin. Tämä tehdään ennen kuin nykyinen formaatti haurastuu lukukelvottomaksi tai lukemiseen tarvittava laitteisto jää vanhanaikaiseksi niin, ettei sitä ole saatavilla. Paperidokumenteista voidaan jopa ottaa valokopioita, ne voidaan siirtää mikrofilmille tai lukea optisesti digitaaliseen muotoon. Digitaaliset tiedostot voidaan kopioida nauhakasetilta toiselle. LP-levyn voi kopioida CD-levylle. Tiedon siirrossa ovat ongelmana lukuisat eri tiedostomuodot. Kuvia voidaan pakata useilla eri tavoilla ja pakatut tiedostot voivat olla eri käyttöjärjestelmissä. Näin ollen on varmistuttava siitä, että myös ohjeet tiedoston lukemiseen säilytetään.

Tiedon säilymistä voidaan merkittävästi edesauttaa, jos tieto tallennetaan jonkinlaisessa standardimuodossa. Jos tekstidokumentti on WP5.1-muodossa, pitää olla saatavilla WP5.1-ohjelma tai tiedostomuotoa tukeva muunnin. Mikäli teksti tallennetaan alunperin tai muunnetaan yleisempään muotoon (jonka määrittelyt ovat nyt ja tulevaisuudessa

tavalla tai toisella yleisemmin saatavilla), on todennäköisempää että tekstiin päästään myös myöhemmin käsiksi. Sama koskee tietenkin esimerkiksi kuva- ja äänitiedostoja. Tällöin kuitenkin voidaan joutua tinkimään dokumenttien ohjelmistokohtaisista yksityiskohdista. (Lesk, M., 1990 ; Lesk, M., 1992 ; Bullock, A., 1999)

Prosessoitavien eli muokattavissa olevien aitojen sähköisten tallenteiden ylläpito tarkoittaa, että varmistetaan tallenteiden luettavuus, tietokoneella tulkittavuus, ymmärrettävyys ihmisille, looginen ja fyysinen rakenne sekä sisältö sellaisena kuin se oli luomistai vastaanottohetkellä. Ylläpito voidaan jakaa kolmeen erilliseen, toisiinsa liittyvään osa-alueeseen: 1) uudistus (reformat), 2) kopio ja 3) muunnos. (Dollar 1998, 28)

4.3.2 Tallenteiden uudistus

Yksinkertaisimmillaan sähköisten tallenteiden uudistus tapahtuu silloin, kun tulee joku muutos tai joitakin muutoksia bittivirtaan, mutta ei tallenteiden esitykseen tai sisältöön. Tavallisesti uudistus liittyy bittivirran siirtoon yhdestä muistivälineestä toiseen erilaiseen välineeseen. Esimerkiksi 9:lle raidalle 6250 bpi:n (bits per inch) nauhalle tallennettujen sähköisten tallenteiden siirtäminen 18:lle raidalle 3480 bpi:n nauhalle merkitsee uudistusta siitä huolimatta, ettei uudistettuja tallenteita katsellessa voida havaita mitään muutoksia niiden ulkonäössä tai sisällössä, vaikka tässä toimenpiteessä tapahtuu muutos tallenteiden fyysisessä tallennusmediassa. Useimmissa arkistointipaikoissa sähköisten tallenteiden uudistaminen tapahtuu todennäköisesti seuraavissa olosuhteissa:

- Muistivälineen fyysisen tallennusmedian muutos
- Kirjainkoodin siirto Binary Coded Decimal-muodosta (BCD) Extended Binary Code for Data Interchange Code-muotoon (EBCDIC) tai Unicode:sta American Standard Code for Information Interchange-muotoon (ASCII) tai poistamalla tekstin ohjausmerkit
- Korvaamalla kiinteäpituiset tallenteet muuttuvapituuisilla tallenteilla tai päinvastoin
- Siirtämällä tallenteita yhdestä ohjelmistosovelluksesta toiseen (Dollar 1998, 13-14)

Sähköisten tallenteiden uudistamiseen liittyy seuraavia kysymyksiä, jotka pitäisi pitkäaikaisen pääsyn strategiassa ottaa huomioon.

- 1) mitä muistivälinettä käytetään sähköisten tallenteiden uudistamisessa?

- 2) Milloin sähköiset tallenteet tulisi uudistaa?
- 3) Kuinka varmistetaan uudistettujen sähköisten tallenteiden aitous?
- 4) Kuinka suojellaan uudistettuja sähköisiä tallenteita katastrofaaliselta vahingolta tai luonnolliselta tuholta? (Dollar 1998, 28-30, 49)

Paras aika muuttaa sähköisten tallenteiden muistiväline uuteen on seuraavissa tapauksissa:

- 1) siirron aikana
- 2) kun varastointisäilytyspaikkaan asennetaan uudet muistilaitteet ja -välineet
- 3) etukäteen päätetyssä käytössä olevan muistivälineen eliniän vaiheessa.

Sähköisten tallenteiden uudistaminen/muuttaminen standardin mukaiseen koodauksen esitykseen ja muistivälineeseen tulisi tapahtua silloin, kun sähköisiä tallenteita siirretään varastosäilytykseen. Kun sähköiset tallenteet siirretään, ne tulisi muuttaa seitsemän- tai kahdeksanbittiseen ASCII-esitystapaan. (Dollar 1998, 28-30, 49)

Sähköisten tallenteiden aitous voi kärsiä useiden uudistusvaiheiden aikana ja silloin kun prosessia ei riittävästi dokumentoida. Kaikki uudistuksen vaiheet tulisi dokumentoida, kuten myös uudistusvaiheiden tekijän henkilöllisyyden tunnistaminen, päivämäärä, tiedon formaatti, alkuperäisten tallenteiden bittien/tavujen lukumäärä, uudistettujen tallenteiden bittien/tavujen lukumäärä sekä visuaalinen vertailu useiden uudistettujen ja niiden alkuperäisten vastineiden välillä. Uudistettujen sähköisten tallenteiden suojaaminen katastrofaaliselta vahingolta tai median rappeutumiselta olisi tehtävä siten, että luodaan kaksi kopiota, joista toista pidettäisiin off-line-medialle varastoivana varmuuskopiona. (Dollar 1998, 28-30, 49)

4.3.3 Tallenteiden kopiointi

Kopio tarkoittaa yleisesti fyysisten ja loogisten rakenteiden ja tieto-objektin sisällön duplikoimista eli kaksoiskappaleeksi tekemistä. Kopiolla on kolme eri määritystä riippuen mm. sen laadusta: 1), kopiolla on kaikki samat yksityiskohdat, kuten päivämäärä, osoitetiedot jne. kuin alkuperäiselläkin kappaleella, paitsi että se on tuotettu myöhempanä ajankohtana (copy in the form of the original). 2) alkuperäistä tallennetta jäljittelevä kopio (an imitative copy), joka uudelleen tuottaa sisällön ja suurimman osan alkupe-

räisen muodosta, kuten valokopio. 3) Alkuperäistä matkivalla yksinkertaisemmalla kopiolla (a simple copy) ei ole samaa uskottavuutta kuin näillä muilla kopioilla. (Dollar 1998, 14)

Sähköisten tallenteiden kopioinnin tavoitteena on ylläpitää niiden luettavuutta siirtämällä ne vanhasta muistivälineestä uuteen muistivälineeseen samoilla muotoilumäärittelyillä ilman häviötä rakenteessa, sisällössä ja kontekstissa. Kopioidut sähköiset tallenteet ovat ”binäärisiä kaksosia” Ulf Anderssonin lauselmaa lainaten (Dollar 1998, 14), koska bittivirta duplikoidaan eli tehdään siitä identtinen kaksoiskappale ykkösine ja nollineen. (Dollar 1998, 14)

Kopioita sähköisistä tallenteista tehdessä on tarkoitus varmistaa niiden jatkumo siirtämällä ne vanhasta muistivälineestä uuteen välineeseen samoilla formaattimäärittelyillä (esimerkiksi vanhalta 3480 nauhakasetilta uudelle 3480 nauhakasetille) ilman, että aiheutuu häviötä rakenteeseen, sisältöön ja kontekstiin. Toisin sanoen kopioinnissa ei tapahdu mitään muutoksia tallenteiden pohjalla olevaan bittivirtaan. Kopioinnin yhteydessä tulevat esiin samat kysymykset kuin uudistamisenkin yhteydessä, eli 1) mitä muistivälinettä olisi paras käyttää sähköisten tallenteiden kopioinnissa? 2) Milloin sähköiset tallenteet tulisi kopioida? 3) Kuinka taataan kopioitujen sähköisten tallenteiden aitous? 4) Kuinka suojellaan kopioituja sähköisiä tallenteita tuhoutumiselta? (Dollar 1998, 30-32, 50)

Kopioinnissa ei tapahdu muutosta tallenteeseen, joten ei tarvitse miettiä sitä, mikä väline olisi paras. Sähköisten tallenteiden kopiointi tapahtuu todennäköisimmin seuraavissa esimerkkitapauksissa:

- 1) Tallenne siirretään pitkäaikaiseen säilytykseen samanlaiselle nauhalle, jolla se on jo säilytetty (kuten vanhemmalta 3480 nauhakasetilta uudemmalle 3480 nauhakasetille). Tällöin kopioinnissa tapahtuu vain tallenteesta tehtävä identtinen kopio.
- 2) Tallennusmedia, jolla tallenne on säilössä, on vanhenemassa ja tarvitaan siirto uudemmalle tallennusmedialle. Ajankohta vanhenemisen vuoksi tapahtuvalle kopioinnille on puolessavälissä arvoitua tallennusmedian elinikää, joka on useimmilla optisilla ja magneettiseen tallennukseen perustuvilla tallennusmedioilla 10-20 vuoden välillä riippuen säilytysolosuhteista.

- 3) Vuosittaisessa tarkastuksessa paljastuu tallenteiden luettavuudessa väliaikaisia ja ohimeneviä ongelmia ja virheitä. (Dollar 1998, 30-32, 50)

Vaikka kopioinnissa tallenteiden pohjalla olevaan bittivirtaan ei tule muutoksia, saattaa kuitenkin olla mahdollista, että bitit ja tavut turmeltuvat. Kopioinnin yhteydessä suositellaan luomaan kaksi kopiota, joista toinen olisi ns. varmuuskopio, joka varastoidaan, kuten suositeltiin myös tallenteiden uudistamisenkin yhteydessä. (Dollar 1998, 30-32, 50)

4.3.4 Muunnos

Muunnos Mackenzie Owenin ja Wallen (Mackenzie Owen ja Walle, 1996, 50) mukaan on perinteisesti tapahtunut siten, että elektroninen tallenne on tulostettu paperille tai siirretty mikrofilmille. Tämä on ollut käyttökelpoinen menetelmä yksinkertaisille, ei-interaktiivisille ja staattisille tallenteille ja asiakirjoille. Mutta tänä päivänä monet tallenteet ovat monimutkaisia – ne sisältävät linkkejä yms., jolloin ne koostuvat useista osista, eivätkä muodosta selvää yksittäistä kokonaisuutta. Muunnos nykyaikaisille monimutkaisille tallenteille täytyy tehdä muulla tavalla kuin paperille tulostamalla tai mikrofilmille siirtämällä.

Mackenzie Owenin ja Wallen (Mackenzie Owen ja Walle, 1996, 49) mukaan elektronisen materiaalin säilyttämisessä tulee ennen pitkää vastaan tarve muuntaa tietoa toiseen formaattiin tai välineeseen. Muunnos on tarpeellinen:

- välineen fyysisen huonontumisen vuoksi
 - Elektronisen tiedon levittämiseen ja varastointiin käytetty media ei säily virheettömänä, vaan media täytyy kopioida (virkistää) säännöllisin väliajoin. Disketit virkistetään eli niiden sisältämät tiedot siirretään uusille disketeille kahden vuoden välein tai ne siirretään suoraan optiselle medialle. CD-ROM -levyjä ei tarvitse virkistää lyhyellä aikavälillä, eli ei aiemmin kuin yli 10 vuoden välein.
- teknisen ympäristön vanhentumisen vuoksi
 - Parhailtaan käytettäviä laitteistoja ja ohjelmistoja ei jatkossa tueta teknisesti ja uusi teknologia tulee ne korvaamaan. Ilman muunnosta uuteen tekniseen ympä-

ristöön vanhentuneita formaatteja, medioita ja käyttöympäristöjä käyttävät materiaalit eivät ole enää käyttökelpoisia.

- taloudellisten ja hallinnollisten syiden vuoksi
 - Kustannustekijät, saatavilla oleva tietotaito, yhteensopivuus muiden järjestelmien ja muiden tahojen formaattien kanssa.

Dollarin (1998, 15) mukaan aitojen sähköisten tallenteiden konversio eli muunnos sisältää niiden automaattisen siirtämisen yhdestä ohjelmistoympäristöstä toiseen ilman, että toimenpiteestä aiheutuu jotain häviötä rakenteeseen, sisältöön tai kontekstiin, vaikka bittivirtaan tulee todennäköisesti hiukan muutoksia. Tässä tapauksessa termi automaattinen viittaa siihen, että käyttäjä on mukana toimenpiteessä vähän tai ei ollenkaan, koska ohjelmiston siirtämisessä toiminnallisuus automaattisesti tunnistaa lähde- tai kohdeohjelmiston ja suorittaa luotettavan konversion eli muunnoksen. Tavallisesti muunnos tapahtuu siten, että

- sähköiset tallenteet siirretään ohjelmistosta riippuvaisesta ympäristöstä ohjelmistosta riippumattomaan ympäristöön
- siirron aikana sähköiset tallenteet siirretään siihen sovellusympäristöön, josta ne ovat peräisin ja joka on hyväksytty standardiksi
- ohjelmistoympäristö siirretään vanhemmasta versiosta uudempaan versioon eli sen hetkiseen teknologiseen ympäristöön
- uusi ohjelmistosovellusympäristö asennetaan.

Jeffrey D. Morellin (Morelli, Jeffrey D., *Defining Electronic Records*. Teoksessa Higgs, Edward, toim., *History and Electronic Artefacts*, 1998, 178-180) mukaan ohjelmistosta riippuvainen ympäristö tarkoittaa ympäristöä, jossa ohjelmistosta riippuvaisia sähköisiä tallenteita voidaan käyttää vain tietyn ohjelmistosovelluksen sisällä. Pitkäaikaiselle saatavuudelle tämä voi olla pääasiallinen este, jos valmistaja ei ohjelmiston uudemmissa versioissa tarjoa tukea tai ei anna jatkuvuutta uudempiin versioihin tai jos valmistajaa ei olekaan enää. Hyväksytty ratkaisu, erityisesti numeerisille tietokannoille, on muuntaa tallenteet niin, että niistä tulee ohjelmistoista riippumattomia.

Muunnos tapahtuu edellä selostettujen tapausten lisäksi silloin, kun ohjelmistoversio vaihtuu ja kun valmistaja tarjoaa yhteensopivuuden uudemman ja vanhemman ohjelmis-

ton välille. Tämän avulla on mahdollista siirtää tallenteita niiden perustalla olevan fyysisen esitysuunnitelman, sisällön ja toiminnallisuuden kanssa uuteen ympäristöön osana asennusprosessia. Tavallisesti tällainen muunnos vaatii vähän tai ei ollenkaan käyttäjän työpanosta tapahtuman kulussa, koska ohjelmiston kehittäjä ymmärtää sekä ohjelmistosovellusten että kehitettyjen ominaisuuksien sisäisen rakenteen, jotka automaattisesti kartoittavat lähdeohjelmistoympäristössä olevat sähköiset tallenteet kohdeohjelmistoympäristöön. Esimerkkinä tästä on tekstitalienteiden muunnos Microsoft Word 2.0-versiosta 4.0-versioon tai taulukkolaskentatallenteiden muunnos Access 2.0-versiosta 4.0-versioon. (Dollar 1998, 15-16)

Dollarin (1998, 32-33, 51-52) mukaan muunnos tapahtuu tavallisimmin silloin, kun tallenteet pitää muuttaa ohjelmistoista riippumattomiksi. Muunnos eroaa uudistamisesta siten, että siinä kohdistetaan huomio siihen ohjelmistosovellusympäristöön, jossa tallenteet luotiin. Muunnos vaatii siten pääsyn siihen ohjelmistoon, jota käytettiin tallenteen luomisessa, ylläpidossa ja käytössä tai siihen ohjelmistoon, jossa tallenteet ovat sillä hetkellä. Kohdeohjelmistosovelluksen pitää pystyä hakemaan tallenteita lähdeohjelmistosovelluksesta.

Margaret Hedstromin (Hedstrom, M., Digital preservation: a time bomb for Digital Libraries, 1996) mukaan erilaisissa formaateissa olevat tallenteet tulisi arkistoissa muuntaa tiettyihin formaatteihin, jotka sopivat yhteen ei-patentoitujen standardien kanssa. Suositeltu ratkaisu tekstitalienteille olisi muuntaa ne SGML -formaattiin, mutta silloin kun se ei olisi mahdollista, voitaisiin käyttää tekstinkäsittelysovelluksia, joilla on jatkuvuutta markkinoilla. Kolmas ratkaisu on poistaa ohjausmerkit, jotka ovat tekstinkäsittelyohjelmiston erikoisuuksia. Relaatiotietokannat voidaan jakaa litteiden taulukoiden sarjoihin tai rajoitettuihin tiedostoihin. Toinen vaihtoehto relaatiotietokannoille on siirtää ne SQL-yhteensopivaan tietokantaan. Tallenteet tulisi muuntaa niiden säilytyspaikkaan vastaanotossa. Muunnos voitaisiin yhdistää tallenteiden siirtoon uuteen muistivälineeseen. Sähköisten tallenteiden standardiformaatin pitäisi sisältää ei-patentoitu sisäinen nimike, joka identifioi tallenteiden yleisen sisällön, eri tiedostojen nimet, ”omistajan” nimen, viimeisen toiminnon päivämäärän sekä tekniset piirteet, kuten tekijän, kompression jne. Tämä sisäinen nimike tulisi kapseloida SGML:ään. Dokumenttityyppimäärittely (Document Type Definition) identifioisi tallenteiden kokonaisuuden loogisen raken-

teen, joka välittäisi tietoa kohteesta, joka loi ja ylläpiti tallenteita, kun ne siirrettiin säilytyspaikkaan, uudistushistorian, kopioinnin ja muunnoksen.

4.3.5 Analogisten tallenteiden kopioiminen digitaalisille medioille

Analogiset mediat aiheuttavat Leskin (1992) mukaan paljon ongelmia. Niillä on samat ongelmat teknologisen vanhenemisen kanssa kuin digitaalisilla medioilla (esim. vinyylilevyt ja soittamiseen tarkoitettujen laitteiden tuhoutuminen). Analogisilla medioilla on teknologisen vanhenemisen lisäksi toinenkin ongelma: analogisia medioita ei voida kopioida ilman, että aiheutuu pienehköä laadun huonontumista, eikä kopiointi ole välttämättä kovin pysyvä keino. Esimerkiksi sekä ääni- että videonauhat ovat käytössä suhteellisen haavoittuvia. Kirjaa voidaan lukea useampia kertoja kuin pyörittää videonauhaa ilman laadun huonontumista. Koska analogiset materiaalit menettävät arvoaan kopioinnissa, ne tulisi muuntaa digitaalisiksi heti ensimmäisen tilaisuuden tullen ja käyttää siitä lähtien digitaalisessa muodossa. Digitaalinen materiaali ei ole ainoastaan kopioitavissa ilman virheitä, vaan se tulee myös ajan mittaan halvemmaksi varastointimuodoksi. Mitä nopeammin muunnos tehdään, sitä parempi. Digitalisoinnissa on myös huonoja puolia, kuten muunnoksen korkeat kustannukset ja vaara muunnoksen huonolaatuisesta tuloksesta. Esimerkiksi Michael Ester Getty Conservation -Instituutissa on havainnut taidehistorioitsijoiden tarkan huomiokyvyn saman valokuvan eri laatuksille digitaalisille muunnoksille ja painottaa sen vuoksi hyvän laadun tärkeyttä. Valitettavasti joissakin paikoissa esimerkiksi kuvat digitalisoidaan vain matalalla resoluutiolla. Näin ei tehdä ainoastaan rahan säästämiseksi, vaan siksi, että helposti laittomasti kopioitavissa olevaa digitaalista kuvaa ei kopioitaisi huonon laadun vuoksi.

4.3.6 Elektronisten tallenteiden asianmukainen säilytys ja tallenteiden kunnan tarkkailu

Magneettinauhut (nauhakasetit, äänikasetit, videonauhat) olisi hyvä kelata alusta loppuun silloin tällöin. Toinen suositeltava tapa on kelata nauhat ajoittain lopusta alkuun, jotta saataisiin mekaaninen rasitus jaettua uudelleen tasapainoisemmin ja ehkäistäkseen rypyyntymistä ja nauhan epämuodostumia. Suositeltava aika kelata nauhat taaksepäin

on joka kolmas vuosi. Magneettiset signaalit heikentyvät yleensä ajan myötä, mutta tämä heikentyminen nopeutuu korkeammassa lämpötiloissa. Korkeassa suhteellisessa kosteudessa magneettiset partikkelit ja niihin äänitetty bittivirta vaurioituvat. Vahvassa magneetikentässä magneettiset mediat saattavat myös rappeutua. (Dollar 1998, 46)

Optiset välineet, kuten CD-levyt, ovat vähemmän alttiita vaurioitumiselle kuin magneettiset välineet, koska magneettisten välineiden nauhoitetut signaalit eivät ole palautettavissa. Kuitenkin optiset välineet ovat yhtä haavoittuvaisia, kun ne joutuvat korkeisiin lämpötiloihin ja suhteellisen kosteaan ilmatilaan. (Dollar 1998, 46) Optisia välineitä pidetään Mackenzie Owenin ja Wallen (Mackenzie Owen et al, 1996, 49-50) mukaan käyttökelpoisempina elektronisten tallenteiden pitkäaikaisessa arkistoinnissa kuin magneettisia välineitä. Seuraavista syistä:

- optisille välineille mahtuu enemmän tietoa ja välineet ovat melko halpoja
- Jukeboxin (vähint. 100 levyn varasto) tapaan varastoimalla voidaan pitää yhtä aikaa saatavilla suuria määriä tallenteita
- suhteellisen pitkä elinaika ennen mahdollisia virheiden ilmaantumisia (10 vuotta tai jopa 25-100 vuotta erityisen korkealaatuisille levyille).

Ditzlerin (Ditzler, C. et al., 1994) mukaan tietokoneiden ja tietotekniikan toivotaan kehittyvän tulevaisuudessa siten, että sähköisestä säilyttämisestä tulee helpompaa. CD-levylle kirjoitettaessa on syytä käyttää standarditiedostomuotoja sekä loogista, yksinkertaista tiedostorakennetta. On hyvä ajatus tallentaa kuvaus levyn sisällöstä ASCII-tiedostona esimerkiksi README -hakemistoon tai tiedostoon.

Dollarin (1998, 46-47) mukaan elektroniset tallenteet on säilytettävä oikeassa ympäristössä. Tämä tarkoittaa, että noudatetaan ohjeita, jotka koskevat säilytyslämpötilaa, ilman kosteutta, altistumista valolle ja magneetikentille. Myös tallennevälineen käsittelyssä on oltava huolellinen: naarmut, hankaumat ja pöly voivat vahingoittaa sitä. Myös sormenjäljet ja yleensäkin käyttäminen vähentävät median käyttöikää ja kuluttavat sitä. Käyttö merkitsee yleensä myös väliaikaisia muutoksia säilytysoloihin. Ideaalitulanteessa on otettava huomioon myös säilytysasento (yleensä pysty- eikä vaakasuorassa). Ilmassa ei saisi olla hiukkaspartikkeleita, savua eikä kaasuja kuten otsonia tai orgaanisia yhdisteitä. Esimerkiksi käsittelemättömistä puuhyllyistä irtoaa haitallisia kaasuja. Tallenne on säilytettävä sellaisessa paikassa, jossa se on mahdollisimman vähän altis onnettomuuks-

sille, varkauksille yms. Tulipalot ja vesivahingot aiheuttavat usein korvaamattomia vahinkoja erilaisille arkistoille. Vahinkoja voidaan ehkäistä tekemällä kopioita tai esimerkiksi tekemällä mahdollisimman paloturvallisia ratkaisuja.

Tallenteen käsittelyssä on noudatettava sovittuja, järjestelmällisiä periaatteita. Arkistoidusta tallenteesta on oltava tieto jossakin. Myös arkiston sisältöä koskevan metatiedon säilyttämisessä on otettava huomioon periaatteet, jotka koskevat itse tiedon säilyttämistä. Jos kellarikomeroon on arkistoitu oikeaoppisesti 20 vuoden asiakirjat, siitä ei ole paljon iloa, mikäli arkistoluettelo on tehty sinikantiseen vihkoon tai levykkeelle, joka on hukkunut. Tallenteen käsittelyssä on myös pidettävä huolta siitä, että tiedetään esimerkiksi, kuka on lainannut jotakin arkistokappaletta ja milloin. Mikäli arkistoa voi käsitellä kontrolloimattomasti, vuosien kuluessa arkistokappaleita voi kadota tai tulla tuhoutuneita.

Sähköisille tallenteille ihanteelliset pitkäaikaisen säilytyksen olosuhteet olisivat noin 10 asteen lämpötila ja 25%:n suhteellinen kosteus. Tutkimusten mukaan jo erittäin vahingollisia ovat olosuhteet, joissa on 26 astetta lämmintä ja suhteellinen kosteus 70%. Suositusten mukaan sähköiset tallenteet tulisi tarkastaa vuosittain, jotta varmistetaan oikeanlaiset olosuhteet ja niiden jatkuminen. Tallenteista tulisi ottaa pieni otos tarkastusta varten. Tallenteiden kunto tarkastetaan lukemalla ne läpi. Kuinka paljon tallenteita tähän otokseen valitaan, riippuu siitä, kuinka paljon tallenteita on yhteensä kyseisessä säilytyspaikassa. Dollar (1998, 47) esittää ohjeita tallenteiden kunnan tarkastamiseen: Jos tallenteita on yhteensä vähemmän kuin 50, ne tarkastetaan kaikki. Kun tallenteita on 50-1800, tarkastetaan tallenteista 20%:n otos. Kun tallenteita on 1801 tai enemmän, tarkastetaan 384 tallenteen otos. Tallenteiden säilytykseen suositellaan vakaita olosuhteita, joissa lämpötila on ± 15 astetta ja suhteellinen kosteus 40%. Säilytyspaikassa tulisi olla myös suodatinjärjestelmä, joka poistaa pölypartikkelit ja kaasut yms. Säilytyspaikassa ei tulisi käsitellä ruokaa ja juomia eikä tupakoida.

4.4 Tutkimuksessa suositeltavat elektronisen arkistoinnin menetelmät

4.4.1 Standardien kehittäminen arkistointia varten

Arkistointia on yritetty kehittää ja helpottaa erilaisten mallien avulla. Yhdysvalloissa on kehitetty viime vuosien aikana viittausmalli, Open Archival Information System (OAIS). OAIS -mallia suunniteltaessa on määritelty arkistolle tarpeellisia ominaisuuksia. Mallin mukaan arkiston pitäisi tarjota pitkäaikainen säilytyspaikka digitaaliselle tiedolle ja sen tulisi helpottaa eri organisaatioiden elektronisten arkistojen ymmärtämistä ja rakentamista. OAIS -mallissa määritellään tiettyjä osia, jotka muodostavat arkistolle abstraktin mallin. Siinä kuvaillaan, miten elektronisia aineistoja käsitellään ja mitä toimintoja arkistointiprosessiin kuuluu. (Stenvall, Jani, 2001, Metatieto elektronisten julkaisujen pitkäaikaissäilytyksessä, 29) Tällä hetkellä mallista on käsittelyssä ns. ISO Draft International Standard (Red Book, Recommendation) eli suositusluonnos kansainväliseksi standardiksi (ISO Archiving Standards - Overview, 2001 ; CCSDS Recommendations (Red Books and Pink Books), 2001).

Stenvallin (2001, 29-30) mukaan elektronisten julkaisujen arkistoinnissa on kuusi erilaista vaihetta ja näiden vaiheiden voidaan olettaa esiintyvän yleensä kaikenlaisessa arkistoinnissa. Vaiheet kuuluvat EU:n Nedlib -projektin (Networked European Deposit Library) OAIS-mallin pohjalta suunniteltuun malliin, joka on nimeltään DSEP -malli (Deposit System for Electronic Publications). Arkistoinnin kuusi vaihetta ovat:

- Informaatiopakettien syöttäminen järjestelmään (Ingest)
- Varastointi (Archival Storage)
- Tiedonhallinta (Data Management)
- Julkaisujen saatavuus (Access)
- Toimintojen hallinnointi (Administration)
- Pitkäaikaissäilytys (Preservation)

Näiden vaiheiden avulla voidaan jäsentää arkistointijärjestelmän toimintaa. Arkistoitavaksi tarkoitettu julkaisu tms. siirretään arkistoon (syöttö), varastoidaan ja säilytetään. Järjestelmän tiedonhallintaan kertyy samalla tietoja arkistoon syötetyistä julkaisuista. Nämä tiedot ovat myös käyttäjän saatavilla, kun hän etsii tietoja eri julkaisujen saatavuudesta (access). Kaikista järjestelmän toiminnoista kertyy tietoja, jotka toimintojen hallinnointi- (administration) toiminnon avulla tuodaan raportteina arkiston ylläpitäjille.

Tällaiseen arkistointimalliin sisältyy olennaisena osana tietojen (metatiedon) siirtyminen eri toimintojen välillä. OAIIS -malliin liittyy myös informaatiomalli, joka jäsentää elektroniseen aineistoon liittyviä tietoja. Informaatiomallin mukaisesti arkistojärjestelmään tuleva informaatiopaketti jaetaan sisältöinformaatioon, säilytyksen kuvausinformaatioon, pakkausinformaatioon ja kuvailuinformaatioon. Pitkäaikaissäilytyksen kannalta katsottuna tärkeitä tietoja sisältyy sisältöinformaatioon. Siihen kuuluvat itse elektroninen objekti ja esitysinformaatio. Säilytyksen kuvausinformaatioon sisältyy neljä alaryhmää: viittaustiedot (reference), alkuperäistiedot (provenance), kontekstitiedot (context) ja muuttumattomuustiedot (fixity).

Dollarin (1992, 60-61) mukaan 1980-luvun lopulla arkistoalalla sekä Pohjois-Amerikassa että Euroopassa alettiin työskennellä arkistoinnin kansainvälisten standardien kehittämisen puolesta. Tavoitteena oli luoda standardoidut arkistoinnin määritelmien kuvauksen ohjeet. Vuonna 1990 International Council on Archives Ad Hoc Commission on Descriptive Standards loi luonnoksen arkistoinnin määritelmien periaatteista, joissa keskeisessä osassa on alkuperän käsite. Alkuperän käsite on, kuten aiemmin tuli esille, hyvin keskeisellä sijalla arkistotieteessä. Tallenteiden alkuperäisyys tulisi pystyä takaamaan pitkäaikaisen arkistoinnin eri vaiheissa ja tallenteiden käytettävyyden turvaamiseksi tehtävistä toimenpiteistä huolimatta.

4.4.2 Tietojärjestelmien suunnittelu arkistointia varten

Tietojärjestelmät ja ohjelmistot tulisi Higgsin (1998, 190) mukaan suunnitella siten, että ohjelmisto pystyisi poimimaan sähköisen tallenteen kontekstuaalisen informaation. Ohjelmiston tulisi itsessään luokitella asiakirjat niiden luomisvaiheessa niin, että niihin sisällytetään metatietoa eli tiedot asiakirjojen tekijöistä, ajankohdasta, tarkoituksesta ja siitä, kenelle ne on lähetetty, jne. Tietojärjestelmien ja tieto-ohjelmistojärjestelmien suunnittelussa tulisi myös huomioida tallenteiden arvioinnin ja tallenteiden säilyttämiseen liittyvät tekijät.

Dollarin (1992, 58-59) mukaan arkistoinnin asiantuntijoiden tulisi tulla mukaan tietojärjestelmien suunnitteluun. Hänen mielestään yksi hyödyllisimmistä tavoista myötävaikuttaa arkistointinäkökohdat huomioonottavan tietojärjestelmän ja/tai ohjelmiston

suunnitteluun on liittää siihen tallennetun informaation elinkaaren hallinta -toimenpiteet. Tämä voi edesauttaa sekä organisaation että business toimintojen vaatimusten ja arkistointia koskevien ongelmien ratkaisemista. Vaikka elinkaaren käsite on hyvin tuttu arkistoihmisille, he eivät ole kuitenkaan analysoineet, kuinka se oikeasti toimisi käytännön tasolla. Dollar viittaa tällä siihen, että arkistoihmiset eivät ole tuoneet selvästi esille tallennetun informaation elinkaaren toiminnallisia vaatimuksia, jotka voisivat olla osa monimutkaisen tietojärjestelmän suunnitteluprosessia. Elektronisen informaation elinkaaren hallinnan toiminnalliset vaatimukset voidaan ottaa mukaan tehokkaimmin, kun määritellään elinkaaren tietoelementit metatietojärjestelmissä. Kuten tietosanakirjat ja tietoresurssisanastot, metatietojärjestelmät tyypillisesti kuvaavat järjestelmissä olevaa tietoa, tukemaansa toimintoa, palvelemaansa käyttäjää sekä katselmuksia ja raportteja. Elektronisten tallenteiden elinkaaren hallinnassa tarvittavien tietoelementtien mukaan ottaminen tietoresurssisanastojen järjestelmiin voisi auttaa varmistamaan sen, että tallenteet identifioidaan, säilytetään ja ne saadaan käyttäjien ulottuville. Avainasemassa pitkäaikaisten tallenteiden saatavuuden turvaamisprosessissa on kansainvälisen tietoresurssisanastojärjestelmään (Information Resource Dictionary Systems) liittyvän standardin kehittäminen. Jotta saataisiin standardin kehittämisestä paras mahdollinen hyöty, olisi tärkeää suunnitella uudet monimutkaiset tietokannat yhteensopiviksi kyseisen standardin kanssa. Tällöin myös kokonaisten tietojärjestelmien eli varsinaisen tietojärjestelmän ja siihen liittyvän metatiedon migraatio eli vanhemmasta järjestelmästä uudempaan järjestelmään siirtäminen olisi suurempi toimenpide, joka edesauttaisi pitkäaikaisen saatavuuden varmistamista sekä auttaisi lieventämään teknologian vanhentumisen aiheuttamia vaikutuksia. Jos arkistoinnin asiantuntijat tulisivat mukaan tietojärjestelmien suunnitteluun samalla pyrkimällä edellä mainittuihin tavoitteisiin, heidän tähänastisen työnsä kuva tulisi muuttumaan. Yhdeksi tärkeimmäksi tavoitteeksi ja tekijäksi nousisi tällöin metatiedon suunnittelu ja kehittäminen tietojärjestelmiä varten. Tällöin metatieto olisi jokseenkin yhdenmukainen perinteisten arkistojen tietojen etsinnässä ja haussa käytettävien apukeinojen kanssa. Monimutkaisia tietojärjestelmiä varten tarkoitettu metatieto poimittaisiin tietoresurssisanastojärjestelmien sisällä. Arkistoinnin asiantuntijoita tarvittaisiin myös tällaisten tietoresurssisanastojärjestelmien standardoinnissa.

4.4.3 Elektronisten tallenteiden arvioiminen säilytystä varten

Kansallisarkistossa työskentelevän Raimo Pohjolan (1999) mukaan sähköisten asiakirjojen ja arkistojen suhteen saattaa olla hiukan väärää käsityksiä, kuten esimerkiksi, ettei niitä ei tarvitse ollenkaan arvioida eikä hävittää ei-elektronisten dokumenttien tapaan, koska ne vievät vähän tilaa. Vaikka elektronisen tiedon varastointivälineet ovat halpoja ja niiden kapasiteetti jatkuvasti kasvaa, ei tämän tiedon varastointiin käytettävä tila tule ikuisuuksia riittämään. Tämän vuoksi sähköiset asiakirjat tulee ei-elektronisten asiakirjojen tapaan seuloa, eli erottaa toisistaan pysyvään säilytykseen tarkoitettut dokumentit tietyn ajan kuluttua hävitettäväksi määrätyistä dokumenteista. Elias Orrman`in (Orrman, Elias, Structural Complexity of Electronic Records as a Factor Guiding Decisions on Permanent Retention, 1999) mukaan sähköisiä asiakirjoja ja tallenteita ei voida arkistoida perinteisten asiakirjojen tapaan yhdellä päätöksellä ja arviointikerralla. Sähköiset tallenteet ja asiakirjat täytyy todennäköisesti arvioida uudelleen vielä sen jälkeen kun ne on valittu pysyvään säilytykseen, mikä on yksi syy kustannusten kumulatiiviseen kasvuun. Strategia, jolla tieto säilytetään hyvin monimutkaisissa sähköisissä järjestelmissä ja jolla alkuperäisten järjestelmien dataa prosessoivat ominaisuudet mahdollisimman suuressa määrin säilytetään, tulee todennäköisesti olemaan mahdoton toteuttaa arkistojen palveluresursseilla. Perinteisiin asiakirjoihin verrattuna sähköiset tallenteet ja asiakirjat ovat paljon monimutkaisempia. Suurempi monimutkaisuus merkitsee korkeampia säilytys- ja käyttökustannuksia.

Sähköisessä muodossa olevat asiakirjat tulisi arvioida jo varhaisessa vaiheessa. Tämä on yleisestikin suositeltavaa, jotta voidaan välttää arkistointiin liittyviä riskejä, kuten epätäydellisyyttä tai epäluotettavuutta. Myös muutokset IT-järjestelmissä voivat aiheuttaa puutteellisen suunnittelun vuoksi uuteen järjestelmään vielä siirtämättömien arkistojen menetyksiä, häviämisiä tai turmeltumista. On myös mahdollista, että vanhojen IT-järjestelmien hävittämistoimenpiteet voivat saada aikaan järjestelmille vahingollisia muutoksia, kuten esimerkiksi tietokannoissa tapahtuvaa säilytettäväksi tarkoitettua informaation häviämistä. (Pohjola 1999)

Sähköisten asiakirjojen arviointi eroaa Pohjolan (1999) mukaan perinteisten asiakirjojen arvioinnista. Sähköiset asiakirjat tehdään informaatioteknologiaa hyväksikäyttäen ja erilaisissa käyttöjärjestelmissä ja eri ATK-ohjelmilla. Tämän vuoksi niiden arvioinnissa

täytyy ottaa huomioon seuraavia seikkoja: 1) IT-järjestelmät, joissa sähköisiä asiakirjoja pidetään yllä, 2) näiden järjestelmien vastaavuus liike-elämän ja säilyttämisen tarpeisiin, 3) kaiken metatiedon tarpeellisuus arviointia ajatellen sekä tiedostorakenteiden riittävä organisointi, 4) IT-järjestelmien yhteensopivuus toivottuihin standardeihin, periaatteisiin ja menettelytapoihin ja 5) asiakirjojen fyysinen sijainti ja hakumahdollisuudet. Sähköisille asiakirjoille tulisi luoda oma hallintajärjestelmänsä tai arkistonmuodostussuunnitelmansa, ja jo sen suunnitteluvaiheessa on hyvä ottaa huomioon toiminnalliset vaatimukset, kuten sisäänrakennetut välineet, jotka tukevat asiakirjahallintaa ja arviointiprosessia. Näiden ominaisuuksien avulla voidaan säilyttämisajat liittää tiettyihin sähköisten asiakirjojen tyyppeihin. Järjestelmän pitäisi myös tallentaa muistiin asiakirjojen kehitysprosessi niiden luomisesta lähtien sekä myös niiden kunto ja ikä.

Suomalaista arviointiperinnettä voitaisiin Pohjolan (1999) mielestä soveltaa myös sähköisten asiakirjojen arviointiin. Suomalainen arviointiperinne on esiarviointia luonteeltaan, mikä tarkoittaa tallenteen säilytysajan määrittämistä elinkaaren aikaisessa vaiheessa, jopa ennen kuin varsinainen tallenne on luotu. Alun perin perinteisten asiakirjojen arkistointia varten luotua arkistonmuodostussuunnitelmaa voidaan käyttää arvioinnissa seuraavasti: Tallenteen luoja tekemiä toimenpiteitä tutkimalla voidaan tunnistaa yksilön toimintaan liittyvät operatiiviset prosessit. Arkistonmuodostussuunnitelmaa tehtäessä on yleisenä käytäntönä käydä läpi, kuinka tehtävät ja prosessit sopivat tallenteen luojaorganisaatorakenteeseen. Tämä arviointi auttaa huomaamaan tallenteiden hallintaja prosessiongelmiä. Arkistonmuodostussuunnitelman avulla tuodaan esiin, minkä sarjojen ja muiden tallenteen tapaisten tietomateriaalien volyyymi kasvaa sitä mukaa, kun tallenteiden luoja hoitaa tehtäviään ja yrittää hallita niitä. Suunnitelma tuo myös esiin materiaalien rekisteröinti- ja arkistointitavat ja se kertoo myös, mitkä tallenteet säilytetään pysyvästi ja mitkä tuhotaan erilaisten säilytysaikojen kuluttua umpeen. Arkistonmuodostussuunnitelmia voidaan käyttää myös sähköisten tallenteiden arvioinnissa, jolloin suunnitelmiin tulee tallentaa mukaan tiettyä perustietoa IT-järjestelmistä, kuten:

- järjestelmän nimi
- onko järjestelmä henkilörekisteri
- onko tieto luokiteltu vai ei
- yleinen kuvaus (tarkoitus; siihen liittyvät tehtävät ja prosessit; järjestelmän perustoiminnot; järjestelmän käyttöön oton jälkeen tehdyt päämuutokset)

- järjestelmän tietosisältö
- tiedon lähteet
- yksikkö/henkilö, joka on vastuussa järjestelmässä olevasta tiedosta
- viranomaisille järjestelmästä annettu tieto
- ohjelmistot
- laitteistot ja välineet
- yksikkö/henkilö, joka on vastuussa järjestelmän teknisestä ylläpidosta.

Sähköiset asiakirjat ovat ohjelmistoista, käyttöjärjestelmistä ja laitteistosta riippuvaisia, minkä vuoksi niiden elinkaaren hallintaa on vaikeaa vahvistaa oikeaksi. Muuttuvassa teknologisessa ympäristössä koko elinkaari-ajatus on epätydyttävä lähestymistapa Greg O`Shean (1996) mukaan. Perinteisten asiakirjojen kohdalla päätökset arkistointiarvosta on usein tehty vasta elinkaaren loppuvaiheessa. Jos tätä periaatetta noudatetaan sähköisten asiakirjojen ja tallenteiden kohdalla, voitaisiin arvioida ainoastaan ne sähköiset tallenteet, jotka ovat onnistuneet selviämään aktiivisen ajan loppuun. Siksi sähköisten tallenteiden elinkaaren alussa niihin on kiinnitettävä huomiota. Jo järjestelmien kehitys- ja tallenteiden luomisvaiheessa tulisi kiinnittää huomiota tallenteisiin.

1990-luvulla on luotu joitakin periaatteita elektronisten asiakirjojen ja tallenteiden pitkäaikaiselle arkistoinnille. Nämä periaatteet ovat seuraavia:

- 1) Arkistoinnin ammattilaisten tulisi omaksua uusi lähestymistapa, koska elektronisia asiakirjoja ei ole enää mahdollista arvioida perinteisten asiakirjojen tapaan, mikä tarkoittaa itse asiakirjan ja sen tulevan säilytysajan arvioimista. Sen sijaan pitäisi alkaa arvioida elektronisten asiakirjojen tuottamisessa ja luomisessa tehtäviä toimenpiteitä, suorituksia ja taitoja ja pätevyyttä.
- 2) Arkistoidun informaation elinkaaren hallintaa varten pitäisi määritellä toiminnalliset vaatimukset.
- 3) Arkistoinnin ammattilaisten pitäisi tulla mukaan asiakirjan luomisprosessiin sen mahdollisimman varhaisessa vaiheessa. Arviointitoiminta tulisi ottaa huomioon jo IT-järjestelmiä suunniteltaessa.
- 4) Metatiedon tulisi sisältää tietoa elektronisten asiakirjojen elinkaaren eri vaiheista. (Arkistolaitoksen suositus arkistonmuodostussuunnitelman laadinnan, käytön ja ylläpidon periaatteiksi) Koska sähköiset asiakirjat tulisi mahdollisimman var-

haisessa vaiheessa eli mielellään jo luomisen yhteydessä arvioida tulevan säilytysajan suhteen, on arkiston määritelmäänkin sen vuoksi tullut hieman muutoksia. Arkisto/arkistot tarkoittavat uuden näkemyksen mukaan niitä tallenteita, jotka on arvioitu sisältävän jatkuvaa käyttöarvoa (O`Shea 1996).

Dollarin (1992, 54) mukaan tulisi myös hylätä ajatus sähköisten tallenteiden pysyvistä säilytyksestä ja pysyvistä arvosta. Sen sijaan pitäisi omaksua ajatus jatkuvasta arvosta, mikä sisältää näkökulman, että säilytyskysymys tullaan käsittelemään uudestaan tulevaisuudessa kaikkine kustannuksineen ja hyötyineen. Organisaatiot säilyttävät sähköiset tallenteensa niin kauan kuin se on mahdollista ja vasta aivan viimeisessä vaiheessa tallenteet siirrettäisiin arkistoon lopulliseen säilytykseen. Tämä tapahtuisi vain siinä tapauksessa, että organisaatiot eivät haluaisi enää huolehtia niiden ylläpidosta ja migraatiosta eri teknologioiden välillä.

4.4.4 Metatiedon käyttö

Metatieto pitkäaikaisessa arkistoinnissa

Metatietoon eli tietoa tiedosta liittyvät kysymykset ovat ajankohtaisia. Metatietoa tarvitaan sähköisen asiakirjan luotettavuuden ja aitouden varmistamiseen, koska se sitoo asiakirjan aikaan, paikkaan ja toimenpiteeseen sekä kertoo asiakirjan rakenteesta ja tietosisällöstä. Metatieto auttaa myös tiedon haussa ja löydettävyydessä. Tällä hetkellä yhtenä ongelmana on, etteivät ohjelmistovalmistajat ole juuri vielä tuottaneet kaupallisia ohjelmia, joissa metatieto ja toiminnalliset vaatimukset olisi otettu huomioon. (Arkistoviesti 2/1998)

Sähköisten asiakirjojen yhteydessä mainitaan usein metatiedot tai Dublin Core. Metatieto on yleisnimitys asiakirjojen sisältämille kuvailutiedoille. Se voi olla asiakirjaan sisällytettyä tietoa, jolla tuetaan asiakirjojen tunnistamista, kuvailua, hallintaa, löytymistä, luokittelua jne. Tieto voi olla asiakirjassa joko näkyvänä, kuten aihe ja diaarinumero, tai asiakirjan ulkopuolella, kuten hakusanat. Asiakirjoihin liittyvä, omaksi tietokannaksi koottu metatieto voi olla esimerkiksi diaari tai erillisrekisteri. Dublin Core on eräs metatietomalli ja yksi sen periaatteista on joustavuus ja laajennettavuus, koska se tarjoaa vaihtoehdon tarkempien kuvailumallien luomiselle. Monissa Dublin Corea käyttäneissä

projekteissa formaattia on sovellettu laajentamalla sitä määrittelemällä omia tarkennuksia tai elementtejä. Yksi esimerkki tästä on Suomessa toteutettu Julkishallinnon (valtio-neuvoston kanslian, valtiovarainministeriön ja kauppa- ja teollisuusministeriön) tietohallinnon neuvottelukunnan (JUHTA) suositus asiakirjojen metatietoformaatiksi, jossa asiakirjapohjiin liitettiin mukaan metatietoja. JUHTA on laatinut metatietoa koskevan suosituksen ja ensimmäinen asiakirjojen kuvailuformaatti, versio 1.02 on hyväksytty 18.12.2001. JUHTA:n suosituksessa metatietomalli perustuu Dublin Coreen, johon on lisätty viisi erityisesti asiakirjojen kuvailussa tarvittavaa elementtiä. Dublin Corea voidaan käyttää mahdollisena välineenä myös pitkäaikaissäilytyksen ongelmia ratkaistaessa. (Stenvall, J., 2001 ; JUHTA: JHS 143)

Metatietoa on eri tyyppisiä sen mukaan minkälaiseen tallenteeseen, tiedostoon, asiakirjaan yms. se liittyy. Metatietoa on kolmenlaista Cornellin yliopiston (Cornell University Library, Digital Imaging Tutorial, 2000) mukaan:

- kuvaileva metatieto helpottaa käyttäjää hakutoiminnoissa, kun se kuvailee ja tunnistaa tietoresursseja
- rakenteinen metatieto helpottaa navigaatiota ja sähköisten resurssien esittämistä, koska se sisältää tietoa tallenteen sisäisestä rakenteesta
- hallinnollinen metatieto helpottaa sekä lyhyt- että pitkäaikaista digitaalisten kokonien hallintaa ja prosessointia, koska se sisältää tietoa luomis- ja laatu- ja laatukontrollista, saatavuuden kontrolloinnista ja säilytyksestä.

Sähköisten tallenteiden pitkäaikaissäilytyksen ongelmien ratkaisemiseksi tutkijoilla on erilaisia ja myös toisistaan eroavia mielipiteitä. Joka tapauksessa yksi asia on varma ja siitä ollaan samaa mieltä. Se on metatieto tai metatieto ja sen käyttö. Metatieto on tietoa tallenteen tekijästä, luomisorganisaatiosta, luomisajankohdasta jne. Metatiedon luominen ja ylläpito ovat mukana pitkäaikaisessa arkistoinnissa huolimatta pitkäaikaisen säilytyksen strategiasta. Seuraavassa lainauksessa on mainittu mitä tietoa metatieto voi sisältää:

”Arkistossa metatieto liittyy ja viittaa kuhunkin digitaaliseen objektiin ja tuo esiin niihin liittyvää kuvailevaa, rakenteista, hallinnollista, oikeuksien hallintaan liittyvää yms. informaatiota. Tätä metatietoa myös ylläpidetään ja sitä siirretään formaatista ja standardista toiseen itsenäisesti riippumatta

objektista, jota se kuvaa.” (Day, Michael, *Metatieto for Digital Preservation: an update*, 1999)

Metatieto voi sisältää edellisen mukaisesti kuvailevaa, rakenteellista, hallinnollista ja oikeuksien hallintaan liittyvää tietoa yms. Metatiedon avulla viitataan kuhunkin digitaaliseen objektiin. Tätä metatietoa ylläpidetään ja se siirretään (migraatio) formaatista ja standardista toiseen standardiin. (Day 1999)

Kuten kappaleissa 4.4.1 ja 4.4.2 tuli esille, standardeja voidaan kehittää ja tietojärjestelmiä suunnitella siten, että arkistoinnin asiantuntijat ovat mukana prosessissa. Elektronisten tallenteiden pitkäaikaisessa arkistoinnissa tavoitteena on turvata tallenteiden autenttisuus eli alkuperäisyys, eheys eli muuttumattomuus ja yleensä käyttökelpoisuus. Metatiedon ja tietoresurssisanastojärjestelmien avulla tämä on kenties mahdollista. Tietoresurssisanastot toimivat myös metatiedon kaltaisesti edesauttamalla siinä, että tallenteet ovat käytettäviä ja käyttäjien saatavilla tulevaisuudessa. (Dollar 1992, 59) Metatieto välittää tietoa tietojärjestelmästä ja ohjelmistosta, jossa tallenne on luotu ja sitä on käsitelty. Tällöin on paremmin mahdollista saada siirretyksi tallenne vanhasta järjestelmästä uuteen järjestelmään. Metatiedon avulla saadaan tietoa tallenteen luoneesta tietojärjestelmä- ja ohjelmistoympäristöstä. Metatieto sisältää tietoa tallenteen luomishetkestä ja vastaa tallenteeseen liittyviin kysymyksiin kuka, mitä, miksi ja missä. (Higgs 1998, 145)

Australia edelläkävijänä sähköisten tallenteiden hallinnassa

”Elektronisten tallenteiden ympäristö antaa tämän päivän arkistoinnin asiantuntijoille mahdollisuuden vaikuttaa ei ainoastaan siihen, kuinka tallenteita luodaan, vaan itse asiassa siihen, *mitä* tallenteita luodaan. Metatieto -teorian ja käytännön toteutuksesta saadut kokemukset auttavat arkistoihmisiä käyttämään aikaisempaa enemmän know how – tietämystään jo siinä vaiheessa, kun tallenteita luodaan.” Näin on kirjoittanut Australian Archives`n apulaisjohtaja Steve Stuckey. (Reed 2000) Australiassa on tutkittu metatiedon käyttöä sähköisten tallenteiden pitkäaikaisessa arkistoinnissa. Jo 1990-luvun alku- puolelta lähtien Australiassa on tutkittu, kehitetty ja suunniteltu ohjeistuksia sähköisten tallenteiden hallitsemiseksi. Vuonna 1995 hyväksyttiin ensimmäinen eri alan ammattilaisten yhteistyönä syntynyt standardi, joka liittyy Australiassa luotuun käsitteeseen ”tallenteiden jatkumo”. Standardi on nimeltään ”the Australian records management standard, AS 4390”. Standardin suunnittelivat arkistoinstituutioiden ammattilaiset, tal-

lenteiden hallinnan asiantuntijat, asiakirjahallinnon ja työnkulun asiantuntijat, kouluttajat ja harjoittelijat. Standardin hyväksymisen jälkeen Australian Council of Archives (ACA) järjesti kaikille Australian johtaville sähköisen tallenteiden ylläpidon asiantuntijoille sekä julkiselta että yrityssektorilta tapaamisen, jonka tuloksena saatiin yleinen tavoiteohjelma: ”Vuoteen 2000 mennessä kaikki australialaiset organisaatiot noudattavat yleisiin ja yhteisiin periaatteisiin, käsityksiin ja kriteereihin perustuvia elektronisten tallenteiden käsittelyyn liittyviä ohjeita ja standardeja.” (Reed 2000)

Mikä Australian esimerkissä on hyvää? Australiassa ollaan varsin pitkällä pitkäaikaisen arkistoinnin ongelmien tutkimisessa ja strategioiden kehittämisessä. Sähköisten tallenteiden ylläpitoon ja arkistointiin liittyvät projektit, jotka Australiassa ovat olleet käynnissä, ovat suurimmaksi osaksi liittyneet metatiedon käyttöön. Metatieto liittyy myös ns. jatkumo-käsitteeseen. Ainakin australialaiset tutkijat puhuvat mieluummin tallenteiden jatkumosta (records continuum) kuin perinteisestä elinkaari-ajattelusta. Tallenteiden jatkumo-käsitteen pohjalta australialaiset tutkijat ovat kehittäneet arkistoinnin hallintakäytäntöjä. Jatkumo-käsitteen myötä joudutaan ajattelemaan arkistointiakin uudesta näkökulmasta, joka näkyy australialaisten tutkimuksissa ja projekteissa. (Day 1999) Sue McKemmish ja Dagmar Parer (McKemmish et al 1998) vertailevat perinteistä arkistointikäsitettä ja jatkumo-käsitettä seuraavasti:

”Jos *arkistointi* määritellään *myöhemmin tapahtuvaksi siirroksi*, jolloin kontrolloidaan arkistokokoelmia valmistelemalla tallenteita koskevaa tietoa, tämä tieto itse asiassa toimii tallenteiden luetteloina, sijaisina, joiden ensisijainen tarkoitus on auttaa oikeiden tallenteiden löytämisessä. *Jatkumossa* arkistoinnin määritelmää tarkastellaan sen sijaan osana tallenteita ylläpitäviä monimutkaisia prosesseja, joissa on mukana tallenteiden luomishetkellä alkunsa saanut hallinnollinen metatieto.”

Edellä kuvattu *metatieto* on yleisesti tunnettu *tallenteita ylläpitäväksi metatietoksi* tai miksi tahansa muun tyyppiseksi tiedoksi/dataksi, joka auttaa meitä hallitsemaan tallenteita ja ymmärtämään niiden tietosisältöä. (McKemmish, Cunningham, Parer, 1998)

Australiassa ovat olleet meneillään mm. projektit SPIRT Recordkeeping Metadata Project ja Australian kansallisarkiston projekti. Kansallisarkiston projektin tuloksena kehitettiin toukokuussa 1999 standardi ”the Recordkeeping Metadata Standard for Com-

monwealth Agencies”. (Day 1999) SPIRT tulee sanoista Strategic Partnership with Industry and Research & Training. Tutkimusprojektin tarkoituksena on ollut kehittää runkoa metatiedon standardoimiselle tallenne- ja arkistoyhteisöissä sekä määrittellä korkeatasoinen tallenteita ylläpitävän metatiedon sarja. Projektin pääpaino on tallenteita koskevan metatiedon määrittelyssä ja käsitteellisessä kartoituksessa. (McKemmish, S, Acland G., 1999)

Victorian osavaltiossa on ollut käynnissä projekti, jonka loppuraporttikin on jo julkaistu jokin aika sitten. Projekti on nimeltään VERS eli Victorian Electronic Records Strategy, jossa myös metatieto on varsin keskeisessä asemassa. Metatieto on projektissa luodussa prototyyppi-sovelluksessa luomisen jälkeen muutettu lopulta XML-muotoon. Kyseessä on tallenteen ja sen sisältämän metatiedon kapselointi yhteen ja samaan tallenteeseen, XML-muodossa olevaksi tallennekokonaisuudeksi. (Chapter 6, Implementing an Electronic Archiving System. VERS, Final Report, 2000).

4.4.5 Tallenteiden luotettavuuden, aitouden varmistaminen

Kansallisarkiston asiakirjahallintoyksikön johtajan Jari Lybeckin mukaan esimerkiksi tietojärjestelmissä liikkuvan tiedon integriteetti (luotettavuus ja aitous) on keskeinen ongelma (Arkistoviesti 2/1998). Säilyykö tieto elinkaarensa kaikissa vaiheissa luotettavana ja aitona? Voidaanko olla todella varmoja siitä, että aineisto tulee siitä lähteestä, mistä sen väitetään tulevan ja ettei aineistoa ole asiattomasti manipuloitu?

Aitouden ja luotettavuuden varmistamiseen on olemassa Dollarin (1998, 26-27) mukaan erilaisia tekniikoita. Toisin kuin paperitallenteet sähköiset tallenteet ovat erittäin alttiita muuttamiselle ja turmeltumiselle. Sähköiset tallenteet voidaan yrittää turvata salasanalla ja lukuoikeuksilla, mutta niin kauan kuin tallenteet ovat niiden luojien ja vastaanottajien saatavilla, ne ovat alttiina muutoksille. Kun tallenteet siirretään ns. luotetun kolmannen osapuolen huostaan, joka on sitoutunut suojelemaan tallenteita, on otettu ensimmäinen kunnan askel uskottavuuden suojelemiseksi. Vakaa säilytysympäristö ja sen ylläpitäminen on ensisijaisen tärkeää tallenteiden suojelemisessa. Bearmanin ja Trantin (Bearman, D. and Trant, J., Authenticity of digital resources, 1998) mukaan sähköisten tallenteiden suojelemisessa aiheuttaa vaikeuksia se, että niitä voidaan muuttaa ilman mitään fyysi-

sesti näkyvää seikkaa. Tahallisille tai tahattomille muutoksille sähköiset tallenteet ovat alttiina aina, kun ne säännöllisin väliajoin joudutaan siirtämään uudelle muistivälineelle. Siten sama teknologia, jonka avulla on suhteellisen helppoa luoda, käyttää, varastoida ja hakea saataville sähköisiä tallenteita, tarjoaa myös mahdollisuuden muuttaa totuutta jättämättä mitään näkyvää jälkeä muutoksesta.

Tallenteiden aitouden ja luotettavuuden pysyvyyttä voidaan Dollarin (1998, 26-27) mukaan yrittää varmistaa mm. seuraavilla digitaalisilla tekniikoilla:

- Turvallinen asiakas-palvelin arkkitehtuuri, joka estää suoran pääsyn sähköisiin tallenteisiin. Tallenteet ovat paikassa, jossa niihin ei pääse käsiksi kuin ainoastaan tiedostojen lukuoikeudella. Asiakas-palvelin arkkitehtuurissa käyttäjä tekee hakuja tietokannasta tai antaa hakuohjeita tietyn sähköisen tallenteen etsimiseksi työasemalla. Työasemalta on pääsy palvelinlaitteistoon, joka oikean tulkinnan jälkeen ohjaa haut suoraan sähköisiin arkistoihin. Tietokantaa tukeva järjestelmä ottaa vastaan ja prosessoi hakuja ja hakuohjeita, ja päästää sitten tulokset takaisin palvelimelle ja käyttäjälle. Koska vain tiedoston luvun sallivassa asiakas-palvelin arkkitehtuurissa alkuperäiset sähköiset tallenteet eivät ole suoraan käyttäjien ulottuvilla, tallenteet ovat turvassa ehjässä säilytyspaikassa.
- Tiedonsiirron virheettömyyden tarkastusmenetelmää (CRC, Cyclical Redundancy Checksum, kts. liite 1) käytetään televiestinnässä virheistä vapaiden viestien siirron ja lähetyksen varmistamiseen. Tässä tekniikassa bittien lukumäärä paketissa jaetaan (käyttäen jakoprosessia ilman jakojäännöstä) ennalta määritellyllä 16-bittisellä tai 32-bittisellä polynomilla ja siitä saatava tulos lisätään pakettiin CRC:nä. Paketin siirron lopussa paketin tarkistussumma jaetaan samalla 16-bittisellä tai 32-bittisellä polynomilla. Jos kaksi CRC:tä ovat toisiinsa sopivia, on korkea todennäköisyys siitä, että mitään virhettä ei ole tapahtunut. CRC-algoritmia voidaan käyttää säilyttämään sähköisten tallenteiden yhtenäisyys tuottamalla yksi tai useampi CRC kullekin tallenteelle joko siten, että upotetaan ne tallenteeseen eräänlaisena digitaalisena allekirjoituksena tai siten, että liitetään ne osaksi tallenteen profiilia. Uudistamisen, muunnoksen tai migraation (järjestelmästä/ohjelmistosta toiseen siirtämisen) teon jälkeen on mahdollista laskea CRC:t kullekin tallenteelle ja verrata alkuperäisen CRC:n kanssa minkä

tahansa bittivirrassa tapahtuneiden muutosten tarkastamiseksi.

- Yksisuuntainen hajautus (one-way hashing) on toiminto, joka pakkaa digitaalisen objektin (kirjan tekstin, raportin, kirjeen, kuvan, äänitiedoston jne.) kiinteäpituisiksi tallenteiksi. Kaksi eniten käytettyä hajautettua tiivistysalgoritmia ovat MD5 ja SHA, jotka vähentävät lähdedokumentin kokoa mistä tahansa koosta 128 tai 160 bittiin. Toimintoa kutsutaan yksisuuntaiseksi hajautukseksi, koska hajautettu tiivistelmä itsessään on peruuttamaton siinä mielessä, että alkuperäisiä tallenteita ei ole mahdollista palauttaa hajautetusta tiivistelmästä. Huolimatta siitä kuinka monta kertaa dokumentti ”hajautetaan”, identtiset hajautetut tiivistelmät lähdedokumentista tuotetaan niin kauan kuin mitään muutosta ei ole tapahtunut. Siten laskemalla toinen hajautettu tiivistelmä yhdestä tai useammasta sähköisestä tallenteesta ja vertaamalla sitä alkuperäiseen hajautettuun tiivistelmään voidaan paljastaa mikä tahansa muutos sähköisissä tallenteissa.
- Neljäs digitaalinen tekniikka on yhdistää hajautettu tiivistelmä digitaaliseen aikaleimaan. Tallenteen yksisuuntainen hajautettu tiivistelmä tiettyä ajankohtana rekisteröidään kolmannen osapuolen todistuspalvelun avulla, kuten esimerkiksi Surety Technologies-yhtiön tuotteella nimeltään Digital Notary Record Authentication System. Käyttämällä Digital Notary Record Authentication (DNRA)-ohjelmistoa asiakas paloittelee minkä tahansa pituisen sähköisen tallenteen (teksti, kuva, piirustus) 288 bittiseksi hajautetuksi tiivistelmäksi, joka on ainoa laatuaan. Sitten se lähetetään Suretyn tai jonkun kolmannen osapuolen koordinoivalle palvelimelle, jossa päivämäärä-aika lähimmälle sekunnille liitetään mukaan. Tähän hajautettuun tiivistelmään yhdistetään kaikki muut samalla sekunnilla saadut hajautetut tiivistelmät ja kootaan siitä niin sanottu ”root superhash”, jonka jälkeen järjestelmä lähettää tarvittavan tiedon takaisin asiakkaalle niin, että ohjelmisto voi valmistaa aitoustodistuksen. Aitoustodistus sisältää hajautetun tiivistelmän kustakin tallenteesta, lasketun polun, joka syntyi kun palat yhdistettiin ”root superhash”-muotoon sekä sen päivämäärän. Tämä aitoustodistus voidaan tallentaa ja varastoida kunkin tallenteen tai tietokannan kanssa asiakkaan vaatimusten mukaisesti.

4.4.6 Tallenteiden migraatio

Migraation määritelmä eri tutkijoiden mukaan

Dollarin (1998, 16) mukaan migraatio tarkoittaa tiedonsiirtoprosessia tietokonejärjestelmästä, ohjelmistosta tai ohjelmasta toiseen muuntamalla tietoa. Kelly Russelin (Russel, K., 1999) mukaan migraation tarkoituksena on säilyttää digitaalisten objektien yhtenäisyys ja säilyttää asiakkaille mahdollisuus hakea, näyttää ja muuten käyttää tallenteita jatkuvasti muuttuvassa teknologisessa maailmassa. Dollarin (1998, 16) mukaan migraatio eroaa muista edellä mainituista sähköisille tallenteille tehtävistä toimenpiteistä siten, että se sisältää piirteitä vaatimusten uudistamisesta, kopioinnista, muuntamisesta luettavalle medialle, käytettäville laitteistoille ja kohdevälineille. Erona näihin muihin toimenpiteisiin migraatiossa ovat mukana vanhat järjestelmät, joilla ei ole ohjelmiston siirtotoiminnallisuutta. Morellin (1998, 180) mukaan on erityisen tärkeää se, että migraatiossa ei aiheuteta mitään muutosta informaatioon ja tallenteeseen. Migraatiosta voidaan suppeammassa mielessä käyttää myös nimitystä muunnos eli konversio, koska muunnoksestahan siinä on myös kyse. Migraatioon sisältyy kuitenkin muitakin toimenpiteitä kuin pelkkä muunnos.

Migraation suunnittelu

Dollarin (1998, 52-53) mukaan aitouden säilyttämisen lisäksi migraatiossa tulisi säilyttää myös tallenteiden luettavuus, tunnistettavuus, saatavuus, selvyys, palautettavuus ja ymmärrettävyys. Näiden seikkojen säilyttäminen tulisi olla migraation tavoitteena. Migraatiolle tulisi määritellä oikeat toteutusolosuhteet. Lisäksi tulisi päättää siitä, kuinka prosessi dokumentoidaan. Migraation toteutuksesta vastuussa oleva taho tai henkilö nimetään sekä suunnitellaan laatukontrollitoimenpiteet. Jos rahoitusresurssit, saatavissa oleva teknologia tai muut olosuhteet puoltavat täyttä ja perusteellista migraatiota, päätös sähköisten tallenteiden migraatiosta pitäisi dokumentoida täydellisesti kuten myös päätös migraation käyttämättä jättämisestä. Laatukontrolliasiakirjoissa tulisi määrätä migraatiosta vastuussa oleva yksikkö tai henkilö, jolla ei ole mitään näkemyksellisiä ristiriitoja tallenteiden sisällön suhteen. Koska migraatio voi aiheuttaa todennäköisesti jossain määrin vahinkoa tallenteille, ennen migraation toteutusta tulisi testata sitä pienelle tallenteiden joukolla, jotta nähtäisiin, aiheutuuko niille vahinkoa vai ei. Varsinaisen migraation aikana aiemmin migraation läpi käyneitä tallenteita voidaan verrata vanhaan järjestelmään upotettuihin tallenteisiin, jotta voidaan nähdä mahdolliset virheet. Sekä

kirjoitettu migraatiotoimenpideluettelo että laatukontrolliohjeet ovat erittäin merkittäviä tekijöitä silloin, kun migraatiossa aiheutuu joitain turmelusta tallenteiden rakenteelle, sisällölle tai kontekstille. Arkistotieteen näkökulmasta katsottuna tällaiset migraatiossa muuttuneet tallenteet ovat uusia tallenteita, joiden aitous ja luotettavuus tulisi määrittää uudelleen. Kun on olemassa laatukontrolliasiakirjat, joissa on kerrottu kaikki migraatioon liittyvät päätökset ja toimenpiteet, voidaan tarjota perusta migraatiossa muutoksia kokeneiden tallenteiden aitouden tutkimista varten. Laatukontrolliasiakirjoilla on tärkeä rooli myös silloin kun päätetään, siirretäänkö tallenteet paperille vai mikrofilmille. Kun dokumentoidaan koko migraatioprosessi ja pidetään kiinni tiukoista menettelytavoista, jotka tulisi ajoittain tarkastaa, niin tehdään tärkeää työtä sähköisten tallenteiden aitouden ja luotettavuuden suojelemisessa.

Mackenzie Owenin ja Wallen (1996, 52-53) mukaan migraatiota suunniteltaessa ensimmäinen askel on miettiä onko kyseessä lyhytaikainen vai pitkäaikainen migraatio. Lyhytaikainen migraatio käsittää viiden vuoden kuluttua toteutettavat toimenpiteet. Pitkäaikainen migraatio käsittää toimenpiteet, jotka tulee toteuttaa tavallisesti 10-25 vuoden kuluttua. Lyhytaikaisen migraation suunnittelussa on huomioitava erilaisia migraatio-toimenpiteitä tarvitsevien materiaalien osuus ja toimenpiteestä aiheutuvat kustannukset, koska migraation toteuttaminen ei ole halpaa. Migraation toteuttavan organisaation on mietittävä tarkoin omat tekniset resurssinsa ja selvitettävä nykyisten laitteistojen ja ohjelmistojen toimivuus 5-10 vuoden päästä. Tallenteet on järkevää varastoida sellaiselle medialle, jolla on mahdollisimman pitkä kestävyys. Sen sijaan pitkäaikaisen migraation suunnittelu on erittäin vaikeaa, koska on melkein mahdotonta ennustaa, missä tilassa teknologian suhteen ollaan joskus 10-25 vuoden päästä ja mitä toimenpiteitä tallenteiden siirtämiseen vanhasta uuteen teknologiaan on silloin tarjolla. Epävarmuutta aiheuttaa myös epätietoisuus teknologisen kehityksen vauhdista – jatkuuko se samaan malliin vai hidastuuko tai nopeutuuko se entisestään?

Migraation toteuttaminen käytännössä

Alison Bullockin (1999) mukaan migraatiossa tulee aina tapahtumaan jossain määrin tiedon häviämistä, mitä pyritään välttämään. Tietoarkistoilla on pitkä historia migraation käytössä, mutta ne ovat kuitenkin olleet tekemisissä melko homogeenisen tiedon kanssa, joka on tallennettu vain hyväksytyihin formaatteihin. Muutamat arkistot muuntavat ei-standardin mukaiset formaatit yhteen tai kahteen standardiin tallenteita vastaan-

otettaessa. Yksi migraatioesimerkki on Australian kansalliskirjastosta, joka siirsi pienen otoksen kaupallisia julkaisuja disketeiltä CD-ROM-levylle. Noin 35%:ia näistä ei voitu testata tai käyttää, koska kirjastolla ei ollut tarvittavaa laitteistoa tai ohjelmistoa. Eri syistä johtuen vain puolet noin 40 CD-ROM-levylle kopioidun disketin tiedoista oli toimintakuntoisia kopioinnin jälkeen. Bullockin mukaan on olemassa keinoja parantaa migraation käyttöä seuraavilla tavoilla:

- Kontekstin poimiminen dokumentoimalla laitteisto ja ohjelmisto, jotka vaaditaan pääsyyn kyseisiin digitaalisiin objekteihin ja käytetyn teknologian myöhemmän saatavuuden varmistaminen
- Ohjelmiston yhteensopivuuden arvioiminen myöhempien ohjelmistojen kanssa migraation ajankohdan määrittämiseksi
- Pienen standardiformaattien määrän valitseminen ja ei-standardiformaattien muuntaminen standardiformaatteihin minimoimaan myöhemmin muunnettavien tallenteiden määrää.

Mackenzie Owenin ja Wallen (1996, 50-52) teoksessa annetaan ohjeita elektronisille julkaisuille tehtävästä migraatiosta, mutta ohjeet voivat olla melko käyttökelpoisia myös muillekin elektronisille tallenteille tehtävää migraatiota varten. He suosittelevat seuraavia strategioita:

- Välineen virkistäminen:
 - Tiedon kopioiminen fyysiseltä välineeltä (esim. disketiltä) toiselle samantyyppiselle välineelle (esim. uudelle disketille) on välttämätöntä, kun median rappeutuminen uhkaa. 20-25 vuoden arvioidun eliniän omaaville medioille virkistäminen ei luultavasti ole välttämätöntä ennen kuin teknologian vanhentuminen tekee sen välttämättömäksi.
- Välineen muuntaminen:
 - Elektronisten julkaisujen siirto yhdeltä (tavallisesti vähemmän pysyvältä/vakaalta tai standardoidulta) välineeltä toiselle (pysyvämpi/vakaampi tai standardoidumpi) välineelle. Välineen muuntaminen on tarpeellista myös silloin, kun media vanhentuu ja sitä eivät laitteistojen valmistajat enää tue.

- Formaatin muuntaminen:
 - Dataformaatin muuntamistavoitteena toiseen on normaalisti organisaatiossa käytettävien erilaisten formaattien lukumäärän vähentäminen ja tallenteen formaatin muuttaminen organisaatiossa käytettävien formaattien mukaiseksi. Muuntamisessa käytettävän ohjelmiston tulee olla myös koformaatteihin ja standardeihin sopiva.
- Teknisen ympäristön migraatio:
 - Tallenne muutetaan siten, että se toimii erilaisessa teknologisessa ympäristössä (laitteisto & ohjelmisto) kuin mihin se oli alun perin tarkoitettu. Tarve tekniseen migraatioon kasvaa sitä mukaa kun yhteys informaation ja ohjelmoitujen toimintojen välillä käy ahtaaksi. Tavallisen interaktiivisen elektronisen multimediajulkaisun osalta jouduttaisiin toteuttamaan osittain uudelleen ohjelmointi. Mackenzie Owenin ja Wallen mielestä on liian aikaista alkaa arvioida suurten, hyvin erilaisten elektronisten julkaisujen teknisessä migraatiossa eteen tulevia vaikeuksia ja mahdollisuuksia. Joka tapauksessa voidaan todeta, että tämän tapainen migraatio tulee maksamaan hyvin paljon, eikä se välttämättä ole mahdollista ilman kyseisen organisaation toiminnassa aiheutuvia heikentymisiä.
- Teknisen ympäristön emulaatio:
 - Teknisen ympäristön migraatio tulee hyvin kalliiksi tai on käytännössä jopa mahdotonta. Vaihtoehto sille voisi olla emulaattorien (joista enemmän kappaleessa 4.4.6) kehittäminen.

Dollar (1998, 33-34) puhuu migraation yhteydessä lähinnä vanhoista, ns. legacy – järjestelmistä, jotka olivat käytössä 1980-luvun alussa ennen avointen järjestelmien kehittymistä, ennen toisiinsa kytkettävyyden mahdollisuutta sekä ennen siirrettäviä tallenteita ja sovelluksia. Hänen mukaansa vanhoille legacy –järjestelmille tehtävän migraation onnistumisen edellytyksiä uuteen järjestelmään vaihe vaiheelta voidaan parantaa seuraavasti:

- Vanhan (legacy)-järjestelmän analysointi
- Vanhan tietojärjestelmän rakenteen jakaminen osiin/hajottaminen
- Kohdejärjestelmän käyttöliittymän suunnittelu
- Kohdejärjestelmän sovellusten suunnittelu

- Kohdejärjestelmän tietokannan suunnittelu
- Kohdejärjestelmän ympäristön asennus
- Tarvittavien porttien ja väylien asentaminen ja luominen
- Vanhan järjestelmän tietokannan migraatio
- Vanhan järjestelmän sovellusten migraatio
- Vanhan järjestelmän käyttöliittymän migraatio

Edellä mainittuja kymmentä vaihetta ei voida kaikissa migraatioissa noudattaa, koska jokaisessa migraatiotapauksessa ovat erilaiset olosuhteet, jotka voivat vaihdella huomattavastikin.

- Vanhan tietojärjestelmän (legacy-järjestelmä) analyysissa yritetään oppia mahdollisimman paljon vanhan järjestelmän toiminnasta ja siihen upotetuista tallenteista. Erilaisten toimintojen selitysten ja perustelujen ymmärtämiseen tulee kiinnittää paljon huomiota. Huomiota tulee kiinnittää myös metatiedon poimimiseen, sen ja tallenteiden sekä eri tallenteiden väliseen suhteeseen. Vanhaa legacy -järjestelmää analysoitaessa saadaan selville määrittelyjä, joita käytetään toimintojen, metatiedon ja tallenteiden uudelleen järjestelyssä uuteen järjestelmään. Analyysi toimii myös tarkistuslistana sille, ettei mitään migraatiossa jätetä huomiotta.
- Vanhan järjestelmän tietoarkkitehtuurin jakaminen osiin mahdollistaa kaikkien yksittäisten komponenttien, käyttöliittymien, sovellusten ja tietokantojen käsittelemistä kutakin erikseen. Brodien ja Stonebrakerin mukaan vanhan järjestelmän tietojärjestelmäarkkitehtuuria voidaan havainnoida osiin jaettuna, puoliksi osiin jaettuna ja osiin jakamattomana. Osiin jaettu arkkitehtuuri on sellainen, jossa kaikki komponentit, niin järjestelmän kuin käyttäjienkin käyttöliittymät, sovellusmodulit ja tietokannat ovat erillisiä ja itsenäisiä komponentteja. Puoliksi osiin jaetussa arkkitehtuurissa käyttöliittymät ja tietokanta ovat itsenäisiä, mutta sovellukset ja tietokantapalvelut ovat yhtenä modulina. Osiin jakamattomassa arkkitehtuurissa käyttöliittymät, sovellukset ja tietokantapalvelut ovat tiiviisti yhdessä modulissa. Kaikki riippuvuudet järjestelmän ulkopuolelle täytyy eliminoida valmistauduttaessa migraatioon.
- Kolmanteen, neljänteen ja viidenteen vaiheeseen sisältyy uuden tietojärjestelmäympäristön tarkastelua ja suunnittelua. Kohdejärjestelmän käyttöliittymien, sovellusten

ja tietokannan tulee tarjota samanaikaiset yhteydet vanhan tietojärjestelmän käyttöliittymiin, sovelluksiin ja tietokantaan.

- Kuudennessa vaiheessa identifioidaan ja valitaan avoin kohdeympäristö kunnollisine asennustyökaluineen.
- Seitsemännessä vaiheessa painopiste on porteissa ja väylissä, joilla on tavallisesti kaksi roolia. Yksi rooli on eristää valitut komponentit muutoksista, jotka on tehty muille komponenteille. Esimerkiksi porttien ja väylien kohdalla pidetään yllä vanhan tietojärjestelmän käyttöliittymää, jonka käyttäjä saattaa odottaa näkevänsä, vaikka vanhassa järjestelmässä tapahtuukin kaiken takana samaan aikaan muutoksia. Toinen rooli on toimia pyyntöjen ja tiedon kääntäjänä neuvottelevien komponenttien välillä. Portit ja väylät tulee suunnitella, luoda ja asentaa varmistamaan johdonmukaisuutta ja tarkkuutta vanhan tietojärjestelmän toimintojen jäljentämisessä kohdetietojärjestelmässä ja tietojen (tallenteiden) siirrossa. Portteja ja väyliä tulee testata tarpeeksi vanhan järjestelmän sisältävien tallenteiden otoksilla varmistamaan johdonmukaisuutta ja tarkkuutta.
- Kahdeksannessa, yhdeksännessä ja kymmenennessä vaiheessa tapahtuu itse tallenteiden, sovelluksen toimintojen ja käyttöliittymien siirto vanhasta tietojärjestelmästä uuteen kohdetietojärjestelmään. Tallenteiden siirto on yksinkertainen toimenpide, kun taas vanhan järjestelmän sovellusten ja käyttöliittymien siirto on aika ongelmallista. Brodien ja Stonebrakerin mukaan vanhan järjestelmän koodin migraatio kohdetietojärjestelmään on mahdotonta, joten uusi sovelluskoodi pitää kirjoittaa kohdetietojärjestelmää varten uudelleen. Melko todennäköisesti vanhan järjestelmän käyttöliittymät (merkkipohjaiset menut/komennot ja näytöt) täytyy korvata graafisilla kuvakkeilla ja kuvilla. (Dollar 1998, 33-34)

Migraation toteuttamisessa olevat ongelmat

Migraatio on kieltämättä tärkeä strategia digitaalisten objektien ja tallenteiden säilyttämisessä. Kuitenkaan sitä ei ole Bullockin (1999) mukaan vielä testattu eikä sen käyttökelpoisuutta ole varmistettu monimutkaisten multimediaobjektien hallintaan pitkällä ajanjaksolla. Migraation arvioidaan Mackenzie Owenin ja Wallen (1996, 53-54) mukaan olevan mutkikasta ja tulevan kalliiksi jo senkin vuoksi, ettei siitä ole käytännön kokemusta sen aiheuttamien tarkkojen kustannuksien laskemiseksi. Kustannuksiin vaikuttavat muutettavien tallenteiden muoto, formaatti, ikä jne. Vuosittaisiin migraatiokustannuksiin vaikuttavat 1) alkuperäisten tietorakenteiden monimutkaisuus, koska migraa-

tio on halvempaa tehdä ASCII –tekstille kuin monimutkaiselle interaktiiviselle multimediajulkaisulle. 2) migraation esiintymistaajuus, johon vaikuttavat median, ohjelmiston, laitteiston ja käyttöjärjestelmän elinkaari. 3) sisällön lisäksi tietokoneen toimivuuden, haettavuuden, esittämisen, linkityksen toimivuuden yms. vuoksi tehtävän migraation määrä. 4) uuden median, laitteiston ja käyttöjärjestelmän hinta. Varastointikustannukset megatavua kohden todennäköisesti pienenevät verrattuna muihin edellä mainittuihin kustannuksiin.

Dollarin (1998, 17) mukaan edellä mainittujen tekijöiden lisäksi kustannuksien kasvuun vaikuttaa tarve suunnitella portteja, käytäviä ja käyttöliittymiä ja kirjoittaa ohjelmistoja tallenteiden siirtoa varten, koska tarvitaan tiettyjä ohjelmistotoimintoja uuteen laitteistojen ja ohjelmistoympäristöön. Kustannuksia tulee myös monista testauskerroista. Migraation itsensä tulee sallia osalle käyttäjistä samanlainen näkymä tallenteista kuin oli aikanaan niiden luojilla ja vastaanottajilla, eli toisin sanoen migraation tulee tukea aitojen sähköisten tallenteiden olemassaoloa. Arkistotieteen näkökulmasta katsottuna aitoudesta puhutaan monesti alkuperäisyys-sanon yhteydessä. Alkuperäisenä tallenteena voidaan pitää myös alkuperäistä jäljittelevää kopiota (imitative copy), koska siinä on kaikki olennaiset yksityiskohdat, jotka tekevät siitä tehokkaan ja täydellisen tallenteen verrattuna alkuperäiseen. Dollarin mukaan kaikissa migraatio-tapauksissa ei ole kenties mahdollista siirtää tallennetta jäljittelevää kopiota, koska tuloksena saattaisi olla erilainen esitystapa, tulkinta tai käänös, jolloin aitouden kanssa tulisi ongelmia. Voidaan mahdollisesti myös kadottaa jotain käyttökontekstiin olennaisesti liittyvästä ohjelmiston toiminnallisuudesta. Tämä uhka voitaisiin kuitenkin ennakoida riskejä arvioimalla jo ennen tehtävää toimenpidettä.

Migraatio ja ei-migraatio

Arkistotieteen näkökulmasta tallenteiden aitouden ja luotettavuuden säilyminen on niin merkityksellinen asia, että migraation jälkeen olisi välttämätöntä varmistua uudelleen tallenteiden aitoudesta Dollarin (1998, 17) mukaan. Nimittäin tallenteet, joille on tehty migraatio, ovat uusia lähteitä. Kun dokumentoidaan koko migraatioprosessi ja pidetään kiinni ajoittain tarkastettavista menettelytavoista, tehdään tärkeää työtä sähköisten tallenteiden aitouden ja luotettavuuden suojelemisessa.

Migraatiota suunniteltaessa voidaan tehdä myös sellainenkin päätös, että päätetään olla tekemättä migraatiota. Englanniksi tätä toimenpidettä kutsutaan nimellä ”non-migration”. Joissakin tapauksissa on teknisesti mahdollista säilyttää tallenteiden toiminnallisuus ja yhtenäisyys, mutta migraation kustannukset tulisivat tuolloin ylittämään käytettävissä olevat resurssit. Joskus taas on myös teknisesti mahdotonta toteuttaa migraatiota säilyttämällä tallenteiden toiminnallisuus ja yhtenäisyys ja olla aiheuttamatta tallenteiden turmeltumista ja tietojen häviämistä. Näissä tapauksissa tulevat harkittaviksi seuraavat ”non-migration”-toimenpiteet:

- Jatketaan tallenteiden ylläpitoa prosessoitavassa muodossa kopioimalla ja uudistamalla niitä olettaen, että tulevaisuuden teknologia tekee migraation paljon halvemmaksi tai että voidaan kehittää käyttöjärjestelmä ja sovellukset, jotka matkivat alkuperäistä käyttöjärjestelmää ja sovellusohjelmia. Sähköisten tallenteiden ylläpito prosessoitavassa muodossa tulee kalliiksi, vaikka niitä ei pystyttäisikään palauttamaan alkuperäiseen muotoonsa ohjelmistoriippuvuuksien vuoksi.
- Luodaan sähköisille tallenteille ns. sijainen, joka poimii tietosisällön, mutta ei koko alkuperäistä esitystä tai käyttökontekstia. Yksi esimerkki tästä olisi siirtää tekstitalenteet tietyistä tekstinkäsittely-ympäristöstä ASCII-tekstiksi tai sitä vastaavaksi. Toinen esimerkki olisi kopioida tietokanta tietyinä ajankohtana niin, että kopio toimii kuvana tai ns. snapshot`ina dynaamisesta tietokannasta.
- Siirretään sähköiset tallenteet mikrofilmille tai paperille niin, että tallenteiden rakenne, sisältö ja konteksti olisi saatavilla, mutta ei kuitenkaan prosessoitavassa muodossa. Toimenpiteellä on seuraavia hyötynäkökohtia:
 - paperille tai filmille siirrettyjen sähköisten tallenteiden aitous on suhteellisen helppoa todeta, jos kunnollista metodologiaa vain noudatetaan.
 - paperilla tai mikrofilmillä olevien tallenteiden muuttaminen on paljon vaikeampaa toteuttaa ilman näkyviä jälkiä kuin sähköisessä muodossa olevien tallenteiden muuttaminen.
 - sähköisten tallenteiden paperille tai mikrofilmille siirtämisessä on myös se etu, että voidaan eliminoida ohjelmisto vanhenemisen ongelma, koska tulostettuja tallenteita voidaan lukea ilman ohjelmistoja.

Sähköisten tallenteiden paperille ja mikrofilmille siirtämisestä on myös merkittäviä haittoja. Tallenteita ei voida enää nopeasti hakea, vaikka kylläkin voidaan säilyttää tietokoneella etsimistä ja löytämistä helpottavia paperi- ja mikrofilmitallenteiden hakemistoja. Paperille tai mikrofilmille siirrettäessä hypertekstilinkeistä, multimediasta, relaatiotietokannoista, äänisignaaleista, taulukkolaskentaohjelmista ja maantieteellisistä tietojärjestelmistä (Geographic Information Systems) koostuvat sähköiset tallenteet kadottavat ne kaikki toiminnallisuudet, jotka tukevat niiden yksilöllisiä muotoja ja rakenteita. Yleisesti voidaan sanoa, että paperille tai mikrofilmille siirto on viimeinen pelastus tai keino. (Dollar 1998, 35-36)

4.4.7 Tallenteiden emulaatio

Emuloinnilla voidaan parhaassa tapauksessa mahdollisesti ratkaista vanhenevien digitaalisten tiedostomuotojen aiheuttamia ongelmia. Esimerkiksi vanhan kotimikron toiminnot voidaan toteuttaa nykyaikaisen koneen "sisällä". Emulaatio tulee sanasta emulaattori, joka on vanhaa ohjelmaa/järjestelmää matkiva ohjelma. Vanhan Amstrad CPC-koneen kolmituumaisten levykkeiden sisältö voidaan siirtää tällä hetkellä käyttökelpoisista Windows-käyttöjärjestelmää käyttävän koneen muistiin. CPC:n ROM:in sisältö voidaan kopioida ja muuttaa emulaattoriohjelmaksi. Tämän jälkeen CPC-ohjelmia voi käyttää uuden koneen sisällä virtuaalisesti. Ilman varsinaista emulointia voidaan vastavasti käyttää hyväksi eri tiedostomuotoja. (Rothenberg, J., 1995 ; Morelli 1998, 181) Emulointitarve voidaan myös ratkaista siten, että dokumentin mukana tallennetaan dokumentin käytössä tarpeellisten ohjelmien rakennusohjeet. Tätä menettelyä kutsutaan kapseloinniksi. (Leinonen, A., 1998)

Jeff Rothenberg on yksi niistä tutkijoista, jotka kannattavat ja suosittelevat emulointia pitkäaikaisen säilytyksen ja arkistoinnin ongelmien ratkaisemiseksi. Sähköisen tallenteen käyttö tulevaisuudessa edellyttää Rothenbergin (1995) mukaan sen luomisessa käytetyn ympäristön simulointia, varsinkin silloin, kun aineiston muuntamisessa häviäisi joitakin olennaisia piirteitä aineiston luonteesta. Emulointi tarkoittaa vanhan tietokoneympäristön laitteiston toimintaa imitoivan ohjelman luomista. Tämä edellyttää, että tiedot ja kuvaukset vanhasta laitteistosta ja käyttöympäristöstä talletetaan tallenteen mukana. Emulointi tarkoittaa siten vanhan käyttö- ja ohjelmistoympäristön matkimista

uudessa ympäristössä. Emuloinnin toteuttaminen tulevaisuudessa vaatii hyvin yksityiskohtaisen kuvauksen vanhentuneesta käyttöympäristöstä ja laitteistosta. Huonona puolella tässä on se, että kuvauksen tarkkuus voi olla ongelmallista toteuttaa tarpeeksi hyvin. Emulointi on menetelmänä aika työläs, koska jokainen uusi laitteistoympäristö tarvitsee uuden emulaattorin edellisiin ympäristöihin ja koko ajan melkein vuosittain julkaistaan tälläkin hetkellä uusia käyttöjärjestelmiä ja ohjelmistoja. Sitä paitsi kaikkia vanhan järjestelmän ominaisuuksia ei saada välttämättä rakennettua emulaattoriin. Tulevaisuudessa emulointia tekevän henkilön tulee hallita vanhat ohjelmistot ja käyttöjärjestelmät.

Jeffrey D. Morelli (1998, 181-182) käsittelee artikkelissaan myös emulaatiota. Etuna esimerkiksi migraatioon emulaatiota tehdessä ei tarvitse muuntaa tietoa, vaan yksinkertaisesti vain ajetaan emulaatio-ohjelmaa, jonka avulla taataan pääsy kyseiseen tietoon. Morelli ei kuitenkaan juurikaan kannata emulaatiota, koska hän näkee sen olevan melko tehoton ja paljon ongelmia aiheuttava vaihtoehto. Esimerkkinä emulaation toteuttamisvaikeuksista hän mainitsee: ”Kuka tahansa, joka on yrittänyt saada oman henkilökohtaisen tietokoneensa lukemaan ulkomaalaisessa formaatissa olevaa levyä, ymmärtää vaikeudet, koska siinä vaaditaan puuhastelua isäntätietokoneen käyttöjärjestelmäsovelluksen ja levynpään hallintamekanismien mekaanisten elementtien kanssa”. Tämä kuulostaa niin monimutkaiselta ja vaikealta, että emulaatio ei vaikuta kovin järkevältä vaihtoehdolta. Toinen emulaation mukanaan tuoma ongelma tulee Morellin mukaan esiin siinä, että emulaattoria suunnittelevan henkilön tulee täysin ymmärtää vanhan järjestelmän ohjelmiston koodia voidakseen luoda tätä ohjelmistoa matkivan emulaattorin. Kuinka hän siihen pystyy, kun välttämättä kyseisen ohjelmiston suunnittelijatkaan eivät ole täysin oman ohjelmistonsa koodia ymmärtäneet. Morellin mielestä kysymys, onko emulaation toteuttaminen mahdollista, on lähinnä akateeminen. Pitkällä aikavälillä emulaatio aiheuttaa todennäköisesti ongelmia. Saatetaan esimerkiksi tarvita kuitenkin muunnoksen toteuttamista ja päivittämistä uuteen mediaan, laitteistoon, käyttöjärjestelmään jne., koska emulaattorin luonut laitteisto on tullut vanhaksi. Tässä on yksi pätevä syy, miksi Morelli (1998, 182) artikkelinsa lopussa toteaa, että ”on luultavasti parempi jättää laitteiston emulointi huomiotta ja keskittyä muuntamaan ja ylläpitämään kohteena olevaa tietoa elinkykyisessä ympäristössä.” Avoimeksi jää, pitääkö hän lopulta emulaatiota lainkaan järkevänä vaihtoehtona.

David Bearman on kritisoinut emulaatioteoriaa sillä perusteella, että sitä on joissakin piireissä pidetty ratkaisuna sähköisten tallenteiden pitkäaikaisen arkistoinnin ongelmiin, vaikka sitä ei ole vielä edes testattu. Bearmanin mukaan Rothenbergkaan ei ole painottanut emulaation olevan ehdottomasti käyttökelpoinen ratkaisu. Bearman toteaa, että "sähköiset tallenteet, joita ei siirretä pois vanhentuneista laitteisto- ja ohjelmistoympäristöistä, hyvin todennäköisesti tulevat kuolemaan näiden vanhentuneiden laitteistojen ja ohjelmistojen mukana" (Bearman, D., 1999). Rothenberg on Bearmanin mukaan väärillä linjoilla siinä, kun hän yrittää säilyttää tietojärjestelmät tallenteiden säilyttämisen sijasta. Emulaation avulla ei voitaisi säilyttää tallenteita todisteena jostakin toimenpiteestä (aitous), koska se ei ole emulaatiossa vaatimuksena. (Bearman 1999) Rothenbergin ja Bearmanin kohdalla tuleekin hyvin esiin Dollarin esittämä jaottelu kahteen koulukuntaan. Rothenberg edustaa sitä koulukuntaa, jonka mukaan sähköisten tallenteiden pitkäaikaisen säilytyksen ongelmat ovat teknologisia. Bearman edustaa taas sitä koulukuntaa, jonka mukaan ongelma on sekä filosofinen että teknologinen. Bearmanin edustaman koulukunnan näkemysten mukaan tallenteiden luettavuuden varmistaminen ei ole riittävä strategia. Siinä jää kokonaan huomiotta aitouden ja luotettavuuden varmistaminen, mikä arkistotieteen näkökulmasta on tärkeää. Vaikka tallenne olisi lukukelpoinen, se ei ole arkistotieteen näkökulmasta käyttökelpoinen todisteena, jos sen aitoutta ei pystytä varmistamaan.

5 Elektronisen arkistoinnin käytännön ratkaisut arkistolaitoksessa

5.1 Elektronisten tallenteiden käsittelyohjeiden suuntaviivoja

Pohjolan (1999) mukaan vuonna 1992 Suomen kansallisarkisto laati sähköisten materiaalien arvioinnille joitakin suuntaviivoja, joiden avulla materiaaleja voitaisiin lainsäädännölliset tekijät ja tieteellisen tutkimuksen tarpeet huomioon ottaen tehokkaammin käyttää. Vuonna 1994 kansallisarkistossa perustettiin työryhmä, joka tutki sähköisiä materiaaleja keskushallinnossa ja tutkimusinstituutioissa. Tutkimuksessa käytetty materiaali oli pääasiassa erilaisia rekistereitä ja tietokantoja. Tutkimuksen tuloksena todettiin, että edellä mainittujen materiaalien määrä oli 200 GB tietoa ja vuosittainen materiaalien kasvuvauhti oli n. 10 %. Työryhmä antoi tutkimustulostensa lisäksi myös muutamia suosituksia koskien näiden materiaalien arviointia ja hävittämistä:

- Kansallisarkiston pitäisi tehdä ero toisaalta pysyvän säilytyksen ja jatkuvan säilytyksen ja toisaalta määräaikaisen säilytyksen välillä
- Pysyvä säilytys tarkoittaisi, että kansallisarkisto on sitoutunut säilyttämään materiaalit kaikissa olosuhteissa. Taloudellisin tapa säilyttää laaja tietokanta olisi ottaa se COM -mikrofilminä. Työryhmän mukaan vuosittainen 1000 sivun määrä sopisi COM -ratkaisuun. Siinä tapauksessa, että materiaalien tutkimuskäyttö olisi vaikeaa, COM -mikrofilmit säilytettäisiin jatkuvasti myös sähköisesti.
- Jatkuva tai määräaikainen säilytys sähköisessä muodossa tapahtuisi silloin, kun todistusvoimaisuus ei olisi pääasiallinen tekijä.

Työryhmän kokemusten mukaan sähköisten materiaalien pysyvä säilytys on välttämättömyyttä tietyissä tapauksissa. Pysyvästi säilytettävät materiaalit tulisi säilyttää erillään niistä materiaaleista, jotka aiotaan säilyttää toistaiseksi. Työryhmä ehdotti kansallisarkiston sisäiseen sähköisten materiaalien arviointiin tiettyjä käytäntöjä: Tietojärjestelmän parissa työskentelyn tulisi kuulua osana tallenteiden luojan valtuutettuihin toimintoihin. Viitattavaa tietoa sisältävät tietokannat tulisi säilyttää vain tapauksissa, joissa ne muodostavat osan tallenteiden luojan rekisterijärjestelmää. Työryhmä tutki myös tarvetta tutkimukseen sisältyneiden järjestelmien pysyvään tai jatkuvaan säilytykseen. Oli yllättävää, että monissa tapauksissa arkistonmuodostussuunnitelman muodostusprosessi oli tuonut mukanaan tietoa, joka tarjosi arviointipäätösten teolle pohjan. Itse tutkimuksessa tehtiin jo olemassa olevien aktiivisten järjestelmien ja materiaalien pohjalta arviointipäätöksiä.

Siinä vaiheessa, kun järjestelmät olivat olleet suunnitteluvaiheessa, arviointiin ei ollut kiinnitetty ollenkaan huomiota. Mutta suunnitteluvaiheen jälkeen otetut arviointinäkökohdat auttoivat tuomaan esiin uutta tietoisuutta materiaalien pitkäaikaisen säilytyksen tarkoituksista. Tutkimuksessa mukana olleet instituutiot olivat suuria tutkimusinstituutioita, joilla oli hyvä erilaisia sähköisiä rekistereitä tuottava IT-infrastruktuuri. Kansallisarkisto päätti siltä pohjalta, että tutkimuksessa mukana olleet ja niiden kaltaiset tutkimusinstituutiot säilyttävät pysyvään tai jatkuvaan säilytykseen tarkoitetut sähköiset materiaalinsa itse niin kauan kuin mahdollista, koska niillä on parhaat tiedot kyseisten materiaalien käytöstä.

Raimo Pohjola kertoo, minkälaisia elinkaaren vaiheita on Suomessa ajateltu elektronisilla asiakirjoilla olevan. Elektronisten asiakirjojen käsittelyn tulevaisuutta arvioineen vuonna 1996 valmistuneen raportin sähköisillä asiakirjoilla on samat elinkaaren vaiheet kuin ei-elektronisillakin asiakirjoilla, eli aktiivinen, passiivinen ja historiallinen vaihe:

- Aktiivinen vaihe:
 - Asiakirja tuodaan ensin osaksi asiakirjojen ylläpitojärjestelmää rekisteröinnin kautta, jolloin se saa mahdollisesti myös elektronisen allekirjoituksen.
 - Asiakirjan yhtenäisyyttä, eheyttä ja aitoutta tarkastellaan uudelleen organisaation sisäisessä informaatiojärjestelmässä.
 - Järjestelmää tarkastetaan ja kontrolloidaan riippumattoman/puolueettoman kolmannen osapuolen tarkastajien toimesta ja kaikki tässä ilmaantuvat toiminnalliset ongelmat kirjataan ylös.
 - Lopuksi määritellään se, kuinka kauan asiakirja säilytetään järjestelmässä – aktiivinen vaihe kestää n. 10 vuotta (todellisuudessa kylläkin 6-7 vuotta). Järjestelmät kuitenkin muuttuvat niin nopeasti, että pitkä aktiivisen vaiheen säilytys on ongelmallista.
- Passiivinen vaihe (puoli-aktiivinen vaihe):
 - Asiakirjojen säilytys off -line –muodossa.
 - Asiakirjat ja niiden metatieto ovat ASCII-muodossa.
 - Jokaisesta asiakirjasta on kaksi kopiota, jotka kopioidaan uudelleen 5 vuoden välein.
 - Niiden asiakirjojen hävittäminen, joiden säilyttämisen aika loppuu.

- Historiallinen vaihe:
 - Pysyvästi säilytettäväksi määrätty asiakirjat siirretään ko. organisaation arkistointipalveluiden huostaan tai varsinaiseen arkistoon.
 - Asiakirjat säilytetään datana niin kauan kuin se on teknisesti ja taloudellisesti mahdollista. Pienessä määrin rakenteisia materiaaleja voidaan hyödyntää käytämällä luokittelevaa/järjestävää hakukriteeriä, ja ne voidaan ottaa pysyvään säilytykseen sähköisessä muodossa. Paperiset tulosteet saattavat tulla kysymykseen silloin, kun kyseessä olevat materiaalit eivät ole valtavan laajoja.
 - Tulostustoiminta täytyy suunnitella ja säännön mukaistaa huolellisesti.

Asiakirjan ilmiselvä arvo perustuu säilytyspaikkaan ja asiakirjan käsittelyprosessiin, jota tulee tukea kunnollisen dokumentaation avulla. (Pohjola 1999) Vuonna 2001 arkistolaitoksessa saatiin valmiiksi ohje sähköisten tallenteiden ja materiaalien käsittelystä (Arkistolaitos 2001).

5.2 Sähköisten tallenteiden käsittelystä arkistolaitoksessa

5.2.1 Arkistolaitoksen haastattelu

Arkistolaitos, johon haastatteluni tein, on johtava organisaatio pitkäaikaisen arkistoinnin hoitamisessa Suomessa. Arkistolaitos muodostuu Helsingissä toimivasta kansallisarkistosta ja maakunnissa piirihallintoviranomaisina toimivista maakunta-arkistoista. Arkistolaitoksen pääpaikka on Helsingissä, jossa myös kansallisarkisto toimii. Haastatteluni kohdetta arkistolaitosta ja sen toimintaa on tarkemmin selostettu tutkimukseni luvussa 2.

Tutkimustani varten otin joulukuussa 2001 yhteyttä arkistolaitokseen ja kysyin sieltä haastateltavia henkilöitä. Saamani vastauksen perusteella lähetin sähköpostitse tietotekniikkayksikön johtajalle Markku Mäenpäälle laatimani kysymykset (Arkistolaitos, Markku Mäenpään sähköpostivastaus kyselyyni, 21.12.2001). Haastattelin myös puhelimitse tietotekniikkayksikön ylitarkastajaa Jukka Ojalaa ja esitin hänelle samat kysymykset. (Arkistolaitos, Jukka Ojalan puhelinhaastattelu, 21.12.2001). Sain kahdet hie-

man erilaiset, toisiaan täydentävät vastaukset esittämiini kysymyksiini. Näin minulla oli mahdollisuus saada monipuolisempi kuva tilanteesta arkistolaitoksessa.

Haastatteluni tavoitteena oli saada selville, kuinka paljon arkistolaitoksessa on oltu tekemisissä elektronisten tallenteiden kanssa, miten niitä on käsitelty, mitä toimenpiteitä niiden säilyttämiseksi on suunniteltu toteutettavan ja mitä ongelmia elektronisten tallenteiden pitkäaikaisessa arkistoinnissa on ilmennyt. Laatimani kysymykset olivat:

1. Kuinka paljon teillä on sähköisiä tallenteita suhteessa perinteisiin paperitallenteisiin ja mikrofilmeihin?
2. Missä muodossa ja formaateissa saatte arkistolaitokseen sähköiset tallenteet, ja joudutteko muuttamaan niiden formaatteja säilytystä varten?
3. Kuinka usein siirrätte tallenteet uudemmille välineille (esim. magneettinauhat uudemmille magneettinauhoille)?
4. Teettekö joitakin aitouden tarkastustoimenpiteitä sähköisille tallenteille?
5. Tehdäänkö teillä sähköisille tallenteille uudistamistoimenpiteitä, kuten toiseen formaattiin tai standardiin siirtämistä tai migraatiota eli tallenteen siirtämistä vanhasta ohjelmistosta uudempaan ohjelmistoon?
6. Oletteko olleet tekemisissä metatietoon liittyvien asioiden kanssa?
7. Minkälaiset tulevaisuuden näkymät teillä on sähköisten tallenteiden ja niiden käsittelyn suhteen?
8. Kuinka paljon teillä on tehty tutkimusta sähköisistä tallenteista ja niiden käsittelystä?

5.2.2 Haastatteluni vastaukset ja niiden analysointi

Sähköisten tallenteiden osuus arkistolaitoksen kaikista tallenteista

Sähköisten tallenteiden käsittelyn suhteen arkistolaitoksessa ollaan varsin alkuvaiheessa kuten suurimmaksi osaksi koko Suomessa. Kokonaishyllymetrimäärä arkistolaitoksessa (Kansallisarkisto ja 7 maakunta-arkistoa) on noin 100 hkm (hyllykilometriä). Näistä sähköisiä tallenteita on vain vähäinen osuus verrattuna paperitallenteiden määrään. Sähköinen aineisto koostuu sekä analogisesta että digitaalisesta aineistosta. Analogisen ja digitaalisen aineiston osuutta sähköisestä aineistosta ei voida verrata, koska mediat ja

tallenteet ovat niin erilaisia. Sähköisen aineiston määrästä ei voida siis antaa mitään lukua tavuina.

Sähköisestä aineistosta digitaalista on noin 15 GB (rekisteriaineistoa), joka ei ole muussa muodossa. Osa digitaalisista tallenteista on eri valtion virastojen tallenteiden varmuuskopioita. Arkistolaitoksessa on myös digitoitu omaa originaalista aineistoa kuvamuotoon. Formaateina ovat olleet Tiff ja Jpeg. Yhteensä tällaista kuvamuodossa olevaa aineistoa noin 120 GB.

Analogisia tallenteita on Kansallisarkistossa noin 300 avokelanauhaa, 300 C-kasettia, noin 40 videonauhaa ja noin 5 vanhaa äänilevyä. Lisäksi on 70 digitaalista CD-R-levyä, jotka on konvertoitu analogisesta materiaalista. Jyväskylän maakunta-arkistossa on noin 10000 avokelanauhaa, noin 500 C-kasettia ja noin 35 videokasettia, Vaasan maakunta-arkistossa on noin 50 avokelanauhaa, muissa maakunta-arkistoissa vain muutamia kappaleita. Kansallisarkiston videonauhat on tallennettu Elokuva-arkistoon asianmukaiseen säilytykseen. Maakunta-arkistojen videonauhat ovat arkistomakasiineissa. Äänitteiden kunto on vielä kartoittamatta, joten niiden sisältö ja säilytysarvo on selvittämättä. Tähän on ymmärrettäviäkin syitä: Dokumentaatiotiedot puuttuvat tai ovat vajavaiset. Sisältöä ei ole kuunneltu, koska tarvittavat laitteet puuttuvat. Arvonmääritys on tekemättä. Sen vuoksi ei ole mahdollista laittaa tallennenauvoja tai tallennelevyjä ja käyttökopioita tärkeysjärjestykseen eikä ole mahdollista seuloa "arvottomia" nauhoja. Arkistolaitoksessa on näin ollen varsin kirjavaa materiaalia, vaikka 40 vuotta vanhempaa aineistoa ei olisi pakko ottaa arkistolaitokseen säilytykseen.

Hyvin mielenkiintoinen kysymys on, kuinka paljon digitaalisia tallenteita on paperitallenteiden määrään nähden. Kuten tuli jo mainittua, paperitallenteita on koko arkistolaitoksessa (kansallisarkisto ja 7 maakunta-arkistoa) 100 hyllykilometriä. Digitaalista aineistoa on 15 GB. Kun arvioidaan, kuinka paljon yksi A4 -arkki on tavuina, saadaan muutettua digitaalisten tallenteiden määrä hyllykilometreiksi. 15 GB digitaalista aineistoa vastaa n. 4 miljoonaa A4-arkkia tekstinä. Digitaalisten tallenteiden osuus olisi arviolta noin 11/2 -2 hyllykilometriä, joka on noin 1,5 – 2,5 % arkistolaitoksessa olevien paperitallenteiden määrästä.

Koska sähköisten aineistojen osuus paperitallenteista on näin pieni, ei ole siinä mielessä yllättävää, että sähköisten aineistojen käsittelyn ja arkistoinnin suhteen ollaan varsinkin alussa. Edellä kappaleessa 5.1 selvitettiin, että 1990-luvun alkutaipaleelta lähtien on arkistolaitoksen piirissä tehty jonkin verran tutkimusta sähköisistä tallenteista. Nämä tutkimukset ovatkin olleet pohjana sille, että vuonna 2001 arkistolaitos sai valmiiksi ohjeen sähköisten aineistojen ja tallenteiden käsittelystä. Ohje tuli juuri hyvään aikaan, koska sähköisten tallenteiden osuus tulee koko ajan kasvamaan. Suuri osa tällä hetkellä valtionhallinnossa olevista tallenteista on sähköisessä muodossa ja monet näistä tallenteista arkistoidaan vain sähköisessä muodossa. Tutkimusten ja tämän ohjeen avulla voidaan yrittää helpottaa sähköisten tallenteiden tulevaisuudessa aiheuttamia ongelmia, kuten miten voidaan turvata niiden säilyminen tuleville sukupolville.

Sähköisiä asiakirjoja koskevat ohjeet

Valtionhallinnollisissa virastoissa ja laitoksissa on määrätty tietty tallennusmuoto, jossa asiakirjat ym. tallennetaan. Tämä toimii aina siihen asti, kun asiakirjaa käsitellään vain kyseisen viraston sisällä. Mutta sitten kun se lähetetään oman viraston tai laitoksen ulkopuolelle, saattaa tulla ongelmia. Vastaanottaja ei välttämättä saa asiakirjaa auki. Nykyään kylläkin suurimmassa osassa valtionhallinnollisia organisaatiota on käytössä Microsoft Office-paketti ja sen Word 97- tai Word 2000-versio, jotka ovat yhteensopivia keskenään. Jos asiakirja on näitä vanhemmassa Word-tiedostossa tai jossain muussa kuin Word-tiedostossa, voi yhteensopivuusongelmia esiintyä. Näiden ongelmien vähentämiseksi arkistolaitos on antanut ohjeensa käytettävästä formaatista jne.

Arkistolaitoksen ohjeessa kehoitetaan valtion viranomaisia luovuttamaan aineistonsa pääasiassa ISO 8859_1 tai ISO 8859-15 formaatissa. Tällöin arkistolaitoksen ei tarvitse muuttaa aineiston formaatteja, kun tulee aika siirtää ne arkistolaitoksen huostaan. Valtionhallinnollinen aineisto on ohjeen mukaan myös pääasiallisesti ASCII-muodossa, jolloin sillä on mahdollisimman suuri laitteistoriippumattomuus.

Arkistolaitoksen kohtaamia ongelmia

Osa tallenteista on kuitenkin tietokantoja, jotka ovat peräkkäistiedostoina. Peräkkäistiedostot ovat tiedostoja, jotka sisältävät merkkejä tai niiden lisäksi binäärisesti koodattuja numeroita. Peräkkäistiedostoissa yksittäisen merkkijonon tai numeron merkitys riippuu sen sijainnista. Peräkkäistiedostot ovat joko kiinteämittaisia, jolloin jokainen tietokenttä

on määritellyn mittainen, tai erotettuja, jolloin jokainen tietokenttä on erotettu toisesta erotinmerkillä. Ongelmallisia arkistolaitokselle ovat relaatiotietokannat, koska niistä ei ole standardeja. Tutkimuksessani käyttämässäni lähdeaineistossa todetaan myös, että tietokannat, erityisesti relaatiotietokannat ovat ja tulevat olemaan aikamoinen ongelma pitkäaikaisessa arkistoinnissa.

Arkistolaitos on kohdannut ongelmia myös analogisten tallenteiden sekä yksityisten arkistonmuodostajien tallettamien tai lahjoittamien aineistojen kanssa. Analogisissa tallenteissa ja erityisesti yksityisluonteisten arkistojen aineistossa dokumentaatio on monesti olematonta. Tällöin arkistolaitoksessa ei voida tietää minkälaista aineistoa esim. korpuissa tai lerpussa on, missä formaatissa ne ovat ja millä ohjelmalla ne on tehty. Tämän kysymyksen ratkaisemisen suhteen on vielä tekemistä.

Sähköisten tallenteiden käsittely arkistolaitoksessa

Arkistolaitoksessa varsinkin vanhempien tallenteiden ja aineiston kohdalla joudutaan tekemään jonkun verran konvertointia yms. Tallenteiden konvertointia tehdään UNIX:lla toimivalla AIX-kääntöohjelmalla. Toinen keino on siirtää windows-pohjaiset dokumentit toiseen ympäristöön, eli tehdä niille migraatio, jolloin konversiot kopioidaan CD-ROM -levylle. On mielenkiintoista havaita, että arkistolaitoksessa tehdään jonkun verran käytännössä migraatiotakin, sillä joissakin käyttämässäni lähteissä koettiin ongelmallisena, kun migraatiota ei ole kokeiltu ja testattu käytännössä.

Kuten joissakin lähteissä arvioitiin, hyvä aikaväli siirtää aineisto uudemmalle medialle on noin viiden vuoden välein. Arkistolaitoksessa valtion viranomaisten ja yleensäkin muukin digitoitu aineisto tuoreutetaan (virkistetään) joka viides vuosi.

Tutkimuskirjallisuudessa kuten myös arkistolaissa ja arkistolaitoksen käytännöissä tallenteiden aitous ja luotettavuus ovat hyvin tärkeitä tekijöitä. Jos tallenteen aitouteen ei voida luottaa, ei sitä voida käyttää tutkimuksellisessa eikä monesti muussakaan käytössä. Aitous, autenttisuus eli alkuperäisyys ja eheys eli muuttumattomuus voidaan käytännössä tarkistaa. Arkistolaitos tekee tarkistustoimenpiteitä valtion viranomaisten tallenteille vertaamalla siirrettyä aineistoa dokumentaatioon. Sen sijaan yksityisluontoisten arkistojen materiaaleille ei tehdä tarkistustoimenpiteitä, monesti sen vuoksi, että se ei ole mahdollista puuttuvan dokumentaation vuoksi.

Kuten edellä on todettu, arkistolaitos tekee jonkin verran konvertointia, muunnostoimenpiteitä aineistolle. Ohjeessaan arkistolaitos on antanut suosituksen siitä, missä formaatissa valtion viranomaisilta tuleva aineisto saa olla. Tällöin ei tarvitse tehdä muunnostoimenpiteitä. Suurin osa valtion viranomaisten aineistosta säilytetään sellaisenaan, koska arkistolaitoksen periaatteena on säilyttää data, ei ohjelmistoa eikä laitteistoa. Tulevaisuutta ajatellen tämä ei ole kuitenkaan riittävä toimi, koska tallenteita ei välttämättä saada enää auki uudemmilla laitteilla ja ohjelmilla joskus myöhemmin. Arkistolaitokseen siirrettäessä tallenteille ei heti tehdä mitään uudistamis-, muunnos- tai migraatiotoimenpiteitä. Ne tehdään vasta myöhemmin. Uudistustoimenpide tehdään viiden vuoden välein, kuten edellä mainittiin.

Arkistolaitoksessa tehdystä muunnoksesta ja migraatiosta on seuraavat esimerkit: Koikeiluprojektina yhden yksityisen arkiston vanhoja tekstinkäsittelytiedostoja siirrettiin uuteen formaattiin. Prosessi oli tavattoman työläs. Uuteen formaattiin siirtäminen ei ole helppoa ja siinä tulee eteen monia vaikeuksia, kuten oli vanhan TEKO-tekstinkäsittelyjärjestelmän sisältämien tiedostojen konvertoinnissa uudempiin käyttöjärjestelmiin. Arkistolaitoksen lisäksi konversiota on yritetty myös muuallakin. Esimerkiksi Helsingin yliopiston kirjastossa TEKO-tekstinkäsittelyjärjestelmässä olevien tekstien konversio ei onnistunut muutoin kuin skannaamalla tekstit, mikä on erittäin työläs prosessi. Arkistolaitoksessa on edellisten esimerkkien lisäksi konvertoitu analogisia tallenteita CD-R-levylle sopivaan muotoon.

Metatieto ja sen kehittäminen pitkäaikaisen arkistoinnin tarpeisiin esiintyy tutkimuskirjallisuudessa tärkeänä asiana tulevaisuuden kannalta. Maailmalla on Dublin Core-metatieto ollut käytetyin metatieto joko sellaisenaan tai pohjana eri projekteissa kehitellyille metadatoille. Arkistolaitoksessa on myös oltu Dublin Core-metatiedon kanssa tekemisissä. Sitä on käytetty esimerkiksi valtionhallinnossa JUHTA:n projektissa, jota on käsitelty jo aiemmin metatiedon yhteydessä, kappaleessa, 4.4.4. Arkistolaitos on todennut selväksi puutteeksi sen, että Windows-ohjelmistojen tiedostoissa ei ole valmiina metatietoa. Metatieto ja sen lisääminen sellaisiin aineistoihin, joissa sitä ei ole, on koettu tarpeelliseksi. Mutta arkistolaitoksella ei ole kuitenkaan tarpeeksi resursseja metatiedon lisäämiseen. Näin metatietoa ei valitettavasti pystytä lisäämään arkistolaitokseen vastaanotettuihin tallenteisiin, joissa sitä ei ole valmiina. Metatieto tulisikin laittaa

tallenteisiin mukaan jo niiden luontivaiheessa eikä vasta arkistosäilytykseen siirrettäessä, kuten todettiin myös eri tutkijoiden artikkeleissa.

Tutkimus ja tulevaisuus

Arkistolaitos on ollut mukana sähköisiin tallenteisiin liittyvissä tutkimushankkeissa. Siellä on siten tutkittu muutoinkin kuin vain sähköisten tallenteiden käsittelystä annetun ohjeen aikaansaamisen vuoksi. Arkistolaitos on ollut mukana CD-R-levyjen kestävydestä tehdyssä tutkimuksessa vuonna 1995. VTT testasi hankkeessa eräitä levyjä. Tällä hetkellä arkistolaitoksella on meneillään sähköisten tietoaineistojen käsittelyn ja pitkäaikaissäilytyksen kehittämishanke, joka kestää vuoden 2003 loppuun saakka. Yleensä tehdystä tutkimuksesta ja meneillään olevista hankkeista voidaan sanoa sen verran, että tutkimusta on tehty arkistolaitoksessa varsin yleisellä tasolla, koska ei ole olemassa resursseja tehdä tutkimusta teknisemmällä tasolla.

Tulevaisuudessa sähköisten aineistojen osuus tulee koko ajan kasvamaan myös arkistolaitoksessa. Sähköinen asiankäsittely, nimenomaan digitaalinen asiankäsittely, lisääntyy koko ajan. Arkistolaitoksen näkökulmasta on haastavaa erityisesti tiedon pitkäaikaissäilytykseen liittyvät asiat. Miten voidaan taata asiakirjojen alkuperäisyyden, eheyden ja käytettävyyden säilyminen koko elinkaaren ajan huolimatta siitä, että tekninen infrastruktuuri vaihtuu nopeasti? Haasteita riittää. Arkistolaitoksen, kuten koko yhteiskunnan näkökulmasta voidaan sanoa, että tilanne karkaa hallinnasta, ellei kehittämistoimenpiteisiin panosteta riittävästi. Tulevaisuudessa tulee todennäköisesti enemmän ohjelmistolaitteista riippumattomia tallenteita. XML ja sen käyttö tulevaisuudessa on edelleen arvoitus. XML:ää on suunniteltu käytettäväksi Turun kaupungissa ja valtionvarainministeriössä, joissa kummassakin on meneillään kehittämissuunnitelma, jonka tavoitteena on ottaa käyttöön Linux-käyttöjärjestelmä ja XML-pohjainen Open Office-ohjelmisto.

6 Yhteenveto

Tutkielmassani olen käsitellyt pitkäaikaisen arkistoinnin käytäntöjä ensin paperi- ja mikrofilmien osalta. Arkistolaissa määritellyt arkistointiin liittyvät tehtävät sekä arkistolaitoksen pitkäaikaisen arkistoinnin käytännöt ovat pohjana myös sähköisten tallenteiden arkistoinnille, jolle vasta ollaan suunnittelemassa järkeviä ja käyttökelpoisia keinoja ja menetelmiä. On selvää, että jo pitkään harjoitetussa perinteisten materiaalien arkistoinnissa käytetyt menetelmät ovat lähtökohtana sähköisten tallenteiden arkistoinnille. Jostain on lähdettävä liikkeelle ja tutuista asioista on helpompaa aina aloittaa. Arkistolaissa on huomioitu myös sähköisten asiakirjojen ja aineistojen olemassaolo, joten laki itsessään on myös pohjana sähköisten tallenteiden pitkäaikaiselle arkistoinnille. Valtionhallinnossa tapahtuvalle sähköisten tallenteiden käsittelylle ja arkistoinnille on arkistolaitos antanut oman ohjeensa. Ne tallenteet ja asiakirjat, jotka arkistolaitos ottaa säilytettäväkseen, tulee tallentaa tietyssä formaatissa ja tiettyjen standardien mukaan. Näin ongelmana ei ole sopivaan formaattiin ja standardiin muuttaminen, vaan kuinka tallenteet ja asiakirjat säilytetään tuleville sukupolville. Arkistolaitoksen olisi mahdotonta omilla resursseillaan alkaa muuttaa kaikkien sinne säilytykseen otettavien tallenteiden ja asiakirjojen formaatteja ja standardeja. Se ei olisi mahdollistakaan, koska arkistolaitoksessa ei voi olla kaikille mahdollisille formaateille muunnostyökaluja.

Monet tutkijat, joiden artikkeleita olen käyttänyt, ovat arkistoinnin kanssa työskennelleitä ja työskenteleviä ammatti-ihmisiä. Heistä puhdas tietotekniikan asiantuntija on vain Jeff Rothenberg, joka on kannattanut emulaatiota ratkaisukeinona pitkäaikaisen arkistoinnin ongelmiin. Muut tutkimuksessani esiin tulleet tutkijat eivät ole juurikaan kannattaneet emulaatiota, vaan suosivat migraatiota, jos ovat ottaneet ylipäättään kantaa näihin asioihin ollenkaan.

Sähköisten tallenteiden pitkäaikaiseen arkistointiin on suositeltu erilaisia toimenpiteitä sen mukaisesti, mitä katsantokantaa kyseinen tutkija edustaa. Kanadalainen Charles Dollar jakaa tutkijat kahteen koulukuntaan sen mukaan, minkälaisena he näkevät sähköisten tallenteiden pitkäaikaisen arkistoinnin ongelmat. Puhtaasti tietoteknisenä kysymyksenä ongelman näkevät tekniikan näkökulmasta asiaa tarkastelevat, kuten Jeff Rothenberg. Enemmistönä ovat tutkijat, jotka näkevät ongelman sekä tietoteknisenä että

humanistisena/ideologisena kysymyksenä. He kannattavat myös migraatiota enemmän kuin emulaatiota. Enemmistö on sitä mieltä, että sähköisten tallenteiden pitkäaikaisen arkistoinnin ongelmat ovat sekä tietoteknisiä että humanistisia/ideologisia kysymyksiä. Arkistoinnin näkökulmasta katsottuna on tärkeää tallenteiden ja asiakirjojen säilyttäminen tuleville sukupolville ja tutkijoille. Arkistolaitoksessa säilytetään kaikki valtionhallinnollinen asiakirjallinen tieto, oli se muodoltaan millaista tahansa. Arkistolaitoksen sähköisten tallenteiden ja asiakirjojen käsittelyohjeen tarkoituksena on taata asiakirjallisen aineiston ja kulttuuriperinnön eheys, autenttisuus, laatu, käytettävyys, saatavuus, säilyminen ja luottamuksellisuus. Alkuperäisestä aineistosta pysyvästi säilytettäviä asiakirjoja on yleensä vain enintään 5-15%. Tallenteet ja asiakirjat säilytetään todisteena jostakin toimenpiteestä, päätöksestä, päätöksen toimeenpanosta jne. Kun niitä pidetään todisteena, niiden täytyy olla aitoja ja luotettavia.

Kun pitkäaikaista arkistointia pidetään vain teknisenä ongelmana, jää aitous ja luotettavuus yms. seikat huomioimatta. Sen sijaan sellaiset seikat kuin luotettavuus, ymmärrettävyys ja selvyys ovat ongelmia, joihin myös teknisestä näkökulmasta ongelmaa tarkastelevat kiinnittävät huomiota, koska ne liittyvät bittivirran tulkintaan, selkeyteen ja ymmärrettävyyteen. Teknisestä näkökulmasta katsottuna ei saada tarpeellista selvitystä koko ongelmaan. Jos pidetään huolta vain siitä, että tallenteet ja asiakirjat ovat teknisesti selviä, luettavia jne., menetetään tutkimuksen kannalta tärkeä tallenteiden aitous, autenttisuus, eheys jne.. Pitkäaikaisen arkistoinnin toteuttaminen sähköisten tallenteiden osalta olisikin yksinkertaisempi ja helpompi asia, jos ei tarvitsisi ottaa huomioon mitään muuta kuin teknisiä näkökohtia. Migraatiota ja muunnosta eli konversiota toteutettaessa esimerkiksi siirrettäisiin vain tallenteet vanhemmasta uudempaan ympäristöön eikä tarvitsisi olla huolissaan siitä, että tallenteet muuttuisivat jonkin verran ja niistä tulisi arkistotieteen näkökulmasta uusia tallenteita, jotka eivät olisi enää aitoja, vaan muutettuja tallenteita.

Aitona tallenne säilyy silloin, kun sen sisällössä, rakenteessa ja kontekstissa ei tapahdu muutoksia. Tallenteiden aitouden takaamiseksi, turvaamiseksi ja toteamiseksi voidaan tehdä erilaisia toimenpiteitä, kuten turvallinen asiakas-palvelin-arkkitehtuuri, jaksottaisen tarkistussumman (CRC) käyttäminen, yksisuuntaisen hajautuksen käyttäminen (one-way hashing) ja hajautetun tiivistelmän yhdistäminen digitaaliseen aikaleimaan. Nämä ovat sellaisia toimenpiteitä, joita vain tietotekniikan ammattilaiset pystyvät teke-

mään. Arkistotalalla työskentelevät ihmiset eivät yksin pärjää pitkäaikaisen arkistoinnin ongelmien kanssa.

Tallenteen muunnos eli konversio on esimerkiksi tallenteen siirtämistä yhdestä ohjelmistoympäristöstä toiseen tai tallenteen siirtoa ohjelmistosta riippuvaisesta ympäristöstä ohjelmistosta riippumattomaan ympäristöön, jotka ovat samantapaisia toimenpiteitä kuin migraatiossa. Muunnoksessa ei todennäköisesti tapahdu muutoksia tallenteen sisältöön, rakenteeseen eikä kontekstiin, mutta sen sijaan bittivirtaan saattaa tulla hieman muutoksia. Migraatiossa arvioidaan tapahtuvan enemmän muutoksia tallenteeseen kuin muunnoksessa, koska migraatiossa tehdään usein muutakin kuin vain muunnetaan jotain tallennetta. Migraation on arvioitu onnistuvan parhaiten (aitous säilyy) silloin, kun dokumentoidaan koko migraatioprosessin kaikki vaiheet ja kaikki toimenpiteet. Arkistolaitoksen mukaan pelkästään migraation dokumentoiminen ei ole välttämätöntä, vaan kattavaa dokumentaatiota tarvitaan kaikille sähköisille tallenteille, riippumatta siitä, mitä toimenpiteitä niille on tehty. Tämä koskee erityisesti niitä tallenteita, jotka on määrätty säilytettäväksi kymmeniä tai satoja vuosia. Tietojärjestelmissä olevista tallenteista ja asiakirjoista tulisi luoda asiakirjallista dokumentaatiota, joka on kontekstuaalista (erottaa asiakirjallisen tiedon pelkästä informaatiosta), loogista (asiakirjallisten tietojen rakenne ja niiden suhteet) ja teknistä (muodostuneen sähköisten aineiston fyysiset ominaisuudet). Arkistolaitoksessa on kokeiltu käytännössä migraation onnistumista mm. vanhojen TEKO -ympäristössä luotujen tallenteiden muuttamiseen nykyaikaiselle tietokoneympäristölle ymmärrettävään muotoon.

Arkistonmuodostussuunnitelma, joka jokaisen viraston ja laitoksen tulee tehdä, auttaa selvittämään tallenteiden käsittelyprosesseja niiden luomisesta lähtien. Arkistonmuodostussuunnitelmien avulla tallenteita voidaan hallita ja ylläpitää. Tallenteiden ylläpitoon liittyy myös Australiassa kannatusta saanut tallenteiden jatkumo-käsite, joka on elinkaari-käsitteen vastine. Jatkumo-käsitteeseen liittyy myös tallenteita ylläpitävä hallinnollinen metatieto. Metatietoa on kehitelty Australiassa eri projekteissa, kuten SPIRT ja VERS. Suomessa metatietoa on kehitelty JUHTA:n projektissa. Kaikissa näissä projekteissa on ollut pohjana Dublin Core-metatieto. Metatietoa ja sen kehittämistä pitkäaikaisen arkistoinnin tarpeisiin pidetään tärkeänä. Sen avulla saadaan säilytettyä sähköisten tallenteiden konteksti. Ilman kontekstia tallenne on varsin irrallinen, eikä se ole niin todistusvoimainen – sitä ei voida pitää niin hyvänä todisteena jostakin tapahtumas-

ta, toimenpiteestä tms. kuin kontekstin sisältävää tallennetta voidaan pitää. Tämänhetkiset ohjelmistot, kuten Windows-ohjelmistot eivät tiedostoineen sisällä metatietoa, vaan metatieto tulisi lisätä niihin erikseen. Tämän vuoksi on suositeltu, että tietojärjestelmiä ja ohjelmistoja suunniteltaessa otettaisiin huomioon pitkäaikaisen arkistoinnin vaatimukset, kuten metatiedon lisääminen.

Rakenteisen dokumentaation käyttö pitkäaikaisessa arkistoinnissa esiintyy muutamien tutkijoiden artikkeleissa, kuten esimerkiksi Charles Dollarin artikkelissa, jossa mainittiin SGML-formaatti yhtenä vaihtoehtona, mutta ei vielä XML-formaattia, joka oli artikkelien kirjoittamisen aikana vielä aika tuntematon. XML on syrjäyttämässä SGML-formaatin. XML on hyvin käyttökelpoinen formaatti laitteistoriippumattoman dokumentaation luomisessa. JUHTA -projektissa Dublin Coreen perustuvaa metatietoformaattia on kehitetty siten, että se määrittelee asiakirjojen kuvailuelementtejä, kuten tekijä, nimi, aihe sekä syntaksit kuvailutietojen tallentamiseen HTML- ja XML-dokumentteihin. Kun asiakirjan tiedostoformaatti on joku muu kuin HTML- ja XML, metatieto voidaan tallettaa erilliseen HTML- ja XML-tiedostoon, johon tallennetaan URL -linkki kuvailtuun asiakirjaan. Metatieto ja rakenteinen dokumentaatio nähdään yhdessä ja erikseen hyvin käyttökelpoisina vaihtoehtoina sähköisten tallenteiden pitkäaikaisessa arkistoinnissa. Rakenteiseen dokumentaatioon on helpompaa saada metatietoa mukaan kuin perinteiseen Word-dokumenttiin. Rakenteinen dokumentti, erityisesti XML, on laitteisto- ja ohjelmistoriippumattomampi kuin muilla tavoilla aikaansaadut dokumentit. Arkistolaitoksessa on todettu käytännössä, että tallennetta konvertoidessa HTML-formaatista ASCII-formaattiin, metatieto yleensä katoaa. Mutta silloin, kun metatieto talletetaan HTML- tai XML-formaatissa olevaan tiedostoon, se voisi kenties säilyä konversion ja/tai migraation jälkeenkin, kun metatietotiedosto pidettäisiin tallenteesta erillisenä.

Tällä hetkellä arkistolaitoksessa on hyvin vähän sähköisiä tallenteita ja niistä osa on eri virastojen asiakirjojen ja tietokantojen varmuuskopioita, mutta tulevaisuudessa sähköisten tallenteiden määrät tulevat kasvamaan huomattavasti. Onko silloin resursseja hoitaa esiin tulevat ongelmat? Arkistolaitoksessa ja tutkijoiden keskuudessa pelätään sähköisten tallenteiden synnyttämien ongelmien tulevan aiheuttamaan melkoisia vaikeuksia koko yhteiskunnassa lähitulevaisuudessa. Kukaan ei pysty ennustamaan tulevaa eli tulee esimerkiksi XML:stä, tallenteita ylläpitävästä metatiedosta ja migraatiosta säh-

köisten tallenteiden pitkäaikaisen arkistoinnin pelastaja, mutta tutkimalla eri vaihtoehtoja voidaan huomattavasti vähentää tulevien vaikeuksien uhkaa. Useat tutkijat ovat sitä mieltä, että nykyisillä resursseilla arkistoissa ei pystytä hoitamaan sähköisten tallenteiden pitkäaikaisen arkistoinnin ongelmia. Kyselyssäni arkistolaitokseen tuli selväksi, ettei tarvittavaan tekniseen tutkimukseen ole resursseja. Tutkijat, kuten Edward Higgs ja Charles Dollar, ovat ehdottaneet, että sähköisten tallenteiden arkistoinnista huolehtisivat ne organisaatiot, jotka niitä luovat ja tallenteet otettaisiin arkistoon kenties vasta lopulliseen säilytykseen. Arkistonhoitajan tulisi olla mukana tietojärjestelmien ja ohjelmistojen suunnittelussa. Hänestä tulisi kenties yksi ohjelmistoprojektissa työskentelevä ja/tai informaation ja dokumenttien hallintaa hoitava henkilö. Näin arkistointikin kokisi aikamoisia mullistuksia. Perinteinen arkistonhoitaja voisi olla muinaismuisto jonkin ajan kuluttua. Tarvitaan informaatioteknologian parissa työskentelevien ja arkistoinnin ammattilaisten yhteistyötä, jotta tulevaisuudessa eteen tulevista ongelmista selvitään. Pitkään eri teillä kulkeneet teknologia ja humanismi yhdistäisivät vihdoin voimansa, mikä olisi ennennäkemätöntä.

Lähteet

Arkistolaitoksen www-sivut:

- *Arkistolaitoksen www-sivu:* <http://www.narc.fi>
- *Asiakirjojen seulonta ja hävittäminen:* <http://www.narc.fi/seulo/seufr.htm>
- *Arkistolaitoksen suositus arkistonmuodostussuunnitelman laadinnan, käytön ja ylläpidon periaatteiksi:*
<http://www.narc.fi/ams/>
- *Asiakirjojen säilyttäminen poikkeuksellisissa oloissa:* <http://www.narc.fi/suoj898.htm>
- *Arkistolaitoksen päätös arkistokelpoisista materiaaleista ja menetelmistä:*
<http://www.narc.fi/paatos.html>
- *Papereiden ja kirjoitustarvikkeiden käyttäminen pysyvästi säilytettävien asiakirjojen valmistuksessa:*
<http://www.narc.fi/pap298.htm>
- *Arkistolaitoksen ohje: Sähköisten tietojärjestelmien ja -aineistojen käsittely. 2001:*
<http://www.narc.fi/sahk>
- *DLM-EXPERTS' MEETING IN TAMPERE, FINLAND, November 11-12, 1999. PRINCIPLES OF APPRAISAL AND THEIR APPLICATION IN ELECTRONIC ENVIRONMENT. European models and concepts. PROCEEDINGS. Arkistolaitos 2000:* <http://www.narc.fi/dlm/index.html>
- *Documenta-asiakirjahallinnon opas:* <http://www.narc.fi/docu/esittely.htm>

Muita lähteitä:

Agricola – Suomen historiaverkko. <http://www.utu.fi/agricola/> [Luettu 13.4.2002].

Agricola – Suomen historiaverkko. Yksityiset keskusarkistot.
<http://www.utu.fi/agricola/inst/arksyksk.html> [Luettu 13.4.2002].

Arkistolaitos, ATK-yksikön Jukka Ojalan puhelinhaastattelu, 21.12.2001.

Arkistolaitos, Markku Mäenpään sähköpostivastaus kyselyyni, 21.12.2001.

Arkistolaki 23.9.1994/831.

Arkistoviesti 2/1998, Sähköisen asiakirjan problematiikkaa ja megatrendejä.

Bearman, D., Reality and Chimeras in the Preservation of Electronic Records. *D-Lib Magazine*, 5 (4), April 1999. <http://www.dlib.org/dlib/april99/bearman/04bearman.html> [Luettu 14.3.2001]

Bearman, D. and Trant, J., Authenticity of digital resources: towards a statement of requirements in the research process. *D-Lib Magazine*, June 1998. <http://www.dlib.org/dlib/june98/06bearman.html> [Luettu 19.12.2001]

CCSDS Recommendations (Red Books and Pink Books). July 2001. http://www.ccsds.org/red_books.html [Luettu 11.12.2001]

Cornell University Library, Digital Imaging Tutorial, 2000. <http://www.library.cornell.edu/preservation/tutorial/contents.html> [Luettu 6.12.2001]

Day, Michael, Metatieto for digital preservation: an update, Ariadne Issue 22, 1999. <http://www.ariadne.ac.uk/issue22/metatieto/intro.html> [Luettu 26.4.2001]

Ditzler, C. et al., The Electronic Information Initiative. Phase 1 Final Report, Chapter 7 - Preservation Issues for Optical Media. U.S. Department of Agriculture, National Agricultural Library, Beltsville, Maryland 1994. http://www.nalusda.gov/services_and_products/other_nal_products/eii/nal-eii7.html [Luettu 11.12.2001]

DLM (Données Lisibles par Machine, Machine Readable Data)–Forum in [www-sivu](http://www.sivu.org): <http://www.dlmforum.eu.org> (Euroopan komission ISPO-palvelimella). [Luettu 11.11.2000]

Dollar, C.M., Archival theory and information technologies: the impact of information technologies on archival principles and methods. Macerata: University of Macerata Press, 1992.

Dollar, C.M., Authentic electronic records: strategies for long-term access. University of British Columbia, June 1998.

Evans, R. J. Toner Stability: A Continuing Detective Story.

<http://www.staff.uiuc.edu/~r-evans4/toner.html> [Luettu 26.11.2000]

Gavitt, Sharon. (2000) Computer Output Microfilm (COM). New York State Archives Technical Information Series #52. Revised February 2000.

<http://www.archives.nysed.gov/pubs/recmgmt/mfilm/com.htm> [Luettu 12.11.2000]

Hedstrom, M., Digital preservation: a time bomb for Digital Libraries. School of Information and Library Studies at the University of Michigan, 1996.

<http://www.uky.edu/%7Ekiernan/DL/hedstrom.html> [Luettu 29.10.2001]

Higgs, Edward, Historians, Archivists, and Electronic Record-Keeping. Teoksessa Higgs, Edward, toim., History and Electronic Artefacts. Oxford University Press Inc., New York, 1998.

Higgs, Edward, The Role of Tomorrow`s Electronic Archives. Teoksessa Higgs, Edward, toim., History and Electronic Artefacts. Oxford University Press Inc., New York, 1998.

ISO Archiving Standards - Overview. 2001.

<http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html> [Luettu 11.12.2001]

JHS 143 1999. Julkisen hallinnon tietohallinnon neuvottelukunnan suositus 143 - Asia-kirjojen kuvailuformaatti. JUHTA . <http://www.intermin.fi/juhta/suosituksset/jhs143.htm> [Luettu 20.12.2001]

Lagus, Antti J., Elektroninen arkistointi vaatii huolellisuutta. Kauppalehti Extra, Informaatiotekniikka. 8. huhtikuuta 2002.

Leinonen, Anne. XML-standardiperhe ja elektronisten dokumenttien arkistointi, Jyväskylän yliopisto, Tietojenkäsittelytieteiden laitos, 1998.

Lesk, M., *Image formats for preservation and access: a report of the Technology Advisory Committee to the Commission on Preservation and Access*. Washington, D.C.: Commission on Preservation and Access, 1990. Available from:
<http://www.clir.org/pubs/reports/lesk/lesk.html> [Luettu 6.12. 2001]

Lesk, M., *Preservation of new technology: a report of the Technology Advisory Committee to the Commission on Preservation and Access*. Washington, D.C.: Commission on Preservation and Access, 1992. Available from:
<http://www.clir.org/pubs/reports/lesk/lesk2.html> [Luettu 6.12. 2001]

Lesk, M., *Preserving digital objects: recurrent needs and challenges*. Available from: <URL:<http://www.lesk.com/mlesk/auspres/aus.html> [Luettu 6.12. 2001]

Mackenzie Owen, J.S., Walle, J.v.d., *Deposit collections of electronic publications*. European Commission, DG XIII-E/4. Libraries in the Information Society. Luxembourg 1996.

McKemmish, S, Acland G., *Accessing essential evidence on the Web: Towards an Australian Recordkeeping Metatieto Standard*, 1999.
<http://ausweb.scu.edu.au/aw99/papers/mckemmish/paper.html> [Luettu 19.12.2001]

McKemmish, Sue, Cunningham, Adrian, Parer, Dagmar, *Metatieto Mania*. The 1998 Annual Conference of the Australian Society of Archivists, Fremantle, Western Australia, 1998. <http://rcrg.dstc.edu.au/publications/recordkeepingmetatieto/sm01.html> [Luettu 25.10.2001]

Moisio, Riku, *Asiankäsittelyjärjestelmä osana sähköisen asioinnin viitearkkitehtuuria*. Kansallisarkisto, 2002.

Morelli, Jeffrey D., *Defining Electronic Records: Problems of Terminology*. Teoksessa Higgs, Edward, toim., *History and Electronic Artefacts*. Oxford University Press Inc., New York, 1998.

National Archives of Australia, Electronic Records.

<http://www.naa.gov.au/recordkeeping/er/summary.html> [23.3.2001]

O`Shea, Greg, Keeping Electronic Records: Issues and Strategies. Provenance. The Electronic Magazine, Vol.1, No.2 , 1996.

<http://www.netpac.com/provenance/vol1/no2/features/erecs1a.htm> [Luettu 10.10.2001]

Orrman, Elias, Structural Complexity of Electronic Records as a Factor Guiding Decisions on Permanent Retention. DLM-Forum, Tampere 1999.

<http://www.narc.fi/dlm/8.htm> [Luettu 13.10.2000]

Pohjola, Raimo, Appraisal and Disposal of Electronic Records. DLM-Forum, Tampere 1999.

<http://www.narc.fi/dlm/6.htm> [Luettu 13.10.2000]

Reed, Barbara, Electronic Records Management in Australia. Records Continuum Research Group, Publications, 2000.

<http://www.sims.monash.edu.au/rcrg/publications/breema.html> [Luettu 23.3.2001]

Rothenberg, Jeff, Ensuring the Longevity of Digital Documents. In Scientific American, Jan 1995, Vol. 272 Issue 1. <http://ehostvgw19.epnet.com/> [14.3.2001]

Stenvall, Jari, Metatieto elektronisten julkaisujen pitkäaikaissäilytyksessä. Pro gradu-tutkielma. Informaatiotutkimuksen laitos. Tampereen yliopisto, 2001.

Tampereen yliopiston historian laitos, Arkistotietokannat ja -luettelot, kuvatietokannat, 1999. <http://www.uta.fi/laitokset/historia/sivut/ih99/ih4.html> [31.1.2002]

VERS, Victorian Electronic Records Strategy, Final Report, Chapter 6: Implementing an Electronic Archiving System.

<http://www.prov.vic.gov.au/vers/final/6implemene.htm> [19.12.2001]

Zweig, Ronald W., *Beyond Content: Electronic Fingerprints and the Use of Documents*.
Teoksessa Higgs, Edward, toim., *History and Electronic Artefacts*. Oxford University
Press Inc., New York, 1998.

Liitteet

Liite 1: Pitkäaikaisen arkistoinnin käsitteitä

Lähteet: Tampereen yliopiston historian laitos, Arkistotietokannat ja -luettelot, kuvatie-
tokannat ja Arkistolaitoksen ohje: Sähköisten tietojärjestelmien ja -aineistojen käsittely,
2001.

Arkisto

Yhteisön tehtävien hoitamisesta tai henkilön toiminnasta kertyneiden asiakirjojen koko-
naisuus.

Arkistointi

Asiakirjojen liittäminen arkistoon ennalta laaditun suunnitelman mukaisesti.

Arkistoluettelo

Asiakirjoista laadittu luettelo, jossa yhden arkistonmuodostajan asiakirjat luetellaan
arkistoyksikön (esim. kotelo tai sidos) tarkkuudella. Yleensä arkistoluettelo noudattaa
jotain arkistokaavaa eli asiakirjojen luokitusjärjestelmää, jossa määritellään asiakirjasar-
jat.

Arkistonmuodostaja

Yhteisö tai yksityinen henkilö, jonka toiminnan tuloksena arkisto syntyy, esim. Tampe-
reen kaupunginhallitus (ei Tampereen kaupunki), Alahärmän säästöpankki tai J.K. Paa-
sikivi.

Arkistonmuodostussuunnitelma

Arkistolain (831/94 8§) arkistonmuodostajalta edellyttämä luettelo sen tehtävien hoidon
tuloksena kertyvistä asiakirjoista ja niiden säilytysajoista ja -tavoista.

Arkistotoimi

Toimi, jonka tehtävänä on varmistaa asiakirjojen käytettävyys ja säilyminen, huolehtia
asiakirjoihin liittyvästä tietopalvelusta, määritellä asiakirjojen säilytysarvo ja hävittää
tarpeeton aineisto. Arkistotoimi on hoidettava siten, että se tukee arkistonmuodostajan
tehtävien suorittamista sekä yksityisten ja yhteisöjen oikeutta saada tietoja julkisista
asiakirjoista, että yksityisten ja yhteisöjen oikeusturva samoin kuin tietosuoja on otettu
asianmukaisesti huomioon ja että yksityisten ja yhteisöjen oikeusturvaan liittyvien asia-

kirjojen saatavuus on varmistettu sekä että asiakirjat palvelevat tutkimuksen tiedon lähteenä.

Asiakirjahallinto

Toiminto, joka kehittää ja ohjaa yhteisön toimintaa koskevien asiakirjojen vastaanottamista, tuottamista, käsittelyä, tallentamista, säilytysarvon määrittelyä ja käyttöön saattamista. Suomen hallinnossa asiakirjahallinto ja arkistotoimi ovat pitkälti yhteneväisiä toimintoja.

Asiakirjallinen tieto

Yhteisön tehtävien hoitamisen tai henkilön toiminnan tuloksena kertynyt tieto. Samaa tarkoittaen voidaan puhua toiminnallisesta tiedosta.

Asiankäsittelyjärjestelmä

Tietojärjestelmä, jonka avulla organisaation koko päätöksenteko prosessia voidaan hallita ennalta määriteltyjen käsittelysääntöjen mukaisesti. Järjestelmä sisältää tai siihen on integroitu eri sovelluksia ja toimistotyökaluja (yleensä ainakin diaari ja dokumenttien hallinta sekä tekstinkäsittelysovellus ja sähköpostiohjelma).

Cartidge Tape

Nauhakasetti. Cartridge = kasetti. Magneettinauhaa sisältävä kasetti, esim. 8-raitakasetti, ei kuitenkaan C-kasetti (cassette). Tape = nauha.

COM (Computer Output Microfilm)

Tietokoneella voidaan tuottaa mikrofilmejä suoraan, tarvitsematta tehdä erikseen tulostetta, joka sitten valokuvataan. Laitteistolla dokumentit tulostetaan valottamalla filmi suoraan. Tekniikkaa kutsutaan COM:ksi (Computer Output Microfilm). Mikrofilmauksen etuna on laitteistoriippumattomuus lukemisessa ja tulosteen suhteellisen pieni koko. 30*30*30cm arkistolaatikkoon mahtuu 6000 filmikorttia, joihin taas mahtuu n. 230 sivua kullekin. Näin ollen laatikkoon mahtuu n. 1,4 miljoonaa sivua tietoa. Vastaavaan laatikkoon mahtuisi vain noin 3000 sivua paperia. Mikrofilmaus on myös paperitulosten arkistointia huomattavasti edullisempaa.

CRC (Cyclical Redundancy Checksum)

Tiedonsiirron virheettömyyden tarkastusmenetelmä tai virheenpaljastusmenetelmä, joka perustuu lähetettävän sanoman muokkaukseen sellaiseen muotoon, että se on jaollinen tietyllä polynomilla. Vastaanotossa bittijono jaetaan samalla polynomilla. Jos jakojäännös on nolla, sanoma on vastaanotettu virheettömästi.

DAT (Digital Audio Tape)

DAT (Digital Audio Tape)–nauha käyttää ns. helical kuvaa/kuvausta (scan) magneetti-nauhalle diagonaalisesti kirjoitetun tiedon kirjoittamiseen ja lukemiseen. DAT/DDS -nauhalle mahtuu 2-4GB,

DAT/DDS2 8-12GB, DAT/DDS3 12-24GB.

Diaari

Tietyn periaatteen, tavallisesti diaarikaavan, mukaan ryhmitelty luettelo vireille laite-tuista asioista ja niiden käsittelyvaiheista. Arkistonmuodostajan kaikki (tai lähes kaikki) tehtäväalueet kattavaa diaaria kutsutaan myös hallintodiaariksi (erotukseksi ns. erityis-rekistereistä). Diaarin eräänä tehtävänä on toimia siihen merkittyjen asiakirjojen arkis-tohakemistona.

DLT (Digital Linear Tape)

DLT -nauhuri siirtää tietoa nopeammin kuin DAT -asetat, ja DLT -kasettien kapasiteetit ovat myös suuremmat. Toisaalta DLT on huomattavasti DAT :ia kalliimpaa, mutta tarjoaa paremman kasvunvaran. Nauha-asetat pakkaavat tiedon reaaliajassa. Valmistajasta ja tekniikasta riippuen asemille luvataan kaksinkertaista tai parempaakin pakkaus-suhdetta. Kaikki työtiedostot eivät kuitenkaan pakkaannu laisinkaan, mutta esimerkiksi tekstitiedostot pakkaantuvat hyvin. Tämän johdosta nauhojen kaksoiskapasiteetteja voi pitää suuntaa-antavina lukuina. Nauhalle siirrettävä data pakataan tietovirtana. Pakkausalgoritmi ei "näe" kaikkea dataa kerralla, joten pakkaussuhde on periaatteessa heikompi kuin perinteisen tiedostopakkaajan aikaansaannoksessa. DLT/DLT40 –nauhalle mahtuu 20-40GB ja DLT/DLT70:lle 35-70GB.

Eheys

1. (Tietojen tai tietojärjestelmän) aitous, väärentämättömyys, sisäinen riidattomuus, kattavuus, ajantasaisuus, oikeellisuus ja käyttökelpoisuus.
2. Ominaisuus, että tietoa tai viestiä ei ole valtuudettomasti muutettu, ja että mahdolliset muutokset voidaan todentaa kirjausketjusta.

Elinkaari

Termillä kuvataan asiakirjojen käsittelyvaiheita organisaatiossa niiden laatimises-ta/saapumisesta pysyvään säilytykseen tai hävittämiseen. Elinkaaren aktiivivaiheessa asiakirjoja tarvitaan jatkuvasti organisaation tehtävien hoidossa. Passiivivaiheessa asiakirjojen käyttö organisaation tarpeisiin on vähäistä, mutta niitä saatetaan edelleen tarvita tehtävien hoidossa tai oikeusturvaan liittyvistä syistä. Sekä aktiivi- että passiivivaiheessa tapahtuu asiakirjojen seulontaa. Historiallisen vaiheen asiakirjat ovat seulottuja, py-

syvään säilytykseen määrättyjä asiakirjoja, joiden säilytys perustuu pääasiassa niiden merkitykseen tutkimuksen lähdeaineistona.

Erityisrekisteri

Yhteisön tiettyyn tehtävään liittyvien asioiden vireille tuloa, käsittelyä ja seurantaan varten ylläpidettävä rekisteri. Erityisrekisteri on kattavuudeltaan suppeampi kuin diaari, johon kirjataan yhteisön kaikkiin (tai lähes kaikkiin) tehtävälajeisiin liittyviä asiakirjoja/asioita. Diaarin tapaan erityisrekisteri voi toimia arkistohakemistona.

Formaatti

1. Jokaisella ohjelmalla on oma tapansa tallettaa levyille sillä luodut tiedostot. Yleensä ohjelmat tallettavat tiedot itselleen sopivassa muodossa. Eri ohjelmilla on erilainen kyky lukea toistensa tiedostomuotoja. Usein myös saman ohjelman eri versioilla on omat tiedostomuotonsa.
2. Metatiedon esittämisen rakenne ja säännöt esim. Dublin Core.

Hyllymetri

Arkistokokonaisuuden laajuuden mitta, arkistokokonaisuuden viemä tila arkistohyllyssä.

Hyvä tiedonhallintatapa

Huolehtiminen asiakirjojen ja tietojärjestelmien sekä niihin sisältyvien tietojen saatavuudesta, käytettävyydestä, suojaamisesta, eheydestä ja muista tietojen laatuun vaikuttavista tekijöistä.

Kirjaaminen

Yhteisön käsittelemiin asioihin liittyvän kirjeenvaihdon ja päätösten sekä asioiden käsittelyvaiheiden merkitseminen diaariin. Kirjaamisen ohella käytetään myös termiä diariointi.

Konvertointi

Tiedon esitysmuodon tai rakenteen muuttaminen toiseen formaattiin tietosisältöä muuttamatta.

Kuvailu

Tiedot, jotka identifioivat ja kuvaavat arkistonmuodostajan ja sen toiminnasta kertyneen aineiston.

Käytettävyys

1. Ominaisuus, että tieto, tietojärjestelmä tai palvelu on siihen oikeutetuille saatavilla ja hyödynnettävissä haluttuna aikana ja vaaditulla tavalla.
2. Helppokäyttöisyys.

Laatu (tiedon)

Tiedolle asetettavien vaatimusten kokonaisuus, käyttökelpoisuus.

Luottamuksellisuus

1. Tietojen säilyminen luottamuksellisina ja tietoihin, tietojenkäsittelyyn ja tietoliikenteeseen kohdistuvien oikeuksien säilyminen vaarantumiselta ja loukkaukselta.
2. Se, missä määrin luottamuksellisuutta pidetään tärkeänä.

Metatieto

Yleisnimitys tietoverkon tai muiden tietovarantojen sisältämien dokumenttien kuvailutiedoille. Metatieto on tietoa tiedosta, tietosisältöä kuvaavaa ja sen merkitystä selittävää informaatiota. Metatieto helpottaa aineiston hakua, paikallistamista, tunnistamista ja säilyttämistä sähköisessä muodossa. Metatiedolla on tärkeä merkitys myös dokumenttien käsittelyvaiheessa. Kuvailutiedot voidaan tallentaa asiakirjaan jo dokumenttia kirjoitettaessa.

Optinen väline

Optisia, digitaalisia levyjä on kolmenlaisia, joita voidaan käyttää digitaalisen tiedon varastointiin: WORM (Write-Once-Read-Many), ReWritable ja CD-ROM. WORM'lle kirjoitetaan kerran ja sitä voidaan lukea monta kertaa. ReWritable-levyyn voidaan kirjoittaa ja siitä voidaan poistaa tietoa useita kertoja. CD-ROM tulee sanoista Compact Disc-Read Only Memory eli pelkästään lukemiseen ja kuuntelemiseen tarkoitettu levy.

Pitkäaikainen säilytys

Asiakirjojen säilyttäminen arkistonmuodostussuunnitelman mukaisesti yli 10 vuotta.

Provenienssiperiaate

Asiakirjat kuuluvat sen yhteisön tai henkilön arkistoon, jonka haltuun ne ovat asian käsittelyn yhteydessä päätyneet, eikä arkistoinnin yhteydessä syntyneitä asiakirjojen järjestystä muuteta, kuitenkin epäjärjestykseen joutunut arkisto voidaan järjestää. Eli ei esimerkiksi siirretä tietyn lähettäjän lähettämiä kirjeitä vastaanottajan arkistokokonaisuudesta lähettäjän arkistoon tai järjestetä kirjeitä uudelleen aiheen mukaan

Pysyvä säilytys

Pysyvä säilytys on asiakirjojen arvonmäärittelyssä ja arkistolaitoksen päätöksissä käytetty ilmaus, jolla tarkoitetaan asiakirjan säilyttämistä ilman, että säilyttämiselle on katsottu voitavan asettaa mitään aikarajaa (säilytysperspektiivi satoja vuosia). Sen vuoksi sähköisistä tietojärjestelmistä pysyvään säilytykseen valittujen asiakirjojen tulee säilyä käytettävissä satoja vuosia laitteistojen ja ohjelmistojen muutoksista huolimatta.

Rajavuodet

Vuosiluvut joiden väliseltä ajalta arkistokokonaisuudessa on aineistoa (ei välttämättä kaikilta vuosilta).

Rekisteröinti

Yhteisön käsittelemien asioiden ja asiakirjojen merkitseminen rekisteriin. Diaariin tapahtuvalle rekisteröinnille on vakiintunut oma terminsä, kirjaaminen (diariointi).

Sarja, asiakirjasarja

Yhden arkistokokonaisuuden aineisto jaetaan sarjoihin aineiston muodon (esim. pöytäkirjat, saapuneet asiakirjat) tai organisaation toimintojen perusteella (yleishallinto, taloushallinto).

Seulonta

Asiakirjojen säilytysajan määrittäminen (arvonmääritys) ja säilytettävän ja hävitettävän aineksen erottelu arkistossa.

Sähköinen allekirjoitus

Tietokoneen luettavassa muodossa oleva henkilön nimikirjoitus tai sen vastine, esimerkiksi digitaalinen allekirjoitus, todisteena nimikirjoitukseen liittyvän asiakirjan tai viestin yhteydestä tiettyyn henkilöön.

Sähköinen asiointi

Asian hoitaminen tai tuotteen hankkiminen tietoverkossa tarjottua palvelua käyttäen. Asiointipalvelun tarjoaja voi olla esimerkiksi viranomainen, rahalaitos tai verkkokauppa.

Tiedosto

Yhden kokonaisuuden muodostama tietojoukko, jolle on annettu yhteinen nimi.

Tietoaineisto

Joukko tietoa, kuten asiakirjoja, tiedostoja tai arkistoja, joka muodostaa joltakin kannalta kokonaisuuden.

Tietosuoja

Henkilötietolain mukaisten henkilötietojen käsittelyä koskevien vaatimusten huomioon ottaminen yksityiselämän suoja ja muita yksityisyyden suojaava turvaavien perusoikeuksien toteuttamiseksi henkilötietoja käsiteltäessä.

Tietoturvallisuus

1. Asiointi, jossa tietojen, tietojärjestelmien ja tietoliikenteen luottamuksellisuuden, eheyden ja käytettävyyden kohdistuvat uhat eivät aiheuta merkittävää riskiä.

2. Keinojen ja toimenpiteiden kokonaisuus, joiden avulla pyritään varmistamaan tietoturvallisuus niin normaali- kuin poikkeusoloissa

Yleisluettelo

Luettelo, jossa on kuvattu yhden arkisto-organisaation säilyttämät aineistot yleisellä tasolla, mm. mainittu arkistonmuodostajat