

Minna Kivioja

**MENETELMIÄ HENKILÖKOHTAISEEN
TIEDONHALLINTAAN**



JYVÄSKYLÄN YLIOPISTO
TIETOJENKÄSITTELYTIETEIDEN LAITOS
2012

TIIVISTELMÄ

Kivioja, Minna Kaarina

Menetelmiä henkilökohtaiseen tiedonhallintaan

Jyväskylä: Jyväskylän yliopisto, 2012, 92 s.

Tietojärjestelmätiede, pro gradu -tutkielma

Ohjaaja(t): Salminen, Airi

Henkilökohtainen tiedonhallinta sisältää erilaisia toimintoja tietojen käsittelemiseen, jota tehdään henkilökohtaisella tietokoneella sekä töissä että kotona. Tässä tutkielmassa käsitellään henkilökohtaisen tiedonhallinnan aihealuetta ja etenkin siihen sisältyviä henkilökohtaisen tiedonhallinnan toimintoja ja niiden suorittamisessa käytettyjä yleisimpiä menetelmiä. Tutkimuksen lähtökohtana oli selvittää, mitä menetelmiä henkilökohtaisessa tiedonhallinnassa käytetään. Tutkimus rajattiin käsittelemään yhdellä kotikoneella tehtävää tiedonhallintaa.

Tutkielman ensimmäisessä osassa tehtiin kirjallisuuskatsaus, jossa lähdekirjallisuuden pohjalta esiteltiin aiheen keskeistä käsitteistöä. Henkilökohtaiseen tiedonhallinnan määritelmien lisäksi esiteltiin myös aiheeseen sisältyvät käsitteet henkilökohtainen tieto ja tietoavaruus, tietoyksikkö, henkilökohtaisen tiedon kokoelma ja henkilökohtaisen tiedonhallinnan toiminnot. Tutkielmassa käsiteltiin löytämisen, säilyttämisen, järjestämisen ja ylläpidon toimintoja. Toimintoihin kuuluvina menetelminä käsiteltiin löytämisessä hierarkkista navigointia ja hakua, säilyttämisessä hierarkkista kansiorakennetta ja sähköpostia tiedon säilytyspaikkana, järjestämisessä tiedon sijoittamisen tapoja ja luokittelua, lajittelua ja nimeämistä sekä ylläpidosta varmuuskopioita ja päivittämistä.

Tutkielman toisessa, empiirisessä osassa laadittiin kyselytutkimus toimintoista ja menetelmistä teorian pohjalta. Saadut tulokset olivat suurimmalta osin yhteneväiset aikaisemmissa tutkimuksissa kerättyjen tulosten kanssa. Huomattiin, että ihmiset mieluummin etsivät tietoja navigoimalla kansiorakenteessa kuin käyttävät hakua. Tietokokoelmista havaittiin tiedostokokoelmaa käsiteltävän aktiivisemmin kuin sähköposti- tai kirjanmerkkikokoelmaa. Todettiin myös, että varmuuskopioita tehdään useasti liian harvoin, koska tietojen katoamisen riskiä aliarvioidaan. Tutkielmassa muokattiin aikaisemmassa tutkimuksessa muodostettua järjestelemisen luokittelua jakamalla yksi luokka kahteen uuteen luokkaan.

Asiasanat: tiedonhallinta, henkilökohtainen tiedonhallinta, terminologia, kirjallisuuskatsaukset, menetelmät, kysely

ABSTRACT

Kivioja, Minna Kaarina

Methods for Personal Information Management

Jyväskylä: University of Jyväskylä, 2012, 92 p.

Information Systems, Master's Thesis

Supervisor(s): Salminen, Airi

Personal information management involves different activities for data processing that is performed in personal computer both work and in home. This thesis concerns the topic of personal information management and including in it the personal information management activities and the most common methods that are used to accomplish them. The starting point for this study was to determine what kinds of methods are used. Study was outlined to cover the information management done in one home computer.

In the literature review that was made based on source materials in the first part of the thesis concerns the presentation of topics terminology. In addition to the definitions of personal information management this thesis also presented the terms including this topic which are personal information and space of information, information item, personal information collection and personal information management activities. This thesis concerns the finding, keeping, organizing and maintaining activities and their methods. The methods that were explored in finding activity are hierarchical navigation and search, keeping activity hierarchical folder structure and email as a repository for information, organising activity ways on placing information and categorizing, sorting and naming and in maintaining activity backups and updating.

In the second, empirical part of this thesis was constructed a query study based on theory concerning the activities and methods. The results observed in this study were in most part consistent with the results found in previous research. It was noticed that people like to find their information through navigating in their folder structure more than using the search. About the information collection it was found that the file collection was the most actively managed collection than the email- or bookmark collections were. It was also noticed that backups are made too infrequently because the risk of losing information is underestimated. In this study was one of the earlier organization strategy classifications modified in a way that one category was divided into two new categories.

Keywords: information management, personal information management, terminology, literature reviews, methods, query

KUVIOT

KUVIO 1 Henkilökohtainen tietoavaruus (Jones, 2008, s. 43).....	17
KUVIO 2 Perinteinen näkökulma tiedonhallintaan (Jones, 2008, s. 60)	20
KUVIO 3 Tiedostojen määrien prosenttijakaumat luokittain	47
KUVIO 4 Sähköpostiviestien kokonaismäärien prosentti jakaumat luokittain .	48
KUVIO 5 Sähköpostiviestien määrien prosenttijakaumat luokittain kansioissa	48
KUVIO 6 Kirjanmerkkien määrien prosenttijakaumat luokittain.....	49
KUVIO 7 Tiedostokokoelman järjestämisen aikaväli.....	51
KUVIO 8 Sähköpostikokoelman järjestämisen aikaväli	52
KUVIO 9 Kirjanmerkkikokoelman järjestämisen aikaväli	54

TAULUKOT

TAULUKKO 1 Aikaisemmissa tutkimuksissa havaitut järjestämisen strategiat. (Henderson & Srinivasan, 2011, s. 2)	33
TAULUKKO 2 Tietoavaruuden järjestämisen tapojen jakaumat tiedekunnan ja sukupuolen mukaan.....	46
TAULUKKO 3 Tietojen järjestämiseen käytetty aika päivässä.....	49
TAULUKKO 4 Tiedostojen järjestämisen tavat.....	50
TAULUKKO 5 Sähköpostin järjestämisen tavat	52
TAULUKKO 6 Kirjanmerkkien järjestämisen tavat	53
TAULUKKO 7 Luokittelun ja nimeämisen uusien luokkien frekvenssit.....	55
TAULUKKO 8 Tiedostojen etsimisen tapojen järjestyksen jakaumat ja keskiarvot.....	56
TAULUKKO 9 Vastaajilla käytössä olevat varmuuskopiointiohjelmat	57
TAULUKKO 10 Varmuuskopioitavat tiedot.....	58
TAULUKKO 11 Varmuuskopioiden määrä yhdestä tiedostosta	58
TAULUKKO 12 Varmuuskopioiden tallennuspaikat.....	59
TAULUKKO 13 Varmuuskopioiden tekemisen aikaväli	59
TAULUKKO 14 Tietojen häviämisen todennäköisyys laitteistovian sattuessa..	60
TAULUKKO 15 Taustamuuttuja vertailu käyttökokemus - todennäköisyys tietokatastrofiin.....	61
TAULUKKO 16 Taustamuuttuja vertailu tiedekunta - todennäköisyys tietokatastrofiin.....	62
TAULUKKO 17 Taustamuuttuja vertailu sukupuoli - todennäköisyys tietokatastrofiin.....	63
TAULUKKO 18 Kansiorakenteen tasojen vertailu	63
TAULUKKO 19 Tietoavaruuden järjestämisen strategioiden vertailu.....	64
TAULUKKO 20 Tietokokoelmien järjestämisen aikavälin vertailu.....	66
TAULUKKO 21 Tietokokoelmien tietoyksiköiden keskiarvojen vertailu	66

SISÄLLYS

TIIVISTELMÄ	2
ABSTRACT	3
KUVIOT	4
TAULUKOT	4
SISÄLLYS.....	5
1 JOHDANTO.....	7
1.1 Tutkielman tavoite ja rajaus	7
1.2 Tutkimusongelmat ja tutkimusmenetelmä.....	9
1.3 Tutkimuskenttä ja aikaisempaa tutkimusta.....	9
1.4 Keskeisimmät huomiot	11
1.5 Tutkielman rakenne	12
2 KÄSITTEISTÖ.....	13
2.1 Tiedonhallinta	13
2.2 Henkilökohtainen tiedonhallinta	14
3 LÖYTÄMINEN.....	22
3.1 Navigoiminen.....	24
3.2 Haku	25
4 SÄILYTTÄMINEN.....	27
4.1 Hierarkkinen kansiorakenne.....	28
4.2 Sähköposti tiedon säilytyspaikkana.....	29
5 JÄRJESTÄMINEN	31
5.1 Tietojen sijoittelu tapoja	31
5.2 Luokittelu, lajittelu ja nimeäminen	36
6 YLLÄPITO.....	37
6.1 Varmuuskopiot	38
6.2 Päivittäminen	39
7 EMPIIRINEN TUTKIMUS.....	41
7.1 Tutkimuksen tavoite ja menetelmä.....	41
7.2 Kohdejoukko ja kyselylomakkeen laadinta.....	42
7.3 Aineisto ja analyysi.....	44

7.4	Tulosten vertailua	60
7.5	Tutkimuksen luotettavuus	66
8	JOHTOPÄÄTÖKSET JA POHDINTA.....	68
9	YHTEENVETO	73
	LÄHTEET	76
	LIITE 1 KYSELYLOMAKE	80
	LIITE 2 TAULUKOT KÄYTTÖKOKEMUKSESTA JA TIETOKOKOELMIEN JÄRJESTÄMISEN TAVOISTA	87
	LIITE 3 TAULUKOT VARMUUSKOPIOINTI MUUTTUJISTA.....	89

1 JOHDANTO

Ihmiset etsivät ja tallentavat monenlaista tietoa elämänsä aikana sekä työssä että arjessa. Tiedonhallinnan kanssa toimivat jollakin tavalla kaikki, koska jokaisella on jonkinlaista tietoa kerättyinä. Tiedot voivat olla paperisia tai digitaalisia dokumentteja, musiikkia, kuvia ja erilaisia muita tiedostoja ja informaatiota tallennettuna eri muodoissa. Näitä tietoja on tallennettuna koneella ja eri tallennusvälineillä, kuten cd:llä/dvd:llä, muistitikulla, ulkoisella kovalevyllä tai verkossa.

Henkilökohtainen tiedonhallinta on laaja aihealue ja pitää sisällään monia tietojenkäsittelyyn liittyviä tehtäviä ja ongelmia. Useat ihmiset kamppailevat erilaisten tiedonhallinnan tehtävien kanssa yrittäessään pitää yhä kasvavat tietojen määrät hallinnassa. Nykyään myös kotikoneilla olevien digitaalisten tietojen määrä on kasvanut, kun erilaisia tietoja on mahdollisuus saada elektronisessa muodossa (esimerkiksi e-laskut, sähköiset pankkitiliotteet ym.). Digitaalikaameroiden yleistyessä on kotikoneille tallennettuna yhä suurempi määrä digitaalisia valokuvia. Kaikkia erilaisia kerättyjä tietoja on myös pystyttävä käsittelemään mahdollisimman tehokkaasti, jotta niiden käyttäminen olisi helppoa. Tähän tiedonhallinnan hoitamiseen ihmiset ovat kehittäneet erilaisia menetelmiä ja käyttäytymistapoja.

Henkilökohtaisella tiedonhallinnalla tarkoitetaan niitä tapoja, joilla ihmiset hankkivat, varastoivat, järjestävät ja hakevat digitaalista tietoa henkilökohtaisessa tietokoneen käyttöympäristössään (computing environment) sekä töissä että kotona (Boardman 2004, 1, 16). Menetelmät ovat niitä tapoja, joita käytetään näistä eri henkilökohtaisen tiedonhallinnan toiminnoista suoriutumiseen.

1.1 Tutkielman tavoite ja rajaus

Tutkielmassa on tarkoitus käsitellä joitakin keskeisimpiä henkilökohtaisen tiedonhallinnan toimintoja ja niihin liittyviä menetelmiä. Tutkielmassa tarkastellaan myös henkilökohtaisen tiedonhallinnan aihealuetta keskeisen käsitteistön

osalta. Tarkoituksena on ensin perehtyä henkilökohtaisen tiedonhallinnan peruskäsitteisiin ja niiden kautta pyrkiä ymmärtämään henkilökohtaisen tiedonhallinnan keskeisintä olemusta. Tutkielmassa tarkasteltavassa henkilökohtaisen tiedonhallinnan keskeisessä käsitteistössä käydään läpi henkilökohtaisen tiedonhallinnan määritelmiä eri tutkijoilta sekä aiheeseen läheisesti liittyvät muut käsitteet: henkilökohtainen tieto, tietoyksikkö, henkilökohtainen tietoavaruus, henkilökohtaisen tiedon kokoelmat sekä henkilökohtaisen tiedonhallinnan toiminnot. Nämä kaikki mainitut käsitteet ovat ominaisuuksia, jotka muodostavat henkilökohtaisen tiedonhallinnan keskeisen sisällön. Nämä käsitteet muodostavat pohjan, joka on perustana tutkielman muulle käsittelylle, jossa käydään tarkemmin läpi henkilökohtaisen tiedonhallinnan toimintoja. Määritelmien pohjalta on tähän tutkielmaan valittu käsiteltäviksi toiminnot löytäminen, säilyttäminen, järjestäminen sekä ylläpito. Tutkielmassa käsiteltyjen toimintojen osalta on jokaiselle esitelty niiden suorittamisessa käytettyjä yleisimpiä menetelmiä.

Aihepiiri on rajattu käsittelemään henkilökohtaisen tiedonhallinnan menetelmiä, joita käytetään jokapäiväisessä tiedonhallinnassa, ja jotka ovat tietokoneen käyttäjille tuttuja. Tässä työssä keskitytään käyttäjän yhdellä koneella digitaalisessa muodossa oleviin tietoihin ja lähinnä arjessa tiedonhallintaan käytettäviin menetelmiin. Menetelmät jaotellaan toimintojen mukaan. Tarkoituksena ei ole käydä läpi kaikkia käytettävissä olevia menetelmiä, vaan työssä keskitytään esittelemään 1-2 menetelmää kullekin käsiteltävänä olevalle toiminnolle. Menetelmät ovat valikoituneet aikaisempien tutkimuksien pohjalta. Tässä tutkielmassa ei tarkemmin syvennytä menetelmissä käytettäviin ohjelmistoihin. Henkilökohtaisen tiedonhallinnan pohjana olevaa tiedonhallinnan käsitettä esitellään melko lyhyesti, ja sen tarkoituksena on lähinnä pohjustaa itse henkilökohtaisen tiedonhallinnan käsittelyä.

Tutkielman teorialuvuissa käsitteiden, toimintojen ja menetelmien esittelyssä käytetään alan akateemista lähdekirjallisuutta, kuten aiheeseen liittyviä kirjoja tai muita tutkielmia sekä myös alan lehdissä olevia artikkeleita julkaisuja. Artikkeleita ja tutkimuksia on haettu eri artikkelitietokannoista, joista käytössä ovat olleet ainakin ACM, IEEE Digital Library ja ScienceDirect. Kirjoja ja muita materiaaleja on haettu kirjaston tietokantojen kautta käyttämällä Jyväskylän yliopiston kirjaston omaa JYKDOKK-tietokantaa sekä muiden kirjastojen materiaaleja LINDA-yhteistietokannasta. Lähteiden hakemisessa on käytetty myös Internetin hakupalveluita, kuten Google scholaria, jolla on haettu aiheita käsitteleviä julkaisuja. Tietoja haettiin muun muassa hakusanoilla personal information management, information management, personal information management study, information keeping, information finding ja information organizing. Empiirisessä tutkimuksessa tiedonkeruu tapahtuu laadittavan kyselylomakkeen kautta.

1.2 Tutkimusongelmat ja tutkimusmenetelmä

Tutkielmassa lähdetään hakemaan vastauksia käsiteltävään aiheeseen seuraavien kysymysten kautta:

Tutkimusongelma:

Mitä menetelmiä henkilökohtaisessa tiedonhallinnassa käytetään?

Osaongelmat:

Mitä ovat henkilökohtaisen tiedonhallinnan menetelmät?

Mitkä menetelmät ovat yleisimmin käytössä?

Miten ihmiset käyttävät menetelmiä?

Tutkielman alussa tehdään kirjallisuuskatsaus alan lähdekirjallisuuden sekä aikaisempien tutkimuksien pohjalta. Kirjallisuuskatsauksen tarkoituksena on ensin käydä läpi käsiteltävää aihepiiriä teorian kautta. Teoriaosassa perehdytään aiheen keskeisimpään käsitteistöön ja henkilökohtaisen tiedonhallinnan piirteisiin. Tämän jälkeen käydään lähdemateriaalien kautta läpi teoriaa henkilökohtaisen tiedonhallinnan toiminnoista sekä niiden kohdalla käytettävistä menetelmistä. Tutkielmassa esiintyvät käsitteet on lähteistä kääntänyt tekijä itse.

Tutkielman empiirisessä osassa tehdään kyselytutkimus, joka pohjautuu kirjallisuuskatsauksessa esiteltyyn aineistoon toiminnoista ja menetelmistä sekä osaksi myös aikaisempiin tutkimuksiin, joita ovat tehneet muun muassa Boardman & Sasse (2004) sekä Bergman, Byeth-Marom, Nachmias ym. (2008). Tutkimuksessa selvitetään osallistujien käytössä olevia tiedon kokoelmia niiden kansiorakenteiden suhteen, ja tapoja, joiden kautta he tietoa järjestelivät, etsivät ja ylläpitävät. Lomakkeessa esitetyt kysymykset ja niissä annetut vastausvaihtoehdot on muodostettu tutkielman teoriaosassa esitettyjen ja aikaisemmissa tutkimuksissa annettujen tietojen pohjalta. Tutkimuksen tuloksia verrataan aikaisempien tutkimusten tuloksiin. Kohderyhmäksi valittiin Jyväskylän yliopiston kahden eri tiedekunnan opiskelijat Informaatioteknologian tiedekunnasta ja Yhteiskuntatieteellisestä tiedekunnasta. Koska yleensä tietotekniikka-alaa opiskelevilla on lähtökohtaisesti paremmat tiedot ja taidot tietotekniikan osaamisessa, on tämän vuoksi kyselyyn otettu mukaan myös toisen tiedekunnan opiskelijoita tulosten vääristymisen välttämiseksi.

1.3 Tutkimuskenttä ja aikaisempaa tutkimusta

Tutkimuksella on oma tehtävänsä nostaa esiin henkilökohtaisessa tiedonhallinnassa esiintyviä ongelmia ja etsiä ratkaisuja niihin. Tutkimuksen kautta on arvioitu toimintojen hoitamiseen kehiteltyjä työkaluja ja tulevaisuudessa keksitään uusia ratkaisuja helpottamaan henkilökohtaista tiedonhallintaa. Seuraavaksi perehdytään hieman henkilökohtaisen tiedonhallinnan tutkimuskenttään ja aiheesta tehtyyn aikaisempaan tutkimukseen.

Kuten Jones (2008, s. xii) kertoo, on henkilökohtaisen tiedonhallinnan tutkimus omana tutkimusalanaan suhteellisen uutta ja sen tutkijayhteisö kehittänyt melko äskettäin. Henkilökohtainen tiedonhallinta on kehittymässä merkittäväksi tutkimus-alueeksi ja prosessi henkilökohtaisen tiedonhallinnan tutkimusyhteisön luomiseksi on käynnissä. Tämän prosessin voi katsoa alkaneen osana CHI 2004 ihminen ja tietokone vuorovaikutus konferenssia pidettyä erityisintressiryhmän istuntoa koskien henkilökohtaista tiedonhallintaa. Käännekohtana henkilökohtaisen tiedonhallinnan yhteisön luomisessa voidaan varsinaisesti kuitenkin pitää PIM (Personal Information Management) 2005 NSF:n (National Science Foundation) sponsoroimaa työpajaa, joka järjestettiin tammikuussa 2005 Seatlessa. Henkilökohtaista tiedonhallintaa tutkiva tutkijayhteisö on tieteenalojen suhteen ja tekniseltä taustaltaan monipuolinen. PIM 2005 työpajaan osallistui tutkijoita useilta tieteenaloilta, kuten kognitiivinen tiede (cognitive science), ihminen-tietokone vuorovaikutus, informaatiotekniikka, tekoäly, tietokannan hallinta, tietämyksen hallinta (knowledge management) sekä tiedonhaku. (Mts. xii–xiii.) Voidaan kuitenkin sanoa, että henkilökohtaisen tiedonhallinnan tutkimuksesta suurin osa, varsinkin varhaisimmista, on tehty lähinnä ihminen-tietokone vuorovaikutus tieteenalan alaisuudessa. Ihmisen ja tietokoneen vuorovaikutuksen tutkimukseen omistautunut yhteisö on muodostunut 1980-luvun aikana (Jones, 2008, s. 11).

Henkilökohtaiseen tiedonhallintaan liittyvässä tutkimuksessa tärkeimpinä alueina voisi pitää empiirisiä tutkimuksia peruskäyttäjien käyttäytymisestä (user behavior) sekä ohjelmistojen ja työkalujen prototyyppejä. Käyttäytymisestä tehdyt tutkimukset tarjoavat monia olennaisia havaintoja käyttäjien strategioista ja tarpeista sekä tarjoavat suosituksia suunnitteluun. (Boardman 2004, s. 3) Suunnittelun ja prototyyppien kautta kehitetään yhä toimivampia ratkaisuja henkilökohtaisen tiedonhallinnan toimintojen suorittamiseen. Henkilökohtaisen tiedonhallinnan käyttäytymiseen liittyvissä tutkimuksissa tarkastellaan yleisesti eri toimintojen suorittamiseen liittyviä käyttäytymismalleja ja -tapoja. Esimerkiksi Bergmanin, Byeth-Maromin, Nachmiasin, Gradovitchin ja Whittakerin (2008) tutkimuksessa tutkittiin käyttäjien tiedon hakemisen tapoja. Käyttäjien järjestämisen ja käyttäytymisen strategioita eri tietokokoelmissa ovat tutkineet sähköpostien osalta muun muassa Whittaker ja Sidner (1996), tiedostoista Henderson ja Srinivasan (2011) ja kirjanmerkeistä Abrams, Baecker ja Chignell (1998) sekä kaikista kokoelmista Boardman ja Sasse (2004). Sähköpostin osalta on myös tutkittu sen käyttämistä kommunikoinnin lisäksi henkilökohtaisessa tiedonhallinnassa ja sen toimintojen suorittamisessa (esimerkiksi Whittaker, Bellotti & Gwizdka, 2006; Ducheneaut & Bellotti, 2001).

Tutkimuksissa on myös yhdessä keskeisessä roolissa ollut työkalujen ja ohjelmistojen arviointi, joita tehdään joko jo markkinoilla ja käytössä oleville tai tutkijoiden kehittämille prototyyppityökaluille. Esimerkiksi Bellotti, Ducheneaut, Howard ja Smith (2003) ovat tutkimuksessaan arvioineet aikaisempien tutkimusten tuloksien pohjalta kehittämänsä Taskmaster -sähköpostityökaluprototyypin ja Dumais, Cutrell, Cadiz, Jancke, Sarin ja Robbins (2003) arvioivat tutkimuksessaan kehittämänsä Stuff I've Seen -

työpöytähakutyökalua. Tutkimuksissa on yleensä työkaluista tarkasteltu niiden toimivuutta siinä tehtävässä, johon ne on suunniteltu sekä arvioitu työkalujen eri ominaisuuksia ja mahdollisesti vertailtu niitä aikaisemmin käytössä olleisiin työkaluihin. Kelly ja Teevan (2007, s. 190) toteavat, että arviointi on tärkeää, jotta tiedettäisiin, mitkä työkalut ovat parhaita, ja mikä tekee niistä parempia kuin muista vastaavista työkaluista. Erilaisten arviointimenetelmien ja mittausten avulla tutkijat voivat luoda periaatteelliset, objektiiviset ja tasapainoiset arvioinnit kehittämillensä työkaluille ja siten paremmin ymmärtää, mikä toimii ja mikä ei. (Mts. 190.)

Useimmat tutkimukset tarkastelevat henkilökohtaisen tiedonhallinnan suorittamista työympäristössä eli käytetään naturalistista tutkimustapaa. Hendersonin ja Srinivasan (2011, s. 3) mukaan tutkimusten keskittyminen työtilanteisiin johtuu lähinnä siitä, että työpaikoilla tiedonhallintaa suoritetaan suuremmissa mittakaavassa ja siellä on enemmän negatiivisia seurauksia, jos tiedonhallinnan menetelmät ovat tehottomia. Kotiympäristössä henkilökohtaisen tiedonhallinnan menetelmien toimimattomuus ei ehkä aiheuta yhtä suuria seurauksia, mutta silti on toimivien tiedonhallinnan menetelmien käyttäminen sielläkin yhtä tärkeää. Vaikka tutkimukset ja niiden näkökulmat usein koskevatkin töissä tapahtuvaa henkilökohtaista tiedonhallintaa, voidaan saatuja tuloksia ja huomioita soveltaa myös kotikoneilla tapahtuvaan tiedonhallintaan. Ihmiset noudattavat usein samoja käyttäytymismalleja tiedonhallinnassa riippumatta siitä käytetäänkö niitä töissä vai kotona. Esimerkiksi luonteeltaan järjestelmällinen henkilö ei todennäköisesti järjestele ahkerasti työkoneensa tietoja ja sitten kotikoneella jätä tietoja epäjärjestykseen. Työkoneella ja kotikoneella tapahtuvassa tiedonhallinnassa käytetään samanlaisia menetelmiä tietojen etsimiseen, järjestämiseen, säilyttämiseen ja ylläpitoon. Ainoastaan käsiteltävänä olevien tietojen määrässä on eroja, sillä työpaikoilla tietomäärät ovat suurempia kuin kotona. Tässä tutkielmassa keskitytään nimenomaan selvittämään kotona tapahtuvaa henkilökohtaisen tiedonhallintaa ja sen menetelmien käytettävyyttä.

Tutkimuksiin yleensä valitut osallistujat ovat eri yritysten työntekijöitä tai akateemisen alan osalta yliopistojen henkilöstöä ja joskus myös opiskelijoita. Yleisimpiä tutkimusmenetelmiä ovat olleet osallistujien haastattelut, joiden lisäksi on kerätty tietoja osallistujien koneiden työpöydältä, lähinnä kansiorakenteisiin liittyen, ottamalla näytönkaappauksia tai valokuvaamalla näyttöjä. Näiden lisäksi tutkimusmenetelmänä on käytetty kyselyjä.

1.4 Keskeisimmät huomiot

Tulokset ovat pääosin yhteneväiset aikaisemmissa tutkimuksissa havaittujen tulosten kanssa. Navigoiminen on hakua suositumpi tapa etsiä tiedostoja ja tiedostokokoelmassa on eniten kansioiden tasoja ja se on aktiivisimmin järjestetty kokoelma verrattuna sähköposti- ja kirjanmerkkikokoelmiin. Varmuuskopioita tehdään liian vähän, mikä on myös havaittu jo aiemmin. Lisäksi tietokatastrofin mahdollisuutta pidetään varsin epätodennäköisenä, mikä taas vahvistaa

aiemmin tehdyn toteamuksen, että negatiivisen tapahtuman todennäköisyyttä usein aliarvioidaan. Tutkielmassa muokattiin aikaisemmassa tutkimuksessa esitettyä järjestämisen tapojen luokittelua, jossa suuren kansion arkistoinen luokka jaettiin yhden suuren kansion arkistoinen ja usean suuren kansion arkistoinen luokkiin.

Jatkossa voidaan tämän tutkielman esittämän aineiston pohjalta laajentaa ja tutkia tarkemmin tietojen jakamista ja synkronointia useiden eri koneiden ja mobiililaitteiden välillä ja niissä tapahtuvaan tietojen hallintaan sekä tarkastella niissä käytettäviä menetelmiä. Lisäksi jatkotutkimuksissa voisi tarkastella menetelmille käytössä olevia ohjelmistoja ja tutkia niiden toimivuutta menetelmien osalta. Tutkielman empiirisessä tutkimuksessa tarkasteltiin opiskelijoiden tiedonhallintaa, joten jatkossa voitaisiin lisäksi laajentaa tutkimuksen kohdejoukkoa tutkimalla myös eri yhteiskunta- ja ikäryhmiä, mikä parantaisi tutkimuksen yleistettävyyttä.

1.5 Tutkielman rakenne

Tutkielma etenee siten, että Luvussa 2 esitellään tutkielmaa koskeva keskeinen käsitteistö siten, että alaluvussa 2.1 käydään ensin läpi tiedonhallinnan käsitettä ja alaluvussa 2.2 keskitytään itse henkilökohtaisen tiedonhallinnan käsitteeseen ja sen ominaispiirteisiin. Luvut 3-6 käsittelevät henkilökohtaisen tiedonhallinnan toimintoja ja niiden menetelmiä siten, että luku 3 käsittelee löytämisen toimintoa, luku 4 säilyttämisen toimintoa, luku 5 järjestämisen toimintoa ja luku 6 ylläpidon toimintoa. Luvussa 7 käsitellään tutkielman empiiristä tutkimusta, jossa esitellään suoritettua kyselytutkimusta. Alaluku 7.1 esittelee tutkimuksen toteuttamiseen valittua menetelmää ja tutkimuksen tavoitteita, alaluku 7.2 käsittelee kyselylomakkeen rakennetta ja kysymyksiä sekä esitellään tutkimukseen valitut kohdejoukot, alaluvussa 7.3 käydään läpi saadut tulokset ja niiden analyysi ja alaluvussa 7.4 vertaillaan ristiintaulukoinnilla taustamuuttujia sekä tämän tutkimuksen tuloksia aikaisemmissa tutkimuksissa mainittuihin tuloksiin. Alaluvussa 7.5 tarkastellaan tutkimuksen luotettavuutta. Luvussa 8 esitellään tutkimusta koskevat johtopäätökset ja pohdinta.

2 KÄSITTEISTÖ

Käsitteiden ja termien kautta perehdytään käsiteltävään aiheeseen ja niiden avulla luodaan pohja, jonka kautta aihetta voidaan tarkastella syvemmin. Ennen peruskäsitteiden tietämistä ei voi kunnolla oppia tuntemaan ja ymmärtämään tutkittavaa aihetta. Tässä luvussa on lyhyesti esiteltynä henkilökohtaisen tiedonhallinnan aihepiirin keskeisimmät käsitteet. Luku etenee siten, että aluksi on esiteltynä perinteisempi tiedonhallinnan käsite, joka toimii pohjana seuraavana olevalle varsinaiselle henkilökohtaisen tiedonhallinnan käsittelylle. Henkilökohtaisen tiedonhallinnan osalta keskitytään tarkemmin esittelemään henkilökohtaiseen tiedonhallintaan sisältyviä eri käsitteitä. Tarkasteltavana ovat muun muassa käsitteet henkilökohtainen tieto, henkilökohtainen tietoavaruus, tietoyksikkö ja henkilökohtaisen tiedon kokoelmat. Lopuksi tarkastellaan lähemmin henkilökohtaisen tiedonhallinnan toimintoja ja määritellään menetelmä tämän tutkielman käsittelyn kohdalla. Esiteltävät käsitteet on suomentanut lähteistä tutkielman tekijä. Käsitteiden kohdalla on pyritty käyttämään vakiintuneita ja hyväksytyjä käännöksiä, mutta joillekin termeille voi olla olemassa tarkempi käänнос, kuin tässä tutkielmassa käytetty.

2.1 Tiedonhallinta

Boardman (2004, s. 13) huomauttaa monen henkilökohtaisen tiedonhallinnan määritelmän olevan lähtöisin siitä perinteisen tiedonhallinnan lähtökohdasta, että tietoa tallennetaan, jotta sitä voidaan noutaa myöhempänä ajankohtana. Henkilökohtaisen tiedonhallinnan ja tiedonhallinnan erona on henkilökohtaisen tiedonhallinnan suuntautuminen yksilöön ja yksilön omien tietojen hallintaan.

Detlorin (2010, s. 103) määritelmän mukaan ”*tiedonhallinta* (Information management) on niiden prosessien ja järjestelmien hallintaa, jotka luovat, hankkivat, järjestävät, säilyttävät, jakavat ja käyttävät tietoa. Tiedonhallinnan päämääränä on auttaa ihmisiä sekä organisaatioita saamaan, muokkaamaan ja käyttämään tietoa tehokkaasti.” Tiedonhallintaa voidaan myös kuvata hallin-

nan periaatteiden soveltamiseksi kaikenlaisten organisaatioiden tehokkaan toiminnan kannalta oleellisten tietojen hankkimiseen, järjestämiseen, hallintaan, levittämiseen ja käyttöön (Boardman, 2004, s. 17–18; Macevičiute & Wilson, 2005, s. 21). Detlor (2010) käsittelee tiedonhallintaa kolmen näkökulman mukaan. Yksi näistä näkökulmista on *henkilökohtainen näkökulma* (personal perspective), joka koskee henkilökohtaista tiedonhallintaa, johon perehdytään tarkemmin luvussa 2.2. Muut tiedonhallinnan näkökulmat ovat organisatorinen näkökulma ja kirjastonäkökulma.

Suosituin ja hallitsevin näkökulma on *organisatorinen näkökulma* (organizational perspective), jossa tiedonhallinnan kautta pyritään parantamaan organisaatioiden kilpailukykyä ja tehokkuutta (Detlor, 2010). Kirk (2005, s. 3–5) mainitsee artikkelissaan, että ”tiedonhallintaa harjoitetaan organisaatioissa”, ja usein termi tiedonhallinta liitetäänkin organisaatioihin ja sitä määritellään organisaation näkökulman mukaan. Kirk määrittelee tietoa ja tiedonhallintaa Bramanin (1989) esittämän hierarkian pohjalta, jossa tiedon määritelmät on jaettu neljään ryhmään. Tämä lähestymistapa soveltuu organisaatioille muun muassa siksi, että se huomioi kvalitatiiviset eroavaisuudet tiedon määritelmässä sekä sen makronäkökulman vuoksi, joka on soveliaampi organisaatioille kuin määritelmät, jotka perustuvat pelkästään yksilöihin tiedon käyttäjinä. (Kirk, 2005, s. 5) Organisatorisessa näkökulmassa tietoa pidetään ja käsitellään strategisena resurssina. Tiedonhallintaan organisaatioissa kuuluu monenlaisten tietoresurssien kokonaisuuksien hallinnointi aina datasta informaatioon. (Detlor 2010.)

Kirjastonäkökulma (library perspective) tiedonhallintaan huomioi sen roolin, joka on tietoa toimittavilla organisaatioilla, kuten kirjastoilla, joiden keskeinen tehtävä on tarjota asiakkailleen pääsy tietoresursseihin ja -palveluihin. Näissä organisaatioissa tiedonhallinta on pääasiassa tiedon kokoelmien, kuten kirjojen ja aikakausjulkaisujen, hoitamista. Kirjastoissa tiedonhallinnan tehtäviin lukeutuu muun muassa hankkia kokoelmien kohteita mahdollisimman edullisesti, kokoelmien järjesteleminen, varastointi ja tietojen hakeminen kokoelmista. Lisäksi tehtävänä on huolehtia tietojen saatavuudesta ja levittämisestä loppukäyttäjille. (Detlor 2010, s. 103, 106.)

2.2 Henkilökohtainen tiedonhallinta

Itse henkilökohtaisen tiedonhallinnan termiä on käytetty ensimmäisen kerran 1980-luvulla, jolloin oli suuria odotuksia henkilökohtaisten tietokoneiden tuomasta mahdollisuudesta parantaa ihmisten kykyä käsitellä ja hallita tietoa (Jones & Teevan, 2007, s. 3; Jones, 2008, s. 11). Henkilökohtaisesta tiedonhallinnasta on olemassa monenlaisia määritelmiä. Eri tutkijoilla on erilaisia näkökulmia henkilökohtaisen tiedonhallinnan tarkasteluun ja siten myös määritelmät painottavat eri asioita. Boardman (2004, s. 13) mainitsee Whittakerin ym. (2000) huomioimasta systemaattisuuden puutteesta ihminen ja tietokone vuorovaikutuksen tutkimuksissa, joissa käsitellään henkilökohtaisen

tiedonhallinnan jokapäiväisiä toimintoja. Lisäksi Boardman tuo esille Whittakerin ym. tekemän huomion, että yhtenä ongelmakohtana kyseisessä asiassa on yksimielisyyden puute keskeisten termien määrittelyssä. Boardmanin mukaan tämä tarkoittaa sitä, että tutkijat tehdessään uutta tutkimusta lisäävät uusia määritelmiä laajaan joukkoon jo olemassa olevia vaihtoehtoja. Jones (2008, s. 59) huomauttaakin, että henkilökohtainen tiedonhallinta on hankala määritellä siten, että se säilyttäisi huomion henkilökohtaisen tiedonhallinnan keskeisiin haasteisiin. Seuraavana tarkastellaan muutamia eri tutkijoiden antamia määritelmiä henkilökohtaisesta tiedonhallinnasta.

Yksi aikaisimmista määritelmistä koskien henkilökohtaista tiedonhallintaa on Lansdalen (1988, s. 55) määritelmä, jonka mukaan *henkilökohtainen tiedonhallinta* (personal information management) on toiminta- ja menettelytapa, jolla me käsittelemme, luokittelemme ja haemme tietoa jokapäiväisellä tasolla. Barreau (1995) taas tarkastelee henkilökohtaista tiedonhallintaa järjestelmänä. Hän kuvailee henkilökohtaista tiedonhallintaa yksilöä varten luoduksi tai yksilön kehittämäksi järjestelmäksi, joka on tehty henkilökohtaista käyttöä varten työympäristössä. Tähän järjestelmään kuuluvat toiminnot tiedon hankkimiseen, tallentamiseen ja järjestämiseen, ylläpitoon, noutamiseen ja tulosteiden tuottamiseen sekä kaikki näihin toimintoihin liittyvät säännöt ja menettelytavat. (Barreau 1995, s. 328.) Bellotti, Ducheneaut, Howard, Smith ja Neuwirth (2002, s. 182) kuvaavat henkilökohtaista tiedonhallintaa tiedon järjestämiseksi luokittelun tai sijoittamisen kautta tavalla, joka tekee tiedosta helpommin haettavan, kun sitä tarvitaan.

Boardmanin (2004, s. 1, 13, 16) määritelmän mukaan henkilökohtainen tiedonhallinta tarkoittaa niitä tapoja, joilla ihmiset hankkivat, varastoivat, järjestävät ja hakevat digitaalista tietoa henkilökohtaisessa tietokoneen käyttöympäristössään (computing environment) sekä töissä että kotona. Tämän määritelmän Boardman on muokannut Barreaun määritelmän pohjalta, minkä lisäksi hän on ottanut määritelmäänsä pohjaksi termit tieto, henkilökohtainen tieto, tietoyksikkö ja tiedon kokoelma. Näiden tuloksena saadusta määritelmästä on Barreaun toiminnoista jätetty pois tulosteiden tuottaminen, ja näkökanta henkilökohtaiseen tiedonhallintaan järjestelmänä on muutettu käyttäjän toimintojen tarkasteluksi. Jones (2008, s. 5) taas määrittelee henkilökohtaisen tiedonhallinnan siten, että se viittaa sekä käytäntöön että tutkimukseen niistä toiminnoista, joita henkilö suorittaa täyttääkseen elämän monet roolit ja vastuut (vanhempana, puolisona, ystävänä, työntekijänä, yhteiskunnan jäsenenä jne.) ja toiminnoista, joita tarvitaan kohtaamaan elämän monet päämäärät (joka päivä ja pidemmällä aikavälillä, työhön liittyvät tai liittymättömät). Toimintoihin kuuluvat tiedon luominen, tallentaminen, järjestäminen, jakaminen, noutaminen, käyttäminen ja ylläpito. Henkilökohtainen tiedonhallinta asettaa erityistä painotusta henkilökohtaisen tiedon kokoelmien järjestämiseen ja ylläpitoon, jossa tietoyksiköt, kuten paperiset ja elektroniset dokumentit, sähköpostiviestit, Web-viittaukset (Web references) ym. on tallennettu myöhempää käyttöä ja uudelleen käyttöä varten. (Mts. 5.)

Yllä mainituissa määritelmässä pyritään kuvaamaan eri tavoin henkilökohtaisen tiedonhallinnan keskeistä olemusta. Määritelmässä esiintyy niitä eri tekijöitä ja piirteitä, jotka henkilökohtainen tiedonhallinta sisältää. Yksi useimmissa määritelmässä esiintyvistä piirteistä on painotus tiedonhallinnan henkilökohtaisuuteen niin henkilön toimintaympäristön suhteen kuin tiedon omistamisen suhteen. Toinen keskeinen tekijä, joka määritelmistä käy ilmi ovat ne monet eri toiminnot, joilla henkilökohtaista tiedonhallintaa suoritetaan. Kun ottaa huomioon nämä piirteet voidaan sanoa, että näistä aikaisemmin esitellyistä määritelmistä tarkimmin henkilökohtaisen tiedonhallinnan olemuksen ilmaisee kuitenkin Jonesin (2008) ja Boardmanin (2004) määritelmät. Näitä kahta määritelmää ja niissä ilmeneviä toimintoja käytetään jatkossa tämän tutkielman tarkastelun pohjalla.

Tämän henkilökohtaisen tiedonhallinnan termin käsittelyn jälkeen keskitytään seuraavaksi esittelemään siihen liittyviä muita käsitteitä. Seuraavat käsitteet ovat olennainen osa henkilökohtaisen tiedonhallinnan aihepiiriä ja ne yhdessä kokonaisuutena muodostavat henkilökohtaisen tiedonhallinnan keskeisen sisällön.

Henkilökohtainen tieto ja tietoaavaruus

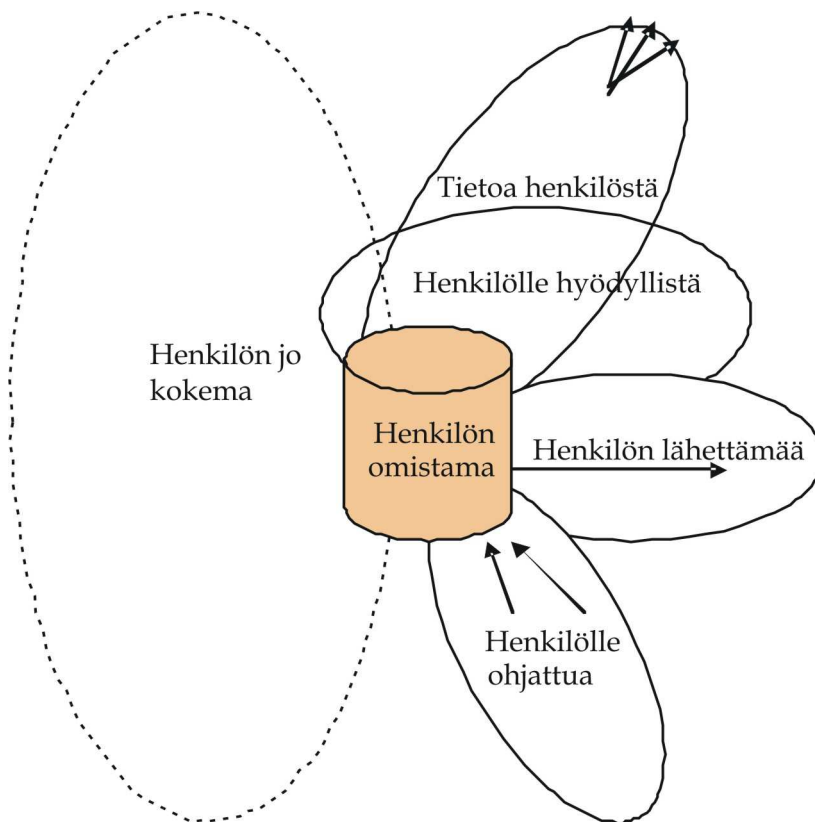
Henkilökohtainen tieto (personal information) on laaja käsite ja sillä on monta eri merkitystä, jotka riippuvat sen suhteesta henkilöön ja siitä miten henkilökohtaista tietoa käsitellään. Seuraavana on Jonesin (2008, s. 33–35) esittämä luettelo henkilökohtaisen tiedon eri merkityksistä:

1. Henkilön omistama ja hallitsema tieto. Tietoa, jota henkilö pitää henkilökohtaista käyttöä varten joko suoraan tai epäsuorasti (esimerkiksi ohjelmisto sovellusten kautta), ja joka on ainakin nimellisesti henkilön hallittavana. Joidenkin tietojen omistusoikeuksien osalta voi joskus olla erimielisyyttä, esimerkiksi tiedot henkilön työpaikalla tai muiden kanssa yhteistyössä tehdyt tiedot.
2. Tietoa henkilöstä. Tieto on toisten pitämää ja muiden hallitsemaa. Tähän merkitykseen lukeutuu se henkilökohtainen tieto henkilöstä, jota säilyttävät lääkärit ja terveydenhuolto organisaatiot sekä verovirasto.
3. Henkilön kokema tieto. Tietoa, jonka henkilö kokee, mutta ei välttämättä ole henkilön hallittavissa. Esimerkkinä tästä on vaikkapa web-sivu, jota henkilö katselee.
4. Henkilölle merkityksellinen tai hyödyllinen tieto. Tietoa, joka määrittyy sen mukaan, onko se henkilölle hyödyllistä tai merkitsevää. Tämä on tietoa, joka on "tuolla jossain" ja jonka henkilö saattaisi haluta kohdata.
5. Henkilölle ohjattava tieto. Tämä merkitys sisältää sähköpostiviestit, jotka tulevat Saapuneet -kansioon sekä ne ponnahdusvalikko ilmoitukset, jotka ilmoittavat saapuneesta viestistä. Lisäksi esimerkkinä ohjatusta tiedosta ovat henkilön tietokoneen nostamat hälytykset ja vierailtavan web-sivun ponnahdusikkuna mainokset sekä soiva puhelin. Tieto voi olla tai olla olematta henkilökohtaisesti merkityksellistä, mutta ohjatun tiedon

vaikutus on henkilökohtaista, sillä se voi häiritä henkilön meneillään olevaa tehtävää, viedä henkilön huomion ja vakuuttaa henkilön kuluttamaan aikaa ja rahaa, muuttamaan mielipidettä tai saada henkilön toimimaan.

- Henkilön lähettämä tai järjestämä tieto. Tieto, jonka henkilö julkaisee tai lähettää on myös henkilökohtaista. Yritämme usein kontrolloida, tosin puutteellisesti, sitä milloin ja mitä kukakin näkee siitä tiedosta jonka lähettämme, toimitamme tai muuten järjestämme. Tämä tehdään esimerkiksi sähköpostin kautta käyttämällä jakelulistoja ja sähköpostiviesteissä ilmoituksilla, kuten ”luottamuksellinen”. (Jones & Teevan, 2007, s.9; Jones, 2008, s. 33–35.)

Nämä kaikki yllä mainitut merkitykset yhteen liitettynä kokonaisuutena muodostavat *henkilökohtaisen tietoavaruuden* (personal space of information), joita jokaisella henkilöllä on yksi. Henkilökohtainen tietoavaruus on yksilöllinen ja se sisältää kaikki henkilön hallinnassa, ainakin nimellisesti, olevat tietoyksiköt sekä sovellukset, työkalut ja käsitteet (constructs) (esimerkiksi yhdistettävät ominaisuudet (associated properties), kansiot ja pinot eri muodoissa), jotka henkilö on hankkinut tai luonut avustamaan tietojen hallinnassa. (Jones & Teevan, 2007, s. 10.) Jones (2008) havainnollistaa henkilökohtaista tietoavaruutta graafisella kuviolla, joka on esitetty kuviossa 1 suomenkielisin termein.



KUVIO 1 Henkilökohtainen tietoavaruus (Jones, 2008, s. 43)

Kuviossa esitettyinä olevassa henkilökohtaisen tietoavaruuden keskiössä ovat ne tietoyksiköt, jotka henkilö omistaa, ja jotka ovat henkilön omassa hallinnassa, ainakin nimellisesti. Tähän kuuluvat kaikki dokumentit sekä paperisessa että digitaalisessa muodossa sekä muut mahdolliset tiedostot useilla koneilla, web-kirjanmerkit ja sähköpostiviestit useilla sähköpostitileillä sekä lisäksi ne sovellukset ja työkalut, jotka mahdollistavat tietojen hallinnan. Henkilökohtaisen tietoavaruuden reunamilla taas sijaitsevat tiedot, jotka henkilö voisi haluta tietää ja hallita, mutta jotka ovat muiden hallittavina. Näihin tietoihin sisältyvät myös muiden pitämät tiedot henkilöstä. Lisäksi reuna-alueilla sijaitsevat tiedot julkisista paikoista, kuten kirjastosta tai Internetistä, jotka voivat olla tai olla olematta henkilölle merkityksellisiä. (Jones & Teevan, 2007, s. 11; Jones, 2008, s. 44.)

Tietoyksikkö

Tiedonhallinnassa puhutaan usein dokumenteista ja niiden käsittelemisestä, mutta tosiasiaassa tiedonhallinta koskee myös kaikkia muita tietoja ja eri tiedostomuotoja. Jones (2008, s. 36) toteaa, että viitattaessa yleisiin henkilökohtaisen tiedonhallinnan objekteihin, kuten sähköpostiviestiin, web-kirjanmerkkiin ja ladattuun musiikkiin, on tavallisen dokumentin merkityksen kannalta hieman hankalaa käyttää dokumentti-termiä näistä objekteista puhuttaessa. Tämän vuoksi henkilökohtaisessa tiedonhallinnassa käytetään mieluummin termiä *tietoyksikkö* (information item). Tietoyksikkö on tiedon esitystapa (packaging), jossa tiedot tiivistetään yhtäläiseen muotoon. Tietoyksikköjä voidaan luoda, muunnella, varastoida, noutaa, niille voidaan antaa nimi ja muita ominaisuuksia, siirrellä, kopioida, jakaa, poistaa ja muuten manipuloida. Esimerkit tietoyksiköistä sisältävät 1) paperiset dokumentit, 2) elektroniset dokumentit, digitaaliset valokuvat, digitaalisen musiikin, digitaalisen filmin ja muut tiedostot, 3) sähköpostiviestit, 4) web-sivut ja 5) viittaukset (esimerkiksi pikakuvakkeet ja aliakset) mihin tahansa yllä olevista. Tietoyksiköllä on siihen liitettävä tietomuoto (information form), joka määrittyy niiden työkalujen ja ohjelmistojen mukaan, jotka tukevat näitä aiemmin mainittuja manipulointi operaatioita. Tavalliset tietomuodot sisältävät paperiset dokumentit, elektroniset dokumentit ja muut tiedostot, sähköpostiviestit ja web- kirjanmerkit. (Jones 2008, s. 37.) Boardman (2004, s. 15) taas esittää asian siten, että tietoyksikkö on itsenäinen tiedon yksikkö, joka esiintyy digitaalisessa ympäristössä vaihtelevassa teknisessä muodossa, kuten tiedosto, sähköpostiviesti tai kirjanmerkki.

Tavat, joilla tietoyksikköä manipuloidaan riippuvat suuresti siitä, mikä sen muoto on ja mitä työkaluja tälle muodolle on käytettävissä. Esimerkiksi vuorovaikutuksessa digitaalisten tietoyksiköiden kanssa ollaan riippuvaisia erilaisten tietokone-pohjaisten työkalujen ja sovellusten tuesta, kuten sähköpostisovellukset, tiedostonhallintaohjelmat ja web-selaimet. Kaikki tietoa sisältävät asiat eivät kuitenkaan sisälly tietoyksikön kategoriaan. Esimerkiksi kahden henkilön välinen keskustelu, joka sisältää tietoa ei itsessään ole tietoyksikkö. Kasettitalenne samasta keskustelusta taas luetaan tietoyksiköksi,

koska sitä voidaan kopioida, varastoida, lähettää jne.. Toisena esimerkkinä voidaan ottaa vaikkapa kakkuresepti, jota voidaan itsessään pitää tietona. Tämän tiedon esitystapa, oli se sitten kirjoitettu paperinpalalle tai esitettynä web-sivulla, on tietoyksikkö. Tietoyksiköt voivat myös sisältää toisia tietoyksiköitä. Esimerkiksi sähköpostiviesti, joka on tietoyksikkö, voi sisältää liitteenä vaikkapa skannatun kuvan reseptistä, joka myöskin on tietoyksikkö. (Jones 2008, s. 37-38.)

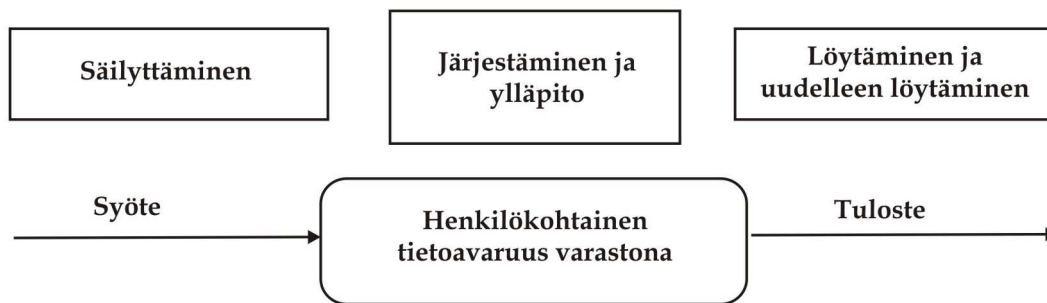
Henkilökohtaisen tiedon kokoelmat

Boardmanin (2004, s. 15) määritelmän mukaan *"henkilökohtaisen tiedon kokoelma* (personal information collection) on itsenäinen joukko tietoyksiköitä. Tyypillisesti kokoelman osilla on tietty yhteinen tekninen muoto ja niitä käytetään tietyn sovelluksen kautta." Jones (2008, s. 47) puolestaan määrittää henkilökohtaisen tiedon kokoelmia niiden toimintojen kautta, joita ihmiset tekevät suhteessa heidän kokoelmiinsa ja antaa kokoelmille seuraavanlaisen määritelmän: *"henkilökohtaisen tiedon kokoelmat ovat henkilökohtaisesti hoidettuja henkilökohtaisen tietoavaruuden osajoukkoja. Henkilökohtaisen tiedon kokoelmat ovat 'saarekkeita' henkilökohtaisessa tietoavaruudessa, jossa ihmiset ovat tehneet tietoisia pyrkimyksiä kontrolloida sekä tietoa, joka tulee sisään, että miten tätä tietoa on järjestelty."* Lisäksi Jones huomauttaa, että henkilökohtaisen tiedon kokoelmat voivat vaihdella paljon lukumäärän, muodon ja kohteiden sisällön johdonmukaisuuden suhteen (Mts. 47).

Kokoelmat saattavat olla vahvasti yhdistettyjä tiettyihin sovelluksiin, esimerkkinä vaikkapa sovellukset, joilla käsitellään digitaalisia kuvia tai musiikkia (Jones, 2008, s. 47). Kokoelmat, joita henkilöt luovat ovat heille ainutlaatuisia ja tiedot kokoelmien sisällä ovat automaattisesti yhdistettyjä omistajansa henkilökohtaisiin kokemuksiin (Elsweiler & Ruthven, 2007, s. 23).

Henkilökohtaisen tiedonhallinnan toiminnot

Kun ottaa huomioon Boardmanin (2004, s. 13) toteamuksen, että perinteisen tiedonhallinnan käsityksen mukaan tieto on tallennettu, jotta se voitaisiin noutaa myöhempanä ajankohtana, ja kun tarkastellaan Barreaun määritelmää, jonka mukaan tietoa kerätään järjestelmään säilytettäväksi ja uudelleen käytettäväksi, voidaan tämän perusteella henkilökohtaista tiedonhallintaa tarkastella suhteessa henkilön toimintaan ja vuovaikutukseen henkilökohtaisen tietoavaruuden kanssa. Henkilökohtainen tietoavaruus sisältää kaikki henkilön tiedot tallennettuna eri kokoelmiin, joten tässä suhteessä henkilökohtaista tietoavaruutta voisi kuvailla suureksi varastoksi. Tällaisen varaston näkökulmasta voidaan keskeiseksi operaatioksi määritellä syöte (input), varasto (store) ja tuloste (output). (Jones, 2008, s. 59.) Tätä jakoa on tarkasteltu kuviossa 2, joka on suomennettu Jonesin (2008, s. 60) esittämästä kuvioista.



KUVIO 2 Perinteinen näkökulma tiedonhallintaan (Jones, 2008, s. 60)

Kuviossa kuvataan perinteistä näkökulmaa tiedonhallintaan, jossa henkilökohtainen tietoavaruus on varasto ja henkilökohtaisen tiedonhallinnan toiminnot vaikuttavat varaston syötteisiin, tulosteisiin sekä sisäiseen rakenteeseen. (Jones, 2008, s. 60.)

Yllä mainittua varastoon liitettyä toimintojen jakoa syöte-varasto-tuloste vastaavasti on Jones (2008, s. 59; Jones, 2007, s. 464) ryhmitellyt keskeiset henkilökohtaiseen tiedonhallintaan liittyvät toiminnot seuraavasti:

- Säilyttämisen toiminnot (Keeping activities) vaihtuvat (move) tiedosta tarpeisiin ja vaikuttavat henkilökohtaiseen tietoavaruuteen meneviin syötteisiin (input)
- Löytämisen/uudelleen löytämisen toiminnot (Finding/re-finding activities) vaihtuvat tarpeesta tietoon ja vaikuttavat henkilökohtaisesta tietoavaruudesta tuleviin tiedon tulosteisiin (output)
- Meta-tason toiminnot (Meta-level activities) keskittyvät itse henkilökohtaiseen tietoavaruuteen sekä sen sisällä olevien henkilökohtaisten tietokokoelmien hoitamiseen ja järjestelmiseen. (Jones, 2007, s. 464.)

Jones (2008, s. 59–60) kuitenkin huomauttaa, että varasto näkökulmassa on rajoituksensa suhteessa tiedon käyttämiseen sekä siihen, mitä itse varastossa tapahtuu. Esitelty varasto näkemys kuitenkin antaa pohjan henkilökohtaisen tiedonhallinnan ja sen toimintojen syvemmälle tarkastelulle.

Tarkemmalle henkilökohtaisen tiedonhallinnan toimintojen (personal information management activities) tarkastelulle voidaan ottaa pohjaksi näkemys, jonka mukaan toimintojen avulla pyritään yhdistämään tiedon tarpeita oleelliseen tietoon (Jones 2008, s. 60). Tietojen ja tarpeiden yhdistämien vaikuttaa toimintojen käyttämiseen siten, että tarpeet käynnistävät löytämisen/uudelleen löytämisen toiminnot, kun taas tiedot käynnistävät säilyttämisen toiminnot (Jones & Teevan, 2007, s. 14–15). Seuraavaksi hieman kuvausta eri toiminoista ja niiden tehtävästä.

Löytämisen toiminnot (Finding activities) vastaavat ilmaantuneeseen tarpeeseen. Löytämiseen kuuluvat kaikki toiminnot, joilla selataan ja etsitään tietoa, joka vastaa henkilön tarpeita. Esimerkkinä voidaan ottaa kansiodien ja niiden sisällön tai hakutulost-listan läpi käyminen pyrittäessä löytämään tarpeellinen tieto. Löytäminen sisältää sekä uuden tiedon etsimisen että

toiminnot uudelleen löytämiselle. (Jones, 2008, s. 62–63.) Näitä toimintoja käsitellään lisää luvussa 3.

Säilyttämisen toimintoihin (Keeping activities) kuuluvat henkilön kohtaamaan tietoon liittyvät päätökset sopivimmista toimista, joita tiedon suhteen tehdään. Säilyttämisen toiminnoissa pyritään arvioimaan sitä, mikä löydetty tai kohdattu tieto on tarpeellista ja mitä ei ole tarpeen säilyttää. Päätösten jälkeen tieto joko hylätään tai sitten ryhdytään toimenpiteisiin tiedon säilyttämiseksi tallentamalla se siten, että sen löytää myöhemmin uudelleen sopivaa käyttöä varten. (Jones, 2008, s. 63–64.) Näitä toimintoja käsitellään lisää luvussa 4.

Meta-tason toiminnot (Meta-level activities) ovat laajalti käytössä tietokokoelmissa, ohjaamassa tiedon tarpeiden ja tiedon yhdistämistä näiden kokoelmien osalta (Jones, 2007, s. 478). Meta-tason toiminnot tehdään kun toiset toiminnot on saatu suoritettua. Meta-tason toimintoihin lukeutuvat järjestely, ylläpito, yksityisyyden ja tiedon kulun (flow) hallinta sekä määrittäminen (measuring) ja arvionti. Yleensä meta-tason toiminnoille ei tavallisessa arjessa ole mitään erityistä virikettä, joka käynnistäisi nämä toiminnot. Usein kyseiset toiminnot siirretäänkin myöhemmäksi ja suoritetaan joskus paremmalla ajalla. Toimintoja saatetaan siirtää viikkojenkin päähän. (Jones, 2008, s. 64–66.)

Meta-tason toiminnoista tässä tutkielmassa käsitellään vain järjestämistä ja ylläpitoa. Tämä jako perustuu Boardmanin (2004, s. 16-17) tulkintaan Barreaun (1995) toimintojen jaosta, jossa tarkasteluun otetaan ryhmittelyn neljä ensimmäistä toimintoa. Tällöin käyttäjän suorittamiksi toiminnoiksi tulevat 1. tietoyksiköiden hankkiminen kokoelman muodostamiseksi, 2. tietoyksiköiden tallentaminen ja järjestäminen, 3. kokoelmien ylläpito ja 4. tietojen noutaminen uudelleenkäyttöön. Jonesin (2008, s. 59; Jones, 2007, s. 464) tekemään toimintojen ryhmittelyyn Boardmanin toiminnot yhdistetään siten, että löytämisen toimintoihin liitetään hankkimisen ja noutamisen toiminnot, säilyttämisen toimintoihin sisältyy tiedon varastoimisen, tai tallentamisen, toiminnot ja meta-tason toiminnot jaetaan ylläpidon ja järjestämisen toiminnoiksi. Tässä toimintojen ryhmittelyssä on Boardmanin toimintojen jaottelusta tietojen tallentaminen yhdistetty säilyttämisen toimintojen yhteyteen ja järjestämisen toiminto esitellään erikseen. Toimintoja tarkastellaan lähemmin luvuissa 3-6, joissa käydään useimpien toimintojen osalta läpi myös kyseiseen toimintoon liittyviä yleisimpiä kotikoneilla käytettyjä menetelmiä.

Menetelmä

Tämän tutkielman osalta *menetelmiksi* luetaan ne tavat, joilla henkilökohtaisen tiedonhallinnan toimintoja lähdetään suorittamaan. Toiminnoille esiteltävät menetelmät on pyritty valitsemaan käyttäjän näkökulmasta siten, että valitut menetelmät ovat tavalliselle käyttäjälle tuttuja, yleisiä ja käytettävissä ilman suurempaa tietotekniikan osaamisen vaatimista. Kaikille tässä tutkielmassa käsiteltäville toiminnoille on pyritty valitsemaan esiteltäväksi kaksi kyseisessä toiminnossa tavallisimmin käytettävää menetelmää.

3 LÖYTÄMINEN

Tässä luvussa tarkastellaan Jonesin (2008, s. 59; Jones, 2006, s. 464) toimintojen jakoon kuuluvia löytämisen ja uudelleen löytämisen toimintoja. Näihin toimintoihin on sisällytetty Boardmanin toimintojen jaottelusta kohdat 1 ja 4, joita ovat tiedon hankkiminen ja noutaminen.

”Henkilökohtaisessa tiedonhallinnassa yksi keskeinen peruskyky on *etsiminen* (search), eli tapa, jolla löydetään tietoyksikkö, joka on tallennettu johonkin tietokoneen kovalevyille kaikkien tiedostojen, kansioiden ja kaikenlaisen epämääräisesti nimetyn ”krääsän” keskelle, jota ajan mittaan on kertynyt (Russell & Lawrence, 2007, s. 153).” Tiedon etsimisen (seeking) tutkimusalue on laaja, ja osana tähän tutkimuksen laajaan piiriin sisältyy myös tutkimus löytämisestä (finding) henkilökohtaisen tiedonhallinnan toimintona (Teevan, Capra & Pérez-Quiñones, 2007, s. 23; Jones 2008, s. 81).

Löytämisen toiminnot (Finding activities) vastaavat ilmaantuneeseen tiedon tarpeeseen. Löytäminen sisältää ne toiminnot, joilla selataan ja etsitään tietoa, joka vastaa henkilön tarpeita. (Jones, 2008, s. 62–63.) Tiedon löytämisen (information finding) määrittelemiseen ovat Jones (2007, s. 465; Jones 2008, s. 82) sekä Teevan, Capra ja Pérez-Quiñones (2007, s. 23) soveltaneet Wilsonin (2000) määritelmää tiedon etsimisestä (information seeking). Määritelmän mukaan tiedon etsiminen tarkoittaa ”tiedon tarkoituksellista etsimistä seurauksena tarpeelle täyttää jokin päämäärä. Etsimisen aikana yksilö voi olla vuorovaikutuksessa manuaalisten tietojärjestelmien (kuten sanomalehtien tai kirjaston) tai tietokone-pohjaisten järjestelmien kanssa (kuten World Wide Web) (Jones 2007, s. 465; Jones 2008, s. 82; Teevan ym., 2007, s. 23).”

Jonesin (2007, s. 465) mielestä tämä etsimisen määritelmä soveltuu yhtä hyvin löytämiseen, ja se korostaa löytämisen tarkoituksellista luonnetta, eli löytämistä tiedon tarpeen täyttämiseksi (Jones, 2008, s. 82). Myös Teevanin ym. (2007, s. 24) mielestä määritelmä tiedon etsimisestä sopii hyvin tiedon löytämisen määrittelemiseen, ja heidän mukaansa termejä usein käytetäänkin toistensa vaihtoehtoina (interchangeably). Tosin Teevan kumppaneidensa kanssa tästä huolimatta käyttäisi mieluummin termiä löytäminen.

Tiedon tarpeen täyttäminen tarkoittaa usein käsillä olevaan tehtävään liittyvien tietokokoelmien kokoamista tai uudelleen kokoamista. Silloin löytämisen toimintoa täytyy toistaa siihen asti, kunnes täysi joukko tietoyksiköitä on kerätty. Varsinkin kun kohteena on aikaisemmin koettu tieto, voidaan löytämisen tapahtumia tarkastella neljävaiheisena prosessina, jossa jokaisessa vaiheessa on mahdollisuus epäonnistumiseen. Prosessin vaihteita ovat:

1. Muistaa etsiä.
2. Muistaa yksityiskohtia tiedosta, joka voisi auttaa rajaamaan myöhempää selausta (scan).
3. Tunnistaa halutut tietoyksiköt.
4. Toistaa tarvittaessa, jotta saadaan uudelleen kerättyä joukko tietoyksiköitä, joita tarvitaan täyttämään nykyinen tarve. (Jones, 2007, s. 468.)

Löytäminen sisältää sekä toiminnan uuden tiedon löytämiseen, josta käyttäjällä ei ole aikaisempia muistoja tarvittavasta tiedosta, että toiminnan tiedon *uudelleen löytämiseen* (re-finding) (Jones, 2007, s. 465). Löytämisen toimintoon yhdistyy Boardmanin toiminnoista tiedon hankkiminen. Tiedon hankkimiseen kuuluu tiedon hakeminen omien tiedon kokoelmien ulkopuolelta, kuten web-sivuilta (Boardman, 2004, 18). Boardmanin toiminnoista tiedon noutaminen taas on toimintaa, jossa tietoa etsitään omista tallennetuista kokoelmista. Noutamisessa tietoja etsitään ja haetaan sen varastointipaikasta uudelleen käytettäväksi (Boardman, 2004, 18), ja siten sisältyy uudelleen löytämisen toimintoon. Tarkemmin määriteltynä uudelleen löytäminen on löytämisen prosessi, jossa haetaan aiemmin nähtyä tietoa. Täten uudelleen löytäminen on tavallaan täydentävä toiminto säilyttämiselle. (Teevan, Capra & Pérez-Quiñones, 2007, s. 24.) Tietoja on kerättyä henkilökohtaiseen tietoavaruuteen aikaisempien säilyttämisen toimintojen seurauksena. Tällöin henkilö voi muistaa tiedon olevan olemassa, mutta ei välttämättä muista tarkkaa sijaintia tälle tiedolle. Henkilöllä voi olla jokin suurpiirteinen idea haluttuun tietoon liittyen, jonka pohjalta lähteä etsimään sitä. Kun tarvittavaa tietoa sitten etsitään, joko hakukoneilla tai selailemalla kansioita, voidaan haluttu tietoyksikkö tunnistaa muiden tietojen joukosta ja noutaa käytettäväksi. (Jones, 2007, s. 466–468.) Uudelleen löytäminen keskittyy enemmän noutamaan aikaisemmin koottua ja koettua tietoa, kuin toistamaan aikaisempia löytämisen toimintoja. Yksi uudelleen löytämiseen vaikuttava tekijä on se, miten tietoja on säilytetty ja järjestetty. (Teevan, Capra & Pérez-Quiñones, 2007, s. 24–25, 30.) Tietojen löytämiseen käytettäviä yleisimpiä menetelmiä ovat hierarkkinen navigointi ja haku (search). Seuraavaksi esitellään tarkemmin näitä kahta yleisintä menetelmää. Menetelmien esittelyn pohjalla käytetään niitä koskevia aikaisempia tutkimuksia ja muuta kirjallisuutta.

3.1 Navigoiminen

Hierarkkinen navigointi, tai lyhyesti navigointi, on kaksivaiheinen prosessi, jossa ensimmäisessä vaiheessa käyttäjä manuaalisesti käy läpi organisatorista hierarkiaansa, kunnes saavuttaa halutun tietoyksikön varaston sijainnin, joka voi olla hakemisto tai kansio. Toisessa vaiheessa käyttäjä paikantaa tietoyksikön hakemistosta tai kansioista, joko aktiivisesti lajittelemalla yksiköitä attribuuttien mukaan tai käyttämällä järjestelmän oletuksia. (Bergmanin, Byeth-Marom, Nachmias, Gradovitch & Whittaker, 2008, s. 2.)

Useissa tutkimuksissa (muun muassa Bergman ym., 2008; Boardman & Sasse, 2004; Barreau & Nardi, 1995; Henderson & Srinivasan, 2011) on osoitettu käyttäjien vahvasti suosivan henkilökohtaisten tietojensa saavuttamiseksi navigoimista hakemisen (search) sijaan silloin, kun molemmat menetelmät ovat käytettävissä. Bergmanin, Whittakerin, Sandersonin, Nachmiasin ja Ramamoorthyn (2010) tutkimus osoittaa kansio-pohjaiselle navigoimiselle suurta onnistumisprosenttia ja kohtuullista haku-aikaa, joka heidän mielestään voi osaltaan selittää käyttäjien mieltymystä navigaatioon hakemisen sijasta. Hakemista käytetään lähinnä viimeisenä keinona silloin, kun käyttäjä ei muista tiedostojen sijaintia (Bergman, Whittaker, Sanderson, Nachmias & Ramamoorthy, 2010, s. 2427). Myös Teevan ym. (2007, s. 26) ovat havainneet käyttäjien suosivan tietojen löytämisessä kansiorakenteidensa selailemista mieluummin, kuin hakusana etsintäpalveluiden (search facility) käyttämistä. He tosin käyttävät navigoimisen sijasta termiä suunnistus (orienteeing). Barreau ja Nardi (1995) taas käyttävät navigoimisesta termiä sijainti-pohjainen löytäminen (location-based finding).

Bergman ym. (2010) toteavat, että heidän tutkimuksensa tulokset osoittavat selkeää yhteyttä tietojen noutamisen ja muistin välillä. Onnistuneeseen tiedon löytämiseen vaikuttaa se, muistaako käyttäjä tiedon tarkan sijaintipaikan. Navigoimisen onnistumiseen vaikuttaa muun muassa käyttäjän tuttuus omien henkilökohtaisten tietojensa rakenteeseen. Tuttuuteen vaikuttaa se, että käyttäjä on itse omien tarpeidensa mukaan varastoinut ja järjestellyt tietoja itse luomiinsa kansioihin. Käyttäjät muistavatkin paremmin itse tekemänsä järjestelmät sekä niiden sijainnin, kuin jonkun muun luomat rakenteet. (Bergman ym., 2010, s. 2434–2437.) Tällöin on luonnollista, että käyttäjän muisti vaikuttaa hakemisen lopputulokseen, kuten Bergman ym. (2010) tutkimuksessaan huomasivat.

Tietojen löytäminen navigoimalla edellyttää olemassa olevaa hierarkkista varastoa, josta tietoja voidaan hakea. Tämä tarkoittaa sitä, että käyttäjän on luotava kansioita tai hakemistoja, joihin varastoida tietoyksiköitä myöhempää käyttöä varten. (Bergman ym., 2008, s.2-3.) Hierarkkisista kansiorakenteista lisää luvussa 4.1.

3.2 Haku

Hakeminen on prosessi, jossa käyttäjä aluksi tekee kyselyn, johon hän määrittelee joitakin ominaisuuksia halutulle tietoyksikölle. Kyselyyn sisältyy ainakin yksi haetun tietoyksikön nimeen tai sisältöön liittyvä sana tai tietoon liittyvä attribuutti. Tämän jälkeen hakukone antaa takaisin tulostilan, josta käyttäjä valitsee asiaankuuluvan yksikön. (Bergmanin, Byeth-Marom, Nachmi-as, Gradovitch & Whittaker, 2008, s. 2.)

Hakukoneen käyttöön tiedon löytämiseksi turvaututaan usein vasta, kun käyttäjä ei löydä haluamaansa tietoa navigoimalla tai silloin, kun käyttäjällä ei ole mitään muistikuvaa siitä, missä tieto sijaitsee (Bergman ym., 2008; Bergman ym., 2010). Ihmisten muisti on tunnetusti rajallinen. Kun henkilökohtaisen tiedon varasto kasvaa hiemankin suuremmaksi kuin mitä lyhytkestoinen (short-term) ihmismuisti pystyy hallitsemaan, vaikuttaa varastoidun tiedon onnistuneeseen käyttöön suuresti mahdollisuus hakea ja löytää tietoa, kun sitä tarvitaan. Hakemisen tarve on olennainen suurten henkilökohtaisten tietovarastojen käyttämisessä ja hoitamisessa. Hakutoiminto on tärkeä varsinkin silloin, kun tietoyksiköiden varasto sisältää merkittävää henkilökohtaista tietoa. (Russell & Lawrence, 2007, s. 153, 155.)

Bernard Cole (2005) huomioi kirjoittamassaan artikkelissa, että tietokoneiden kiintolevyjen koon kasvamisen ja saatavilla olevien uusien tietolähteiden vuoksi tietokoneen käyttäjillä tulee olemaan koneillaan nykyistä paljon enemmän tietoa eri muodoissa, mukaan lukien multimedia tiedostot. Tämä johtaa siihen, että on yhä vaikeampaa löytää dokumentteja, sähköpostiviestejä, äänitiedostoja ja muita tietoyksiköitä koneelta. Colen mielestä senhetkiset työpöytä-pohjaiset hakumahdollisuudet (capabilities) eivät olleet riittäviä tällaisen haasteen selvittämiseen. Koska työpöytä-pohjaiset (desktop-based) dokumentit yleensä ovat rakenteeltaan erilaisia kuin Internetissä olevat, tarvittaisiin uusia lähestymistapoja hakemiseen. (Cole, 2005, S. 14.) Muutamia vuosia sitten esiteltyjen web-pohjaisten ja työpöytä-pohjaisten kokoteksti hakemistojen myötä alkoi suuri muutos hakukäyttöliittymissä. Ne toivat mukanaan nopeamman ja yksinkertaisemman tavan tehdä kyselyjä laajaan epämuodolliseen aineistoon. Avainsana-pohjaisen haun tilalle yleistyi nopeasti vähemmän muodollinen tekstipohjainen haku. (Russell & Lawrence, 2007, s.159.)

Pätevien hakutyökalujen kanssa, jotka voivat hakea tietoyksiköitä koko koneelta, käyttäjä voi hakea tietoa koko hierarkkisen rakenteen läpi. Yleensä järjestelmiin ja sovelluksiin on sisällytetty jonkinlainen hakutyökalu perustoinnoksi. (Russell & Lawrence, 2007, s. 155.) Työpöytähakukoneet luetteloivat ja etsivät tiedostoja henkilökohtaiselta tietokoneelta. Ne hakevat viittauksia tietokoneen kiintolevyllä oleviin tiedostoihin avainsanojen, tiedostotyyppien tai määrättyjen kansioden perusteella. (Lu, Shukla, Subramanya & Wu, 2007, s. 110.) Käyttöjärjestelmiin, sähköpostiohjelmiin ja muihin sovelluksiin rakennetut työpöytähaku ominaisuudet ovat vähemmän tehokkaita kuin web-hakukoneet. Ne tarjoavat yleensä vain yksinkertaisia avainsanahakuja tiedostojoukosta ja

yleensä vain yhdestä tiedostotyypistä. Microsoft Wordin ja muiden sovellusten useilla eri tavoilla muodostamat erityyppiset dokumentit tekevät tietokoneen työpöydästä monimutkaisemmin etsittävän kohteen. Työpöytähakukoneet joutuvat tunnistamaan tiedostoja useiden eri tiedostotyyppien joukosta. Lisäksi niiden pitää pyrkiä olemaan tehokkaita aiheuttamatta kuitenkaan huomattavaa prosessointi- tai muistikuormitusta koneelle. (Cole, 2005, S. 14.)

”Web-haussa nimensä mukaan tietoa etsitään World Wide Webbissä oleviin sisältöihin kohdistetussa hakutoimenpiteessä (Russell & Lawrence, 2007, s.160).” Internetissä hakukoneet voivat käyttää hyväkseen tietoa on, joka yhtäläisessä standardoidussa esitysmuodossa ja järjestetty yleiseen HTML formaattiin. Hakukoneet pystyvät käyttämään tätä tietoa sekä linkkejä muihin dokumentteihin ja parantamaan mahdollisuuksia oleellisten hakutulosten palauttamiseen. (Cole, 2005, s. 14.)

Viime vuosien aikana ovat useat yritykset tuoneet markkinoille yhdistettyjä työpöytä- ja web-hakukoneita. (Cole, 2005, s. 14) Yksi tällainen on Google Desktop Search (GDS), joka on integroitu työpöytä- ja web-hakukone, joka tekee jokaisesta selaimella tehdystä hausta tavallisen osan työpöytähakua. Kun koneelle aktivoidaan GDS, jokaiseen Google-hakuun annetaan mukaan paikalliselta työpöydältä lisättyjä tuloksia hakutulosten sekaan. (Russell & Lawrence, 2007, s.160; Cole 2005, s. 14, 15.) Tällöin käyttäjä voi huomioida hakemaansa aihetta koskevia tietoja, joita hänellä on jo koneellaan, mutta joiden olemassa oloa ei välttämättä muista.

4 SÄILYTTÄMINEN

Tässä luvussa tarkastellaan Jonesin (2008, s. 59; Jones, 2007, s. 464) toimintojen ryhmittelystä säilyttämisen toimintoja, joihin on sisällytetty Boardmanin toimintojen jaottelusta kohta 2, josta tähän ryhmään otetaan mukaan tiedon tallentaminen. Kohtaan 2 kuuluvaa toista toimintoa, tiedon järjestämistä, käsitellään erikseen luvussa 6.

Säilyttämisen (keeping) ja *järjestämisen toiminnot* (organizing activities) kuuluvat yhteen, mutta ovat kuitenkin keskenään erilaisia. Keskeisinpänerona näiden kahden välillä on se, että säilyttäminen keskittyy enemmän itse tietoon ja järjestäminen taas keskittyy tietokokoelmiin. Esimerkiksi tiedoston laittaminen kansioon on säilyttämistä, kun taas järjestämisessä päätetään miten kansiot luodaan ja nimetään tai se miten kansiot liittyvät toisiinsa. (Jones, 2008, s. 125) Säilyttämisen ja järjestämisen toimintoja käytetään usein rinnakkain. Tallentamisen yhteydessä tiedot saatetaan samalla kertaa järjestää niille sopivaan paikkaan.

Jones (2008, s. 125) kuvaa säilyttämisen olevan tehtyjä päätöksiä ja toimia, joilla yhdistetään tämänhetkinen tieto (saatavilla oleva tieto tai harkinnan alaisena oleva tieto) ennakoituihin tarpeisiin. Päätöksiin voi sisältyä vaihtoehtoina 1. "jätä huomiotta, tällä ei ole merkitystä minulle", 2. "jätä huomiotta, voin palata tähän myöhemmin", ja 3. "säilytä tämä erityisessä paikassa tai erityisellä tavalla siten, että voin olla varma, että käytän tätä tietoa myöhemmin."

Kohtaamme joka päivä paljon tietoa joko sattumalta tai sitten tietoisesti etsimällä sitä, ja meidän on tehtävä päätöksiä siitä, miten näitä tietoja käsitellään. Säilyttämisen toimintoon kuuluvat henkilön kohtaamaan tietoon liittyvät päätökset sopivista toimista, joita tiedon suhteen tehdään. Säilyttämisen toiminnoissa pyritään arvioimaan sitä, mikä löydetty tai kohdattu tieto on tarpeellista, ja mitä ei ole tarpeen säilyttää. Päätösten jälkeen tieto joko hylätään tai sitten ryhdytään toimenpiteisiin tiedon säilyttämiseksi tallentamalla se siten, että sen löytää myöhemmin uudelleen sopivaa käyttöä varten. (Jones, 2008, s. 63–64.) Säilytettäväksi tulevaan tietoon liittyy päätöset siitä, missä ja miten tietoa säilytetään, kuten missä muodossa tieto tallennetaan

ja millä koneella, laitteella tai tallennusvälineellä se pidetään. Näin ollen voidaan sanoa, että säilyttämiseen liittyvät päätöset ovat monitahoisia. (Jones, 2007, s. 473.)

Ennen levytilat olivat pieniä ja tiedon säilyttämisestä aiheutui kustannuksia, joten oli tarkasti mietittävä, mitkä tiedot olivat tarpeellisia säilytettäväksi. Nykyään levytilat ovat suurempia ja muutenkin tiedon säilyttämisen mahdollisuudet ovat lisääntyneet. Teknologian kehittyminen lähivuosien aikana on aiheuttanut sen, että säilyttämiseen liittyvät kustannukset ovat suuresti alentuneet tai joissain tapauksissa jopa kadonneet, varsinkin säilyttämisen prosessin yhteydessä tapahtuneiden virheiden suhteen. Tämä tarkoittaa, että tieto, joka osoittautuu tarpeettomaksi ei ole ongelma alentuneiden säilytyksen kustannusten vuoksi sekä sen vuoksi, että lajittelulla, haulilla ja suodattamalla pystyy erottelamaan tarpeellisen tiedon tarpeettomien joukosta. Myöskään ohitettu tieto, jota ei säilytetty, ja joka myöhemmin osoittautuu hyödylliseksi, ei aiheuta suuria ongelmia. Tämä johtuu siitä, että meillä on yhä kasvava määrä tietoa saatavilla Internetissä helposti haettavassa muodossa. (Jones, 2007, s. 476.) Seuraavaksi perehdytään pariin säilyttämiseen kuuluvaan menetelmään.

4.1 Hierarkkinen kansiorakenne

Digitaalisten tietojen käsittelemisessä yleisimmin käytetty hallintajärjestelmä on hierarkkinen kansiorakenne (Ma & Wiedenbeck, 2009, s. 3745). Useimmat ihmiset varastoivat tietojansa tietokoneen käyttöjärjestelmän tarjoamaan hierarkkiseen kansiojärjestelmään (file system) ja käsittelevät tiedostojansa hierarkkisen kansio-selaimen (file browser) kautta (Henderson & Srinivasan, 2011, s. 1). Käyttäjät varastoivat tietoyksiköitä kansiorakenteen eri tasoille niiden luokittelun mukaan. Kansiorakenne koostuu juurikansiosta, jonka alle sijoitetaan eri tiedostoja sekä muita kansioita, joiden alle taas voidaan sijoittaa lisää alikansioita ja tiedostoja. Hierarkiassa olevaan yläkansioon voidaan sijoittaa monta alikansiota, mutta yksittäinen alikansio voi kuulua vain yhteen yläkansioon kerrallaan (Ma & Wiedenbeck, 2009, s. 3746). Käyttöjärjestelmässä on valmiina joitakin oletuskansioita, kuten työpöytä, ja käyttäjäkohtaiset omat tiedostot tai omat kuvatiedostot -kansiot. Osa käyttäjistä tallentaa tiedostojansa näihin valmiina oleviin oletuskansioihin. Toiset käyttäjät taas mieluummin luovat omia kansiohierarkkioita omien kriteereidensä ja tarpeidensa mukaan sovellusten tarjoamien oletussijaintien sijasta. (Bergman ym., 2010, s. 2430.) Henderson ja Srinivasan (2011, s. 7) totesivat, että käyttäjät luovat kansioita joko tietoyksiköiden järjestämisen yhteydessä tai tallentamisen yhteydessä, jos tietoyksikkö ei sovi mihinkään olemassa olevaan kategoriaan. He huomasiivat myös, että osa käyttäjistä luo kansiorakenteita siksi, että ne tarjoavat yleiskuvan omiin tiedostoihin ja niiden rakenteeseen (Mts. 7). Kansiorakenteen selaaminen antaa kuvan siitä, mitä tietoa on saatavilla ja miten eri tiedot liittyvät toisiinsa (Henderson, 2005, s. 76). Hendersonin (2005, s. 76) mukaan kansiorakenteen

muodostaminen vastaa attribuuttien ja avainsanojen määrittämistä dokumenttiin. Esimerkiksi pelkästään tiedoston sijoittaminen tiettyyn kansioon lisää siihen ominaisuuksia ja tietoja. Kansion nimi antaa tiedostolle asiayhteyden ja nimen tai tiedostoformaatin mukaan se erotetaan muista samassa asiayhteydessä olevista tiedostoista. (Mts. 76.)

Tietoyksiköiden laittaminen kansioiden puurakenteeseen on prosessi, joka on erityisen herkkä valinnoille, joita tehdään kun mennään alaspäin hierarkiassa. Virhe kategorian valinnassa voi johtaa siihen, että tietoyksikkö on kaukana sen "oikeasta" sijaintipaikasta. Tämä virhe voi johtaa siihen, että tietoa ei löydetä, kun sitä myöhemmin etsitään käyttöä varten. (Russell & Lawrence, 2007, s. 155.) Tämän vuoksi säilyttämisen päätökset siitä, mitä säilytetään ja mihin tiedot sijoitetaan ovat erityisen tärkeitä.

Tallentaessaan tietoja koneelle, käyttäjät kokoavat eri tietotyypeistä omia kokoelmiaan. Tämän vuoksi on tietokoneella tyypillisesti useita eri tietoja koskevia käyttäjän luomia hierarkkioita. Näihin tietokokoelmien hierarkkioihin kuuluvat tiedostot, sähköposti ja kirjanmerkit. (Boardman, 2001, s. 403.) Tiedostokokoelmaan kuuluvat käyttäjän kansiojärjestelmään tallentamat tiedostot, joihin sisältyvät eri muodoissa tallennetut tekstidokumentit, musiikki ja kuvat. Sähköpostikokoelmaan kuuluvat käyttäjän sähköpostisovellukseen varastoidut viestit ja niiden liitetiedostot. Kirjanmerkkikokoelmat sisältävät käyttäjän Internet-selaimeen tallentamat linkit ja suosikit-listan. (Boardman, 2004, s. 68.) Kirjanmerkkejä käytetään henkilökohtaisena Internet-tietoavaruutena, joka auttaa ihmisiä muistamaan ja hakemaan mielenkiintoisia web-sivuja. Kirjanmerkit ovat yksinkertainen väline henkilökohtaisten tiedon osajoukkojen rakentamiseen, jonne kiinnostavia tai hyödyllisiä web-sivuja voidaan varastoida myöhempää käyttöä varten. Kirjanmerkit toimivat kätevinä oikopolkuina usein käytettyihin sivustoihin sekä eräänlaisena osoittimena hyödylliseen tietoon, joka muuten voisi tulla unohdetuksi. (Abrams, Baecker & Chignell, 1998, s. 41.)

Kansiohierarkian rakenteen syvyys vaihtelee käyttäjien järjestämisen tapojen ja tarpeiden mukaan. Hierarkian syvyydellä tarkoitetaan sitä, paljonko juurikansion alapuolella on alikansioiden tasoja. Usein tutkimuksissa juurikansiolle annetaan arvo 0, alikansiolle arvo 1 ja sen alikansiolle arvo 2 jne.. Sähköpostikansioiden osalta ovat Ducheneaut ja Bellotti (2001, s. 33) huomanneet hierarkioiden olevan melko matalia ja yleensä syvyydeltään kaksi tasoa. Boardmanilla (2004, s. 82) keskimääräinen syvyys sähköpostikansioille oli yksi tasoa, tiedostoille kolme tasoa ja kirjanmerkeille yksi tasoa.

4.2 Sähköposti tiedon säilytyspaikkana

Sähköpostia käytetään yhä useammin myös muihin tarkoituksiin, kuin mihin se on alun perin kehitetty, eli yhteydenpitoon henkilöiden välillä. Whittaker ja Sidner (1996, s. 276) nimittävät tätä käytäntöä sähköpostin ylikuormitukseksi. Ducheneaut ja Bellotti (2001, s. 30) toteavat joka päivä sähköpostin välityksellä

lähetettyjen viestien ja liitetiedostojen huomattavan määrän johtaneen siihen, että käyttäjät ovat omaksuneet sähköpostin henkilökohtaisen tiedonhallinnan työkaluksi. Tutkimuksissa on osoitettu, että ihmiset usein käyttävät sähköpostiansa tapana, jolla he tallentavat (capture) henkilökohtaista tietoansa. Syynä tähän on se, että sähköposti on helposti saatavilla, se on kaikkialla läsnä oleva (ubiquitous) ja useimmissa sähköpostijärjestelmissä on kohtuulliset hakupalvelut saatavilla, jotka mahdollistavat sen, että käyttäjä voi hakea, suodattaa (filtering) ja lajitella tietoa eri ominaisuuksilla. Tämän lisäksi sähköposti on myös usein säilytyspaikka sisällölle, sillä sähköpostin liitteet sallivat käyttäjän yhdistää viestin sisältöön erilaisia tiedostoja, kuten valokuvia, tekstidokumentteja ym.. (Russell & Lawrence, 2007, s. 157–158.) Whittaker, Bellotti ja Cwizdka (2006) mainitsevat kolme henkilökohtaisen tiedonhallinnan tehtävää, joihin sähköpostia useimmin käytetään: tehtävienhallinta, henkilökohtainen arkistointi ja yhteyksienhallinta (contact management).

Whittaker ym. (2006) toteavat sähköpostin roolin tiedonvälittäjänä johtavan useimmilla meistä siihen, että käytämme sähköpostilaatikkaa tehtävienhallintaan. Tehtävienhallintaan kuuluu itsemme muistuttaminen nykyisistä tehtävistä ja tehtävien tilanteen seuraaminen sekä oleellisten tietojen hoitaminen. Jotkut jopa lähettävät postilaatikkoonsa itsellensä viestejä muistutukseksi ja saattavat lisätä niihin linkkeinä tarpeellista tietoa. Koska sähköposti on yleisin tiedonvälitys kanavamme, on yleistä käyttää sitä myös yhteystietojemme varastointiin. Tehtävienhallinnan ja yhteystietojen keräämisen lisäksi sähköpostia käytetään henkilökohtaisen tiedon arkistointiin. Sähköposteissa tulleet liitetiedostot tai tiedot valmistuneista tehtävistä usein päätyvät sähköpostikansioihin talteen myöhempiä käyttöä varten. (Whittaker, Bellotti & Cwizdka, 2006, s. 70; Whittaker & Sidner, 1996, s. 276.) Osa käyttäjistä järjestee viestejä aktiivisiin kansioihin senhetkisten tehtävien mukaan ja palaa niihin tarpeen mukaan hakemaan tietoja. (Whittaker, Bellotti & Cwizdka, 2006, s. 70) Jotkin käyttäjien luomista kansioista taas ovat käyneet tarpeettomiksi, eivätkä ne enää ole käytössä (Ducheneaut & Bellotti 2001, s. 33). Ducheneaut ja Bellotti (2001, s. 33) huomasivat sähköpostikansioiden yleisimmiksi järjestämiskriteereiksi lähettäjän, organisaation, projektin ja henkilökohtaisen mielenkiinnon kohteen.

Käyttäjillä on sähköpostissaan kerättyinä viestejä, jotka sisältävät monentyyppistä tietoa. Suoritettavia tehtäviä sisältävät viestit edellyttävät käyttäjiltä erilaisia toimenpiteitä. Nämä ovat tehtävienhallintaan liittyviä viestejä, jotka toimivat postilaatikossa ollessaan muistutuksena keskeneräisistä tehtävistä. Osa viesteistä sisältää pitkiä tekstejä tai dokumentteja, jotka käyttäjä on ehtinyt vain osittain lukea tai ne ovat vielä kokonaan lukematta. Sähköpostiin saapuu myös viestejä, joiden merkitystä ja tarpeellisuutta käyttäjä ei heti osaa arvioida ja siksi lykkää niiden käsittelemistä myöhemmäksi. Postilaatikkoon sisältyy myös keskeneräisiä yhteydenpito viestejä, joihin käyttäjä ei vielä ole ehtinyt vastata. (Whittaker & Sidner, 1996, s. 278–279.)

5 JÄRJESTÄMINEN

Tässä luvussa tarkastellaan Jonesin (2008, s. 59; Jones, 2007, s. 464) toimintojen ryhmittelyn meta-tason toiminnoista järjestelemisen toimintoa. Boardmanin toimintojen jaottelusta tämä vastaa kohtaa 2, josta tässä luvussa käsitellään toiminnoista jälkimmäinen eli tiedon järjestäminen.

Kuten jo edellisen luvun alussa tuli lyhyesti mainittua, on järjestäminen yhteydessä säilyttämiseen, mutta kuitenkin eroaa tietyssä suhteessa säilyttämisestä. Järjestämisen suhteen tämä ero oli se, että järjestäminen keskittyy enemmänkin tietokokoelmiin, kun taas säilyttäminen keskittyy itse tietoon. (Jones, 2008, s. 124–125.) Järjestäminen ja säilyttäminen voidaan joskus suorittaa samanaikaisesti, jolloin tiedot järjestellään sopivaan paikkaan jo tallentamisen yhteydessä. Joissakin tapauksissa tiedot vain laitetaan talteen johonkin ja tietojen tarkempi järjestäminen voidaan siirtää hoidettavaksi myöhemmin.

Paremmiin määriteltynä "*järjestämistä* (organizing) on se kun tehdään päätöksiä siitä, kuinka kansioita tulisi luoda, nimetä ja kuinka ne liittyvät toisiinsa (Jones, 2008, s. 125)." Järjestelemisen keskeisimpiä pyrkimyksiä on kehittää tietokokoelmaa. Järjestelemisen tavoitteena on yhdistää kokoelmien sisältämät tietoyksiköt ennakoituihin tarpeisiin. Tämän tavoitteen toteutumiseen liittyy kokoelmia, ja sen tietoja, koskevia erilaisia päätöksiä. Päätöksiin kuuluu esimerkiksi se, kuinka kokoelmassa olevat tiedot tulisi nimetä tai mitkä ominaisuudet ovat järkeviä ja auttavat erottamaan kokoelman tietoyksiköt toisistaan, miten tietoyksiköt tässä kokoelmassa tulisi ryhmitellä sekä mihin kasaan tai kansioon ne tulisi sijoittaa. (Jones, 2008, s. 125–126.) Seuraavissa menetelmissä on esiteltyä joitakin niitä keinoja joilla näitä päätöksiä voidaan toteuttaa.

5.1 Tietojen sijoittelu tapoja

Aikaisemmissa tutkimuksissa on henkilökohtaiseen tiedonhallintaan luokiteltu kaksi tyypillisintä strategiaa, joita ovat arkistointi (filing) ja kasaaminen

(piling) (Hardof-Jaffe, Hershkovitz, Abu-Kishk, Bergman & Nachmias, 2009, s. 252). Malone (1983), tutkiessaan toimistojen työpöytien järjestämistä, havaitsi arkistoinen ja kasaamisen yleisimmiksi tavoiksi järjestää paperi dokumentteja. Hänen mukaansa molemmat tavat ovat keinoja kerätä elementtejä suuremmiksi yksiköiksi. Mapit ovat yksiköitä, joissa elementit (kuten yksittäiset kansiot) on selkeästi nimetty ja järjestetty johonkin systemaattiseen järjestykseen. Joissain tapauksissa ryhmät itsessään on selkeästi nimetty ja järjestetty systemaattisesti, kuten esimerkiksi koko arkistokaappi. Kasoissa yksittäiset elementit (kuten paperit, kansiot jne.) eivät välttämättä ole nimettyjä ja ne eivät yleensä ole järjestettyinä mihinkään tiettyyn järjestykseen. (Malone, 1983, s. 105–106.)

Vaikka Malone käsitteli paperisten dokumenttien järjestämistä toimiston työpöydällä, voidaan näitä arkistoinen ja kasaamisen strategioita käyttää myös digitaalisten tietojen järjestämisen strategioina. Nykyisin, kun suurin osa tiedoista on digitaalisessa muodossa, arkistoinen tapahtuu sijoittamalla tiedostot hakemistoihin tai kansioihin nimiöiden (label) kanssa, jotka viittaavat niiden kategoriaan. Kasaamisessa taas kerätään tietoyksiköitä johonkin juurikansioon, kuten sähköpostiviestien kohdalla keräämällä niitä postilaatikkoon tai tiedostojen käsittelemisen yhteydessä vaikka omat tiedostot -kansioon. (Hardof-Jaffe ym., 2009, s. 252.) Kasaamisen ongelmana on, että se vaikeuttaa tietojen löytämistä. Bergman ym. (2010, s. 2434) huomasivat tutkimuksessaan, että mitä enemmän kansiossa on tietoyksiköitä sitä kauemmin kului aikaa oikean tiedoston löytämiseen.

Keräämiensä tulosten pohjalta, ja aikaisempien luokittelujen pohjalta, Hardof-Jaffe ym. (2009, s. 254-256) esittävät oman henkilökohtaisen tietoavaruuden järjestämisen strategioita koskevan luokittelun, jonka strategioina ovat kasaaminen, yhden kansion arkistoinen (one-folder filing), pienten kansioiden arkistoinen (small-folder filing) ja suuren kansion arkistoinen (big-folder filing). Kasaamisen luokassa suurin osa tiedostoista on sijoitettuna juurikansiossa. Yhden kansion arkistoinen luokassa suurin osa tiedostoista sijaitsee yhdessä kansiossa juurikansion alapuolella. Pienten kansioiden arkistoinen luokassa tietoyksiköt on jaoteltu useisiin kooltaan suhteellisen pieniin kansioihin, joissa on keksimäärin noin 6 tiedostoa kansiota kohden. Suuren kansion arkistoinen luokassa tietoyksiköt on jaoteltu kansioihin niin, että noin puolet niistä on sijoitettuna yhdessä suuressa kansiossa. Tässä luokassa on keskimäärin noin 23 tiedostoa kansiota kohden.

Hendersonilla ja Srinivasanilla (2011) strategiona ovat kasaaminen, arkistoinen ja jäsentäminen (structuring). Jäsentämisessä on samoja piirteitä kuin arkistoinisessa, mutta tässä kohdassa kansiorakenteet ovat käyttäjille muutakin kuin pelkkä keino löytää tietojansa. Kansiorakenne tarjoaa käyttäjille kuvan tietojensa rakenteesta ja käyttäjät luovat kansioita etukäteen paikaksi, johon sitten laittaa tiedostoja, mikä tekee luokittelusta helpompaa. (Mts. 7-8.)

Tiedostojen sijoittaminen toimii myös tärkeänä muistamisen apukeinona. Tutkimuksissa on huomattu käyttäjien sijoittavan tiedostoja paikkoihin, josta ne tulevat varmimmin huomatuksi. Sähköpostissa viestejä jätetään postilaatikkoon muistuttamaan tapaamisesta tai keskeneräisistä tehtävistä ja koneella tiedostoja

sijoitetaan kansiohierarkian ylemmille tasoille muistuttamaan tiedostojen sisältämistä keskeneräisistä töistä tai muuten tärkeistä tiedoista. Esimerkkinä voidaan ottaa vaikka poistettaviksi tarkoitettujen tiedostojen sijoittaminen työpöydällä roskakorin läheisyyteen tai tärkeitä tietoja sisältävän tiedoston sijoittaminen työpöydälle kohtaan, josta sen varmasti huomaa. (Barreau & Nardi, 1995, s. 41.)

Ihmisillä on eri lähestymistapoja ja tottumuksia tietojen järjestämisen suhteen. Näitä järjestämisen käyttäytymisen tapoja, tai stretegioita, on monissa aikaisemmissa tutkimuksissa ryhmitelty eri luokitteluihin. Seuraavaksi esitellään taulukossa 1 käännettynä Hendersonin ja Srinivasanan (2011, s.2) aikaisemmista tutkimuksista kokoamia luokitteluja eri tietotyypeille.

TAULUKKO 1 Aikaisemmissa tutkimuksissa havaitut järjestämisen strategiat. (Henderson & Srinivasan, 2011, s. 2)

Tutkimus	Tietotyyppi	Luokittelu
Malone (1983)	Paperi dokumentit	Siisti ja sotkuinen
Mackay (1988)	Sähköposti	Etusijalle laittaja, arkistoija, pyytäjä ja vastaaja
Whittaker & Sidner (1996)	Sähköposti	Arkistoimaton, toistuvasti arkistoiva, kevät siivooja
Bälter (1997)	Sähköposti	Kansioton siivooja, kansioton kevät siivooja, siivooja, kevät siivooja
Gwizdka (2004)	Sähköposti	Siivooja, säilyttäjä
Abrams, Baecker & Chignell (1998)	web-kirjanmerkit	Arkistoimaton, luomisen aikainen arkistoija, istunnon jälkeinen arkistoija, satunnainen arkistoija
Boardman & Sasse (2004)	Dokumentit, sähköposti ja web-kirjanmerkit	Järjestämissielinen, järjestämisneutraali

Malone (1983) kutsuu kahta toimiston järjestämisen tapaa siistiksi (neat) ja sotkuiseksi (messy). Siistissä toimistossa käytetään tietojen järjestämistä kansioihin erilaisten kategorioiden mukaan ja myös paperi pinot on luokiteltu ja järjestetty. Sotkuinen toimisto taas on täytetty sekalaisilla paperipinoilla, jotka on huonosti määritelty ja vailla mitään erityistä luokittelu kerätty kasalle. (s. 104.)

Mackayn (1988) luokitteluun kuuluvat etusijalle laittajat (prioritizers), arkistoijat (archivers), pyytäjät (requesters) ja vastaajat (responders). Etusijalle laittajat mieluusti tunnistavat viestejä eri kriteereiden avulla, jotta voisivat optimaalisesti järjestää töitensä. He myös suosivat toimintoja, jotka eivät vaadi heidän osallistumistaan, kuten automaattisesti lisättävät tapaamiset verkko-

kalenteriin. Tässä luokassa käyttäjät suorittavat tärkeimmät tehtävät ensin ja jättävät huomiotta tehtävät, jotka eivät ole tärkeitä. Arkistojat pääsääntöisesti lukevat kaikki viestit ennenkuin lajittelevat ne johonkin arkistoon myöhempää käyttöä varten. Pyytäjät delegoivat tehtäviä muiden hoidettavaksi ja vastaajat lähinnä ottavat vastaan muiden lähettämiä pyyntöjä, eikä heillä ole ketään jolle delegoida tehtäviä. (Mackay, 1988, s. 350–351.)

Whittaker ja Sidner (1996) tunnistivat kolme sähköpostissa käytettyä strategiaa, jotka oli muodostettu kahden kriteerin pohjalta, 1) oliko käyttäjällä käytössään kansioita ja 2) siivosivatko käyttäjät postilaatikkoansa päivittäin. Näistä muodostui strategiat:

- Arkistoimaton (no-filers): käyttäjä ei käytä kansioita ja luottaa kokotekstihakuun tiedon löytämiseksi. Koska viestejä ei järjestellä kansioihin, postilaatikko on kooltaan huomattavan suuri. Osa viesteistä on monta kuukautta vanhoja ja viestien tulvaa pyritään hallitsemaan ajoittaisella puhdistuksella.
- Toistuvasti arkistoiva (frequent filers): käyttäjät käyttävät kansioita viestien järjestämiseen ja pyrkivät siivoamaan postilaatikkoansa päivittäin. Yritetään ahkerasti minimoida postilaatikossa olevien viestien määrää päivittäisellä arkistomisella ja puhdistamisella. Tästä syystä postilaatikon koko on verrattain pieni ja koostuu lähinnä uusista viesteistä.
- Kevät siivoaja (spring-cleaners): käyttäjät käyttävät kansioita, mutta siivoavat postilaatikkoansa vain ajoittain. Kansioita käytetään paljon, mutta huonolla menestyksellä. Postilaatikossa on paljon viestejä, lähinnä yhteydenpitoon liittyviä, ja useimmat viesteistä ovat yli kolme kuukautta vanhoja. (Mts. 280.)

Whittakerin ja Sidnerin luokittelua on laajentanut Bälter (1997) tutkimuksessaan "Strategies for organising email". Hän jakoi arkistoimattomien luokan kansioton siivoaja -luokkaan (folderless cleaners) ja kansioton kevät siivoaja -luokkaan (folderless spring-cleaners) sen mukaan, kuinka usein käyttäjät poistivat tietoja postilaatikostaan. Muita luokitteluja olivat siivoaja (cleaners) ja kevät siivoaja (spring-cleaners). (Henderson & Srinivasan, 2011, s. 2.)

Gwizdka (2004) tunnisti kaksi lähestymistapaa: siivoajan (cleaners) ja säilyttäjän (keepers). Siivoojilla on tietyt ajat sähköpostien hoitamiseen, eivätkä he pidä sähköpostissaan tapahtumia tai suoritettavien tehtävien tietoja. Säilyttäjät lukevat sähköpostia jatkuvasti ja antavat sähköpostin keskeyttää työskentelynsä. He pitävät tapauhtumia ja suoritettavien tehtävien tietoja sekä hakevat tietoja sähköposti arkistojensa keskeltä. (Mts. 1237.)

Abrams ym. (1998, s. 44) toteavat, että käyttäjät järjestelivät kirjanmerkkien arkistojaan estääkseen epäjärjestyksen ja lisätäkseen rakennetta. He luokittelevat kirjanmerkkien järjestämisen tavoiksi:

- Arkistoimaton (no-filer): käyttäjät, jotka eivät koskaan järjestele kirjanmerkkejä. Heidän kokoelmissaan kirjanmerkit pysyvät siinä järjestyksessä mihin ne on luotu.
- Luomisen aikainen arkistoiija (creation-time filer): käyttäjät, jotka varastoivat uusia kirjanmerkkejä sopiviin kategorioihin silloin kun ne ensikeran luodaan. Nämä käyttäjät mielellään välttävät järjestelemättömien kirjanmerkkien kasaantumista.
- Istunnon jälkeinen arkistoiija (end-of-session filer): käyttäjät, jotka järjestelevät kaikki uudet kirjanmerkit istunnon lopuksi. He varaavat erikseen aikaa järjestämiseen istunnon loppuun.
- Satunnainen arkistoiija (sporadic filer): käyttäjät, jotka järjestävät kirjanmerkkejä satunnaisesti ja aikatauluttavat siivoamis istuntoja. Tämän ryhmän käyttäjät valitsevat, milloin optimoivat työympäristönsä. (Abrams ym., 1998, s. 45.)

Boardman ja Sasse (2004) havaitsivat tutkimuksessaan, että suurin osa vastaajista ei kuulu mihinkään tiettyyn aikaisemmin esiteltyyn järjestämisen tavoista tehtyyn luokitteluun, vaan he kehittivät oman luokittelun, jossa huomioidaan eri kokoelmien välillä järjestämisen suhteen kahdenlaisia taipumuksia, joita he kutsuvat järjestämismieliseksi (pro-organizing) ja järjestämisneutraaliksi (organizing neutral). Järjestämismieliset käyttäjät pyrkivät pitämään kokoelmiaan järjestyksessä ja heillä on laajoja kansiorakenteita käytössään. Tämä taipumus koskee lähinnä tiedostokokoelmien järjestämisestä. Järjestämisneutraalit käyttäjät taas eivät koe järjestämistä niin tärkeäksi tehtäväksi. Tämä käyttäytymismalli koskee suurimmalla osasta käyttäjistä lähinnä kirjanmerkkikokoelmia ja osittain myös sähköpostikokoelmia. (Mts. 587-588.)

Taulukossa lueteltujen luokitteluiden lisäksi ovat Hardof-Jaffe ym. (2009, s. 252) uusien kansioiden luomisaikaan liityen määritelleet ryhmät ennakkorakentajat (pre builder) ja jälkirakentajat (post builder). Ennakkorakentajat luovat uusia kansioita ennenkuin heillä on luotuna mitään tietoyksiköitä, joita sijoittaa sinne. Jälkirakentajat luovat mieluummin kansioita vasta, kun joukko uusia tietoyksiköitä on kerättynä. (Mts. 252.)

Boardman (2004, s. 87) huomasi, että käyttäjät eivät noudata samoja järjestämisen strategioita kaikkien kokoelmiensa suhteen. Vaikka suurin osa käyttäjistä järjesteli aktiivisesti tiedostokokoelmiaan, niin kirjanmerkkikokoelmien suhteen ne käyttäjät, jotka ylipäätään keräsivät kirjanmerkkejä, harvoin järjestelivät niitä millään tavalla. Lisäksi havaittiin, että kaikki käyttäjät eivät sopineet aikaisempiin luokitteluihin, vaan he käyttivät monia strategioita yhdessä, jättäen osan tiedoista arkistoimatta ja taas kuitenkin järjestämällä toisia. (Boardman, 2004, s. 89-91.) Sähköpostissa viestit usein jätetään postilaatikkoon, koska niiden luokittelu ja järjestäminen kansioihin koetaan hankalaksi. Toinen syy viestien kasaantumiseen on se, että käyttäjät pelkäävät etteivät löydä viestejä kansioista, koska eivät muista mihin kansioon viesti on laitettu. (Whittaker & Sidner, 1996, s. 279-280.)

5.2 Luokittelu, lajittelu ja nimeäminen

Järjestämisen keskeisimpiin päätöksiin kuuluu se, kuinka kansioita tulisi luoda, nimetä ja kuinka ne liittyvät toisiinsa. Näihin päätöksiin liittyvät menetelmät ovat lajitteleminen, luokitteleminen ja nimeäminen, joita seuraavaksi käydään lyhyesti läpi. Materiaalien lajittelemisessa käyttäjä valitsee aiheeseen liittyvän kansion tai kansiot jokaiselle uudelle tietoyksikölle ja tällä tavoin määrittää tietoyksiköiden väliset suhteet käyttämällä hierarkiaa. Yhdistämällä monia tietoyksiköitä yhteen kokonaisuuteen, joilla muodostetaan hierarkkista rakennetta, voidaan rakentaa henkilökohtaista tietoavaruutta. Tämä on prosessi, jossa tarvitaan luokittelemisen taitoja. Nämä taidot ovat osa integraatio-prosessia, jossa uutta tietoa yhdistetään aikaisemmin rakennettuun tietoon, kun käyttäjän henkilökohtaiseen tietoavaruuteen lisäämät tietoyksiköt yhdistetään muihin tietoyksiköihin hierarkiassa olevan sijaintinsa mukaan. (Hardof-Jaffe ym., 2009, s. 251.)

Luokitteleminen ja lajitteleminen ovat tärkeä osa tietojen järjestämistä, koska niiden kautta muodostetaan tiedoista yhteensopivia kokonaisuuksia ja laitetaan samaan asiayhteyteen kuuluvat tiedot samaan paikkaan. Tämä taas on äärimmäisen tärkeää tietojen hakemisen kannalta, koska se helpottaa tietojen löytämistä käyttöä varten. Luokitteleminen on vaikea tehtävä, jossa on kyettävä ennakoimaan tulevia asiayhteyksiä tietojen käyttämisessä. Tästä johtuen käyttäjien luokittelut ovat usein epäjohdonmukaisia ja joskus käyttäjät voivat jopa unohtaa joidenkin pitkäaikaisten kansioidensa olemassaolon. (Whittaker & Sidner, 1996, s. 279.) Luokittelun päämääränä on lähinnä pyrkiä järjestelemään tietoyksiköt kansioihin siten, että ne ovat helpommin saatavilla myöhemmin (Whittaker, Bellotti & Gwizdka, 2006, s. 71).

Kansion nimeämisessä käyttäjän pitää muistaa joitain perusominaisuuksia tiedoista, joita tietoyksikkö sisältää tai tunnistaa niiden pääteema, jotta voisi määrittellä ja nimetä kansion tai tiedoston asianmukaisesti (Mts. 251). Kansioiden ja tiedostojen nimeäminen on henkilökohtainen ja erikoinen tehtävä. Tästä syystä joillakin nimillä voi olla merkitystä vain luojallensa jossain tiettyssä yhteydessä. Jotkin kansioiden nimet voivat toistua eri paikoissa kansiohierarkiassa, mikä jossain tapauksissa voi johtua joidenkin kansiorakenteiden kopioimisesta. (Henderson, 2005, s. 76.) Boardman (2004, s. 75) kutsuu tätä kansioiden päällekkäisyydeksi (overlap), millä hän tarkoittaa sitä, missä määrin kansioita, jotka viittaavat samaan asiaan esiintyy eri kokoelmissa (tiedosto, sähköposti ja kirjanmerkit). Esimerkkinä vaikka kansio Seminaarityö joka löytyy sekä tiedostojen kansiohierarkiasta että sähköpostista. Osa kansioista on järjestelmän luomia ja niiden nimet on määrittänyt käyttöjärjestelmä tai ohjelmisto, joka loi ne. Näihin kansioihin kuuluvat oletuskansiot kuten työpöytä, omat tiedostot, omat kuvatiedostot jne. (Henderson, 2005, s. 76–77.) Yleisimmiksi kansioiden nimityypeiksi Henderson (2005) havaitsi genren, tehtävän, aiheen ja ajan. Lähellä seurasivat myös kurssi ja henkilö. (Mts. 78–79)

6 YLLÄPITO

Tässä luvussa tarkastellaan Jonesin (2008, s. 59; Jones, 2007, s. 464) toimintojen ryhmittelystä meta-tason toiminnoista ylläpidon toimintoa. Boardmanin toimintojen jaottelussa tämä on kohta 4.

Kuten säilyttäminen ja järjestäminen ovat usein yhteydessä toisiinsa, niin ovat myös usein järjestäminen ja *ylläpito* (maintaining) yhteydessä toisiinsa. Toisaalta myös nämä kaksi toimintoa eroavat omalla tavallaan toisistaan. Järjestämisessä keskitytään siihen, kuinka kokoelmassa olevat tietoyksiköt liittyvät toisiinsa ja kuinka ne erotetaan toisistaan annettavien nimien ja muiden ominaisuuksien kautta sekä niiden kansioon sijoittamisen kautta. Ylläpito sisältää muutakin, kuin pelkkää tietojen järjestämistä. Tietokokoelmien ylläpitoon sisältyy tiedon varastoimiseen liittyvät asiat, niiden varmuuskopiointi ja tietoyksiköiden säilyttäminen sopivassa järjestyksessä ja toimintakunnossa pitäminen niiden käyttämistä varten. Näiden lisäksi ylläpitoon kuuluvat pyrkimykset puhdistaa kokoelmat tiedoista, jotka eivät enää kuulu sinne. (Jones, 2008, s. 156–157.) Tarkemman määritelmän mukaan, ylläpito tarkoittaa tehtyjä päätöksiä ja toimintoja henkilökohtaisen tiedon säilyttämiseen (preservation) ja kunnossapitoon liittyen. Päätöksiin kuuluu se, kuinka ja missä muodossa tietoja varastoidaan (store) sekä kuinka tietoja päivitetään tai korjataan ja mitä tapahtuu vanhemmalle tiedolle. (Jones, 2008, s. 126, 156.)

Ylläpito on meta-tason toiminto, mikä tarkoittaa sitä, että ylläpidon toiminnot tehdään usein vasta muiden töiden jälkeen, mikä yleensä tarkoittaa harvoin, jos koskaan. Tämä johtuu siitä, että tavallisessa päivässä ei ole mitään, mikä pakottaisi toteuttamaan ylläpidon toimintoja. Ennen levytilan loppuminen aiheutti tarpeen levyn siivoamiseen, eli tarpeettoman tiedon poistamiseen tai sen siirtämiseen muualle säilytykseen. Nykyään levytilojen kasvamisen myötä, monella on käytössään enemmän tilaa henkilökohtaiselle tiedolle, joten levytilan raivaamiselle ei ole enää niin usein tarvetta. Tämä on vapauttanut ihmiset aikaavievistä ja hankalista päätöksistä, jotka liittyvät siihen, mitä tietoa vielä säilytetään, mitkä tiedot poistetaan tai mitkä tiedoista arkistoidaan pitkäaikaista säilyttämistä varten. (Jones, 2008, s. 157–158.)

6.1 Varmuuskopiot

Ihmiset keräävät elämänsä aikana paljon tietoja henkilökohtaisiin arkistoihin. Tietojen säilyttämisen taustalla on usein myös pelko niiden mahdollisesta menettamisestä jonkin katastrofin seurauksena. Pelkona voi olla tulipalo tai varkaus, jonka seurauksena tiedot häviävät. Lisäksi voi olla huolta tulevasta muutosta, jonka aikana tiedot voivat kadota tai sitten on yleistä epäluottamusta tietokoneiden toimivuutta kohtaan sekä pelko kiintolevyn kaatumisesta. Nämä kaikki huolet ja pelot vaikuttavat käyttäjän arkiston rakenteeseen niin säännöllisten varmuuskopiointien kuin erikoiskansioiden muodossa. (Kaye, Vertesi, Avery, Dafoe, David, Onaga, Rosero & Pinch, 2006, s. 278.)

Useimmat henkilökohtaisen tiedon kokoelmissa olevat tiedot on kerättyinä ja tallennettuna tietokoneidemme kiintolevylle. Mutta kuten Kawai & Sandoh (2003, s. 1055–1056) mainitsevat, voi osa näistä tiedoista silloin tällöin kadota inhimillisen tai laitteistossa tapahtuvan virheen seurauksena. Yksinkertaisin tapa suojata tietojamme tällaisilta menetyksiltä, on tehdä säännöllisesti varmuuskopiot tiedostoista ulkoiselle kiintolevylle tai muulle tallennusvälineelle. (Mts. 1056) Säännöllisillä varmuuskopiointilla varmistetaan se, että tiedot ovat ajan tasalla ja vahingon sattuessa ne voidaan palauttaa nopeasti käyttöön, eikä mitään tarpeellista tietoa menetetä. Tähän varmuuskopioiden ajan tasalla pitämiseen on tätä työtä kirjoittaessakin tullut omakohtainen esimerkki niiden tärkeydestä. Tallennuksen yhteydessä sattuneen virhetilan vuoksi viimeisimmät graduun tehdyt muokkaukset hävisivät. Aikaisemman varmuuskopion ansiosta suurin osa tiedoista oli vielä tallessa, mutta muutaman päivän aikana tehdyt muutokset olivat poissa. Tämä aiheutti usean tunnin verran lisätyötä, mutta ei onneksi mitään suurta katastrofia ja pääsin melko nopeasti jatkamaan työtä eteenpäin. Nykyään teen varmuuskopion heti lopetettuani työn muokkaamisen.

Vaikka ihmiset tietävät, että varmuuskopiointi on tärkeä tapa, jolla välttää tietojen häviäminen, tämä tieto itsessään ei johda tarvittavaan määrään tai tasoon varmuuskopiointissa. Monet ihmiset ovat huolissaan tietojensa menettamisestä, mutta silti suurin osa heistä tekee varmuuskopioita vain kerran kuukaudessa tai harvemmin. (Yechiam, Haruvy & Erev, 2002, s. 599–600.) Yechiam, Haruvy ja Erev (2002) havaitsivat, että suurin osa käyttäjistä tekee liian vähän varmuuskopioita, ja vaikka käyttäjät ovat tietoisia tästä epäkohdasta, he silti jatkavat tätä käyttäytymistapaa. Yhtenä mahdollisena selityksenä varmuuskopiointin alisuorittamiseen he pitivät sitä, että käyttäjät aliarvioivat mahdollisuutta tietokatastrofiin, eli tietojen häviämistä laitteistovian, varkauden tai vahingossa tapahtuvan tiedoston poistamisen vuoksi. (Mts. 601)

Internetistä löytyy varmuuskopiointille paljon ohjeita eri lähteistä. Perusohjeiksi varmuuskopiointiin voidaan luetella muutama tärkein seikka, jotka ainakin pitää ottaa huomioon. Tärkeimmäksi ohjeeksi voitaisiin antaa, että varmuuskopioita tulisi tehdä säännöllisesti, ainakin kerran viikossa. Tosin, jos

haluaa varmistaa, että tiedoista on aina ajantasalla oleva varmuuskopio, on varminta kopioida tiedot päivittäin. Varmuuskopioita olisi myös hyvä tehdä useampi kappale eri tallennusvälineille, joita tulisi vielä säilyttää eri paikoissa, jotta esimerkiksi tulipalon sattuessa ei menetetä kaikkia varmuuskopion kappaleita. Yleisesti voisi sanoa, että 2-3 kappaletta on sopiva määrä varmuuskopioita. Yksinkertaisin tapa tehdä varmuuskopioita on manuaalisesti siirtää joistakin tiedostoista tai kansioista kopiot toiselle tallennusvälineelle. Saatavilla on myös erilaisia varmuuskopiointiohjelmiä, niin yrityksille kuin kotikäyttöön, joilla voidaan automatisoida varmuuskopioiden tekeminen. Kun varmuuskopiot tehdään automaattisesti, esimerkiksi kerran päivässä, niin varmuuskopiointi ei ole henkilön muistin tai viitsimisen varassa. Tällöin voidaan olla varmoja siitä, että tiedot ovat tallessa ja ne ovat aina ajantasalla. Varsinkin yrityksissä on yleensä investoitu kattaviin varmuuskopiointiohjelmiin, jotka tekevät automaattisesti yrityksen tiedoista varmuuskopioita määriteltynä ajanjaksoina. Yrityksille tietojen menettäminen voi merkitä koko yrityksen kaatumista, joten tietojen säilyttämisestä yleensä huolehditaan riittävästi. Kotikoneilla olevia tietoja harvemmin varmistetaan yhtä tunnollisesti, eikä kovin monella ole käytössään varmuuskopiointijärjestelmiä ja automatisoitua varmuuskopiointia.

6.2 Päivittäminen

Tässä luvussa käsitellään ylläpitoon kuuluvia tehtäviä, jotka liittyvät tiedon päivittämiseen. Päivittämisen menetelmän alle on tämän luvun käsittelyn puitteissa kerätty tietojen muokkaaminen, tietojen poistaminen ja tietojen arkistointi. Tieto ei aina ole ikuista ja sen käyttö tarkoitukset voivat muuttua ajanmyötä. Tämä johtaa erilaisiin toimenpiteisiin tietokokoelmien kunnossapitämisen suhteen.

Tietojen muokkaaminen ajantasalle on yksi tehtävä, joka usein tulee vastaan tiedonhallinnan hoitamisessa pitkällä aikavälillä. Puhelinnumerot ja sähköpostiosoitteet muuttuvat ajan kuluessa, sukunimet muuttuvat avioliiton tai avioeron myötä ja ihmiset muuttavat, jolloin osoitteet vaihtuvat (Jones, 2008, s. 162). Tämä aiheuttaa tarpeen muokata osoitekirjojen yhteystietoja ajantasalle. Muokkaamiseen liittyy myös tiedostojen ja kansioden uudelleen nimeäminen. Kuten jo aikaisemmin luvussa 5.2 huomioitiin, on tietojen luokittelu vaikeaa ja se voi olla epäjohdonmukaista. Tästä johtuen voivat jotkin tietoyksiköt olla väärin luokiteltuja. Uutta tietoa kertyy koko ajan lisää ja aikaisemmin luodut kansio- ja tiedostonimet eivät välttämättä olekaan tarpeeksi kuvaavia tai tarkkoja tiedon sisällön suhteen, joten ne pitää muokata paremmin asiasisältöön sopivaksi. Myös tiedostomuodon muuttaminen voi tulla ajankohtaiseksi, kun tiedosto pitää voida ottaa käyttöön toisella ohjelmalla. Muutoksia tapahtuu kun ohjelmaversioita päivitetään uudempaan tai vaihdetaan kokonaan käytetty ohjelmisto toisiin ohjelmiin. Tällöin vanhat tiedostomuodot eivät enää toimi kunnolla tai ne eivät toimi enää ollenkaan.

Tietoja yleensä kerätään täyttämään jokin tiedon tarve. Tiedontarpeet kuitenkin vaihtelevat ajanmyötä. Ajankohtaiset tiedot pidetään yleensä aktiivisissa kansioissa sellaisessa paikassa, josta ne on helppo löytää ja ottaa käyttöön. Kun tieto on täyttänyt tehtävänsä, se yleensä siirretään pitkäaikaiseen arkistoon mahdollista myöhempää käyttöä varten, joko kansiorakenteen syrjäisempään osaan tai kokonaan toiselle tallennusvälineelle. Mutta joskus käy niin, että tieto vanhentuu, eikä meillä ole enää käyttöä sille. Tämän vuoksi tietoa ei enää ole tarpeen säilyttää. Aika ajoin tulee tehtäväksi poistaa vanhentuneet ja tarpeettomiksi käyneet tiedot, joita ei enää käytetä ja tehdä tilaa uusille tiedoille. Tällä pidetään tieto kokoelmat ajantasalla ja puhtaana turhasta "roinasta".

7 EMPIIRINEN TUTKIMUS

Tässä luvussa käydään läpi aiheesta toteutettua kyselytutkimusta, jossa keskitytään tarkastelemaan aikaisemmissa teorialuvuissa käsiteltyjä henkilökohtaisen tiedonhallinnan toimintoja ja niiden menetelmiä. Lähtökohtana tutkimukseen oli saada lisätietoja menetelmien käyttämisestä käytännön kannalta. Tutkimuksen kautta lähdettiin hakemaan vastauksia tutkimuskysymykseen: miten ihmiset käyttävät menetelmiä? Tutkimuksessa selvitetään osallistujien tietokokoelmien rakennetta ja niiden kohdalla käytettyjä toimintoja. Säilyttämisen toimintoon liittyen tarkastelussa olivat tiedosto-, sähköposti- ja kirjanmerkkikokoelmien sisältämät kansiorakenteet, joista erityisesti selvitettiin osallistujien luomien kansioden tasojen syvyyttä sekä kansioden sisältämiä tietoyksiköiden määriä. Järjestämisen toiminnon osalta tutkimuksessa käytettiin joitakin aikaisemmin esiteltyjä järjestämisen tapojen luokitteluja selvittämään, millaisia toimintatapoja osallistujilla on tietojensa järjestämiseen eri kokoelmissa. Löytämisen toimintoon liittyen selvitettiin, miten osallistujat etsivät tietojansa, ja etenkin navigoinnin ja haun välistä mieltymystä. Lisäksi tarkastelussa oli, kuinka hyvin tietojen sijainti muistetaan. Ylläpidosta pääpaino tutkimuksessa laitettiin varmuuskopioinnin ja sen suorittamisen tarkasteluun.

Luku etenee siten, että alaluku 7.1 esittelee tutkimuksen tavoitteita ja tutkimuksen toteuttamiseen valittua menetelmää, alaluku 7.2 esittelee tutkimukseen valitut kohdejoukot sekä käsittelee kyselylomakkeen muodostamista ja rakennetta, alaluvussa 7.3 käydään läpi saadut tulokset ja niiden analyysi ja alaluvussa 7.4 vertaillaan ristiintaulukoissa taustamuuttujia sekä omia tuloksia aikaisemmissa tutkimuksissa saatuihin tuloksiin. Alaluvussa 7.5 käsitellään tutkimuksen luotettavuutta.

7.1 Tutkimuksen tavoite ja menetelmä

Tavoitteeni tässä tutkimuksessa oli selvittää, miten kotikoneilla käytetään eri tiedonhallinta menetelmiä. Lisäksi tutkimuksella pyrittiin keräämään aineistoa,

jota voitaisiin verrata aikaisemmissa tutkimuksissa saatuihin tuloksiin. Tarkoituksena on kerätä aineistoa, jolla voitaisiin tarkastella enemmän kotikoneilla tapahtuvaa henkilökohtaista tiedonhallintaa ja sen suorittamisen käytännön kokemuksia.

Tutkimus toteutettiin survey-tutkimuksena. Survey-tutkimuksessa aineistoa kerätään standardoidusti kyselyn, haastattelun tai havainnoinnin kautta kohdejoukolta, joka muodostaa otoksen perusjoukosta. (Hirsjärvi, Remes & Sajavaara, 2009, s. 193.) Tutkimusmenetelmäksi tämän tutkimuksen suorittamiseen valittiin kyselytutkimus. Kyselytutkimuksen etuna on, että sillä voidaan kerätä laaja tutkimusaineisto ja kyselylomakkeella voidaan kysyä monia asioita. Kysely on myös tehokas, koska se säästää tutkijan aikaa ja vaivannäköä, ja mikäli lomake on huolellisesti suunniteltu, voidaan aineisto nopeasti käsitellä tallennettuun muotoon ja analysoida tietokoneen avulla. Lisäksi aikataulu ja kustannukset voidaan arvioida melko tarkasti. Kyselyn heikkoutena on, että aineistoa voidaan pitää pinnallisena ja tutkimuksia teoreettisesti vaatimattomina. Myös tulosten tulkinta voi olla ongelmallista. Lisäksi ongelmallista on myös se, että tutkija ei voi olla varma siitä, kuinka vakavasti vastaajat suhtautuvat tutkimukseen tai kuinka rehellisesti vastataan. (Hirsjärvi, Remes & Sajavaara, 2009, s. 195.)

7.2 Kohdejoukko ja kyselylomakkeen laadinta

Tutkimuksen kohdejoukkona toimivat kahden eri tiedekunnan opiskelijat Jyväskylän yliopiston Informaatioteknologian tiedekunnasta ja Yhteiskuntatieteellisen tiedekunnasta. Koska voidaan olettaa, että yleensä tietotekniikka-alaa opiskelevilla on lähtökohtaisesti paremmat tiedot ja taidot tietotekniikan suhteen, otettiin tämän vuoksi kyselyyn mukaan myös toisen tiedekunnan opiskelijoita tulosten vääristymisen välttämiseksi.

Kysely toteutettiin verkkojulkaisuna ja lähetettiin osallistujille sähköpostin välityksellä opiskelija- ja ainejärjestöjen kautta. Tässä tapauksessa ainejärjestöistä valittiin Informaatioteknologian tiedekunnasta Dumppi ja Linkki sekä Yhteiskuntatieteellisen tiedekunnasta Fokus ja Puolue. Kyselylomake tehtiin webkyselyiden tekoon suunnitellulla mrInterview-ohjelmalla. Lomake testattiin ensin 5 henkilöllä, jonka jälkeen se muokattiin lopulliseen muotoonsa. Kyselylomakkeessa kysymykset on jaettu eri ryhmiin, sen mukaan, mitä aihetta kysymykset koskevat. Kysymysryhmiksi on jaoteltu taustatiedot, kansiorakenne, tietojen järjestäminen, tiedon löytäminen ja ylläpito. Verkkolomakkeessa jokaisen ryhmän sisältämät kysymykset ryhmiteltiin erikseen niin, että ne näkyvät samalla sivulla. Kysely pohjautuu lähinnä esitettyyn teoriaan ja aikaisempiin tutkimuksiin, kuten esimerkiksi Bergman, Byeth-Marom, Nachmias, Gradovitch ja Whittaker (2008), Boardman ja Sasse (2004) sekä Yechiam, Haruvy ja Erev (2002). Kyselylomake kokonaisuudessaan löytyy liitteestä 1.

Kysymyksiin on pyritty antamaan valmiita vastausvaihtoehtoja, joista vastaaja voi valita hänelle parhaiten sopivan vastauksen. Lomakkeessa on käytetty paljon valmiita vastausvaihtoehtoja, jotta vastaajien samaan kysymykseen antamia vastauksia olisi helpompi vertailla. Lisäksi vaihtoehtoja on käytetty lomakkeeseen vastaamisen helpottamiseksi, koska vaihtoehdot auttavat vastaajaa tunnistamaan asian sen muistamisen sijaan. (Hirsjärvi, Remes & Sajavaara, 2009, s. 201.) Vastausvaihtoehtoja on valittu aikaisemmissa tutkimuksissa esiintyneistä aiheista, joita käsiteltiin teorialuvuissa. Tutkimusten pohjalta valittujen vaihtoehtojen lisänä on kohta "muu, mikä", joka lisättiin, koska on mahdollisuus, että kohdehenkilöllä on vastaus, jota tekijä ei ole etukäteen osannut ajatella (Hirsjärvi ym., 2009, s. 199). Aikaisemmista tutkimuksista saatujen vaihtoehtojen lisäksi on myös joitakin itse lisättyjä vaihtoehtoja, jotka tulevat omasta kokemuksesta.

Lomakkeen kysymykset 1-7 koskevat vastaajan taustatietoja. Ne käsittävät yleiset kysymykset kohdehenkilön taustaan liittyen, kuten tiedekunta, sukupuoli, tietokoneen käyttökokemus ja henkilön käytössä olevat ohjelmistot. Lisäksi tiedusteltiin aikaa, jonka henkilö käyttää tietokoneella yhdessä päivässä.

Kysymykset 8-25 koskevat kansiorakennetta ja käsittelevät luvuissa 4 ja 5.1 esiteltyjä aiheita. Kysymyksissä selvitettiin vastaajan tietokokoelmien muodostamista, niiden kansiorakenteiden syvyyttä ja kansiodien ja niiden sisältöjen määriä sekä kansiorakennetta. Kansiorakennetta koskevaan kysymykseen vastausvaihtoehdot otettiin Hardof-Jaffen ym. (2009) tutkimuksesta. Kirjanmerkkejä koskeviin kysymyksiin on otettu mallia Abramsin (1997) kyselylomakkeesta. Muihin tämän ryhmän kysymyksiin ei ollut käytettävissä valmista kyselylomakemallia, joten kysymykset on johdettu aikaisemmissa tutkimuksissa mainittujen haastattelukysymysten ja aihealueiden sekä tutkimustulosten pohjalta.

Kysymykset 26-32 käsittelevät tietojen järjestämistä ja koskevat luvun 5 aiheita tietojen järjestämisen tavoista ja käyttäytymisestä. Useimpien kysymysten vastausvaihtoehdot on otettu aikaisempien tutkimusten (muun muassa Boardman ja Sasse (2004), Whittaker ja Sidner (1996) ja Abrams ym. (1998)) esittelemistä tapojen luokitteluista. Lisäksi kirjanmerkkejä koskeviin kysymyksiin on otettu mallia Abramsin (1997) tutkimuksessa esitellystä kyselylomakkeesta.

Kysymykset 33-37 käsittelevät tiedon löytämistä ja koskevat luvun 3 aiheita. Kysymykset koskevat kohdehenkilön tapoja etsiä tiedostoja koneelta. Kysymykset on muodostettu Bergmanin ym. (2008) tutkimuksen pohjalta, johon on itse lisätty kysymyksiä muun muassa tiedon sijaintipaikan muistamiseen.

Ylläpitokysymyksissä 38-52 käsitellään lähinnä varmuuskopioita. Mallia on kysymyksiin otettu Yechiamin ym. (2002) tutkimuksen kyselylomakkeesta, jonka kysymyksiä on muokattu vaihtamalla avoin vastauskenttä valmiisiin vastausvaihtoehtoihin ja lisäämällä yksi vaihtoehto. Lisäksi on itse lisätty muita varmuuskopiointikysymyksiä, koska mallilomakkeessa varmuuskopioiden osuus oli melko vähäinen. Kysymyksiä lisättiin varmuuskopiointitavoista ja

lukumääristä sekä varmuuskopioitaviin tietoihin liittyen. Lopuksi laitettiin muutama kysymys päivittämiseen liittyen.

7.3 Aineisto ja analyysi

Aineisto kerättiin ajanjaksolla 28.11.2011 - 12.12.2011. Tällä ajalla vastauksia saatiin vain 29 ja kaikki Yhteiskuntatieteellisen tiedekunnasta. Karhuviestillä Dumpin postilistalle aikavälillä 13.12-16-12-2011 saatiin kerättyä vastauksia 27 lisää Informaatioteknologian tiedekunnasta, mikä nosti kokonaismäärän 56 vastaukseen. Tutkimuksesta kerätty aineisto analysoitiin SPSS (Statistical Package for Social Sciences) -tilasto-ohjelmistolla ja osa aineistosta Microsoft Excelillä. Exceliä on käytetty muun muassa joidenkin luokitteluiden tekemiseen sen vuoksi, että se oli SPSS:ssä liian monimutkaista. Aineistoa ja sen analyysia käsitellään tarkemmin seuraavissa kappaleissa kysymysryhmien mukaan, joissa osaa tuloksista käsitellään vielä erikseen eri tietokokoelmien (tiedosto, sähköposti ja kirjanmerkit) välillä. Aineistoa analysoitiin käyttäen tavanomaisia tilastomenetelmiä, kuten frekvenssien laskentaa ja keskiarvoja. Lisäksi mitattiin muuttujien välisiä riippuvuuksia sekä selvitettiin eroja eri taustatekijöiden välillä.

Taustatiedot

Kokonaisuudessaan vastauksia saatiin 56, joista 48,2% (N=27) tuli Informaatioteknologian tiedekunnasta ja 51,8% (N=29) Yhteiskuntatieteellisen tiedekunnasta. Vastaajista 53,6% (N=30) oli miehiä ja 46,4% (N=26) naisia. Vastaajat käyttävät tietokonetta päivässä keskimäärin 6,5 tuntia (min. 1, maks. 14, M 6,57). Tietokoneen käyttökokemukseksi vastaajista 30,4% (N=17) arvioi erittäin hyvän, 48,2% (N=27) hyvän, 14,3% (N=8) melko hyvän, 3,6% (N=2) kohtuullisen, 1,8% (N=1) melko huonon ja 1,8% (N=1) huonon.

Käyttöjärjestelmistä vastaajista 71,4 % (N=40) oli käytössään Windows, 16,1 % (N=9) Macintosh, 8,9 % (N=5) Linux, 1,8 % (N=1) Linux ja Windows ja 1,8 % (N=1) Windows, Linux ja Macintosh. Sähköpostin kohdalla vastaajista 57,1 % (N=32) oli käytössään vain sähköpostitili. Sähköpostin hallintatyökaluista 16,1 % (N=9) vastaajista oli käytössään Mozilla Thunderbird, 8,9 % (N=5) MS Outlook, 1,8 % (N=1) Lotus Notes sekä 3,6 % (N=2) Apple Mail, 3,6 % (N=2) Mac Mail, 1,8 % (N=1) Evolution, 1,8 % (N=1) Opera ja 1,8 % (N=1) gmail. Lisäksi 3,6 % (N=2) vastaajista ilmoitti pelkän Mailin. Internet selaimena vastaajista 50 % (N=28) käyttää Mozilla Firefoxia, 25 % (N=14) Google Chromea, 8,9 % (N=5) Safaria, 7,1 % (N=4) Operaa, 3,6 % (N=2) Safaria ja Chromea ja 3,6 % (N=2) Chromiumia.

Kansiorakenne

Kyselyssä kerättiin tietoja eri tietokokoelmien rakenteesta sekä yleisestä kansiorakenteen muodosta, jossa käytettiin pohjana Hardof-Jaffen ym. (2009) tietoavaruuden järjestelemisen luokittelua kasaaminen, yhden kansion arkistointi, pienten kansioden arkistointi ja suuren kansion arkistointi. Tähän luokitteluun lisättiin oman kokemuksen pohjalta kohta, jossa tietoyksiköitä kerätään useampaan suureen kansioon. Tällöin luokittelun kohta suuren kansion arkistointi voidaan jakaa luokitteluiksi yhden suuren kansion arkistointi ja usean suuren kansion arkistointi.

Taulukossa 2 on esitettyä yhdistelmätaulukossa havaintojen jakaumat mainituista luokista tiedekunnan ja sukupuolen osalta. Vastaajien tapojen jakautumisesta voidaan huomata, että 3,6 % vastaajista oli kasaajia, 3,6 % yhden kansion arkistointia, 32,1 % pienten kansioden arkistointia, 3,6 % yhden suuren kansion arkistointia ja 51,8 % usean suuren kansion arkistointia. Vastaajista 5,4 % antoi kohtaan "muu, mikä" oman kuvauksensa, jotka eivät sopineet mainittuihin luokkiin. Näistä vastaajista yksi ilmoitti, että hänellä oli paljon kansioita juurikansion alapuolella. Toisella oli paljon kansioita, joissa oli muutamia tiedostoja, mutta kuvakansioissa oli paljon kuvia. Viimeinen vastaaja ilmoitti, että hänellä on paljon kansioita aiheen ja tärkeyden mukaan.

Tulosten perusteella huomataan, että vastaajista suurimmalla osalla on kansiorakenteessaan useita suuria kansioita, joihin kerätyt tietoyksiköt on arkistoitu. Tiedekuntien välillä ei useimpien luokkien kesken ole huomattavaa eroa. Suurin ero löytyy pienten kansioden arkistointia, joita on Informaatioteknologian tiedekunnassa puolet enemmän kuin Yhteiskuntatieteellisen puolella. Usean suuren kansion arkistointia on Yt-tiedekunnassa hieman enemmän kuin It-tiedekunnassa. Yhden suuren kansion arkistointia ei It-tiedekunnassa löydy yhtään. Pienten kansioden arkistointia suurin osa on miehiä, mutta muuten sukupuolten väliset erot luokkien kesken ovat melko pieniä. Tiedekuntien osalta on havaittavissa, että Yt-tiedekunnassa on usean suuren kansion arkistointia enemmän naisia kuin It-tiedekunnassa.

Muuttujien välinen korrelaatio laskettiin Spearmanin järjestyskorrelaatiokertoimella (ρ). Järjestämisen tapojen luokittelun ja tiedekunnan välinen korrelaation on ($\rho = 0,154$). Sukupuoleen verrattaessa korrelaatio on ($\rho = 0,184$). Molemmat korrelaatiot ovat positiivisia, mutta yhteys on aika heikko. Koska otoskoko on pieni, ei kummankaan muuttujan kohdalla voida sanoa, että korrelaatiokerroin eroaisi nolasta tilastollisesti merkitsevästi. Kolme luokkaiseksi vähennettyyn käyttökokemus muuttujaan verrattuna järjestelemisen tapojen (ks. taulukko liitteessä 2) korrelaatio on ($\rho = 0,044$), joka ei juuri eroa nolasta, joten havaittavissa ei ole tilastollisesti merkitsevää riippuvuutta.

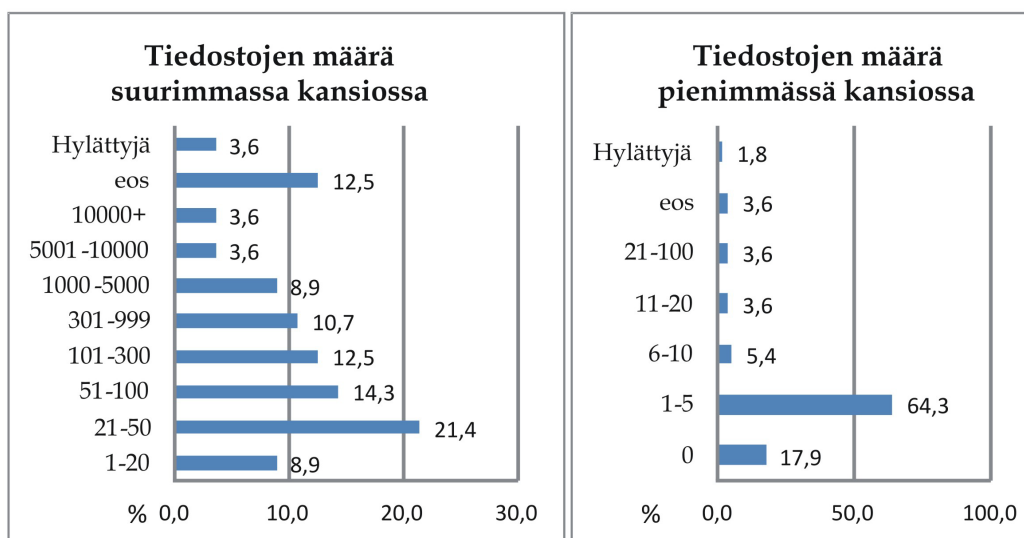
TAULUKKO 2 Tietoavaruuden järjestämisen tapojen jakaumat tiedekunnan ja sukupuolen mukaan

Tietoavaruuden järjestämisen tavat	Tiedekunta						
	IT			YT			Yht.
	n N/ %	m N/ %	yht. N/ %	n N/ %	m N/ %	yht. N/ %	
Kasaaja		1 1,8 %	1 1,8 %		1 1,8 %	1 1,8 %	2 3,6 %
Yhden kansioon arkistoiija		1 1,8 %	1 1,8 %	1 1,8 %		1 1,8 %	2 3,6 %
Pienten kansioiden arkistoiija	4 7,1 %	8 14,3 %	12 21,4 %	3 5,4 %	3 5,4 %	6 10,7 %	18 32,1 %
Yhden suuren kansioon arkistoiija				1 1,8 %	1 1,8 %	2 3,6 %	2 3,6 %
Usean suuren kansioon arkistoiija	4 7,1 %	7 12,5 %	11 19,6 %	11 19,6 %	7 12,5 %	18 32,1 %	29 51,8 %
Muu	1 1,8 %	1 1,8 %	2 3,6 %	1 1,8 %		1 1,8 %	3 5,4 %
Yhteensä	9 16,1 %	18 32,1 %	27 48,2 %	17 30,4 %	12 21,4 %	29 51,8 %	56 100 %

Kansioiden luomisen kohdalla vaihtoehdot oli moudostettu Hardof-Jaffen ym. (2009) ennakkorakentajat (pre builder) ja jälkirakentajat (post builder) ja Abramsin ym. (1998) luomisen aikainen arkistoiija luokittelusta. Tulokset osoittivat, että vastaajista 16,1% (N=9) loi kansioita samalla kun tallensi tietoja, 12,5% (N=7) teki kansioita etukäteen, 30,4% (N=17) teki kansioita jälkikäteen järjestelemisen yhteydessä ja 41,1% (N=23) käyttää näitä kaikkia kansion luomisen tapoja. Kuten vastauksista näkyy, suurin osa vastaajista ei pitäydy tiettyyn tapaan luoda kansioita, vaan käyttää kaikkia kansion luomisen tapoja.

Seuraavaksi tarkastellaan tiedosto-, sähköposti- ja kirjanmerkkikokoelmista kerättyjä tietoja. Tiedostokokoelmissa kansiorakenteen tasojen syvyys oli noin 3 tasoa (min. 0, maks. 10, M 3,5). Tiedostojen määrästä suurimmassa ja pienimmässä kansiossa muodostettiin Excelillä uudet luokat saatujen arvojen pohjalta. Muutamat vastaukset olivat epäselviä, joten ne laskettiin hylätyiksi. Suurimmassa kansiossa olleet kaksi 0 arvoa luettiin ei osaa sanoa -vastausten joukkoon. Näin tehtiin, koska voidaan olettaa, että jossakin kansiossa olisi edes joi-takin tiedostoja. Muokkauksen jälkeen luokkafrekvensseiksi saatiin suurim-massa kansiossa: 1-20 5, 21-50 12, 51-100 8, 101-300 7, 301-999 6, 1000-5000 5, 5001-10000 2 ja 10000+ 2. Ei osaa sanoa - vastauksia tuli 7 ja hylättyjä vastauksia oli 2. Pienimmässä tiedostokansiossa olevien tiedostojen määrän luokka-frekvenssijakauma oli: 0 10, 1-5 26, 6-10 3, 11-20 2 ja 21-100 2. Ei osaa sanoa -

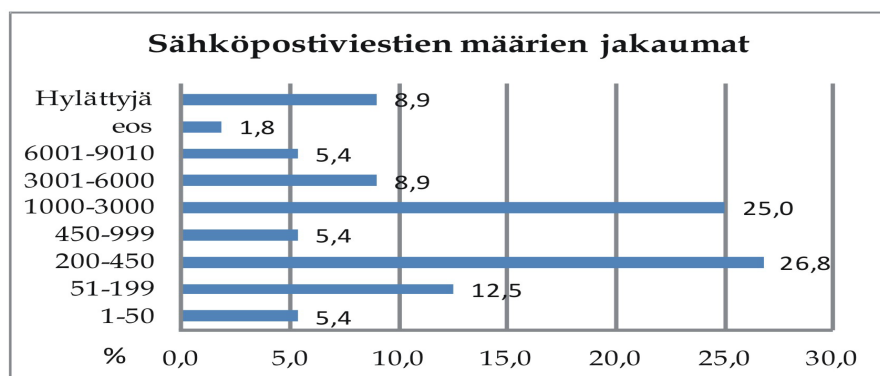
vastauksia tuli 2 ja hylättyjä vastauksia oli 1. Kuviossa 3 esiintyvät tiedosto määrien prosenttijakaumat luokittain suurimmasta ja pienimmästä kansiossa.



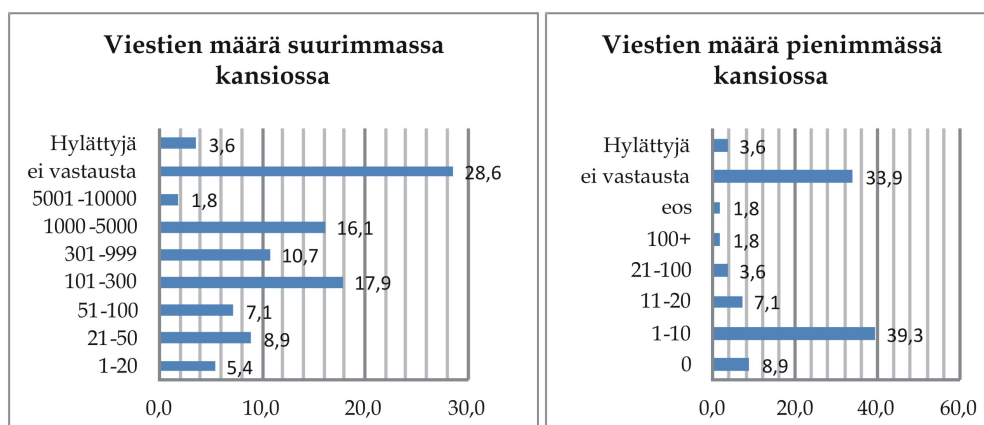
KUVIO 3 Tiedostojen määrien prosenttijakaumat luokittain

Seuraavaksi käsiteltävissä sähköposti- ja kirjanmerkkikokoelmissa, oli kansiorakenteita koskevissa kysymyksissä 20–25 mahdollisuus jättää vastaamatta, mikäli ei käyttänyt kyseisissä kokoelmissa kansioita tietoyksiköiden järjestelmiseen. Tästä johtuen vain osa vastaajista on antanut tietoja kansioiden syvyydestä sekä viestien ja kirjanmerkkien määristä.

Sähköpostikokoelmien suhteen vastaajista 64,3% (N=36) ilmoitti keräävänsä viestejä talteen eri kansioihin ja 35,7% (N=20) ilmoitti, ettei käytä kansioita, vaan jättää viestit postilaatikkoon. Sähköpostiviestien kokonaismäärästä muodostettiin saatujen arvojen pohjalta uudet luokat. Luokkien prosenttijakaumat näkyvät kuviossa 4. Luokkien frekvenssijakaumat ovat: 1-50 3, 51-199 7, 200-450 15, 451-999 3, 1000-3000 14, 3001-6000 5 ja 6001-9010 3. Ei osaa sanoa -vastauksia tuli 1 ja hylättyjä vastauksia oli 5. Sähköpostikansioiden rakenteen syvyys oli noin 1 tasoa (min. 0, maks. 4, M 1,12), kun vastauksia saatiin 73,2 % (N=41) ja vastaamatta jätti 26,8 % (N=15). Myös viestien määristä suurimmassa ja pienimmässä sähköpostikansiossa muodostettiin uudet luokat saatujen arvojen perusteella. Vastauksen antoi 71,4 % (N=40) vastaajista ja vastaamatta jätti 28,6 % (N=16). Frekvenssijakaumiksi saatiin suurimmassa kansiossa: 1-20 3, 21-50 5, 51-100 4, 101-300 10, 301-999 6, 1000-5000 9, 5001-10000 1 ja 10000+ 0. Hylättyjä vastauksia oli 2. Pienimmän sähköpostikansion kohdalla vastauksia tuli 66,1 % (N=37) ja vastaamatta jätti 33,9 % (N=19). Viestien määrän frekvenssijakaumaksi saatiin: 0 5, 1-10 22, 11-20 4, 21-100 2 ja 100+ 1. Vastauksista hylättiin 2. Kuviossa 5 esiintyvät luokkien prosenttijakaumat suurimmassa ja pienimmässä kansiossa.



KUVIO 4 Sähköpostiviestien kokonaismäärien prosentti jakaumat luokittain



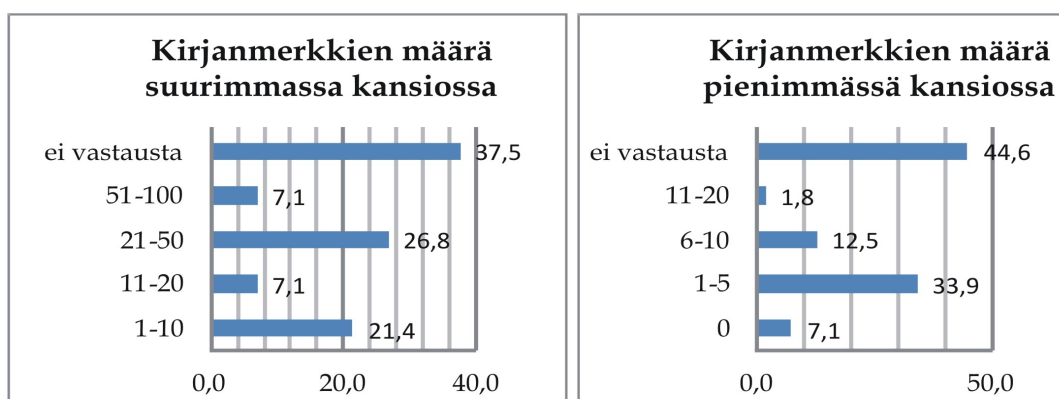
KUVIO 5 Sähköpostiviestien määrien prosenttijakaumat luokittain kansioissa

Kirjanmerkkikokoelmien osalta 98,2% (N=55) vastaajista ilmoitti keräävänsä kirjanmerkkejä ja 1,8% (N=1) ilmoitti, ettei hänellä ole kirjanmerkkejä. Kirjanmerkkien määrissä käytettiin Abramsin (1997) luokkakajoa, johon oli lisätty vaihtoehto "Ei yhtään". Vastauksia tuli siten, että luokissa 1-10 oli vastauksia 17,9% (N=10), 11-25 17,9% (N=10), 26-100 50,0% (N=28) ja 101-300 14,3% (N=8). Mielenkiintoista näissä vastauksissa oli, että vaikka edellisessä kohdassa yksi vastaaja oli ilmoittanut, ettei kerää kirjanmerkkejä, tässä kohdassa kukaan ei kuitenkaan valinnut vaihtoehtoa "Ei yhtään".

Kirjanmerkkikansioita ilmoitti käyttävänsä 58,9 % (N=33) vastaajista ja 41,1 % (N=23) ilmoitti, ettei käytä kansioita kirjanmerkkien järjestämiseen. Kirjanmerkkikansioiden rakenteen syvyys oli noin 1 tasoa (min. 0, maks. 4, M 1,26), kun vastauksia oli 75,0 % (N=42) ja vastaamatta jätti 25,0 % (N=14). Kirjanmerkkien määristä suurimmassa ja pienimmässä kirjanmerkkikansiossa tehtiin uudet luokat saatujen arvojen perusteella. Vastauksen antoi 63,5 % (N=35)

ja vastaamatta jätti 37,5 % (N=21). Havaintomääräksi saatiin suurimmassa kansiossa: 1-10 12, 11-20 4, 21-50 15 ja 51-100 4. Pienimmässä kansiossa olevien kirjanmerkkien määrän jakauma oli: 0 4, 1-5 19, 6-10 7 ja 11-20 1. Vastauksia tuli 55,4 % (N=31) ja vastaamatta jätti 44,6 % (N=25). Kuviossa 6 esiintyvät luokkien prosenttijakaumat suurimmassa ja pienimmässä kansiossa.

Tästä kysymysryhmästä päätettiin jättää käsittelemättä kysymykset 15 ja 16, koska huomattiin, että ne eivät ole asian käsittelyn kannalta kovin oleellisia. Lisäksi kysymykset hylättiin, koska joillakin vastaajilla oli näihin kysymyksiin liittyen joitakin epäselvyyksiä vastaamisen suhteen, joten tulokset eivät havainnoi sitä, mitä niiden piti ja voivat siten olla hieman vääristyneet.



KUVIO 6 Kirjanmerkkien määrien prosenttijakaumat luokittain

Tietojen järjestäminen

Vastaajat käyttivät päivässä aikaa tietojensä järjestämiseen noin 8,5 minuuttia (min. 0,0, maks. 30,0, M 8,61). Arvoista muodostettujen luokkien havaintojen luokkafrekvenssit ja niiden prosenttijakauman näkee taulukosta 3.

TAULUKKO 3 Tietojen järjestämiseen käytetty aika päivässä

Minuutit	f	%
0-1	2	3,6
2-4	12	21,4
5-10	30	53,6
11-20	10	17,9
21-30	2	3,6

Seuraavaksi tarkastellaan tietojen järjestämisen tapoja eri tietokokoelmissa. Tiedostojen järjestämisen tapojen vaihtoehtoisissa käytettiin Whittakerin ja Sidnerin (1996) sähköpostin järjestämisen tapojen luokittelua: arkistoimaton (no-filers), toistuvasti arkistoiva (frequent filers) ja kevätsiivooja (spring-

cleaners). Näihin oli myös lisätty vaihtoehto "muu, mikä". Havainnot luokitte-
luista esitellään taulukossa 4 tiedekunnan ja sukupuolen mukaan ryhmiteltyinä.
Vastaukset jakautuivat siten, että vastaajista 1,8 % oli mielestään arkistoi-
matomia, 46,4 % oli toistuvasti arkistioivia ja 51,8 % oli kevätsiivoojia. Tiedekuntien
kohdalla toistuvasti arkistioivia on hieman enemmän It-tiedekunnassa, kun taas
Yt-tiedekunnassa on hieman enemmän satunnaisesti arkistioivia. Sukupuolen
kohdalla ei tavoissa ole havaittavia eroja. Tiedekuntien sisällä taas on pieniä
eroja sukupuolten välillä, kun It-tiedekunnassa toistuvasti arkistioivista hieman
enemmän on miehiä, kun taas Yt-tiedekunnassa suurin osa on naisia. Satun-
naisesti arkistioivissa on It-tiedekunnassa puolet enemmän miehiä, kun Yt-
tiedekunnassa ei eroa sukupuolten välillä juuri ole.

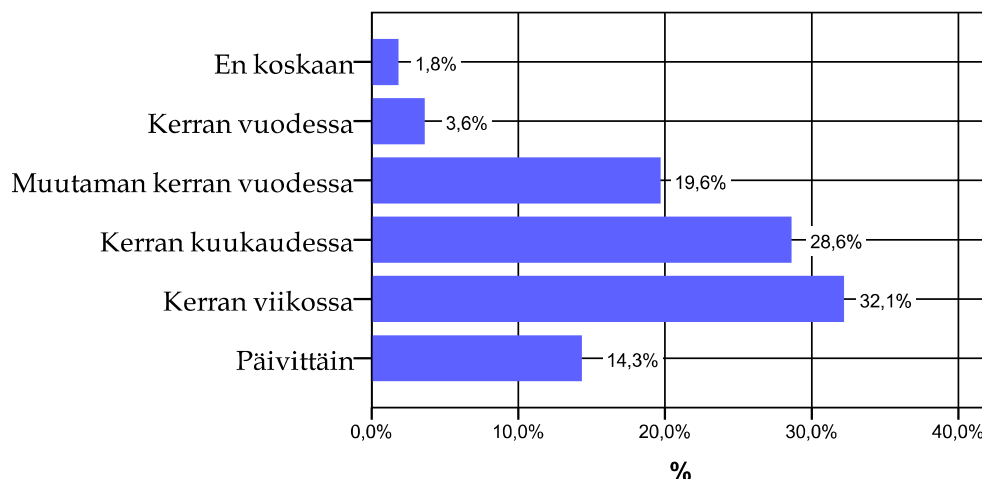
Tiedostojen järjestämisen tapojen ja tiedekunnan välille laskettu Spearman-
nin järjestyskorrelaatiokerroin on ($\rho = 0,156$), joten yhteys on positiivista, mut-
ta heikkoa. Sukupuolen välillä korrelaatio on ($\rho = 0,017$), joten muuttujien vä-
lillä ei näyttäisi olevan riippuvuutta. Kummankaan muuttujan välillä ei ole ha-
vaittavissa tilastollisesti merkitsevää riippuvuutta. Tiedostojen järjestämisen ja
käyttökokemuksen (ks. liite 2) välinen korrelaatio on ($\rho = 0,18$), mikä on hei-
kosti positiivinen, mutta ei kuitenkaan tilastollisesti merkitsevä

TAULUKKO 4 Tiedostojen järjestämisen tavat

		Tiedostojen järjestämisen tavat			
		Toistuvasti			
		Arkistoi- maton	arkistioiva	Kevätsiivooja	Yht.
Tdk	Suku- puoli	N/ %	N/ %	N/ %	N/ %
IT	n		5	4	9
			8,9 %	7,1 %	16,1 %
	m	1	9	8	18
		1,8 %	16,1 %	14,3 %	32,1 %
yht.	1	14	12	27	
		1,8 %	25,0 %	21,4 %	48,2 %
YT	n		8	9	17
			14,3 %	16,1 %	30,4 %
	m		4	8	12
			7,1 %	14,3 %	21,4 %
yht.		12	17	29	
			21,4 %	30,4 %	51,8 %
Yhteensä		1	26	29	56
		1,8 %	46,4 %	51,8 %	100,0 %

Tiedostojensa vastaajista 14,2 % (N=8) ilmoittaa järjestelevänsä päivittäin,
32,1 % (N=18) kerran viikossa, 28,6 % (N=16) kerran kuukaudessa, 19,6 %

(N=11) muutaman kerran vuodessa, 3,6 % (N=2) kerran vuodessa ja 1,8 % (N=1) ei koskaan. (Kuvio 7). Kuviosta voidaan havaita, että vastaajista enemmistö näyttää järjestävän tiedostokokoelmaa melko aktiivisesti.



KUVIO 7 Tiedostokokoelman järjestämisen aikaväli

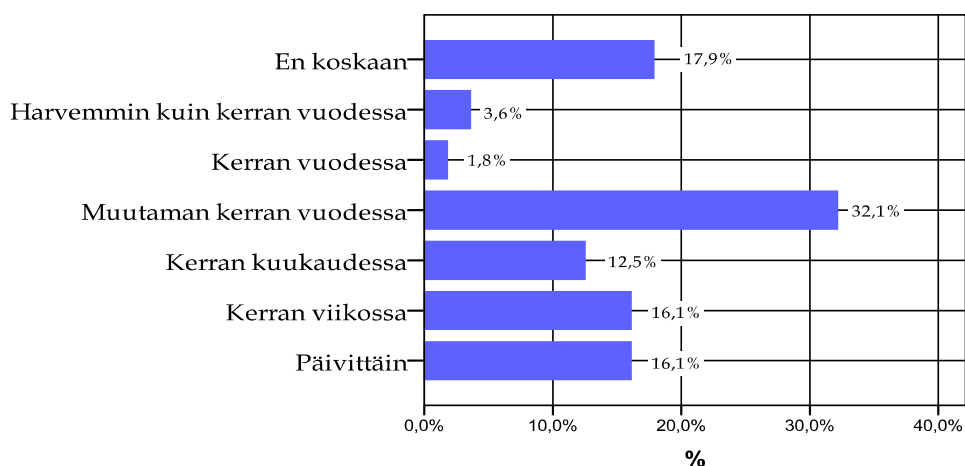
Sähköpostin järjestämisen tapojen havainnot Whittakerin ja Sidnerin (1996) luokittelun pohjalta löytyvät taulukosta 5 ryhmiteltynä tiedekunnan ja sukupuolen mukaan. Luokat jakautuivat vastaajien mukaan siten, että arkistoimattomia oli 33,9 %, toistuvasti arkistoivia 19,6 % ja kevätsiivoojia 44,6 %. 1,8 % vastasi kohtaan muu, mikä, että viestit menevät automaattisesti omiin kansioihinsa. Tiedekunnissa on luokkien kohdalla jonkin verran eroja. Yt-tiedekunnassa on hieman enemmän arkistoimattomia ja satunnaisesti arkistoivia kuin It-tiedekunnassa. Toistuvasti arkistoivia taas on selvästi enemmän It-tiedekunnassa kuin Yt-tiedekunnassa. Sukupuolten välillä ei tiedekuntien kesken ollut kovinkaan paljoa eroa, mutta tiedekuntien sisällä oli joissakin luokissa jonkin verran eroja. It-tiedekunnassa oli arkistoimattomissa ja satunnaisesti arkistoivissa enemmän miehiä kuin naisia. Toistuvasti arkistoivissa ei huomattavaa eroa ollut. Yt-tiedekunnassa taas oli arkistoimattomissa ja satunnaisesti arkistoivissa enemmän naisia. Toistuvasti arkistoivissa määrä oli sama.

Sähköpostin järjestämisen tapojen ja tiedekunnan välinen korrelaatio on ($\rho = 0,04$) ja sukupuolen välinen korrelaatio on ($\rho = 0,024$). Tässäkään tapauksessa muuttujien välillä ei havaittu tilastollisesti merkitsevää riippuvuutta. Käyttökokemuksen ja sähköpostin järjestäminen (ks. liite 2) välinen korrelaatio on ($\rho = 0,145$), mikä on heikosti positiivinen, mutta ei kuitenkaan tilastollisesti merkitsevä.

TAULUKKO 5 Sähköpostin järjestämisen tavat

		Sähköpostin järjestämisen tavat				
Tdk	Sp	Arkistoi- maton N/ %	Toistuvasti arkistoiva N/ %	Kevätsii- vooja N/ %	Muu N/ %	Yhteensä N/ %
IT	n	2	4	3		9
		3,6 %	7,1 %	5,4 %		16,1 %
	m	5	5	8		18
		8,9 %	8,9 %	14,3 %		32,1 %
	yht.	7	9	11		27
		12,5 %	16,1 %	19,6 %		48,2 %
YT	n	7	1	9		17
		12,5 %	1,8 %	16,1 %		30,4 %
	m	5	1	5	1	12
		8,9 %	1,8 %	8,9 %	1,8 %	21,4 %
	yht.	12	2	14	1	29
		21,4 %	3,6 %	25,0 %	1,8 %	51,8 %
Yhteensä		19	11	25	1	56
		33,9 %	19,6 %	44,6 %	1,8 %	100,0 %

Sähköpostikokoelman järjesteleminen ei näyttäisi olevan kovinkaan aktiivista, sillä vastaajista 32,1 % (N=18) järjestelee sähköpostia muutaman kerran vuodessa ja 17,9 % (N=10) ei koskaan järjestele sähköpostia. Vain 16,1 % (N=9) vastaajista ilmoitti järjestävänsä sähköpostiansa päivittäin. Kerran viikossa sähköpostia järjestele 16,1 % (N=9), kerran kuukaudessa 12,5 % (N=7), kerran vuodessa 1,8 % (N=1) ja harvemmin kuin kerran vuodessa 3,6 % (N=2). (Kuvio 8).



KUVIO 8 Sähköpostikokoelman järjestämisen aikaväli

Kirjanmerkkikokoelmissa järjestämisen tapoihin sovellettiin Abramsin (1997) kyselylomaketta. Kirjanmerkkivalikon järjestämisen keinoihin vastaajista 3,6 % (N=2) ilmoitti, ettei kerää kirjanmerkkejä. Muutoin vastaukset jakautuivat niin, että 35,7 % (N=20) vastaajista ei järjestele kirjanmerkkejä, vaan antaa niiden kerääntyä siihen järjestykseen mihin ne on luotu. Vastaajista 28,6 % (N=16) järjestele kirjanmerkkien listaa manuaalisesti, eli vaihtaa kirjanmerkkien järjestystä valikossa, 28,6 % (N=16) luo kansioita, joilla yhdistää samaan asiaan liittyviä kirjanmerkkejä ja 3,6 % (N=2) vastaajista järjestää kirjanmerkkejä hierarkiaan luomalla alikansioita.

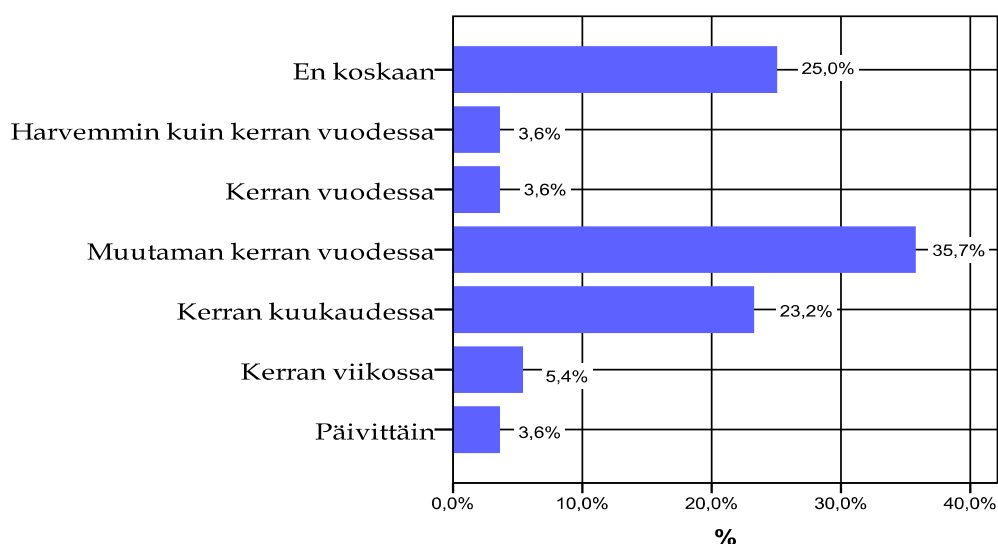
Taulukkoon 6 on kerätty kirjanmerkkien järjestämisen strategiat luokittain tiedekunnan ja sukupuolen mukaan. Havainnot jakautuvat niin, että 21,4 % laitaa uudet kirjanmerkit kansioihin silloin, kun ne on luotu (luomisen aikainen arkistoiija), 48,2 % järjestele kirjanmerkkejä vain ajoittain (satunnainen arkistoiija) ja 30,4 % ei koskaan järjestele kirjanmerkkejä (arkistoimaton). Tiedekuntien osalta suurin ero luokkien kesken löytyy arkistoimattomista, jossa Yt-tiedekunnassa on huomattavasti enemmän havaintoja kuin It-tiedekunnassa. Muissa luokissa erot eivät ole kovin suuria. Sukupuolen osalta ei luokkien kesken ole myöskään suurempia eroja, vaan lukemat ovat melko tasaisia. Eniten eroa löytyy satunnaisista arkistoijista, jossa Yt-tiedekunnassa on hieman enemmän vastaajia, kuin It-tiedekunnassa. Tiedekuntien sisällä on luokissa jonkin verran eroja sukupuolten kesken.

TAULUKKO 6 Kirjanmerkkien järjestämisen tavat

		Kirjanmerkkien järjestämisen tavat			
Tdk	Sp	Luomisen aikainen arkistoiija	Satunnainen arkistoiija	Arkistoimaton	Yht.
		N/ %	N/ %	N/ %	N/ %
IT	n	3	6		9
		5,4 %	10,7 %		16,1 %
	m	5	10	3	18
		8,9 %	17,9 %	5,4 %	32,1 %
	yht.	8	16	3	27
		14,3 %	28,6 %	5,4 %	48,2 %
YT	n	3	5	9	17
		5,4 %	8,9 %	16,1 %	30,4 %
	m	1	6	5	12
		1,8 %	10,7 %	8,9 %	21,4 %
	yht.	4	11	14	29
		7,1 %	19,6 %	25,0 %	51,8 %
Yhteensä		12	27	17	56
		21,4 %	48,2 %	30,4 %	100,0 %

Kirjanmerkkien järjestämisen tapojen ja tiedekunnan välinen korrelaatio on ($\rho = 0,378^{**}$), joka on merkitsevä 99 % luottamustasolla ($p < 0,01$), eli muuttujien välillä näyttäisi olevan jonkinlainen yhteys. Sukupuolen ja järjestämisen tapojen välinen korrelaatio on ($\rho = 0,038$), joka ei ole tilastollisesti merkitsevä. Kirjanmerkkien järjestämisen tapojen ja käyttökokemuksen (ks. liite 2) välillä korrelaatio on ($\rho = 0,367^{**}$), joka on merkitsevä 99 % luottamustasolla ($p < 0,01$). Näin ollen voidaan sanoa, että näiden muuttujien välillä on riippuvuutta.

Kuviossa 9 esiintyvät prosenttijakaumat siitä, kuinka usein vastaajat järjestelivät kirjanmerkkejä. Kirjanmerkkikokoelman osalta voidaan todeta, että suurin osa vastaajista ei järjestele kirjanmerkkejä kovinkaan aktiivisesti. Vastaajista 25 % ei järjestele koskaan kirjanmerkkejään ja 35,7 % vastaajista järjestele vain muutaman kerran vuodessa. Päivittäin kirjanmerkkejä järjestele 3,6 %, kerran viikossa 5,4 %, kerran kuukaudessa 23,2 %, kerran vuodessa 3,6 % ja harvemmin kuin kerran vuodessa 3,6 % vastaajista. Nämä tulokset vastaavat hyvin aikaisemmin käsitellyä luokittelua järjestämisen tavoista, joista todettiin, että suurimmat luokat ovat satunnainen arkistoiija (48,2 %) ja arkistoimaton (30,4 %).



KUVIO 9 Kirjanmerkkikokoelman järjestämisen aikaväli

Yksittäisistä mittarin osioista on mahdollista muodostaa summamuuttujia. Niiden avulla tiivistetään yhteen muuttujaan usean samankaltaista ominaisuutta mittaavan muuttujan sisältämä tieto. (Nummenmaa 2004, 151.) Summamuuttujia muodostettiin lomakkeen Likert-asteikollisista (skaalaus 1-7) muuttujista, jotka koskivat luokittelua ja nimeämistä mittaavia väittämiä. Summamuuttujat muodostettiin käyttäen ns. keskiarvomuttuja -menetelmää (Nummenmaa 2004, 151), jossa kuhunkin summamuuttujaan sisällytettävien muuttujien havaintoarvoista laskettiin keskiarvot. Summamuuttujien reliabiliteettia, eli sisäistä johdonmukaisuutta, arvioitiin käyttäen Cronbachin alfa -kerrointa. Alfakertoimen suuruus perustuu aineiston koon, jakautumisen ja muuttujien määrän suhte-

siin. Summamuuuttujaa voidaan pitää reliaabelina, mikäli alfakerroin on suurempi kuin 0,6. (Metsämuuronen, 2003, s. 47) Luokittelua mittaavat kysymykset korreloivat keskenään (alfa=.867) ja nimeämistä mittaavat kysymykset korreloivat keskenään (alfa=.784), joten summamuuuttujan tekeminen oli mahdollista.

Ennen summamuuuttujien muodostamista muuttujat luokiteltiin uudelleen kolmeluokkaiseksi siten, että luokista 1-2 tuli 1, 3-5 tuli 2 ja 6-7 tuli 3. Muodostuneille uusille luokille annettiin uudet nimet siten, että luokka 1= Huono, 2= Sujuva ja 3= Hyvä. Luokittelua koskevista väittämistä (32a, b, d ja e) muodostettiin uusi muuttuja, jolle annettiin nimi Luokittelu. Vastaavasti nimeämistä koskevista väittämistä (32c ja f) muodostettiin muuttuja Nimeäminen. Uudelleen luokitellut ja muodostetut summamuuuttujat on kuvattu taulukossa 7 luokkien frekvenssien ja prosenttijakaumien kanssa. Taulukosta voidaan havaita, että vastaajat kokevat osaavansa tietojen luokittelun ja nimeämisen pääosin joko sujuvasti tai hyvin.

TAULUKKO 7 Luokittelun ja nimeämisen uusien luokkien frekvenssit

Muuttuja	Huono		Sujuva		Hyvä	
	f	%	f	%	f	%
Luokittelu			30	53,5	26	46,4
Nimeäminen	1	1,8	24	42,8	31	55,4

Tiedon löytäminen

Vastaajat käyttivät päivässä tietojen löytämiseen keskimäärin 7,5 minuuttia (min. 0,00, maks. 60, M 7,66). Muodostettujen luokkien frekvenssit ja prosenttijakaumat jakaantuvat seuraavasti: 0-1 12,5 % (N=7), 2-5 53,6 % (N=30), 6-10 16,1 % (N=9), 11-15 10,7 % (N=6), 16-20 83,6 % (N=2), 21-40 1,8 % (N=1) ja 41-60 1,8 % (N=1). Yhden tiedoston löytämiseen vastaajat käyttivät keskimäärin 1,5 minuuttia (min. 0,00, maks. 10, M 1,57). Luokkiin havainnot jakautuvat siten, että < 1 19,6 % (N=11), 1-2 62,5 % (N=35), 3-4 8,9 % (N=5) ja 5-10 8,9 % (N=5).

Tietojen löytämisen eri tavoista kysyttiin, missä järjestyksessä vastaajat suorittavat mainitut tavat etsiessään tiedostojaan. Vastaukset kustakin löytämisen tavasta ja niiden järjestyksien jakaumista sekä järjestyksestä laskettu keskiarvo näkyvät taulukossa 8. Taulukosta voidaan havaita, että suurin osa vastaajista käyttää ensisijaisesti navigointia etsiessään tiedostoja koneelta. Keskiarvojen mukaan toisena on pikakuvakkeiden käyttäminen, kolmantena hakutoiminto ja neljäntenä tiedoston etsiminen viimeisimmät tiedostot -listasta.

TAULUKKO 8 Tiedostojen etsimisen tapojen järjestyksen jakaumat ja keskiarvot

Tiedon etsimisen tavat	Järjestys				Järjestyksen M
	1	2	3	4	
Navigoiminen	37 (66,1 %)	9 (16,1 %)	7 (12,5 %)	3 (5,4 %)	1,57
Pikakuvake	8 (14,3 %)	22 (39,3 %)	12 (21,4 %)	14 (25,0 %)	2,57
Hakutoiminto	8 (14,3 %)	12 (21,4 %)	15 (26,8 %)	21 (37,5 %)	2,87
Viimeisimmät tiedostot -lista	3 (5,4 %)	13 (23,2 %)	22 (39,3 %)	18 (32,1 %)	2,98

Vastaajilta kysyttiin myös kuinka hyvin he muistavat tiedostojensa sijainnin koneella, johon vastauksena saatiin, että 37,5 % (N=21) muistaa tarkasti sen, missä kansiossa haettu tiedosto on, 58,9 % (N=33) muistaa melko hyvin mistä päin pitäisi etsiä ja 3,6 % (N=2) muistaa jotakin tiedoston sijaintiin liittyvää. Kun kysyttiin sitä käyttävätkö vastaajat hakutoimintoa tiedostojen löytämiseen vaikka he tietäisivät, missä haluttu tiedosto sijaitsee, vastaajista 48,2 % (N=27) ilmoitti ettei käytä ollenkaan, 46,4 % (N=26) ilmoitti käyttävänsä joskus ja 5,4 % (N=3) käyttää aina hakutoimintoa.

Ylläpito

Ensimmäisenä varmuuskopiointista kysyttiin oliko vastaajilta koskaan hävinnyt tietoja laitteistovian tai jonkin muun syyn vuoksi. Tähän vastaajista 67,9 % (N=38) ilmoitti kadottaneensa tietoja ja 32,1 % (N=18) ilmoitti, ettei heiltä ollut hävinnyt tietoja. Näistä 38 vastaajasta, joilta tietoja oli hävinnyt vian seurauksena, 27 oli ollut tiedoistaan varmuuskopiot ja 11 vastaajalla ei ollut tiedoistaan varmuuskopioita. Käytössä oleviin varmuuskopiointiohjelmiin vastauksena saatiin, että 75,0 % (N=42) vastaajista ei ole erityisiä varmuuskopiointiohjelmiä käytössään ja 25,0 % (N=14) ilmoitti käytössään olevan ohjelman. Vastaajilla käytössä olevat varmuuskopiointiohjelmat löytyvät taulukosta 9. Vastauksista Dropbox ja dvd tosin sopivat ehkä paremmin luokiteltaviksi varmuuskopioiden säilytyspaikaksi, kuin varmuuskopiointiohjelmaksi.

Varmuuskopiointitapojen kohdalla 3,6 % (N=2) vastaajista ilmoitti, ettei tee varmuuskopioita. Muuten vastaukset jakautuivat tapojen suhteen siten, että 73,2 % (N=41) vastaajista ilmoitti kopioivansa joistakin tiedostoista ja kansioista kopiot talteen toiseen paikkaan ja 14,3 % (N=8) ilmoitti varmuuskopiointiohjelman hoitavan kopiointin automaattisesti. Muu, mikä -vaihtoehdon vastauksista varmuuskopiointitapoina olivat 1,8 % (N=1) vastaajalla kopioida tiedostot sähköpostiin, 1,8 % (N=1) tekee kopioita kaikista peli- ja mediatiedostoista sekä

koulutöistä, 1,8 % (N=1) polttaa dvd:n ja 1 vastaaja ilmoitti vastauksena RAID 1 + manuaalisesti sekä 1 joka ilmoitti kopioivansa liian harvoin.

TAULUKKO 9 Vastaajilla käytössä olevat varmuuskopiointiohjelmat

Vastaukset	f	%
Cobian Backup	1	1,8
FreeFileSync	1	1,8
Hp advisor	1	1,8
Mac Time Machine	3	5,4
MoxyHome	1	1,8
Norton	1	1,8
Norton Ghost	1	1,8
Windows 7:ssa mukana tuleva	1	1,8
Dropbox	1	1,8
dvd	1	1,8
En muista	2	3,6

Seuraavassa taulukossa 10 esitetään tiivistettynä vastaukset kysymykseen, mitä tietoja vastaajat varmuuskopioivat. Vastaajat saivat valita tässä kysymyksessä useamman vaihtoehdon. Taulukossa 11 näkyvät yhdestä tiedostosta tehtyjen varmuuskopioiden määrät tiedekunnan mukaan jaoteltuina. Suurimmalla osalla vastaajista, 53,6 % (N=30), on vain yksi varmuuskopiokappale tiedoistaan, mikä on melko vähän ja alle yleisten suositusten mukaisen määrän. Tiedekuntien kesken vastaukset jakaantuvat melko tasan, eikä mitään huomattavia eroja ole havaittavissa varmuuskopio määrissä. Ainut isompi ero löytyy niistä, jotka eivät tee yhtään varmuuskopioita, sillä heitä on enemmän yhteiskuntatieteellisen tiedekunnassa. Spearmanin järjestyskorrelaatiokertoimella saatu muuttujien välinen korrelaatio on ($\rho = -0,136$). Muuttujien välillä on heikkoa negatiivista yhteyttä, mutta se ei ole tilastollisesti merkitsevää. Jakaumat kopioiden määristä sukupuolen mukaan näkee taulukosta liitteessä 3.

Vastaajista suurin osa tietää itsekin, että varmuuskopioita on liian vähän, sillä varmuuskopioiden määrän suhteen vastaajista 69,6 % (N=39) tekee mielestään liian vähän varmuuskopioita. Vastaajista 28,6 % (N=16) tekee varmuuskopioita mielestään sopivan määrän ja 1,8 % (N=1) tekee mielestään liian paljon. Tiedekuntien kesken vastaukset jakautuvat hieman eroavasti siten, että mielestään liian vähän varmuuskopioita tekeviä on pieni enemmistö Yhteiskuntatieteellisen tiedekunnassa. Mielestään sopivan määrän varmuuskopioita tekevissä taas on pieni enemmistö Informaatioteknologian tiedekunnan puolella. Spearmanin järjestyskorrelaatiokertoimella saatu muuttujien välinen korrelaatio on ($\rho = -0,153$). Muuttujien välillä on heikkoa negatiivista yhteyttä, mutta se ei ole tilastollisesti merkitsevää. Taulukot mielipiteistä varmuuskopioiden määrän suhteen tiedekunnan, sukupuolen ja käyttökokemuksen mukaan jaoteltuina löytyvät liitteestä 3.

TAULUKKO 10 Varmuuskopioitavat tiedot

Varmuuskopioitava tieto	f	%
Teen muutamista tiedostoista varmuuskopiot	46	82,1
Teen kaikista tiedostoista varmuuskopiot	9	16,1
Teen tietyn ohjelman tiedoista varmuuskopiot	10	17,9
Varmuuskopioin koko järjestelmän tiedot	8	14,3

TAULUKKO 11 Varmuuskopioiden määrä yhdestä tiedostosta

Kopioiden määrä	Tiedekunta				Yhteensä	
	Informaatioteknologiaan tdk		Yhteiskuntatieteellisen tdk		f	%
	f	%	f	%		
0	1	1,8	3	5,4	4	7,1
1	14	25,0	16	28,6	30	53,6
2	9	16,1	8	14,3	17	30,4
3	1	1,8	1	1,8	2	3,6
3+	2	3,6	1	1,8	3	5,4

Taulukossa 12 tarkastellaan eri tallennusvälineitä ja -paikkoja, joissa vastaajat säilyttävät varmuuskopioitansa. Myös tässä kohdassa oli mahdollista vastata useampaan vaihtoehtoon. Taulukosta voidaan havaita, että suurin osa vastaajista säilyttää tietojansa ulkoisella kovalevyllä, muistitikulla tai verkossa. Taulukossa 13 näkyvät tiedot siitä, kuinka usein vastaajat tekevät varmuuskopioita. Tulosten mukaan suurin osa, 35,7 % (N=20) vastaajista, tekee varmuuskopioita vain muutaman kerran vuodessa ja 28,6 % (N=16) tekee niitä kerran kuukaudessa, eli varmuuskopioita tehdään aika harvoin. Myös valtaosa vastaajista itse arvio näin, sillä varmuuskopioiden aikavälistä kysyttäessä vastaajista 73,2 % (N=41) tekee mielestään varmuuskopioita liian harvoin. Vastaajista 26,8 % (N=15) on heidän mielestään varmuuskopioilla sopiva aikaväli. Tiedekuntien välillä ei kummankaan luokan kohdalla ole merkittäviä eroja, vaan lukemat menevät melko tasan. Spearmanin järjestyskorrelaatiokertoimella saatu muuttujien välinen korrelaatio on ($\rho = 0,019$). Muuttujien välillä ei ole tilastollisesti merkitsevää yhteyttä. Taulukot arvioista varmuuskopioinnin aikavälistä tiedekunnan, sukupuolen ja käyttökokemuksen mukaan jaoteltuina löytää liitteestä 3.

TAULUKKO 12 Varmuuskopioiden tallennuspaikat

Tallennuspaikat	f	%
En missään	3	5,4
CD:llä/DVD:llä	15	26,8
Ulkoisella kovalevyllä	38	67,9
Muistitikulla	30	53,6
Verkossa	27	48,2
Muu, mikä	6	10,7
Sähköpostissa	2	3,6
Toisella kovalevyllä	1	1,8
Toisella tietokoneella	1	1,8
Toiselle asemalle	1	1,8
Toisessa kansiossa tai toisella tietokoneella (esim. vanhempinen koneella)	1	1,8

TAULUKKO 13 Varmuuskopioiden tekemisen aikaväli

Aika	f	%
Päivittäin	6	10,7
Kerran viikossa	7	12,5
Kerran kuukaudessa	16	28,6
Muutaman kerran vuodessa	20	35,7
Kerran vuodessa	4	7,1
Harvemmin kuin kerran vuodessa	2	3,6
En koskaan	1	1,8
Yhteensä	56	100,0

Taulukossa 14 näkyvät havaitut frekvenssit ja prosenttijakaumat siitä, kuinka todennäköisenä vastaajat pitävät tietojensa menettämistä laitteistovian tai jonkin muun syyn seurauksena. Taulukosta voidaan havaita, että vastaajista enemmistö painottuu epätodennäköisyyden kannalle, kun 17,9 % mielestä tietokatastrofi on erittäin epätodennäköistä, 28,6 % epätodennäköistä ja 25 % melko epätodennäköistä. Vastaajista 17,9 % taas pitää tietokatastrofia erittäin todennäköisenä. Muuttujan vertailu taustamuuttujiin ristiintaulukoinnilla löytyy alaluvusta 7.4.

TAULUKKO 14 Tietojen häviämisen todennäköisyys laitteistovian sattuessa

Todennäköisyys	f	%
Erittäin epätodennäköistä	10	17,9
Epätodennäköistä	16	28,6
Melko epätodennäköistä	14	25,0
Kohtuullisena	8	14,3
Melko todennäköistä	4	7,1
Todennäköistä	2	3,6
Erittäin todennäköistä	10	17,9
Yhteensä	56	100,0

Vanhojen tai harvemmin käytettyjen tiedostojen osalta vastaajista 51,8 % (N=29) ei tee niille mitään, vaan ne jäävät siihen paikkaan johon ne on viimeksi laitettu. Vastaajista 10,7 % (N=6) siirtää tiedostot kansiorakenteen syrjäisempään osaan varastoon, 8,9 % (N=5) vastaajaa siirtää tiedostot toiselle tallennusvälineelle ja 26,8 % (N=15) vastaajaa poistaa tiedostot koneelta. Muu, mikä -vastauksena yksi vastaaja ilmoitti, että poistaa ne tiedostot, joille ei ole käyttöä, ja jättää muut harvemmin käytetyt tiedostot sinne, missä ne ovat aikaisemminkin olleet. Vanhoja tiedostoja poistaa kansiorakenteesta kerran viikossa 8,9 % (N=5) vastaajista, kerran kuukaudessa 21,4 % (N=12), muutaman kerran vuodessa 44,6 % (N=25), kerran vuodessa 5,4 % (N=3), harvemmin kuin kerran vuodessa 14,3 % (N=8) vastaajista ja 5,4 % (N=3) ei koskaan poista vanhoja tiedostoja kansiorakenteesta.

Kansioiden tai tiedostojen uudelleen nimeämistä tekee päivittäin 3,6 % (N=2) vastaajista, kerran viikossa 3,6 % (N=2), kerran kuukaudessa 30,4 % (N=17), muutaman kerran vuodessa 44,6 % (N=25), kerran vuodessa 5,4 % (N=3), muutaman kerran vuodessa 35,7 % (N=20), harvemmin kuin kerran vuodessa 14,3 % (N=8) vastaajista ja 8,9 % (N=5) ei koskaan muuta tietojensa nimeä. Syitä nimen muuttamiseen on 12,5 % (N=7) nimessä oleva kirjoitusvirhe, 25,0 % (N=14) muuttaa nimeä, koska alkuperäinen ei olekaan sisältöön sopiva, 44,6 % (N=25) muuttaa nimeä tiedoston tai kansion luokittelun muuttuessa ja 14,3 % (N=8) ei muokkaa kansioiden tai tiedostojen nimiä ollenkaan. Muu, mikä-vaihtoehtoon vastasi kaksi, joista toinen ilmoitti uudelleen nimeämisen syyksi eri versioiden lisäämisen ja toinen päivämääräpohjaisen version nimeämisen.

7.4 Tulosten vertailua

Tässä alaluvussa käsitellään aluksi varmuuskopioinnin tietokatastrofimuuttujan ja taustamuuttujien vertailua ristiintaulukoinnin kautta. Tämän jälkeen käsi-

tellään joitakin kerättyjä tuloksia vertaamalla niitä aikaisempien tutkimusten tuloksiin. Lisäksi tehdään joitakin muita vertailuja tämän tutkimuksen tulosten kesken. Tarkastelussa on eri tietokokoelmien kansioden tasojen syvyydet ja kokoelmien kohdalla käytetyt järjestämisen tapojen luokittelut sekä vertailu tietokokoelmien järjestelemisen aikaväleistä. Lisäksi on lyhyt vertailu varmuuskopioinnista.

Ristiintaulukoinnit

Seuraavaksi tarkastellaan ristiintaulukoinnilla varmuuskopiointiin liittyviä mielipiteitä tietokatastrofin todennäköisyydessä, joita on verrattu käyttökokemukseen, tiedekuntaan ja sukupuoleen. Koska aikaisemmin ei kaikissa muuttujien luokissa ollut havaintoja, on analyysin helpottamiseksi luokkia vähennetty seitsemästä kolmeen ja tiedekunnan ja sukupuolen vertailua varten kahteen luokkaan. Ristiintaulukoinnin lisäksi on laskettu muuttujien välinen korrelaatio, jossa käytettiin Spearmanin järjestyskorrelaatiokerrointa (ρ).

Taulukossa 15 vertaillaan käyttökokemuksen vaikutusta vastaajien arviointiin tietokatastrofin todennäköisyydestä, eli tietojen katoamisesta laitteistovian tai jonkin muun tapahtuman vuoksi.

TAULUKKO 15 Taustamuuttuja vertailu käyttökokemus - todennäköisyys tietokatastrofiin

		Todennäköisyys tietokatastrofiin			Yhteensä
		Epätodennäköistä	Jokseenkin todennäköistä	Todennäköistä	
Käyttökokemus	N				
	%				
Huono	N			1	1
	%			100,0 %	100 %
Kohtuullinen	N	1	9	1	11
	%	9,1 %	81,8 %	9,1 %	100 %
Hyvä	N	11	29	4	44
	%	25,0 %	65,9 %	9,1 %	100 %
Yhteensä	N	12	38	6	56
	%	21,4 %	67,9 %	10,7 %	100 %

Khiin neliö-testin käyttöehdot eivät tässä tapauksessa täyttyneet, sillä yli 20 % odotetuista frekvensseistä oli alle viisi, joten tulkintaa ei tässä tapauksessa voida suorittaa "laillisesti". Taulukosta voidaan kuitenkin havaita, että käyttökokemuksensa hyväksi arvioineista vastaajista 25 % (N=11) pitää tietokatastrofia epätodennäköisenä ja 65,9 % (N=29) jonkin verran todennäköisenä. Spearmanin järjestyskorrelaatiokertoimella saatu muuttujien välinen korrelaatio on ($\rho = -$

0,194). Yhteys on negatiivista, mutta heikkoa. Koska otoskoko on pieni, tulosta ei voida pitää riittävästi tilastollisesti merkitseväenä.

Taulukossa 16 verrataan tietokatastrofin todennäköisyyden jakautumista tiedekuntien kesken. Tähän kohtaan on muuttujan tietokatastrofin todennäköisyyden luokat vähennetty kahteen luokkaan. Tässä tapauksessa Khiin neliötestin ehdot täyttyvät ja vähemmän kuin 20 % odotetuista frekvensseistä on alle 5. Khiin neliön p-arvo (0,61) on suurempi kuin 0,05, joten tiedekuntien välillä ei aineiston perusteella näyttäisi olevan eroa todennäköisyyksien arvioissa. Taulukosta voidaan havaita, että vastausten jakaumat menevät puoliksi eli epätodennäköisen ja todennäköisen mielipiteiden kesken vastauksia on saman verran. Informaatioteknologian tiedekunnan osalta 63 % (N=17) vastaajista pitää tietokatastrofia epätodennäköisenä, kun taas yhteiskuntatieteellisestä tiedekunnasta 37,9 % (N=11) on tällä kannalla. Todennäköisyyden kohdalla tämä on taas päinvastoin, sillä 37 % (N= 10) Informaatioteknologiantiedekunnasta pitää tietokatastrofia todennäköisenä ja Yhteiskuntatieteellisen tiedekunnasta 62,1 % (N=18). Spearmanin järjestyskorrelaation mukaan tiedekunnan ja tietokatastrofin todennäköisyyden välillä on yhteyttä ($\rho = 0,328^{**}$, $p < 0,05$). Tuloksen perusteella voidaan sanoa, että tiedekunta vaikuttaa jotenkin arvioon tietokatastrofin todennäköisyydestä.

TAULUKKO 16 Taustamuuttuja vertailu tiedekunta - todennäköisyys tietokatastrofiin

Todennäköisyys tietokatastrofiin				
Tiedekunta		Epätodennä-	Todennä-	
		köistä	köistä	Yhteensä
Informaatioteknologi-	N	17	10	27
an tiedekunta	%	63,0 %	37,0 %	100,0 %
Yhteiskuntatieteellinen	N	11	18	29
tiedekunta	%	37,9 %	62,1 %	100,0 %
Yhteensä	N	28	28	56
	%	50,0 %	50,0 %	100,0 %

Taulukossa 17 tarkastellaan yhteyttä sukupuolen ja tietokatastrofin todennäköisyyden mielipiteen välillä. Tässä Khiin neliötestin ehdot täyttyvät ja vähemmän kuin 20 % odotetuista frekvensseistä on alle 5. Khiin neliön p-arvo (0,07) on hieman suurempi kuin 0,05, mutta ylittää kuitenkin luottamustason kriittisen arvon, joten tiedekuntien välillä ei näyttäisi olevan eroa todennäköisyyksien arvioissa. Spearmanin järjestyskorrelaation mukaan sukupuolen ja tietokatastrofin todennäköisyyden välillä on yhteyttä ($\rho = 0,358^{**}$, $p < 0,01$). Sukupuolella näyttäisi tämän mukaan olevan jotakin vaikutusta siihen, miten todennäköisenä tietokatastrofia pidetään.

TAULUKKO 17 Taustamuuttuja vertailu sukupuoli – todennäköisyys tietokatastrofiin

Sukupuoli		Todennäköisyys tietokatastrofiin		
		Epätodennä- köistä	Todennäköistä	Yhteensä
Nainen	N	8	18	26
	%	30,8 %	69,2 %	100,0 %
Mies	N	20	10	30
	%	66,7 %	33,3 %	100,0 %
Yhteensä	N	28	28	56
	%	50,0 %	50,0 %	100,0 %

Tutkimuksien välinen tulosten vertailu

Kansiorakenteen tasojen syvyydestä on taulukossa 18 esitetty omien tulosten ja Boardmanin (2004, s. 82) saamien tulosten yhteenkerätyt arvot eri tietokokoelmien osalta. Näiden lisäksi Ducheneaut ja Bellotti (2001) havaitsivat 60 sähköpostin käyttäjällä kansioden hierarkian tasojen olevan yleensä syvyydeltään kaksi tasoa. Tiedostokansioissa on molempien tutkimusten vastaajilla eniten kansioden tasoja, ja keskiarvot ovat melko lähellä toisiaan (3,5 ja 3,0). Tiedostokansioden suhteen on maksimin kohdalla näiden tulosten kohdalla eniten eroa, kun maksimi arvot ovat 10 ja 7. Tosin tässäkään ero ei ole mitenkään erityisen huomattava. Sähköpostissa minimi ja maksimit ovat täysin samat ja kirjanmerkeissäkin Boardmanilla on maksimi kansiotaso vain yhtä tasoa pienempi. Keskiarvot ovat sähköposti- ja kirjanmerkkikansioden tasoissa melko lähellä toisiaan. Tämän perusteella voi todeta, ettei Boardmanin havaintoihin verrattuna tuloksissa ole havaittavissa mitään suuria eroavaisuuksia. Sähköpostikansioden osalta on verrattuna Ducheneautin ja Bellottin tulokseen kahdesta tasosta hieman eroa omassa tutkimuksessa havaittuun noin yhteen kansiotasoon. Boardmanin vastaavaan tulokseen, joka on melkein kaksi tasoa, eroa on vähemmän.

TAULUKKO 18 Kansiorakenteen tasojen vertailu

Tutkimus	N	Tiedosto		N	Sähköposti		N	Kirjanmerkki	
		Taso M	min/ maks.		Taso M	min/ maks.		Taso M	min/ maks.
Oma	56	3,5	0/10	41	1,1	0/4	42	1,3	0/4
Boardman (2004)	25	3,0	1/7	23	1,7	0/4	16	1,1	0/3

Tietokokoelmien järjestämisen strategioista käydään vertailussa läpi joitakin aikaisemmin luvussa 5 esitellyistä luokitteluista. Taulukossa 19 näkyvät Hardof-Jaffen ym. (2009) kansioiden järjestämisen luokittelun tulokset verrattuna omiin tuloksiin. Tässä tapauksessa omissa tuloksissa luokittelua on hieman muokattu jakamalla suuren kansion arkistoiijat luokkiin yhden suuren kansion arkistoiijat ja usean suuren kansion arkistoiijat. Lisäksi omassa tutkimuksessa oli 3 vastaajaa, joita ei luokiteltu mihinkään tiettyyn luokkaan ja heidät on merkitty kohtaan muut. Hardof-Jaffen ym. (2009) luokittelussa suurin osa osallistujista on luokiteltu pienten kansioiden arkistoijiksi, kun taas omassa tutkimuksessa vastaajista suurin osa tunnisti itsensä enemmänkin useamman suuren kansion arkistoijaksi.

TAULUKKO 19 Tietoavaruuden järjestämisen strategioiden vertailu

Luokittelu	Hardof-Jaffe ym. (2009)		Omat tulokset	
	f	%	f	%
Kasaaja	141	27	2	3,6
Yhden kansion arkistoija	49	9	2	3,6
Pienten kansioiden arkistoija	262	51	18	32,1
Suuren kansion arkistoija	66	13		
yhden suuren kansion arkistoija			2	3,6
usean suuren kansion arkistoija			29	51,8
Muut			3	5,4
YHTEENSÄ	518	100	56	100

Sähköpostin järjestämisen tavoissa käytettiin Whittakerin ja Sidnerin (1996) sähköpostin järjestämisen luokittelua. Luokittelut jakautuivat oman tutkimuksen 55 vastaajan osalta niin, että arkistoimattomia oli 33,9 % (N=19), toistuvasti arkistoivia 19,6 % (N=11) ja kevätsiivoojia 44,6 % (N=25). Yksi vastaaja ei kuulunut mihinkään mainituista luokista. Whittakerilla ja Sidnerillä (1996) 18 tutkimukseen osallistujan kohdalla arkistoimattomia oli 33,3 % (N=6), toistuvasti arkistoivia oli 27,8 % (N=5) ja kevätsiivoojia oli 38,9 % (N=7). Luokittelujen jakautumisen suhteen ei tutkimuksissa ole paljoa eroa toisiinsa. Molemmissa tapauksissa tulokset ovat luokkien kesken melko samankaltaisia, jossa kevätsiivoojien luokka on suurin, arkistoimattomien on toiseksi yleisin ja pienin luokka on toistuvasti arkistoivat. Tätä luokittelua tukee myös vastaukset siihen kuinka usein sähköpostia järjestellään, jonka mukaan 32,1 % vastaajista järjesteli sähköpostiansa vain muutaman kerran vuodessa ja 17,9 % ei koskaan järjestele sähköpostiansa, kun taas 16,1 % vastaajista järjesteli sähköpostia kerran päivässä ja 16,1 % kerran viikossa.

Kirjanmerkkien järjestämisen tavoissa omia 56 vastaajan tuloksia verrataan Abramsin ym. (1998, s. 45) 299 vastaajan tuloksiin, jotka on merkitty sul-

keisiin. Luokittelujen osalta vastaukset jakaantuivat seuraavasti: arkistoimaton 30,4 % (26,0 %), luomisen aikainen arkistoiija 21,4 % (23,0 %), satunnainen arkistoiija 48,2 % (noin 48 %) ja istunnon jälkeinen arkistoiija 0 % (7 %). Tuloksissa ei ole havaittavissa merkittäviä eroja luokkien jakautumisen kesken, muuten kuin viimeisen luokan kohdalla, jossa omassa tutkimuksessa ei tullut yhtään havaintoja, kun taas Abramsilla ym. tähän luokkaan kuului 7 % vastaajista. Tulosten pohjalta pääteltynä nähtävästi kovin moni ei järjestele kirjanmerkkejä erikseen netin selaamisen lopuksi, vaan kirjanmerkit järjestellään heti kun ne tallennetaan tai sitten myöhemmin erikseen varatulla siivous istunnolla. Kirjanmerkkien määrät oman tutkimuksen ja Abramsilla ym. (1998) tulosten kesken jakaantuivat seuraavasti: ei yhtään kirjanmerkkejä 0 % (6 %), 1-10 17,9 % (10 %), 11-25 17,9 % (24 %), 26-100 50 % (44 %) ja 101-300 14,3 % (14 %) ja 300+ 0 % (2 %).

Tietojen löytämisen suhteen havaittiin taulukossa 8, että navigoiminen on suosittu tapa hakea tietoja kuin hakutoiminnon käyttäminen, joten tämä tulos on yhteneväinen Bergmanin ym. (2008); Boardmanin ja Sassen (2004), Barreaun ja Nardin, (1995) sekä Hendersonin ja Srinivasanin (2011) havaitsemien tulosten kanssa. Navigoimisen suosimiseen luultavasti vaikuttaa se, kuinka hyvin tiedon sijainti muistetaan, sillä vastaajista 37,5 % muistaa tarkasti sen missä kansiossa haettu tieto on, 58,9% muistaa melko hyvin mistä päin tiedostoa pitäisi etsiä ja 3,6 % muistaa jotakin tiedoston sijaintiin liittyvää. Lisäksi vastaajista 48,2 % ilmoitti, ettei käytä ollenkaan hakutoimintoa tiedostojen löytämiseen kun tietävät missä se sijaitsee ja 46,4 % ilmoitti käyttävänsä joskus vaikka he tietäisivät missä haluttu tiedosto sijaitsee. Vain 5,4 % käyttää aina hakutoimintoa, huolimatta siitä muistavatko tiedon sijainnin.

Varmuuskopioiden tekemisen suhteen vastaajista suurin osa oli sitä mieltä, että tekee varmuuskopioita liian harvoin ja liian vähän. Tämä vastaa Haruvyn ja Erevin (2002, s. 605) saamia tuloksia, joissa noin puolet osallistujista teki liian vähän varmuuskopioita. Lisäksi oman tutkimuksen osalta havaittiin, että suurin osa vastaajista piti tietokatastrofia, jossa tietoja häviäisi laitteistovian tai muun ongelman takia, epätodennäköisenä. Havainto vastaa Yechiamin, Haruvyn ja Erevin (2002, s. 601) toteamusta, että usein negatiivisen harvinaisen tapahtuman todennäköisyyttä aliarvioidaan ja positiivisen harvinaisen tapahtuman mahdollisuutta yliarvioidaan.

Boardman ja Sasse (2004, s. 585-586) huomasivat, että tietokokoelmista aktiivisesti kerättiin tiedostoja ja sähköpostia, kun taas kirjanmerkkien kerääminen ei ollut niin tärkeää. Kirjanmerkkikokoelmat ovat kooltaan myös pienempiä kuin tiedosto- ja sähköpostikokoelmat. Samankaltaisia havaintoja tuli myös omasta aineistosta. Esimerkiksi sähköpostin järjestelemisen tavoista voi huomata aikaisemman vertailun mukaan, ettei sähköpostikokoelmia järjestellä kovin ahkerasti. Aikavälien vertailun eri kokoelmien kesken näkee tarkemmin taulukossa 20. Kirjanmerkkikokoelmassakaan järjesteleminen ei ollut kovin aktiivista, sillä suurimmat luokat olivat arkistoimaton 30,4 % ja satunnaisesti arkistoiija 48,2 % ja vastaajista suurin osa järjesteli kirjanmerkkejä kerran kuukaudessa tai vain muutaman kerran vuodessa ja osa taas ei

järjestellyt ollenkaan. Tiedostokokoelmien suhteen taas järjestelemisen tavoista käy ilmi, että 46,4% vastaajista oli toistuvasti arkistoivia ja 51,8 % kevätsiivojia, jotka ajoittain järjestelivät tietojään. Lisäksi havaittiin, että suurin osa vastaajista järjesteli tiedostokokoelmiaan vähintään kerran kuukaudessa tai uammin. Tämän mukaan tätä kokoelmaa järjestellään aika ahkerasti verrattuna kahteen muuhun kokoelmaan. Taulukossa 21 esitellään lasketut keskiarvot sekä minimi ja maksimi tietokokoelmien tietoyksiköiden määristä suurimmassa ja pienimmässä kansiossa. Taulukosta voidaan havaita, että tiedostokokoelma sisältää eniten tietoyksiköitä ja kirjanmerkkikokoelma on kokoelmista kaikkein pienin.

TAULUKKO 20 Tietokokoelmien järjestelemisen aikavälin vertailu

Aika	Tiedosto		Sähköposti		Kirjanmerkki	
	f	%	f	%	f	%
Päivittäin	8	14,3	9	16,1	2	3,6
Kerran viikossa	18	32,1	9	16,1	3	5,4
Kerran kuukaudessa	16	28,6	7	12,5	13	23,2
Muutaman kerran vuodessa	11	19,6	18	32,1	20	35,7
Kerran vuodessa	2	3,6	1	1,8	2	3,6
Harvemmin kuin kerran vuodessa			2	3,6	2	3,6
En koskaan	1	1,8	10	17,9	14	25,0
Yhteensä	56	100	56	100,0	56	100,0

TAULUKKO 21 Tietokokoelmien tietoyksiköiden keskiarvojen vertailu

Kansiot	Tiedosto		Sähköposti		Kirjanmerkki	
	M	min./maks.	M	min./maks.	M	min./maks.
Suurin	12344,42	0/ 260000	118,11	5/ 9001	29,17	6/ 100
Pienin	16,55	0/ 100	19,64	0/ 150	5,80	0/ 20

7.5 Tutkimuksen luotettavuus

Yleistettävyyden kannalta tulee ottaa huomioon tutkimuksen melko pieni otoskoko (N=56). Lisäksi on otettava huomioon tutkimuksen kohdejoukko, joka

koostui kahden eri tiedekunnan opiskelijoista. Opiskelijat ovat yleensä melko nuoria, joten tulokset ovat yleistettävissä lähinnä nuoreen väestöön. Tänä päivänä, varsinkin nuorilla ikäryhmillä on hyvin tietotekniikan osaaminen hallussa, joten tässä tutkielmassa käsitellyt asiat ovat suurelta osalta melko tuttuja. Eri tiedekunnista valituista opiskelijoista osa opiskelee muuta kuin IT- alaa ja osalla on koulutuksen osalta jo enemmän tietämystä tietotekniikan suhteen. Tämän perusteella voisi sanoa, että tulokset ovat yleistettävissä myös eritasoisille tietokoneen käyttäjille. Tosin paremman yleistettävyyden saavuttamiseksi, olisi tulevaisuudessa jatkotutkimuksissa parempi valita kohdejoukkoon eri-ikäisiä ja eri ammateissa olevia henkilöitä.

Tutkimuksen luotettavuuteen vaikuttaa myös käytössä olleet käsitteet, jotka vaikuttavat siihen ovatko tutkimuksessa tarkasteltavat asiat tutkimusalueen ja -aiheen mukaisia. Kyselyssä käytetyt käsitteet ovat pääosin yleisiä tietotekniikassa käytössä olevia termejä, joten niitä voidaan pitää oikeellisina. Myös muita tutkimuksessa käytettyjä termejä (kuten henkilökohtainen tiedonhallinta, tietoyksikkö jne.) on käsitelty teorian kautta, joten niitä voidaan pitää asianmukaisina. Koska tutkimus pohjasi lähinnä teoriassa esiteltyihin aiheisiin, ja aikaisempaan tutkimukseen, voidaan tutkimusta pitää tältä osin luotettavana.

Tutkimuksen sisäistä luotettavuutta voi parantaa se, että kysytyt asiat koskivat osallistujien omia kokemuksia ja tietokokoelmien sisältöjä, jotka ovat osallistujille tuttuja. Tosin tutkijan on vaikea tietää kuinka tosissaan osallistujat ovat kyselyyn suhtautuneet ja kuinka tarkkoja annetut vastaukset ovat. Esimerkiksi suurimmassa kansiossa olevien tiedostojen ja viestien suuri määrä herättää kysymyksen siitä, kuinka tarkkoja vastaukset oikeasti ovat. Näissä määrätietoja mittaavissa kysymyksissä on hieman epätarkkuutta, joka voi vaikuttaa osallistujien antamiin vastauksiin. Tällä kohdalla kysymyksenasettelu ja valittu vastausmuoto eivät ole täysin onnistuneita, ja siten vaikuttavat tutkimuksen luotettavuuteen alentavasti. Lisäksi tutkimukseen luotettavuuteen voi vaikuttaa tutkielman tekijän kokemattomuus tutkimuksen tekemisestä. Tämä tulee näkyviin juuri kyselylomakkeen muodostamisessa ja siinä ilmenneissä epäkohdissa, jotka vaikuttavat koko tutkimuksen luotettavuuteen. Lomaketta muokkaamalla ainakin kysymyksenasettelua tarkentamalla ja vastausmuotoja muuttamalla, voitaisiin tulosten tarkkuutta parantaa, ja siten nostaa tutkimuksen luotettavuutta.

Tutkimuksen reliabiliteettia voidaan myös pitää riittävänä. Tutkimuksessa esitetyt johtopäätökset ja tulokset ovat toistettavissa. Tutkimuksessa käytettyjen mittareiden reliabiliteettia voidaan suurimmalta osin pitää hyvänä, sillä ne pohjautuvat esitettyyn teoriaan ja aikaisempaan empiiriseen tutkimukseen. Tosin jotkin mittarit ovat onnistuneempia kuin toiset mittaamaan sitä, mitä niiden oli tarkoitus mitata.

8 JOHTOPÄÄTÖKSET JA POHDINTA

Teoriaosassa käytiin melko hyvin lävitse henkilökohtaisen tiedonhallinnan keskeinen käsitteistö ja sen pohjalta pystyy saamaan hyvän kuvan siitä aiheesta, jota tässä työssä tarkastellaan. Keskeisimmistä tiedonhallinnan toiminnoista ja niiden menetelmistä annetaan myös melko hyvät kuvaukset, joiden kautta vastataan tutkielmassa annettuun tutkimuskysymykseen siitä, mitä menetelmiä henkilökohtaisessa tiedonhallinnassa käytetään. Lisäksi empiirisellä tutkimuksella saatiin niitä käytännön kokemuksia menetelmistä, mitä tässä työssä myös haettiin.

Tutkielmassa esiteltyjen toimintojen (löytäminen, säilyttäminen, järjestäminen ja ylläpito) lisäksi on kuitenkin muitakin toimintoja, joita ei tässä työssä käyty läpi. Samoin on myös toiminnoille esiteltyjen menetelmien osalta, sillä työssä käytiin lävitse vain kahta menetelmää jokaiselle esitellylle toiminnolle. Koska aihe on niin laaja, ei tässä työssä ole voitu käydä läpi kaikkea sitä, mitä henkilökohtaisesta tiedonhallinnasta on olemassa. Aihetta on rajattu aika tiukasti tutkielman liiallisen laajentumisen estämiseksi ja sen jotenkin käsiteltävässä koossa pitämisen vuoksi, joten tutkielman sisällön voi sanoa olevan vain pintaraapaisu käsiteltävästä aiheesta. Kun tutkimusta tulee vielä jatkuvasti lisää, on raja vedettävä johonkin. Aiheen rajaaminen käsittelemään yhden kotikoneen tiedonhallintaa vaikutti siihen, että työssä ei esimerkiksi käsitelty tiedon hajaantumista (information fragmentation), vaikka se onkin iso ongelma, kun käytössä on useampia koneita sekä lisänä erilaisia mobiililaitteita tai käytetään useampia sähköpostitilejä. Työssä ei myöskään ole yhden koneen rajauksen vuoksi käsitelty tietojen jakamista eri laitteiden välillä, mikä tänäpäivänä myös on melko suuri osa tiedonhallintaa. Tältä osin voidaan todeta, että aiheen käsittely tässä tutkielmassa on hieman kapea. Toiminnoista on olemassa paljon hyvää tutkimusta ja erilaisia lähdemateriaaleja (kuten ainakin kaksi aiheesta kirjoitettua kirjaa), joten siltä osin ei asian käsittely tuo esille mitään erikoisen uutta. Myös menetelmistä löytyy eri lähteistä paljon kirjallisuutta. Tämän tutkielman kannalta voidaan ajatella uudeksi näkökulmaksi menetelmien esittelemistä toimintojen yhteydessä tässä kokonaisuudessa. Lisäksi uutta näkökulmaa tuo oma näkemys järjestelemisen

tapojen aikaisempaa luokittelua koskien, jossa luokittelua muokattiin niin, että yksi vanha luokka jaettiin kahteen luokkaan.

Tutkimuksen kohteena oli tarkastella henkilökohtaisella kotikoneella tapahtuvaa henkilökohtaista tiedonhallintaa. Henkilökohtainen tietokone (PC, Personal Computer) ei nimensä mukaisesti kuitenkaan aina ole niin henkilökohtainen. Tämä johtuu siitä, että kotitaloudessa, jossa asuu useampia henkilöitä, saattaa olla vain yksi tietokone, joka on kaikkien asukkaiden yhteisessä käytössä. Tällöin koneella on tallennettuna useamman henkilön henkilökohtaista tietoa. Tosin tietokoneisiin on mahdollista luoda useammalle käyttäjälle omat käyttäjätilit, jotka voi suojata salasanalla, jolloin kukaan toinen ei pääse käsiksi käyttäjän tietoihin eikä mitään ongelmaa esimerkiksi yksityisyyden suojan kanssa pääse syntymään. Lisäksi tämä takaa sen, että käyttäjän luomat tiedot ja niiden kokoelmat pysyvät siinä muodossa ja järjestyksessä, johon käyttäjä on ne tarkoittanut, kun toiset eivät pääse sekoittamaan niitä. Joissakin talouksissa voi myös olla useampia koneita, jolloin voi olla mahdollista, että useammalla asukkaalla on mahdollisuus henkilökohtaiseen tietokoneeseen. Nykyään monella henkilöllä on myös tavallisen pöytäkoneen lisäksi, tai sijasta, käytössä kannettava tietokone. Kannettavan etuna on, että sen voi ottaa mukaansa mihin tahansa, sillä se kulkee helposti mukana niin kouluun kuin työpaikalle. Kotikoneen käsite on tässä suhteessa hieman kadottanut jotakin aikaisemmasta merkitykustään.

Empiirisessä tutkimus osassa tehdyssä kyselytutkimuksessa tarkasteltiin teoriassa esiteltyjä asioita menetelmistä käytännön kannalta. Kyselyn kautta saatiin vastauksia tutkimuksen osaongelmaan siitä, miten ihmiset käyttävät menetelmiä. Vaikka kyselylomakkeessa havaittiin joitakin puutteita, ja jotkin tulokset eivät välttämättä ole aivan tarkkoja, saatiin kyselyn kautta kerättyä melko kattavaa tietoa tiedonhallinnan menetelmistä käytännössä. Empiirisen tutkimuksen tuloksista ei yleisesti ottaen noussut esiin mitään erityisen poikkeavaa aikaisempiin tutkimuksiin verrattuna. Saadut tulokset ovat melko samansuuntaisia aikaisempien havaintojen kanssa. Tämän perusteella voisi todeta, ettei tiedonhallinta ole vuosien aikana kovin paljon muuttunut, tai ainakaan siinä käytetyt tavat eivät ole juurikaan muuttuneet. Tulosten samansuuntaisuus koskee esimerkiksi tietojen löytämisen tavoissa navigoimisen suosimista haun tekemisen sijaan. Nykyaikaiset hakukoneet ovat tehokkaita, joten navigoimisen suosion syytä voi ehkä hieman ihmetellä. Yksi asiaan vaikuttava tekijä voi olla ihmisten melko tarkka muistikuva siitä minne tiedot on tallennettu koneella, mikä käy ilmi myös tutkimuksen tuloksista. Toinen tekijä voisi mahdollisesti olla, että navigoiminen on vanha tapa, eikä sitä ole koettu tarpeelliseksi vaihtaa, vaikka haku olisikin helpompi tapa etsiä tietoja. Tietojen löytämiseen käytetyt muut tavat, kuten pikakuvakkeiden ja viimeisimmät tiedostot -listan käyttäminen, ovat harvemmin ensimmäisiä tiedon löytämiseen käytettäviä keinoja.

Myös tietokokoelmien hoitamisen ja käsittelemisen suhteen havainnot kävivät yhteen jo aikaisemmin todettujen asioiden kanssa. Tietokokoelmien välillä on havaittavissa eroja niiden järjestelemisen aktiivisuudessa. Myös

kokoelmien sisältämien tietoyksiköiden määrissä on selkeitä eroja kokoelmien välillä. Tiedostokokoelmista todettiin, että suurin osa ihmisistä järjestee näitä tietoja kerran kuukaudessa tai useammin. Sähköpostikokoelmaa taas järjesteltiin lähinnä muutaman kerran vuodessa tai ei koskaan. Kirjanmerkkikokoelmaa järjesteltiin pääosin kerran kuukaudessa ja sitä harvemmin tai sitten ei koskaan. Nämä järjestämisen aikavälejä koskevat tulokset tukevat hyvin havaittuja järjestämisen käyttäytymistapoja. Tietokokoelmien järjestämisen tavoista huomattiin, että tiedostokokoelmissa tavat jakautuivat vain muutaman prosenttiyksikön erolla kevätsiivoojiin ja toistuvasti arkistoiviin. Toisin sanoen ihmiset pääsääntöisesti järjestävät tiedostojansa vain silloin tällöin tai aktiivisesti eri kansioihin kansiorakenteeseen. Sähköpostikokoelmassa suurin osa oli kevätsiivoojia ja toiseksi suurimpana ryhmänä oli arkistoimattomat. Pienin ryhmä tässä kokoelmassa oli toistuvasti arkistoiivat. Eli tässä kokoelmassa ihmiset järjestävät viestejä eri kansioihin vain ajoittain tai sitten toiseksi yleisimmässä tavassa jättävät viestit saapuneet-kansioon, eivätkä järjestele viestejä mitenkään. Vähemmistö ihmisistä järjestee viestejään aktiivisesti eri kansioihin. Kirjanmerkeissä melkein puolet vastaajista kuului satunnaisesti arkistoiviin. Toiseksi suurin ryhmä oli arkistoimattomat ja kolmanneksi tullut ryhmä oli luomisen aikaiset arkistoiijat. Eli ihmiset pääsääntöisesti järjestävät kirjanmerkkejä vain ajoittain tai sitten niitä ei järjestellä ollenkaan, vaan ne jätetään kirjanmerkit-valikkoon siihen järjestykseen, mihin ne on luotu. Tässä tapauksessa vähemmistö ihmisistä näyttäisi laittavan kirjanmerkit kansioihin niiden luomisen aikoihin.

Erot kokoelmien hoitamisessa voivat johtuvat siitä, kuinka paljon mitäkin kokoelmaa käytetään tai sen mukaan, mikä niistä koetaan tärkeimmäksi tietojen kohteeksi. Useimmilla meistä tärkein kokoelma on tiedostokokoelma, johon kuuluu suurin osa koneelle kerätyistä tietoyksiköistä, joista tärkeimpiä ehkä ovat erilaiset dokumentit ja valokuvat. Toiseksi käytetyin kokoelma on sähköposti, johon yleensä kerätään talteen yhteydenpitoon tai tehtäviin liittyviä viestejä ja niiden sisältämää tärkeää tietoa. Kirjanmerkit eivät juuri ole kovinkaan tärkeällä sijalla, kun mietitään kerättyjen kirjanmerkkien määrää tai sitä aikaa, joka käytetään kokoelman järjestelemiseen. Yksi syy kirjanmerkkien vähäiseen keräämiseen voi olla Internetin hyvissä hakukoneissa, joilla tiedot löytää helposti myös myöhemmin.

Kansiorakenteen tarkastelussa käytettiin Hardof-Jaffen ym. (2009) luokittelua, johon tehtiin tässä tutkimuksessa pientä muokkausta. Alkuperäinen luokittelu sisälsi tavat kasaaminen, yhden kansion arkistointi, pienten kansioiden arkistointi ja suuren kansion arkistointi. Tähän luokitteluun lisättiin oman kokemuksen pohjalta kohta, jossa tiedostoja kerätään useampaan suureen kansioon. Tällöin luokittelun kohta suuren kansion arkistointi jaettiin luokitteluiksi yhden suuren kansion arkistointi ja usean suuren kansion arkistointi. Tuloksista käy ilmi, että vastaukset jakautuivat suurimmalta osin kahden luokan välille. Vastaajista selkeästi suurin osa kuului tähän uuteen luokkaan, jossa kansiorakenteessa on paljon kansioita, mutta

suurin osa tietoyksiköistä on sijoitettu muutamaaan suureen kansioon. Toiseksi suurin luokka oli pienten kansioiden arkistointi, jossa kansiorakenne koostuu lähinnä suuresta määrästä pienikokoisia kansioita. Tulokset poikkesivat hieman Hardof-Jaffen ym. (2009) havainnoista, joissa suurin luokka oli pienten kansioiden arkistointi. Useamman suuren kansion arkistointi perustuu osaltaan siihen, että usein eri tiedostomuotojen tietoyksiköt lajitellaan omiin kansioihinsa, kuten esimerkiksi kuvat, musiikki ja asiakirjat. Suuret kansiokokot taas johtuvat tiedon määrän lisääntymisestä. Kansioiden luomisen suhteen havaittiin, ettei suurin osa henkilöistä pitäydy mihinkään tiettyyn tapaan, vaan käyttää useampaa eri tapaa kansioiden tekemiseen. Osa ihmisistä taas mieluummin tekee kansioita samalla, kun tallentaa tietojansa ja osa luo kansioit etukäteen, jotta heillä olisi valmiina paikka, johon laittaa tietoja. Jotkut taas tekevät kansioita vasta jälkikäteen silloin, kun järjestelivät tietojansa. Tavat voivat vaihdella tilanteen tai tarpeen mukaan. Tietojen luokittelamisen ja nimeämisen ihmiset kokevat pääosin sujuvan joko sujuvasti tai hyvin.

Havaintojen mukaan suurin osa teki varmuuskopioita kerran kuukaudessa tai peräti vain muutamana kerrana vuodessa. Useimmat vastaajista totesivatkin tekevänsä liian vähän ja liian harvoin varmuuskopioita, mikä on luultavasti melko yleinen ongelma monen kohdalla. Lisäksi tiedostoista tehtävien varmuuskopioiden määrä oli melko vähäinen, sillä enemmistö vastaajista teki vain yhden kappaleen varmuuskopioita. Toinen suuri ryhmä oli vastaajat, jotka tekivät kaksi varmuuskopioita. Hyvin harva vastaajista teki kolmea tai useampaa varmuuskopioita. Varmuuskopioiden liian vähäinen määrä voi johtua siitä, kuten Yechiam ym (2002) myös mainitsevat, että ihmiset usein aliarvioivat riskin tietokatastrofiin, eivätkä siksi aina varaudu tarpeeksi tehokkaasti tietojen mahdolliseen häviämiseen, jonka voisi ehkäistä tekemällä varmuuskopioita. Usein asiaan havahdutaan vasta, kun vahinko on jo tapahtunut, kuten kävi omallakin kohdalla. Suurimmalla osalla henkilöistä ei ole kotikoneellaan käytössä varmuuskopiointiohjelmaa. Tästä luultavasti johtuu se, että suurin osa ihmisistä tekee varmuuskopiot manuaalisesti. Vain pienellä osalla vastaajista varmuuskopiointi tapahtuu automaattisesti varmuuskopiointiohjelman kautta. Yleisin tapa varmuuskopiointissa on tehdä muutamista tärkeimmistä tiedostoista varmuuskopiot. Toisella sijalla on jonkin ohjelman sisältämien tietojen varmuuskopiointi.

Tietokatastrofin todennäköisyyden arvioissa on yhteneväisyyttä tiedekunnan ($\rho = 0,328^{**}$, $p < 0,05$) ja sukupuolen kanssa ($\rho = 0,358^{**}$, $p < 0,01$). Tämän mukaan tiedekunnalla ja sukupuolella on jotakin vaikutusta siihen, miten todennäköisenä tietokatastrofia pidetään. Varmuuskopioiden määrässä tiedekuntien välillä suurempi ero oli vain siinä ryhmässä, joka ei tehnyt yhtään varmuuskopioita. Tässä ryhmässä enemmistö oli Yhteiskuntatieteellisen tiedekunnan puolella. Arvioissa varmuuskopioiden tekemisen aikavälistä ei tiedekuntien välillä ilmennyt suuria eroja. Tulokset olivat pääsääntöisesti melko tasaisia tiedekuntien kesken. Arvioissa varmuuskopioista tehtävien kopioiden määrästä Yhteiskuntatieteellisen tiedekunnassa oli pieni enemmistö liian vähän varmuuskopioita tekevien ryhmässä. Sopivan määrän varmuuskopioita

tekevien ryhmässä pieni enemmistö taas löytyi Informaatioteknologian tiedekunnan puolelta. Tiedosto- ja sähköpostikokoelmien järjestämisen tavoissa ei havaittu merkittäviä eroavaisuuksia sukupuolen, tiedekunnan tai käyttökokemuksen välillä. Kirjanmerkkikokoelman järjestämisen tavoissa sen sijaan oli merkitsevää korrelaatiota tiedekunnan ($\rho = 0,378^{**}$) ja käyttökokemuksen ($\rho = 0,367^{**}$) välillä luottamustasolla ($p < 0,01$). Toisin sanoen tiedekunnalla ja käyttökokemuksella on jotakin vaikutusta kirjanmerkkien järjestämisen tapoihin. Syitä tulosten jakaantumiseen tiedekuntien välillä on hieman vaikeata arvioida tämän tutkimuksen tulosten valossa, mutta mahdollisella lisätutkimuksella voisi selvittää mitkä tekijät tähän vaikuttavat.

Päivittämiseen liittyen selvisi, että hieman yli puolet vastaajista ei tee vanhentuneille tiedoille mitään, kun taas hieman yli neljännes poistaa vanhat tiedot koneelta. Tietojen poistamisen suorittamisen aikaväleistä havaittiin, että suurin osa poistaa vanhoja tietoja vain muutaman kerran vuodessa. Toiseksi yleisin aikaväli oli kerran kuukaudessa. Myös tietojen uudelleen nimeämistä suurin osa tekee vain muutaman kerran vuodessa. Yleisin syy tiedoston tai kansion uudelleen nimeämiseen on niiden luokittelun muuttuminen.

9 YHTEENVETO

Tämän tutkimuksen tavoitteena oli perehtyä henkilökohtaisen tiedonhallinnan keskeisiin toimintoihin ja niiden menetelmiin. Tutkielmassa myös määriteltiin ja tarkasteltiin henkilökohtaisen tiedonhallinnan keskeisiä käsitteitä. Käsittelyssä pyrittiin etsimään vastauksia tutkimuskysymykseen: Mitä menetelmiä henkilökohtaisessa tiedonhallinnassa käytetään?; sekä osaongelmiin: Mitä luetaan menetelmiksi?, Mitkä menetelmät ovat yleisimmin käytössä? ja Miten ihmiset käyttävät menetelmiä? Tutkielman aihe oli rajattu käsittelemään kotona yhdellä koneella tapahtuvaa digitaalisessa muodossa olevien tietojen tiedonhallintaa.

Tutkielma jaettiin teoriaan ja empiiriseen tutkimus osaan. Teoriaosassa tehdyssä kirjallisuuskatsauksessa käytiin läpi tutkielman aiheen teoriataustaa henkilökohtaisen tiedonhallinnan käsitteistön osalta. Teoriaosuuden jälkeen toteutettiin esitellyn teorian pohjalta kyselytutkimus. Kirjallisuuskatsauksen aluksi perehdyttiin henkilökohtaisen tiedonhallinnan keskeiseen käsitteistöön tarkastelemalla ensin lyhyesti perinteistä tiedonhallinnan termiä, joka toimii pohjana itse henkilökohtaisen tiedonhallinnan käsitteen tarkastelulle. Tiedonhallinta sisälsi kolme näkökulmaa, joista yksi koski henkilökohtaista tiedonhallintaa. Muita näkökulmia olivat organisatorinen ja kirjastonäkökulma. Henkilökohtaisen tiedonhallinnan käsittely aloitettiin tarkastelemalla sitä koskevia eri määritelmiä, joita esiteltiin muutama kappale. Näistä määritelmistä Jonesin ja Boardmanin antamat kuvaukset todettiin parhaimmiksi ja tiedonhallinnan keskeisimmät ominaisuudet huomioiviksi. Määritelmät toimivat myös tutkielman käsittelyn pohjana. Määritelmien jälkeen keskityttiin käymään läpi henkilökohtaiseen tiedonhallintaan liittyviä eri ominaisuuksia ja käsitteitä, jotka yhdessä muodostavat henkilökohtaisen tiedonhallinnan keskeisen sisällön. Tarkastelussa olivat henkilökohtainen tieto, tietoyksikkö, henkilökohtainen tietoavaruus, henkilökohtaisen tiedon kokoelmat sekä henkilökohtaisen tiedonhallinnan toiminnot. Toiminnoista esiteltiin jaottelut Jonesin ja Boardmanin mukaan, joista tutkielmassa otettiin käsiteltäviksi Löytäminen, säilyttäminen, järjestäminen ja ylläpito. Lopuksi esiteltiin tutkielmassa käytettävä menetelmän määritelmä, jossa todettiin niiden olevan

ne tavat, joilla henkilökohtaisen tiedonhallinnan toimintoja lähdetään suorittamaan.

Toimintoja käsittelevissä luvuissa 3-6 saatiin vastauksia siihen millaisia menetelmiä tietokoneen käyttäjillä on yleisimmin käytössään. Löytämisen toimintojen yhteydessä esiteltiin menetelmiksi hierarkkinen navigointi, jonka on useissa tutkimuksissa todettu olevan käyttäjien ensisijainen tapa etsiä tietojansa koneeltaan. Toisena menetelmänä oli haku. Säilyttämisen toiminnoissa menetelmiksi valikoituivat hierarkkinen kansiorakenne ja sähköpostin käyttäminen tiedon säilytyspaikkana. Kansiorakenteen todettiin olevan yleisin käytössä oleva järjestelmä tietojen varastointiin. Sähköpostia todettiin käytettävän myös sen alkuperäisen suunnitellun käytön, eli yhteydenpidon, lisäksi myös eräänlaisena henkilökohtaisen tiedonhallinnan työkaluna, jolla toteutetaan tehtävienhallintaa, tietojen varastointia postilaatikkoon erilaisissa viesteissä sekä yhteystietojen kerääminen. Järjestämisen toimintojen osalta todettiin menetelmiksi tietojen sijoittamisen tavat sekä luokittelu, lajittelu ja nimeäminen. Sijoittelun tavoissa tietojen annetaan joko kerääntyä johonkin koneen määrittelemään oletuskansioon sen kummemmin käyttäjän niitä järjestelemättä tai sitten käyttäjät arkistoivat tietojansa erilaisiin kansiorakenteisiin. Arkistoisissa huomioitiin luokittelu, jossa käytössä on joko pelkät oletuskansiot, niiden alla oleva yksi alikansio, useita pieniä kansioita tai sitten useita kansioita, joista yksi on huomattavan suuri verrattuna muihin. Käyttäjillä todettiin myös olevan erilaisia käyttäytymistapoja sen suhteen miten he tietojansa järjestävät. Osa käyttäjistä ei järjestele tietojansa ollenkaan ja toisilla tavat vaihtelevat aktiivisesta järjestämisestä ajoittaiseen kokoelmien siivoamiseen. Luokittelulla, lajittelulla ja nimeämisellä yhdistetään tietoykiköitä toisiinsa ja muodostetaan niistä omia kokonaisuuksia, joilla rakennetaan tietokokoelmia.

Empiirisessä osassa käsiteltiin kyselytutkimuksen kautta teoriassa läpi käytyjä aiheita henkilökohtaisen tiedonhallinnan toiminnoista sekä niiden menetelmistä. Tutkimuksella haettiin vastausta tutkimuskysymykseen: miten ihmiset käyttävät menetelmiä? Tutkimuksella kautta pyrittiin hakemaan teoriaan lisätietoja menetelmistä ja käytännön kokemuksia niiden käyttämisestä kotikoneilla. Lisäksi tutkimuksella pyrittiin keräämään aineistoa, jota voitaisiin verrata aikaisemmissa tutkimuksissa saatuihin tuloksiin. Tutkimuksessa selvitettiin osallistujien käytössä olevia tiedon kokoelmia ja niiden kansiorakenteita sekä tapoja, joilla he tietoa järjestävät, etsivät ja ylläpitävät.

Tutkimuksen kohdejoukkona toimivat kahden eri tiedekunnan opiskelijat Jyväskylän yliopiston Informaatioteknologian tiedekunnasta ja Yhteiskuntatieteellisen tiedekunnasta. Toisen tiedekunnan opiskelijat otettiin kyselyyn mukaan tulosten vääristymisen välttämiseksi, koska Informaatioteknologiaa opiskelevilla voidaan olettaa olevan paremmat taidot ja tiedot tietotekniikan hallitsemisessa ja osaamisessa. Saadut tulokset ovat pääosin yhtenevät jo aikaisemmissa tutkimuksissa havaittujen asioiden kanssa. Navigoiminen havaittiin tässäkin tutkimuksessa hakua suositummaksi tavaksi etsiä tietoa koneelta, mihin todettiin luultavasti vaikuttaa se, että suurin osa ihmisistä muistaa melko hyvin

sen mihin on tallentanut etsimänsä tiedon. Tietokokoelmien osalta todettiin myös tiedostokokoelmien järjesteleminen aktiivisemmaksi kuin sähköposti- ja kirjanmerkkikokoelmien. Kirjanmerkkikokoelman huomattiin olevan vähiten kerätty ja järjestelty kokoelma. Varmuuskopioiden tekemisessä havaittiin sama ilmiö, että useimmat ihmiset tekevät liian vähän ja harvoin varmuuskopioita. Lisäksi suurin osa pitää tietokatastrofin mahdollisuutta epätodennäköisenä, mikä vahvistaa aikaisemmassa tutkimuksessa todetun ilmiön, että negatiivisen tapahtuman todennäköisyyttä usein aliarvioidaan. Tiedekuntien välillä havaittiin joitakin mielenkiintoisia eroja järjestämisen tapojen ja varmuuskopioinnin yhteydessä, joita voisi jatkossa tutkia vielä tarkemmin.

Tämän tutkielman kohdalla tietoavaruuden järjestelemiseen esiteltyä luokittelua muokattiin siten, että suuren kansion arkistojat luokka jaettiin luokkiin yhden suuren kansion arkistojat ja usean suuren kansion arkistojat. Aikaisemmista tutkimuksista poiketen kansiorakenteiden luokittelussa suurimmaksi luokaksi muodostui usean suuren kansion arkistojat, kun aiemmissä havainnoissa suurin luokka oli pienten kansioiden arkistojat.

Tässä tutkielmassa aihe oli rajattu yhdellä käyttäjän kotitietokoneella tapahtuvaan tiedonhallintaan ja joihinkin siinä käytettäviin menetelmiin. Jatkossa voidaan tämän tutkielman esittämän aineiston pohjalta laajentaa ja tutkia tarkemmin tietojen jakamista ja synkronointia useiden eri koneiden ja mobiililaitteiden välillä ja niissä tapahtuvaan tiedonhallintaa. Lisäksi voitaisiin tarkastella näissä käytettäviä menetelmiä. Jatkotutkimuksissa voitaisiin myös tarkastella menetelmille käytössä olevia ohjelmistoja ja niiden toimivuutta menetelmien käytössä. Tutkielman empiirisessä tutkimuksessa tarkasteltiin opiskelijoiden tiedonhallintaa, joten jatkossa voitaisiin lisäksi laajentaa tutkimuksen kohdejoukkoa tutkimalla myös eri yhteiskunta- ja ikäryhmiä, mikä parantaisi tutkimuksen yleistettävyyttä.

LÄHTEET

- Abrams, D. (1997). *Human factors of personal Web information spaces*. MS Thesis. University of Toronto, Department of Computer Science, Haettu 5.10.2011 osoitteesta <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.16.3348&rep=rep1&type=pdf>
- Abrams, D., Baecker, R. & Chignell, M. (1998). Information archiving with bookmarks: personal Web space construction and organization. Teoksessa *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems* April 18-23 (s.41-48). Los Angeles, USA: Communications of the ACM .
- Barreau, D.K. (1995). Context as a factor in personal information management systems. *Journal of the American Society for Information Science*, 46(5), 327-339.
- Barreau, D.K & Nardi, B.A. (1995). Finding and reminding: file organization from the desktop. *ACM SIGCHI Bulletin*, 27(3), 39-43.
- Bellotti, V., Ducheneaut, N., Howard, M. & Smith, I. (2003). Taking email to task: the design and evaluation of a task management centered email tool. Teoksessa *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems* April 5-10 (s. 345-352). Ft. Lauderdale, Florida, USA: Communications of the ACM.
- Bellotti, V., Ducheneaut, N., Howard, M., Smith, I. & Neuwirth, C. (2002). Innovation in extremis: evolving an application for the critical work of email and information management. Teoksessa *DIS '02: Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques* June 25-28 (s. 181-192). The British Museum, London: Communications of the ACM.
- Bergman, O., Byeth-Marom, R., Nachmias, R., Gradovitch, N. & Whittaker, S. (2008). Improved search engines and navigation preference in personal information management. *Communications of the ACM*, 26(4), Article No. 20, 1-24.
- Boardman, R. (2004). *Improving tool support for personal information management*. PhD Thesis. Imperial College London, Department of Electronic and Electrical Engineering, Intelligent and Interactive Systems Group. Haettu 19.3.2010 osoitteesta <http://www.iis.ee.ic.ac.uk/~rick/thesis/boardman04-thesis.pdf>
- Boardman, R. (2001). Multiple hierarchies in user workspace. Teoksessa *Proceedings of the CHI '01 Conference on Human factors in computing systems. extended abstracts on human factors in computing systems* 31 March-5 April 2001 (s. 403-404). Seattle, Washington.
- Boardman, R. & Sasse, M. A. (2004). "Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management.

- Teoksessa *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems* April 24-29, 2004 (s. 583-590). Vienna, Austria, Communications of the ACM.
- Braman, S. (1989). Defining information: an approach for policymakers. *Telecommunication Policy*, 13 (3), 233-242.
- Bälter, O. (1997). Strategies for organising email. Teoksessa *Proceedings of the HCI '97: HCI on people and computers XII* August 12-15, 1997 (s. 21 - 38). Springer-Verlag London, UK: Communications of the ACM.
- Cole, B. (2005) Search engines tackle the desktop. *Computer*, Volume 38, Issue 3, March 2005, 14-17.
- Detlor, B. (2010). Information management. *International Journal of Information Management*, Volume 30, Issue 2, April 2010, 103-108.
- Ducheneaut, N. & Bellotti, V. (2001). Email as habitat: an exploration of embedded personal information management. *ACM Interactions*, Volume 8 Issue 5, September/October 2001, 30-38.
- Dumais, S., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R. & Robbins, D.C. (2003). Stuff I've seen: a system for personal information retrieval and re-use. Teoksessa *SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval* July 28 - August 1, 2003 (s. 72-79). Toronto, Canada: Communications of the ACM.
- Elsweiler, D. & Ruthven, I. (2007). Towards task-based personal information management evaluations. Teoksessa *SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* July 23-27, 2007 (s. 23-30). Amsterdam, The Netherlands: Communications of the ACM.
- Gwizdka, J. (2004). Email task management styles: the cleaners and the keepers. Teoksessa *CHI '2004: Proceedings of the conference on Human factors in computing systems extended abstracts* April 24-29, 2004 (s. 1235-1238). Vienna, Austria: Communications of the ACM.
- Hardof-Jaffe, S., Hershkovitz, A., Abu-Kishk, H., Bergman, O. & Nachmias R. (2009). How do students organize personal information spaces? Teoksessa *Proceedings of the 2nd International Conference on Educational Data Mining* (s. 250-258). Cordoba, Spain: Educational Data Mining, July 1-3, 2009. Haettu 26.8.2011 osoitteesta <http://www.educationaldatamining.org/EDM2009/uploads/proceedings/hardof.pdf>
- Henderson, S. & Srinivasan, A. (2011) Filing, piling & structuring: strategies for personal document management. Teoksessa *Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS) 2011* . 4-7 January 2011 (s. 1 - 10). Kauai, HI: IEEE Xplore Digital Library
- Hirsjärvi, S., Remes, P. & Sajavaara, P. (2009). *Tutki ja Kirjoita*. (15. uud. painos). Helsinki: Tammi
- Jones, W. (2008). *Keeping found things found: The study and practice of personal information management*. (1. painos). Morgan Kaufmann publishers.
- Jones, W. (2007). Personal information management. *Annual Review of Information Science and Technology*, Volume 41, Issue 1, 2007, 453-504

- Jones, W. & Teevan, J. (2007). Introduction. Teoksessa W. Jones & J. Teevan (toim.), *Personal information management*. Seattle: University of Washington Press.
- Kawai, H. & Sandoh, H. (2003). An efficient backup warning policy for a hard disk. *Computers & Mathematics with Applications*, Volume 46, Issue 7, October 2003, 1055-1063
- Kaye, J., Vertesi, J., Avery, S., Dafoe, A., David, S., Onaga, L., Rosero, I. & Pinch, T. (2006). To have and to hold: exploring the personal archive. Teoksessa *Proceedings of the CHI '06: SIGCHI conference on Human Factors in computing systems* April 22-27, 2006 (s. 275 - 284). Montreal, Canada: Communications of the ACM.
- Kelly, D. & Teevan, J. (2007). Understanding what works: Evaluating PIM tools. Teoksessa W. Jones & J. Teevan (toim.), *Personal information management* (s. 190-204). Seattle: University of Washington Press.
- Kirk, J. (2005). Information in organizations: directions for information management. Teoksessa E. Macevičiute & T. D. Wilson (toim.), *Introducing information management: an information research reader* (s. 3-17). London: Facet Publishing.
- Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics*, 19 (1) 55-66. Haettu 20.3.2010 osoitteesta <http://simson.net/ref/1988/Lansdale88.pdf>
- Lu, C-T., Shukla, M., Subramanya, S.H. & Wu, Y. (2007). Performance Evaluation of Desktop Search Engines. Teoksessa *Proceedings of the IRI 2007, IEEE International Conference on Information Reuse and Integration* August 13-15, 2007 (s. 110 - 115). Las Vegas, IL. IEEE Xplore Digital library
- Ma, S. & Wiedenbeck, S. (2009). File management with hierarchical folders and tags. Teoksessa *Proceedings of the CHI '09: 27th international conference extended abstracts on Human factors in computing systems* April 4-9, 2009 (s. 3745-3750). Boston, USA: Communications of the ACM.
- Macevičiute, E., & Wilson, T. D. (2005). The development of the information management research area. Teoksessa E. Macevičiute & T. D. Wilson (toim.), *Introducing information management: an information research reader* (s. 18-30). Facet Publishing.
- Mackay, W.E. (1988) More than just a communication system: diversity in the use of electronic mail. Teoksessa *Proceedings of the CSCW'88 Conference on Computer-Supported Cooperative Work* September 26-28 (s. 344-353). Portland, Oregon, USA: Communications of the ACM.
- Malone, T.W. (1983). How do people organize their desks? Implications for the design of office information systems. *ACM Transactions on Office Information Systems*, 1 (1), 99-112.
- Metsämuuronen, J. (2003). *Tutkimuksen tekemisen perusteet ihmistieteissä*. (2. uud. painos). Jyväskylä: Gummerus Kirjapaino Oy.
- Nummenmaa, L. (2004). *Käyttäytymistieteiden tilastolliset menetelmät*. (1.-3. painos). Helsinki: Tammi

- Russell, D. & Lawrence, S. (2007). Search everything. Teoksessa W. Jones & J. Teevan (toim.), *Personal information management*. Seattle (s. 153-166). University of Washington Press.
- Teevan, J., Capra, R. & Pérez-Quiñones, M. (2007). How people find personal information. Teoksessa W. Jones & J. Teevan (toim.), *Personal information management*. Seattle: University of Washington Press.
- Whittaker, S., Bellotti, V. & Gwizdka, J. (2007). Everything through email. Teoksessa W. Jones & J. Teevan (toim.), *Personal information management* (s. 167-189). Seattle: University of Washington Press.
- Whittaker, S., Bellotti, V. & Gwizdka, J. (2006). Email in personal information management. *Communications of the ACM*, 49(1), 68-73.
- Whittaker, S., Terveen, L., & Nardi, B. A. (2000). Let's stop pushing the envelope and start addressing it: a reference task agenda for HCI. *Human Computer Interaction*, 15, 75-106.
- Whittaker, S. & Sidner, C. (1996). Email overload:exploring personal information management of email. Teoksessa R., Bilger, S. Guest & M. J. Tauber (toim.), *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground* April 13-18, 1996 (s. 276-283). Vancouver, British Columbia, Canada: Communications of the ACM.
- Wilson, T. (2000). Human information behavior. *Informing Science*, 3(2), 49-55.
- Yechiam, E., Haruvy, E. & Erev, I. (2002) Reinforcement learning and the prevention of data catastrophes. *Journal of Managerial Psychology*, 17(7), 599 - 611

LIITE 1 KYSELYLOMAKE

Saatekirje

Olen Jyväskylän yliopiston tietojärjestelmätieteen opiskelija. Teen graduani henkilökohtaisen tiedonhallinnan menetelmiin liittyen. Tämä kysely kuuluu olennaisena osana graduuni.

Henkilökohtaisella tiedonhallinnalla tarkoitetaan niitä tapoja joilla ihmiset hankkivat, varastoivat, järjestävät ja hakevat digitaalista tietoa henkilökohtaisessa tietokoneen käyttöympäristössään sekä töissä että kotona. Henkilökohtaisen tiedonhallinnan toimintoja ovat mm. tietojen löytäminen, järjestäminen, säilyttäminen ja ylläpito. Menetelmät ovat niitä tapoja, joita käytetään näistä eri toiminnoista suoriutumiseen. Tämä tutkimus pyrkii selvittämään mitä menetelmiä tietokoneiden käyttäjillä on käytössään ja mitä tapoja niiden käyttämiseen liittyy.

Tärkein kohteeni ja tavoitteeni tässä tutkimuksessa on selvittää, miten kotikoneilla käytetään eri tiedonhallinta menetelmiä. Kyselyssä keskitytään tiedostojen ja kansioden rakenteeseen, tietojen järjestämiseen ja löytämiseen koneelta sekä tietojen ylläpitoon.

Pyydän teitä ystävällisesti antamaan hetkisen ajastanne ja vastaamaan kysymyksiin. Vastaaminen kestää noin 15 minuuttia.

Vastaukset käsitellään nimettöminä ja luottamuksellisesti. Vastatkaa kysymyksiin valitsemalla haluamanne vaihtoehto, joka sopii parhaiten omalle kohdalle. Jos mikään valmiista vaihtoehdoista ei ole sopiva, kirjoita lyhyt kuvaus kohtaan "Muu".

Kiitos vastauksistanne jo etukäteen!

Terveisin
Minna Kivioja

Vastausaika on 28.11.2011 - 12.12.2011, 13.12-16-12-2011

Kysely löytyy alla olevasta linkistä:

<http://MRINTERVIEW2.ad.jyu.fi/mrIWeb/mrIWeb.dll?I.Project=KYSELYLOMAKE5>

TAUSTATIEDOT

1. Tiedekuntasi on?

- Informaatioteknologian tiedekunta
 Yhteiskuntatieteellinen tiedekunta

2. Sukupuolesi on?

- Nainen
 Mies

3. Mikä on tietokoneen käyttökokemuksesi?

- Erittäin huono Melko hyvä
 Huono Hyvä
 Melko huono Erittäin hyvä
 Kohtuullinen

4. Mikä käyttöjärjestelmä koneellasi on käytössä?

- En osaa sanoa Macintosh
 Windows Muu, mikä
 Linux

5. Mitä Internet -selainta ensisijaisesti käytät?

- Mozilla Firefox
 Internet Explorer
 Muu, mikä

6. Mitä sähköpostin hallintatyökalua käytät?

- Minulla on käytössä vain sähköpostitili
 MS Outlook
 Mozilla Thunderbird
 Lotus Notes
 Muu, mikä

7. Montako tuntia keskimäärin käytät tietokonetta päivässä? (0-24)

KANSIORAKENNE

Kirjanmerkki on web-selaimeen tallennettava web-sivun URL-osoite, jonka kautta voidaan siirtyä suoraan URL-osoitteessa sijaitsevalle web-sivulle.

8. Keräätkö web-selaimesi kirjanmerkkivalikkoon/suosikit-listaan web-sivujen osoitteita talteen?

- Kyllä, minulla on selaimessani kerättynä lista kirjanmerkkejä
 En, minulla ei ole selaimessani kerättynä kirjanmerkkejä
 En osaa sanoa

En osaa sanoa

9. Paljonko sinulla on kirjanmerkkejä?

- Ei yhtään 26 - 100
 1 - 10 101 - 300
 11- 25 300+

10. Käytätkö kirjanmerkkien järjestämiseen kansioita?

- Kyllä
 En

11. Käytätkö sähköpostiviestien varastointiin kansioita?

- Kyllä, laitan viestejä talteen kansioihin
 En, kaikki viestini ovat postilaatikossa

12. Kuinka paljon sinulla on sähköpostiviestejä yhteensä?

Kansiorakenne on yläkansion ja sen alla olevien kansioden muodostama kokonaisuus. Juurikansio on kansiorakenteen ylin kansio. Oletuskansio on käyttöjärjestelmän määrittämä kansio, kuten omat tiedostot.

13. Mikä seuraavista vaihtoehdoista kuvaa parhaiten kansiorakenteitasi?

- Minulla on lähinnä vain oletuskansioita käytössäni
 Minulla on juurikansion alla yksi kansio, jossa on paljon tiedostoja
 Minulla on paljon pieniä kansioita, joissa on muutamia tiedostoja
 Minulla on useita kansioita, mutta suurin osa tiedostoista on yhdessä suuressa kansiossa
 Minulla on paljon kansioita, joista muutama on melko suuria
 Muu, mikä

14. Milloin luot kansioita?

- Luon kansioita samalla, kun tallennan tiedostoja
 Teen kansioita etukäteen, jotta minulla olisi paikka johon tallentaa tiedostoja
 Teen uusia kansioita silloin, kun järjestelen tietoja eikä minulla ole ennestään tehtynä niille sopivaa kansiota
 Vähän kaikkia edellisiä yhdessä
 Muu, mikä

15. Arvioi kuinka paljon kansiosi sisällöistä keskimäärin on alikansioita? (määrä % 0-100)

16. Arvioi kuinka paljon kansiosi sisällöistä keskimäärin on tiedostoja? (määrä % 0-100)

17. Kuinka syviä kansiorakenteita sinulla on tiedostoillasi?
(kun juurikansio on 0, alikansio 1, sen alikansio 2 jne.)

18. Paljonko tiedostoja on kansiorakenteesi suurimmassa kansiossa?

19. Paljonko tiedostoja on kansiorakenteesi pienimmässä kansiossa?

20. Kuinka syviä kansiorakenteita sinulla on sähköpostikansioillasi?
(kun juurikansio on 0, alikansio 1, sen alikansio 2 jne.. Jos et käytä kansioita, siirry kysymykseen 23)

Ei vastausta

21. Kuinka paljon viestejä sinulla on suurimmassa sähköpostikansiossasi?

Ei vastausta

22. Kuinka paljon viestejä sinulla on pienimmässä sähköpostikansiossasi?

Ei vastausta

23. Kuinka syviä kansiorakenteita sinulla on kirjanmerkkikansioillasi?
(kun juurikansio on 0, alikansio 1, sen alikansio 2 jne.. Jos et käytä kansioita, siirry kysymykseen 26)

Ei vastausta

24. Kuinka paljon sinulla on kirjanmerkkejä suurimmassa kirjanmerkkikansiossasi?

Ei vastausta

25. Kuinka paljon kirjanmerkkejä on pienimmässä kirjanmerkkikansiossasi?

Ei vastausta

[Edellinen](#) [Seuraava](#)

TIETOJEN JÄRJESTÄMINEN

26. Paljonko tietokoneella käyttämästäsi ajasta päivässä kuluu tietojen järjestämiseen? (vastaus minuuteissa)

27. Mikä seuraavista vaihtoehdoista kuvaa tiedostojesi järjestämisen tapaa parhaiten?

- En järjestele tiedostoja, ne kerääntyvät oletuskansioihin
- Järjestelen aktiivisesti tiedostoja kansiorakenteen eri kansioihin
- Järjestelen tiedostoja vain silloin tällöin
- Muu, mikä

28. Mikä seuraavista vaihtoehdoista kuvaa sähköpostisi järjestämisen tapaa parhaiten?

- En järjestele viestejä, ne kerääntyvät postilaatikoon
- Järjestelen viestejä aktiivisesti eri kansioihin
- Järjestelen viestejä vain silloin tällöin
- Muu, mikä

29. Mikä seuraavista vaihtoehdoista kuvaa kirjanmerkkiesi järjestämisen tapaa parhaiten?

- En järjestele kirjanmerkkejä, ne kerääntyvät siihen järjestykseen mihin ne on luotu
- Järjestelen kirjanmerkkejä listaan manuaalisesti (vaihtaa kirjanmerkkien järjestystä valikossa)
- Luon kansioita joilla yhdistän samaan asiaan liittyviä kirjanmerkkejä
- Järjestän kirjanmerkkejä hierarkiaan luomalla alikansioita
- En kerää kirjanmerkkejä
- Muu, mikä

30. Milloin järjestele kirjanmerkkejäsi?

- Laitan uudet kirjanmerkit kansioihin silloin, kun ne on luotu
- Järjestelen kirjanmerkkejä uudelleen Internetin selaamisen loppuksi
- Järjestelen kirjanmerkkejä vain ajoittain
- En koskaan järjestele kirjanmerkkejä

31. Kuinka usein järjesteleet eri tietokokoelmiasi?

	Päivittäin	Kerran viikossa	Kerran kuukaudessa	Muutaman kerran vuodessa	Kerran vuodessa	Harvemmin kuin kerran vuodessa	En koskaan
31a Tiedosto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31b Sähköposti	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31c Kirjanmerkki	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Luokittelu tarkoittaa prosessia jossa yhdistetään tiedostoja eri aiheita koskeviin kansioihin ja luodaan niistä samaa asiaa koskevia ryhmiä ja kokonaisuuksia.

32. Arvioi seuraavia väittämiä

	Erittäin huonosti	Huonosti	Melko huonosti	Kohtalaisesti	Melko hyvin	Hyvin	Erittäin hyvin
32a Osaan luokitella uuden tiedoston sopivaan kansioon	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32b Pystyn luomaan aiheista sopivia kokonaisuuksia	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32c Pystyn nimeämään tiedoston/kansion aiheeseen sopivaksi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32d Onnistun yhdistämään samaa aihetta koskevia tiedostoja yhteen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32e Kaikki kansion sisältö kuuluu samaan aihepiiriin	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32f Tiedoston/kansion nimi vastaa sen sisältöä	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Edellinen

Seuraava

TIEDON LÖYTÄMINEN

33. Paljonko tietokoneella käyttämästäsi ajasta päivässä kuluu tiedostojen etsimiseen? (vastaus minuuteissa)

34. Millä tavoin etsit koneelle tallennettua tiedostoa? Laita tavat järjestykseen 1-4 sen mukaan miten käytät niitä.

- 34a Navigoin siihen kansioon, jossa tiedosto on
- 34b Käytän pikakuvaketta, jonka olen luonut
- 34c Käytän haku-toimintoa
- 34d Haen tiedoston viimeisimmät tiedostot –listasta

35. Kuinka hyvin muistat tiedostojesi sijainnin koneella?

- Muistan tarkasti sen missä paikassa ja kansiossa hakemani tiedosto on
- Muistan melko hyvin mistä päin etsiä tiedostoa
- Muistan jotakin tiedoston sijaintiin liittyvää
- Minulla ei ole mitään muistettavaa tiedoston sijainnista

36. Arvioi kuinka monta minuuttia keskimäärin käytät yhden tiedoston etsimiseen?

37. Käytätkö haku-toimintoa vaikka tietäisit tiedoston sijainnin?

- En ollenkaan
- Joskus
- Käytän aina hakua

Edellinen

Seuraava

YLLÄPITO

38. Onko sinulta koskaan hävinnyt tietoja jonkin laitteistovian tai muun syyn vuoksi?

- Kyllä
 Ei

39. Oliko sinulla varmuuskopioita hävinneistä tiedoista?

- Kyllä
 Ei

40. Onko käytössäsi varmuuskopiointiohjelmaa?

- Ei
 Kyllä, mikä

41. Mikä seuraavista kuvaa parhaiten varmuuskopiointitapojasi?

- En tee varmuuskopioita
 Kopioin joistakin tiedostoista ja kansioista kopiot toiseen paikkaan talteen
 Varmuuskopiointiohjelma hoitaa kopioinnin automaattisesti
 Muu, mikä

42. Mitä tietoja varmuuskopioit? (voit valita useamman vaihtoehdon)

- Teen muutamista tiedostoista varmuuskopiot
 Teen kaikista tiedostoista varmuuskopiot
 Teen tietyn ohjelman tiedoista varmuuskopiot
 Varmuuskopioin koko järjestelmän tiedot

43. Kuinka monta varmuuskopio kappaletta sinulla on yhdestä tiedostosta?

- 0 3
 1 3+
 2

44. Missä säilytät varmuuskopioitasi? (Voit valita useamman vaihtoehdon)

- En missään Muistitikulla
 CD:llä/DVD:llä Verkossa
 Ulkoisella kovalevyllä Muu, mikä

45. Kuinka usein teet varmuuskopioita?

- Päivittäin Kerran vuodessa
 Kerran viikossa Harvemmin kuin kerran vuodessa
 Kerran kuukaudessa En koskaan
 Muutaman kerran vuodessa

46. Mikä seuraavista mielestäsi kuvaa varmuuskopioidesi määrää?

- Teen liian vähän varmuuskopioita
 Teen liian paljon varmuuskopioita
 Varmuuskopioita on sopiva määrä

47. Mikä seuraavista mielestäsi kuvaa varmuuskopioidesi aikaväliä?

- Teen liian harvoin varmuuskopioita
 Teen liian usein varmuuskopioita
 Varmuuskopioilla on sopiva aikaväli

48. Kuinka todennäköisenä pidät tietojesi menettämistä laitteistovian tai virheen vuoksi?

- Erittäin epätodennäköistä Melko todennäköistä
 Epätodennäköistä Todennäköistä
 Melko epätodennäköistä Erittäin todennäköistä
 Kohtuullisena

49. Mitä teet harvemmin käyttämillesi tiedostoille? Valitse seuraavista yleisimmin käyttämäsi vaihtoehto.

- En mitään, ne pysyvät siinä mihin, olen ne aikaisemmin laittanut
- Siirrän tiedostot kansiorakenteessa syrjäisempään paikkaan varastoon
- Siirrän tiedostot toiselle tallennusvälineelle (cd/dvd, ulkoinen kovalevy jne.)
- Poistan ne
- Muu, mikä

50. Kuinka usein poistat kansiorakenteestasi vanhoja tiedostoja?

- Päivittäin
- Kerran viikossa
- Kerran kuukaudessa
- Muutaman kerran vuodessa
- Kerran vuodessa
- Harvemmin kuin kerran vuodessa
- En koskaan

51. Kuinka usein nimeät uudelleen kansioitasi tai tiedostojaasi?

- Päivittäin
- Kerran viikossa
- Kerran kuukaudessa
- Muutaman kerran vuodessa
- Kerran vuodessa
- Harvemmin kuin kerran vuodessa
- En koskaan

52. Minkä vuoksi nimeät uudelleen kansioitasi tai tiedostojaasi?

- Nimessä on kirjoitusvirhe
- Alkuperäinen nimi ei olekaan sisältöön sopiva
- Tiedoston tai kansion luokittelu muuttuu
- En muokkaa tiedostojen tai kansioiden nimiä
- Muu, mikä

[Edellinen](#)

[Seuraava](#)

53. Jos sinulla on jotain lomakkeessa oleviin kysymyksiin liittyvää tarkennettavaa, voit kirjoittaa sen tähän.

Eivastausta

KIITOS VASTAUKSISTASI!

[Edellinen](#)

[Seuraava](#)

LIITE 2 TAULUKOT KÄYTTÖKOKEMUKSESTA JA TIETOKO- KOELMIEN JÄRJESTÄMISEN TAVOISTA

Käyttökokemus - tietojen järjestämisen tavat

Järjestämisen tavat	Käyttökokemus			
	Huono	Kohtuullinen	Hyvä	Yhteensä
	f %	f %	f %	f %
Kasaaja	0 ,0 %	0 ,0 %	2 3,6 %	2 3,6 %
Yhden kansion arkistoija	0 ,0 %	0 ,0 %	2 3,6 %	2 3,6 %
Pienten kansioden arkistoija	1 1,8 %	2 3,6 %	15 26,8 %	18 32,1 %
Yhden suuren kansion arkistoija	0 ,0 %	2 3,6 %	0 ,0 %	2 3,6 %
Usean suuren kansion arkistoija	0 ,0 %	7 12,5 %	22 39,3 %	29 51,8 %
Muu	0 ,0 %	0 ,0 %	3 5,4 %	3 5,4 %
Yhteensä	1 1,8 %	11 19,6 %	44 78,6 %	56 100,0 %

Käyttökokemus - tiedostokokoelman järjestämisen tavat

Tiedostojen järjestämisen tavat	Käyttökokemus			
	Huono	Kohtuullinen	Hyvä	Yhteensä
	f %	f %	f %	f %
Arkistoimaton	0 ,0 %	0 ,0 %	1 1,8 %	1 1,8 %
Toistuvasti arkistoiva	1 1,8 %	5 8,9 %	20 35,7 %	26 46,4 %
Kevätsiivooja	0 ,0 %	6 10,7 %	23 41,1 %	29 51,8 %
Yhteensä	1 1,8 %	11 19,6 %	44 78,6 %	56 100,0 %

Käyttökokemus – sähköpostikokoelman järjestämisen tavat

Sähköpostin järjestämisen tavat	Käyttökokemus			
	Huono	Kohtuullinen	Hyvä	Yhteensä
	f %	f %	f %	f %
Arkistoimaton	1 1,8 %	3 5,4 %	15 26,8 %	19 33,9 %
Toistuvasti arkistoiva	0 ,0 %	0 ,0 %	11 19,6 %	11 19,6 %
Kevätsiivooja	0 ,0 %	7 12,5 %	18 32,1 %	25 44,6 %
Muu, mikä	0 ,0 %	1 1,8 %	0 ,0 %	1 1,8 %
Yhteensä	1 1,8 %	11 19,6 %	44 78,6 %	56 100,0 %

Käyttökokemus – kirjanmerkkikokoelman järjestämisen tavat

Kirjanmerkkien järjestämisen tavat	Käyttökokemus			
	Huono	Kohtuullinen	Hyvä	Yhteensä
	f %	f %	f %	f %
Luomisen aikainen arkistoija	0 ,0 %	0 ,0 %	12 21,4 %	12 21,4 %
Istunnon jälkeinen arkistoija	0 ,0 %	0 ,0 %	0 ,0 %	0 ,0 %
Satunnainen arkistoija	0 ,0 %	5 8,9 %	22 39,3 %	27 48,2 %
Arkistoimaton	1 1,8 %	6 10,7 %	10 17,9 %	17 30,4 %
Yhteensä	1 1,8 %	11 19,6 %	44 78,6 %	56 100,0 %

LIITE 3 TAULUKOT VARMUUSKOPIOINTI MUUTTUJISTA

Varmuuskopioiden lukumäärä - sukupuoli

Kopioiden määrä	Sukupuoli					
	Nainen		Mies		Yhteensä	
	F	%	f	%	f	%
0	1	1,8 %	3	5,4 %	4	7,1 %
1	15	26,8 %	15	26,8 %	30	53,6 %
2	7	12,5 %	10	17,9 %	17	30,4 %
3	1	1,8 %	1	1,8 %	2	3,6 %
3+	2	3,6 %	1	1,8 %	3	5,4 %
Total	26	46,4 %	30	53,6 %	56	100,0 %

Arvio varmuuskopioiden määrästä - sukupuoli

Mieli pide varmuuskopioiden määrästä	Sukupuoli					
	Nainen		Mies		Yhteensä	
	f	%	f	%	f	%
Teen liian vähän varmuuskopioita	20	35,7 %	19	33,9 %	39	69,6 %
Teen liian paljon varmuuskopioita	0	0,0 %	1	1,8 %	1	1,8 %
Varmuuskopioita on sopiva määrä	6	10,7 %	10	17,9 %	16	28,6 %
Yhteensä	26	46,4 %	30	53,6 %	56	100,0 %

Arvio varmuuskopioiden määrästä - tiedekunta

Mieli pide varmuuskopioiden määrästä	Tiedekunta					
	Informaatioteknologian tiedekunta		Yhteiskuntatieteellinen tiedekunta		Yhteensä	
	f	%	f	%	f	%
Teen liian vähän varmuuskopioita	17	30,4 %	22	39,3 %	39	69,6 %
Teen liian paljon varmuuskopioita	0	0,0 %	1	1,8 %	1	1,8 %
Varmuuskopioita on sopiva määrä	10	17,9 %	6	10,7 %	16	28,6 %
Yhteensä	27	48,2 %	29	51,8 %	56	100,0 %

Arvio varmuuskopioiden määrästä - käyttökokemus

Mieli pide varmuuskopioiden määrästä	Käyttökokemus							
	Huono		Kohtuullinen		Hyvä		Yhteensä	
	f	%	f	%	f	%	f	%
Teen liian vähän varmuuskopioita	0	0,0 %	10	17,9 %	29	51,8 %	39	69,6 %
Teen liian paljon varmuuskopioita	0	0,0 %	0	0,0 %	1	1,8 %	1	1,8 %
Varmuuskopioita on sopiva määrä	1	1,8 %	1	1,8 %	14	25,0 %	16	28,6 %
Yhteensä	1	1,8 %	11	19,6 %	44	78,6 %	56	100,0 %

Arvio varmuuskopioiden aikavälistä - sukupuoli

Mieli pide varmuuskopioiden aikavälistä	Sukupuoli					
	Nainen		Mies		Yhteensä	
	f	%	f	%	f	%
Teen liian harvoin varmuuskopioita	20	35,7 %	21	3,7,5 %	41	73,2 %
Teen liian usein varmuuskopioita	0	0,0 %	0	0,0 %	0	0,0 %
Varmuuskopioilla on sopiva aikaväli	6	10,7 %	9	16,1 %	15	26,8 %
Yhteensä	26	46,4 %	30	53,6 %	56	100,0 %

Arvio varmuuskopioiden aikavälistä – tiedekunta

Mieli-pide varmuuskopioiden aikavälistä	Tiedekunta					
	Informaatioteknologiaan tiedekunta		Yhteiskuntatieteellinen tiedekunta		Yhteensä	
	f	%	f	%	f	%
Teen liian harvoin varmuuskopioita	20	35,7 %	21	37,5 %	41	73,2 %
Teen liian usein varmuuskopioita	0	0,0 %	0	0,0 %	0	0,0 %
Varmuuskopioilla on sopiva aikaväli	7	12,5 %	8	14,3 %	15	26,8 %
Yhteensä	27	48,2 %	29	51,8 %	56	100,0 %

Arvio varmuuskopioiden aikavälistä – käyttökokemus

Mieli-pide varmuuskopioiden aikavälistä	Käyttökokemus							
	Huono		Kohtuullinen		Hyvä		Total	
	f	%	f	%	f	%	f	%
Teen liian harvoin varmuuskopioita	0	0,0 %	8	14,3 %	33	58,9 %	41	73,2 %
Teen liian usein varmuuskopioita	0	0,0 %	0	0,0 %	0	0,0 %	0	0,0 %
Varmuuskopioilla on sopiva aikaväli	1	1,8 %	3	5,4 %	11	19,6 %	15	26,8 %
Total	1	1,8 %	11	19,6 %	44	78,6 %	56	100,0 %

Spearmanin järjestyskorrelaatio - Arvio varmuuskopioiden määrästä

			Correlations		
			Arvio varmuuskopioiden määrästä	Käyttökokemus	Sukupuoli
Spearman's rho	Arvio varmuuskopioiden määrästä	Correlation Coefficient	1,000	,134	,139
		Sig. (2-tailed)	.	,324	,309
		N	56	56	56
	Käyttökokemus	Correlation Coefficient	,134	1,000	,218
		Sig. (2-tailed)	,324	.	,107
		N	56	56	56

Sukupuoli	Correlation Coefficient	,139	,218	1,000
	Sig. (2-tailed)	,309	,107	.
	N	56	56	56

Spearmanin järjestyskorrelaatio - Arvio varmuuskopioiden aikavälistä

Correlations

			Arvio varmuuskopioiden aikavälistä	Käyttökokemus	Sukupuoli
Spearman's rho	Arvio varmuuskopioiden aikavälistä	Correlation Coefficient	1,000	-,091	,078
		Sig. (2-tailed)	.	,505	,568
		N	56	56	56
	Käyttökokemus	Correlation Coefficient	-,091	1,000	,218
		Sig. (2-tailed)	,505	.	,107
		N	56	56	56
Sukupuoli	Correlation Coefficient		,078	,218	1,000
	Sig. (2-tailed)		,568	,107	.
	N		56	56	56

Spearmanin järjestyskorrelaatio - Varmuuskopioiden määrä

Correlations

			Varmuuskopioiden lkm	Käyttökokemus	Sukupuoli
Spearman's rho	Varmuuskopioiden lkm	Correlation Coefficient	1,000	,107	-,043
		Sig. (2-tailed)	.	,432	,754
		N	56	56	56
	Käyttökokemus	Correlation Coefficient	,107	1,000	,218
		Sig. (2-tailed)	,432	.	,107
		N	56	56	56
Sukupuoli	Correlation Coefficient		-,043	,218	1,000
	Sig. (2-tailed)		,754	,107	.
	N		56	56	56