

# KUVIEN KUVAILUN JA -HALLINNAN KEHITTÄMINEN

Tapauksena Tampereen yliopiston viestintäyksikön digitaalinen kuva-arkisto

Johanna Nevalainen

Tampereen yliopisto  
Informaatiotieteiden yksikkö  
Informaatiotutkimus ja interaktiivinen media  
Pro gradu -tutkielma  
Toukokuu 2015

Tutkimuksessa selvitetään digitaalisessa muodossa oleviin kuviin ja niiden kuvailuun liittyviä käytäntöjä ja ongelmia. Tutkimuskohteena on Tampereen yliopiston viestintäyksikön digitaalisessa muodossa oleva valokuva-arkisto. Tutkimus on laadullinen tapaustutkimus, jossa kuvaillaan yhden digitaalisen kuva-arkiston piirteitä. Tutkimus on toteutettu osana Tampereen yliopiston viestintäyksikön kuva-arkiston kehittämishanketta.

Tutkielmassa kartoitetaan Tampereen yliopiston viestintäyksikön digitaalisen valokuva-arkiston käyttötarpeita, sisältöä ja kuvan kuvailun periaatteita. Aineisto kerättiin vuosina 2011 ja 2014 haastattelujen ja havainnoinnin avulla. Tutkimukseen osallistui seitsemän henkilöä, jotka toimivat Tampereen yliopistossa tiedotustehtävissä. Lisäksi haastattelujen avulla kerättiin tutkimukseen osallistuneiden kokemuksia kuva-arkiston käytöstä ja heidän toivomuksiaan arkiston kehittämiseksi.

Haastattelujen pohjalta ilmeni, että keskeisiä kehitystoiveita olivat yhtenäinen arkistointikäytäntö, arkistointiohjeet sekä arkiston kuvaus. Lisäksi toivottiin nykyistä helpompaa hakua, vaivatonta kuvien jakamista ulkopuolelle sekä mahdollisesti arkistohallintaohjelmaa. Useimmat haastateltavista toivoivat myös kuviin lisää kuvailutietoja.

Tutkimuksessa ilmeni, että kuva-arkiston keskeiset ongelmat ja kehittämistarpeet liittyvät kuvien löytymiseen ja kuvailutietoihin. Arkistossa ilmenee epäyhtenäisiä käytäntöjä kuvien järjestämisen ja kuvailun suhteen. Kansioihin tallennettujen kuvien löytäminen perustuu paljolti työntekijöiden muistinvaraiseen tietoon, sillä arkistossa ei ole hakutoimintoa.

Tutkimuksessa selvisi, että niissä kansioissa, joissa kuvailu on yhdenmukaisempaa, noudatettiin Shatfordin (1986) lukuohjeessa esittämiä kuvailun attribuutteja: kuka, mitä, missä ja milloin. Kuvailuelementeistä etenkin milloin (päivämäärä) esiintyi paljon kuvailussa. Jotta kuvat löytyisivät paremmin, niihin pitäisi liittää enemmän mitä ja kuka-kysymyksiin vastaavia metatietoja.

Tutkimuksessa vertaillaan perinteisen, ammattimaisesti järjestetyn arkiston ja internetin kuvasovelluksissa tapahtuvan avoimen avainsanoituksen piirteitä. Vertailun pohjalta esitetään kehitysnäkökulmia kuva-arkiston kehityshankkeen ja haastateltavien huomioiden puitteissa. Jatkotutkimusta arkistossa voitaisiin toteuttaa kokeilemalla kuvanhallintaohjelmaa.

Avainsanat: kuvien kuvailu, digitaalinen kuva-arkisto



# Sisällysluettelo

1	JOHDANTO .....	1
2	KIRJALLISUUSKATSAUS .....	4
2.1	Käsitteelliset lähtökohdat ja aiemman tutkimuksen tuloksia .....	4
2.2	Kuvailu ja haku.....	4
2.3	Kuvan kuvailun ja haun menetelmät .....	6
2.4	Valokuvan erityispiirteet dokumenttina .....	9
2.5	Valokuvien kuvailun tutkimus ja kuvailumallit .....	13
2.6	Yhteisöllinen sisällönkuvailu verkossa .....	20
2.6.1	Käyttäjät avainsanoittajina valokuvapalveluissa .....	21
2.6.2	Avoin avainsanoitus kuvapalveluiden kuvatulvan suodattajana ....	22
2.6.3	Käyttäjien avainsanoitus Flickr -kuvapalvelussa .....	23
2.6.4	Avainsanapilvet, tagipelit ja avainsanasuositukset.....	24
3	TUTKIMUSTEHTÄVÄ JA TUTKIMUSMENETELMÄ .....	28
3.1	Tutkimuskysymykset.....	28
3.2	Tutkimuksen lähtökohdat ja tutkimuksen toteutus.....	29
3.3	Empiirisen aineiston keruu .....	31
3.4	Aineiston analyysi .....	33
4	KUVA-ARKISTON NYKYTILANNE .....	36
4.1	Viestintäyksikön digitaalisen arkiston käyttötarpeet.....	36
4.2	Arkiston sisältö .....	38
4.3	Kuvan elinkaari: miten kuva päättyy arkistoon? .....	39
4.4	Kuvatyytit Tay:n viestintäyksikön kuva-arkistossa.....	41
5	KUVIEN KUVAILUKÄYTÄNNÖT ARKISTOSSA .....	44
5.1	Kuvien kuvailun työnjako .....	44
5.2	Kuvien kuvailun elementit .....	45
5.2.1	Esimerkki kuvan kuvailusta arkistossa.....	46
5.2.2	Henkilö .....	48
5.2.3	Päivämäärä.....	48
5.2.4	Tiedostonimi: kuka, mitä, missä ja milloin?.....	49
5.2.5	Kameran metatiedot.....	50
5.3	Kuvan kuvailun kehittäminen käyttäjien silmin.....	51
6	TULOSTEN YHTEENVETO JA TULKINTA .....	54
7	ARKISTON ONGELMIA JA KEHITYSNÄKÖKULMIA .....	57
8	KUVAILUN KEHITTÄMINEN NETTISOVELLUSTEN TAPAAAN.....	62
9	YHTEENVETO JA JOHTOPÄÄTÖKSET .....	68
	Lähteet .....	72

## LIITTEET

# 1 JOHDANTO

Kommunikaation historiassa kuva on tehnyt mielenkiintoisen paluun keskeiseksi viestintävälineeksi. Kuvan sisältö on nopea omaksua ja kuvia on nykyisin nopeaa jakaa. Kuvien sisällönkuvailun kehittämistä tarvitaan kuva-aineiston valtavan lisääntymisen takia. Karhula kirjoitti pro gradu -työssään jo 1990-luvun alussa että kuvien hallinta on lukutaitoon verrattava elinehto. (Karhula 1991, 3.) Karhulan huomion jälkeen on tapahtunut suoranainen kuvien vallankumous: kaikki voivat jakaa kuvia kaikkialla, kaikille kaiken aikaa. Digitaalisessa muodossa olevien kuvien määrä kasvaa koko ajan. Käyttäjystävälliset sovellutukset, kuten digitaalisten kameroiden ja kuvankäsittelyohjelmien kehittyminen, sekä sosiaalisen median palvelut ovat helpottaneet kuvien tuottamista ja jakamista. Yhä useampi ihminen voi tuottaa sekä jakaa kuvia yhä nopeammin.

Digitaalisissa kuva-arkistoissa ja kokoelmissa (kuten esimerkiksi lehtitalojen kokoelmat) kuvia pyritään järjestämään löydettäviksi kokonaisuuksiksi. Kuvien sisällönkuvailu on hajautunut erilaisiin organisaatioihin, joilla kaikilla on erilaiset käytännöt sisällönkuvailuun (esim. Jörgensen 2003.) Digitaalisten arkistojen tehtäväkuvaan on tulossa uutena vaatimuksena lisääntyvä käyttäjien suora pääsy digitaalisiin kuvakokoelmiin, kuten museoiden digitaalisiin kuvakokoelmiin (esim. Markkula & Sormunen 2000).

Kuvien sisällönkuvailun tutkimuksen ja kuvahallinnan (*image management*) kehittämisen tarve on lisääntynyt huomattavasti, sillä digitaalisten kuva-arkistojen pitäisi kyetä hallitsemaan kuvatulvaa sekä parantamaan kuvien löydettävyyttä. Tutkimuksellinen mielenkiinto onkin suuntautunut enemmän verkossa oleviin kuviin ja avoimiin kuvapalveluihin kuin sisäisiin, suljettuihin valokuva-arkistoihin. Tulevaisuudessa on odotetta-

vissa näiden kahden käytännön yhdistymistä verkossa olevien kuvapankkien ja säilytysratkaisujen, kuten pilvipalveluiden myötä.

Tässä tutkimuksessa lähestytään digitaalista valokuva-arkistoa kuvien sisällönkuvailun tutkimuksen näkökulmasta. Tutkimuksessa esitellään Tampereen yliopiston viestintäyksikön digitaalisen valokuva-arkiston käyttötarpeita, sisältöä ja kuvan elinkaarta. Lisäksi esitellään tutkimukseen osallistuneiden kokemuksia kuva-arkiston käytöstä ja heidän toivomuksiaan arkiston kehittämiseksi. Tutkimuksessa esitellään myös digitaalisten valokuvien kuvailuun liittyvää problematiikkaa sekä näkökulmia yhteisölliseen sisällönkuvailuun verkossa.

Tässä tutkimuksessa haetaan vastauksia seuraaviin kysymyksiin:

- Mitkä ovat Tampereen yliopiston viestintäyksikön digitaalisen valokuva-arkiston käyttötarpeet?
- Minkälaisista kuvista (kuvatyyppit) arkiston sisältö koostuu ja millainen on kuvan elinkaari arkistossa?
- Millaisia ovat kuvan kuvailun käytännöt valokuva-arkistossa?
- Millaisia kehitysnäkökulmia internetin kuvanhallintasovellukset voivat tarjota arkistolle?

Tutkimus lähti liikkeelle vuonna 2010 lukiessani sähköpostiviestin Tampereen yliopiston viestintäyksikön tarjoamasta projektista joka kohdistui kuva-arkiston kehittämiseen. Näin siis tutkimukselle oli valmista tilausta organisaatiopuolelta. Aihe vastasi suunnitelmaani graduaihetta. Tutkimus on toteutettu osana Tampereen yliopiston viestintäyksi-

kön kuva-arkiston kehittämishanketta. Aineisto kerättiin vuosina 2011 ja 2014 haastattelujen ja havainnoinnin avulla.

Luvussa 2 esittelen alan aiempaa tutkimusta, teoreettisia malleja sekä keskeisiä käsitteitä. Luvussa 3 esittelen tutkimuksen lähestymistavan, keskeiset tutkimuskysymykset sekä aineistonkeruumenetelmän. Luvussa 4 aloitan analyysiosuuden kuvailemalla Tampereen yliopiston viestintäyksikön digitaalisen kuva-arkiston nykytilanteen. Luvussa 5 käsittelen arkiston työntekijöiden työnjakoa kuvien kuvailun suhteen ja esittelen kuvien kuvailun käytäntöjä. Lisäksi esittelen kuvien kuvailuelementtejä ja niiden esiintymistä arkistossa. Luvussa 6 teen yhteenvedon tuloksista ja luvussa 7 tarkastelen arkiston ongelmia ja kehitysnäkökulmia tutkimuskirjallisuuden näkökulmasta. Luvussa 8 esittelen kehitysnäkökulmia kuvailun nettisovelluksien tutkimuksen näkökulmasta. Luvussa 9 teen yhteenvedon tutkielmasta ja esittelen keskeiset johtopäätökset.

## **2 KIRJALLISUUSKATSAUS**

### **2.1 Käsitteelliset lähtökohdat ja aiemman tutkimuksen tuloksia**

Luvussa 2.2 esittelen kuvailun ja hakemisen peruskäsitteet. Luvussa 2.3 esittelen kuvan sisällönkuvailun menetelmät. Luvussa 2.4 tarkastelen valokuvaa eri näkökulmista tuoden esille sen erityispiirteet dokumenttina. Luvussa 2.5 esittelen valokuvan sisällönkuvailun aiempaa tutkimusta ja malleja. Lopuksi luvussa 2.6 tarkastelen yhteisöllisen sisällönkuvailun menetelmiä verkossa. Alaluvussa 2.6.1 hahmottelen yhteisöllisen sisällönkuvailun menetelmiä sekä niiden tulevaisuuden mahdollisuuksia kuvien sisällönkuvailun kehittämisessä. Alaluvussa 2.6.2 esittelen avoimen avainsanoituksen mahdollisuuksia kuvatulvan suodattajana. Alaluvussa 2.6.3 esittelen käyttäjien avainsanoitusta Flickr-kuvapalvelussa ja lopulta alaluvussa 2.6.4 käyttäjien tekemän avainsanoituksen kehittämistä.

### **2.2 Kuvailu ja haku**

Kuvailu tarkoittaa dokumentin piirteiden kuvaamista määrätyn kaavan asettamissa puitteissa. Bibliografisessa valvonnassa voidaan tehdä erottelu kuvailuun ja hakuelementteihin. Kuvailtavaan dokumenttiin liitetään tietoja eli hakupisteitä, jotka helpottavat dokumentin hakua edistäen sen identifointia ja valintaa. Dokumentin sisältö pyritään siis selittämään sanoin. Sisällönkuvailu on toiminto, joka tukee dokumenttien hakua sisällöllisin perustein. Myös luokitus ja indeksointi ovat sisällönkuvailua. (Suominen ym. 2009, 44-45.)



Sisällönkuvailussa on keskeistä pohtia, mistä dokumentissa on varsinaisesti kysymys. Esimerkiksi kirjan ollessa kyseessä kysytään, mitä teos varsinaisesti haluaa sanoa, tai mikä teoksen yleinen sanoma on. Tällainen sisällön analyysi on eräänlaista karsintaa, jossa dokumentin sisällöstä yritetään saada esiin ydinsanoma tai keskeisin sisältö. Sisällönkuvailulla pyritään siis tuottamaan metadataa, jonka avulla tietyn sisältöiset dokumentit löytyvät. Metadata tai metatieto tarkoittaa tietoa tiedosta, eli dokumentista tai dokumentin sisällöstä. (Suominen ym. 2009, 140–142.)

Käsitte pohjainen kuvailu on ihmisen tekemää dokumentin sisällönanalyysiä ja kuvailutietojen valintaa. Tätä käsitteisiin perustuvaa kuvailua kutsutaan intellektuaaliseksi kuvailuksi (*intellectual/manual indexing*). Käsitte pohjaisessa kuvailussa voidaan käyttää joko luonnollista kieltä eli avainsanoitusta, tai erityistä dokumentaatiokieltä. Käsitte pohjaisessa kuvailussa tutustutaan dokumenttiin, josta poimitaan dokumenttia kuvaavia sanoja. Näistä termeistä analysoidaan olennaiset käsitteet, jotka voidaan vielä kääntää dokumentaatiokielelle. (Karhula 1991, 48–50.)

Dokumentaatiokieli on sanojen hierarkkisiin suhteisiin perustuvaa, strukturoitua metadataa eli kuvailutietoa, jolla on tietty säännöstö ja kontrolloitu sanasto. Dokumentaatiokielten avulla on pyritty antamaan sisällönkuvailijalle ja hakijalle yhteinen kieli, jotta haussa ja tallennuksessa käytetyt ilmaisut kohtaisivat paremmin. Esimerkiksi *thesaurus* on asiasanasto, jossa termien suhteet on määritelty standardoidulla tavalla. (Järvelin & Sormunen 2010, 186–187.)

Avainsanoitus on luonnollisella kielellä tapahtuvaa kuvailua, jossa dokumentista poimitaan sanoja tai fraaseja. Avainsana (*keyword*) on merkityksellinen sana tai termi, joka kuvaa dokumentin sisältöä. (Järvelin & Sormunen 2010, 186–187.)

Digitaalisen kuva-arkiston voi määritellä nollina ja ykkösinä, eli numeerisessa muodossa olevaksi kuvakokoelmaksi. Digitaalisen kuvan edut verrattuna paperilla tai filmillä olevaan aineistoon perustuvat kuvien esittämiseen lukuina (*bit*), jolloin aineistoa voidaan tallentaa, käsitellä ja jakaa tietokoneympäristössä. (Vainio 1996, 4–5.)

Digitaalinen dokumentti on ns. rakenteinen dokumentti, eli siinä on tietoa automaattisesti tunnistettavissa osissa eli kentissä. Kuvan kuvailutiedoista voidaan hakea sanallisesti (merkkijonohaut), koska sanat muodostavat merkkijonoja. (Suominen ym. 2009, 332.)

Haun menetelmät voidaan eritellä komentohauksi ja selailuhauksi. Komentohaun hakukomento muodostuu merkkijonoista. Esimerkiksi yksi sana on merkkijono, jonka perusteella voi hakea. Myös sanojen välit ovat osa merkkijonoa. Merkkijonoja löytyy vain tekstistä, ei kuvista. Komentohaussa saadaan usein tulokseksi suoraan toivottu dokumentti tai sen viite. (Suominen ym. 2009, 39-40.) Selailuhaussa hakija voi silmäillä erityisen sisällön mukaisia listauksia. Hakija voi tehdä selailunäkymän (esim. luettelo) perusteella valintoja mihin dokumenttiin haluaa tutustua paremmin. (Suominen ym. 2009, 39-40.)

### **2.3 Kuvan kuvailun ja haun menetelmät**

Tässä alaluvussa esittelen jaon kuvien kuvailun jaon käsitte pohjaiseen ja sisältöpohjaiseen menetelmään. Tässä kirjallisuuskatsauksessa painopiste on käsitte pohjaisella kuvailulla, sillä intellektuaalisen kuvailun on todettu parantavan hakujen tuloksellisuutta silloin, kun dokumenteissa on luonnostaan vähän tekstiä, kuten valokuvissa. (Järvelin & Sormunen 2010, 188–189.)

Kuvatiedonhaussa voidaan soveltaa niin sanalliseen metadataan, kuin visuaaliseenkin sisältöön perustuvia hakumenetelmiä. Kuvatiedonhaku perustuu enemmän selailuun kuin tekstidokumenttien haku. Koska ihminen omaksuu kuvan keskeisen sisällön nopeasti, ovat pienet selailukuvat (*thumbnails*) kätevä menetelmä etsittäessä tarvittavaa kuvaa tulosjoukosta. Käytännössä kuvien sanalliseen metadataan kohdistetaan kyselyjä ja relevantti kuva etsitään silmäilemällä tulosjoukon pieniä selailukuvia. (Järvelin & Sormunen 2010, 200.) Esimerkiksi Googlen kuvahaku ([www.google.fi/images](http://www.google.fi/images)) on paljon käytetty hakukone.

Käsitteiden indeksointi on ollut ihmisten laatimaa, koska kuvien objektien tunnistaminen automaattisesti on osoittautunut vaikeaksi. Kuva itsessään ei sisällä tekstin tapaan merkkijonoja joita voisi tunnistaa. Kuvan sisältämien käsitteiden kuvailua onkin kehitetty kontrolloidun sanaston sovelluksen avulla käyttämällä luonnollisen kielen kuvailua kuvasta. (Rasmussen 1997, 179.)

Automaattinen hahmopohjainen indeksointi on menetelmä, jossa kuvasta tuotetaan kuvankäsittelyn keinoin automaattisesti kuvan visuaalista sisältöä edustavia piirrevektoreita (*feature vectors*). Tämä sisältöpohjainen kuvailu on suhteellisen uusi tekniikka, joka perustuu kuvan dataan ja siinä operoidaan pikselitasolla. Kyselynä voidaan antaa mallikuva, josta laskettuja piirrevektoreita täsmäytetään tietokannan kuvien vastaaviin vektoreihin. Sisältöpohjaiset hakutavat ovat hyödyllisiä hakiessa kuvia, joita ei osata sanallisin keinoin kuvata. (esim. Järvelin & Sormunen 2010, 202; Jörgensen 2003.)

Sisältöpohjaisen kuvahaun sovelluksissa kysely on tyypillisesti toinen kuva. Sovelluksia on kolmella eri tasolla. Järjestelmä voi toimia matalan tason kuvaominaisuuksien tasolla, kuten väri, tekstuuri ja muodot. Toisen tason järjestelmä toimii myös semanttisella käsitteiden tasolla, jolloin järjestelmä kykenee tunnistamaan kuvasta erityisiä kohteita.

Nykyään pystytään tunnistamaan hevonen, ihmisiä tai puita, jopa tietynlainen ympäristö, kuten kaupunkimaisema. Kolmannella tasolla, jota eivät kuvahakujärjestelmät ole vielä saavuttaneet, järjestelmä tunnistaa abstrakteja seikkoja. Yksi harvoista tutkimuksista tällä tasolla on tutkimukset värien alueelta, jossa yritetty tunnistaa automaattisesti, sopivatko jotkut värit yhteen. (Kaakinen 2008, 7.)

Käytännössä kuvahakukoneet, kuten muutkin hakukoneet keräävät keräilyohjelmien (*harvesting programmes*) avulla indeksin www-sivuilta ja sijoittavat sen hakupalvelun kiintolevyille. Tämän jälkeen sivua tai kuvaa haettaessa hakukone tekee haun omaan tietokantaansa, ja antaa tuloksen sillä perusteella mikä vastaa lähinnä sen indeksissä olevaa kuvaa. Sisältöpohjaisessa haussa koneen on kyettävä vertaamaan hahmotelmaa tai hakukuvaa hakemistonsa kuviin. (Kaakinen 2008.)

Sisältöpohjaisessa kuvailussa on kehitetty niin kutsuttua relevanssipalautetta, joka perustuu siihen, että hakukone täydentää omatoimisesti olemassa olevia kuvailutietoja poimimalla hyvistä dokumenteista hakutietoja sekä muokkaa kyselyä niiden avulla uudelleenlaiseksi. Esimerkiksi Google-kuvahakukoneen käyttämä *find similar pages* -toiminto on yksinkertainen esimerkki relevanssipalautteesta. Relevanssipalautetta on pidetty lupaavana kuvatiedonhaun alalla (Enser 1995).

Uusia mahdollisuuksia kuvien kuvailuun on tuonut internetsovellusten mahdollistama käyttäjien suorittama avoin avainsanoitus, jota tarkastelen menetelmänä ja ilmiönä tarkemmin luvussa 2.6.

## 2.4 Valokuvan erityispiirteet dokumenttina

Tässä luvussa esittelen (digitaalisen) valokuvan olemusta dokumenttina sekä kuvan sisällönkuvailun teoreettista pohjaa. Tutkimuksessani käytän kuvan luokituksista ja indeksoinnista käsitettä sisällönkuvailu, joka sisältää nämä molemmat termit. Luokituksen ja indeksoinnin problematiikkaa ei yleensä erotella toisistaan kuvien sisällönkuvailun pohdinnoissa. Dokumentaatiokieliä, eli luokitusta ja indeksointia käytetään sisällönkuvailussa, mutta myös dokumentaatiokielen laatimisessa tarvitaan sisällönanalyysia, sillä useissa menetelmissä dokumentaatiokielen termit kerätään kuvailemalla dokumenttien sisältöä. Nämä alueet ovat toisiinsa kietoutuneita. (Karhula 1991, 9.)

Tässä tutkielmassa tarkoitan käsitteellä "kuva" kaksiulotteista esitystä, representaatiota henkilöstä, eläimestä tai asiasta. Rajaan kuvan tarkoittamaan valokuvia, jotka ovat digitaalisessa muodossa, jättäen pois liikkuvan kuvan ja multimedian sekä erilaiset tallennusmuodot. Jätän tarkastelun ulkopuolelle abstraktin aiheen omaavat kuvat, koska muutoin siirtyisin jo kuvien tulkinnan puolelle. Tämä on ongelmana jo yksiselitteisiltä näyttävissä kuvissa.

Flexnerin jäsenystä hyödyntäen Jörgensen (2003, 3) on määritellyt kuvan tarkoittamaan fyysistä samankaltaisuutta tai esitystä henkilöstä, eläimestä tai asiasta, joka on valokuvattu, maalattu veistetty tai muuten tehty näkyväksi. Vainio (1996, 20–21) määrittelee kuvan pinnaksi, joka voi optisesti jäljitellä todellisuutta. Kuva ei ole itse todellisuus, vaan sen representaatio eli esitys.

Enser (2008, 4–5) on erottanut sisällönkuvailun näkökulmasta kolmentyyppisiä kuvia:

- ”suora kuva” (*direct picture*). Se voidaan ymmärtää Vainion (1996, 20-21) tavoin todellisuuden representaatioksi

- ”epäsuora kuva” (*indirect picture*), kuten esimerkiksi röntgenkuva
- hybridikuva (*hybrid picture*), kuva jossa on myös tekstiä, kuten esimerkiksi kuvateksti.

Tässä tutkimuksessa rajaan kuvan dokumenttina ja kuvailun kohteena Enserin määrittelemään suoraan kuvaan, jättäen pois epäsuorat tai hybridikuvat. Digitaalisen kuvan dokumentaalaisia ominaisuuksia on kuitenkin hedelmällistä valottaa vertailemalla sitä muihin dokumenttityyppisiin, kuten tekstiin ja paperiseen kuvaan.

Kuva ja teksti eroavat radikaalisti dokumentteina. Chen ja Rasmussen (1999) kiteyttävät artikkelinsa *Intellectual Access to Images -image database systems* johdannossa tekstin ja digitaalisten kuvien eron dokumentteina pääsyn kannalta: tekstidokumentissa on itsessään hakutieto tai hakuelementti (*access point*), jonka perusteella se voidaan hakea. Teksti voidaan löytää sen sisältämän luonnollisen kielen kautta. Pääsy tekstiin on siis mahdollinen ilman ihmisen tekemää sisällönkuvailua, sen sijaan kuvat sisältävät sellaisia merkityksen tasoja, jotka vain ihminen voi kääntää verbaalisiksi. (Chen & Rasmussen 1999.) Yksi keskeinen ero tekstidokumenttien sisällönkuvailuun verrattaessa onkin se, että valokuvan sisällönkuvailussa käännostasoja on enemmän. (Karhula 1991, 48–50.)

Digitaaliset kuvat koostuvat pikseleistä, joiden kuviot vaihtelevat äärettömästi. On mahdotonta ennustaa tiettyä kuviota, joka tyydyttäisi tiedontarpeen. Nykyisin voidaan hakea suoraan vain suhteellisen matalan tason ominaisuuksia tai piirteitä, kuten värit ja yksinkertaiset kuviot, esimerkiksi neliöt ja ympyrät. Matalan tason ominaisuudet eivät kuitenkaan sisällä kuvan merkitystä. Jopa ihmisten tekemässä kuvailussa on vaikea päästä yhteisymmärrykseen kuvan sisällöstä, merkityksestä tai kuvailuun sopivista nä-

kökulmista. Sama kuva voi merkitä eri asioita eri ihmisille eri aikoina. Kuvat saavat eri merkityksen eri asiayhteyksissä, joissa niitä käytetään esimerkiksi kuvituksena. (Chen & Rasmussen 1999.)

Kuvien sisällönkuvailu on huomattavasti tekstin sisällönkuvailua monimutkaisempaa, koska kuvan keskeinen tietosisältö ei ole kuvaan kiinnittyvä asia. Kuvan olennainen tietosisältö jäsentyy paremminkin sen käyttökontekstista, kuin kuvan itsensä ominaisuuksista. (Karhula 1991, 48–50.) Kuvat kantavat siis tietoa ja merkityksiä täysin poikkeavalla tavalla kuin tekstit, jotka kantavat merkitystä sopimuksenmukaisten ja keinotekkoisten symbolien kautta. Kuvat kantavat merkitystä muistuttaen kohteitaan. (Sahavirta & Sormunen, 2001.)

Neugebauer (2005) tiivistää hyvin kuvien sisällönkuvailun teoreettiset ongelmat:

- Kuvat eivät täytä kielen vaatimuksia kuten tekstimuodossa oleva materiaali
- Kuvat sisältävät merkityksen kerroksia, jotka vain ihmiset voivat kääntää kielellisiksi ilmauksiksi
- Kuvien monitieteellinen ja -tulkintainen luonne aiheuttaa sen, että valitut kuvailutermit ovat ainoa pääsy (*access point*) dokumentteihin.

Karhulan (1991) mukaan valokuvan olemus perustuu kohteisiin. Karhula tarkoittaa kohteella valokuvan ulkopuolista, havaittavaa objektia, ominaisuutta, ilmiötä tai toimintaa. Valokuva ennemminkin muuntaa kohteita esittäessään ne eri tavoin, kuin jäljentää niitä. Karhula määrittelee valokuvan elementeiksi kaikki visuaaliset olennot, joita voidaan nimetä, kuten hahmot, hahmon osat tai hahmoja muodostamattomat elementit. Valokuvan tulkinnan perusta on kielen ulkopuolisessa, havaittavassa todellisuudessa. Katsojan tunnistusreaktio mahdollistuu opittujen kuvallisten konventioiden välityksellä. Valoku-

va kuitenkin muuntaa esittämiään kohteita, joiden tulkinta on aina sopimuksenvaraista. (Karhula 1991, 188.)

Shatford (1986) jakaa kuvan tasot siihen mitä kuvassa konkreettisesti on (*of*), ja mistä se kertoo (*about*). *Of*-tasolla nimetään kuvan elementtejä, joita ovat Karhulan (1991) mukaan kaikki visuaaliset olennot, joita voidaan nimetä. Tällaisia olentoja ovat esimerkiksi hahmot, hahmon osat tai hahmoja muodostamattomat elementit. *About*-tasolla tehdään edellä kuvattua sisällönanalyysiä. *About*-taso käsittää siis kuvan ydinsanomana eli keskeisen sisällön. Valokuvaa voidaan kuvailla myös kysymällä kysymykset kuka, mitä, missä ja milloin.

Merkitysten kontrollilla dokumentaatiokieliissä pyritään eliminoimaan puhutun ja kirjoitetun kielen moniselitteisyys siten, että käytetään vain yhtä ilmaisua kutakin käsitettä kohden (Karhula 1991, 7). Kuvan esittävyys aiheuttaa sen, ettei siihen sisältyviä merkityksiä ole mahdollista tyhjentävästi ja yksiselitteisesti kääntää luonnolliselle tai dokumentaatiokielelle. Jokainen kuvahakujärjestelmä on näin väistämättä epätäydellinen, koska kuvan merkitykset ja representaatio varioivat, käytettiin sitten visuaalisia hahmoja tai kielessä muotoiltuja kyselyjä. (Sahavirta & Sormunen 2001.)

Sisällönkuvailun termit poimitaan kohteista, joille ei ole olemassa yhtenäistä tulkintaa. Kuvan elementit eivät ole yksiselitteisiä. Ne edustavat eri asioita ihmisille eri aikoina, ja ne saavat merkityksen asiayhteydessä, jossa niitä käytetään. Henkilökohtaiset tulkinnalliset erot ovatkin merkittävä ongelma kuvien sisällönkuvailussa. Havainto ja tunnistaminen muotoutuvat suhteessa katsojan laajempaan tietämykseen, sekä kulttuurin omaksumiin esitystapoihin. Kuva esittää eri asioita riippuen siitä, mihin huomio kiinnitetään. Tämä aiheuttaa kuvien sumeuden. Kyse on merkityksen varioinnista, joka on kuvahaun kannalta keskeinen ongelma. Kaikkia kuvan merkityksiä ei voida indeksoida, koska



kuvassa on aina sekä esittäviä että käsitteellisiä merkityksiä, jotka eivät ole kuvan itsensä ominaisuuksia vaan tiedostavan ihmismielen antamia merkityksiä. Osa merkityksistä syntyy vasta kun kuva on asetettu johonkin kontekstiin, kuten esimerkiksi sanomalehtiartikkelin yhteyteen. Kuvien sisällönkuvailun suurin ongelma onkin kuvan monitulkintaisuudesta johtuva merkityksen variointi. (Sahavirta & Sormunen 2001; Chen & Rasmussen 1999.)

## **2.5 Valokuvien kuvailun tutkimus ja kuvailumallit**

Kuvien sisällönkuvailun tutkimus on kulkenut käsi kädessä tiedonhakupöytäjärjestelmien kehitysvaiheiden kanssa. Vaikka kuvien sisällönkuvailun ja dokumentaatiokielen laatimisen ongelmia on pohdittu jo 1920-luvulta lähtien, suurin osa varhaisemmasta sisällönkuvailun tutkimuksesta on keskittynyt pääasiassa tekstidokumentteihin. (Karhula 1991, 50.) Kuvien sisällönkuvailun tutkimus on lisääntynyt 1990-luvulta lähtien. Tutkimusta löytyy etenkin isoista kuvakokoelmista.

Yksi varhaisimmista tutkimuksista on Unescon RAMP-hanke (*A Long-Term Records and Archives Management Programme*), joka keskittyi tutkimaan arkistojen ja asiakirjojen pitkäaikaishallintaa. Valokuvien seulontaan liittyvä raportti julkaistiin jo 1985. Hankkeen yhtenä tavoitteena oli luoda ohjeet valokuvien arvonmääritykseen arkistonäkökulmasta sekä valokuvien arkistollisten ominaisuuksien ymmärtäminen (Leary, W.H.1985).

Peter Enser on yksi eniten kuvien sisällönkuvailusta kirjoittanut tutkija. Hän on ollut toteuttamassa laajamittaista käyttäjätutkimusta, jossa Enser ja McGregor (1993) tutkivat Hulton Deutsh-kuvakokoelmaan kohdistuneita kyselyitä, joita tekivät lähinnä lehdistön

edustajat. Enser (1995) on artikkelissaan *Pictorial information retrieval* esittänyt nelita-  
hoisen mallin kuvien tallennuksen ja haun lingvistisistä sekä visuaalisista menetelmistä  
ja niiden yhdistelmistä. Mallissaan hän esittelee erilaisia mahdollisuuksia hakea kuvia,  
kuten hakusanoilla tehtyä hakua kuvan kielellisiin kuvailutietoihin tai kuvaan kohdistu-  
vaa hakua visuaalisin keinoin.

Rahoitusta kuva-arkistojen tutkimukseen on myönnetty Euroopan unionin taholta. EC-  
PA-hanke (*European Commission on Preservation and Access* 1999) toteutettiin järjes-  
tämällä kysely 141 Eurooppalaiselle muistiorganisaatioille, joilla oli valokuvakokoel-  
mia. Tutkimukseen osallistuneiden organisaatioiden kokoelmissa oli yhteensä noin 120  
miljoonaa valokuvaa. Tutkimus on tehty aikana, jolloin vallitsi ylimenovaihe paperisen  
ja digitaalisen rajapinnassa. Kyselyssä digitointiprojektien syyksi mainittiin halu suojel-  
la valokuvien originaaleja toistuvalla käytöllä (Klijn & Lusenet 2004, 7-8.)

Toinen Euroopan unionin rahoittama tutkimushanke on Sepia-hanke (*Safeguarding Eu-  
ropean Photographic Images for Access*), joka toteutettiin vuosina 1999-2003. Tähän  
hankkeeseen osallistui myös Suomen valokuvataiteen museo. Tutkimuksen tavoitteena  
oli luoda suositus asiasanoista, jotka pitäisi kirjata ylös kustakin valokuvakokonaisuu-  
desta. Suosituksella pyrittiin parantamaan valokuvien hallintaa, säilytystä ja käyttöä.  
Hankkeen tuloksena syntyi Sepiades (*Sepia Data Element set*), joka on malli ja väline  
valokuvien luettelointiin (Klijn & Lusenet 2004, 39: Sepiades 2003, 2.)

Vuosituhanen vaihteessa kuvien käyttäjien tarpeista, kuvien käytöstä ja käyttäjien ha-  
kutavoista tiedettiin vielä suhteellisen vähän (Markkula & Sormunen 2000, 259). Tut-  
kimuksen vähäisyyteen on esitetty syyksi kuvien sisällönkuvailun hajautuminen erilai-  
siin organisaatioihin, joilla kaikilla on erilaiset käytännöt sisällönkuvailuun (esim. Jör-  
gensen 2003). 2000-luvulla tutkimuksellinen mielenkiinto onkin suuntautunut enemmän

verkossa oleviin kuviin kuin valokuva-arkistoihin. Verkkoyhteisöissä tapahtuvaa käyttäjien tekemää valokuvien sisällönkuvailua on tutkittu paljon. Tutkimusta aiheesta ovat tehneet esimerkiksi Fokker ym.(2006) sekä Sigurbjörsson & van Zwol (2008).

Sara Shatford (nykyisin Shatford Layne) on myös tutkinut (esim. 1986) kuvien kuvailuun liittyviä ongelmia. Hän on esitellyt kuvien lukuohjeen kuville niiden ominaisuuksien pohjalta sekä korostanut pääsyn tarjoamista kuvien ryhmittelyihin (*groupings*). Corinne Jörgensen (2003) esittelee kirjassaan *Image retrieval. Theory and research* kuvatiedonhaun kenttää: teoriaa, käytäntöä ja käytännön sovellutuksia.

2000-luvulla on tutkittu paljon verkkoyhteisöissä tapahtuvaa käyttäjien tekemää valokuvien sisällönkuvailua. Fokker ym. (2006) ovat artikkelissaan käsitelleet käyttäjien tekemää avainsanoitusta kuvapalveluissa. Artikkelissaan esittelevät avainsanapilven ominaisuuksia haun apuvälineenä avainsanapaljouden keskellä. Avainsanapilvet (*tag clouds*) ovat graafisia esityksiä, joita käytetään kuvahakupalveluissa havainnollistamaan avainsanojen (*tags*) tilastollisia suhteita. Ne pohjautuvat avainsanojen esiintymistiheyteen ja käytännössä kuvaavat avainsanojen suosiota. Mesnage ja Carman (2009) ovat musiikin tiedonhaun saralla vieneet ajatusta eteenpäin tutkimalla avainsanapilvien kehittämismahdollisuuksia.

Sigurbjörsson ja van Zwol (2008) ovat tutkineet *Flickr*-kuvapalvelun käyttäjien avainsanoitustapoja. He ovat esittäneet käyttäjille suunnatun avainsanojen suosittelumenetelmän, jolla sisällönkuvailua voitaisiin yhdenmukaistaa verkon valokuvapalveluissa ja dokumenttien löydettävyyttä verkossa voitaisiin parantaa. Myös Stvilia ja Jörgensen (2009) ovat tutkineet kuvien kuvailua, selvittäen syitä henkilökohtaisten ja julkisten kuvakokoelmien kuvailulle ja nimeämiselle.

Valokuvien sisällönkuvailusta on niukasti kotimaista tutkimusta. Päivikki Karhulan (1991) pro gradu -tutkielma *Sanomalehtivalokuvien sisällönkuvailun kehittäminen* on alan pioneerityö Suomessa. Karhula laatii työssään lukuohjeita valokuvan sisällönkuvailua varten. Karhula lähestyy aihealuetta ensisijaisesti kuva-aineistoa (sanomalehtivalokuva) sisältävien, tiedonhakujärjestelmien kannalta, joihin kuuluvat viitetietokannat sekä kuvatietokannat.

Marjo Markkula ja Eero Sormunen ovat tutkineet kuvanhakukeinoja ja niiden soveltuvuutta jokapäiväiseen käyttöön tietojenkäsittelyopin näkökulmasta esimerkiksi artikkelissaan *Kuvatiedonhaun menetelmät ja niiden evaluointi* (1999). Markkula ja Sormunen (2000) ovat myös tutkineet Aamulehden toimittajien tekemiä kuvahakuja. Toimittajien kuvahaussa selailu koettiin luonnollisena tapana etsiä kuvia ja kuvitusideoita.

Leena Penttinen (2007) on gradussaan *Yrityksen valokuva-arkisto ja arkistointisuunnitelma: tapaus Rautaruukki Oyj* tutkinut yritysten kuva-arkistoja tapaustutkimuksen keinoin. Penttisen pro gradu-työ on toiminut käytännönläheisen ja havainnoivan otteensa vuoksi tämänkin opinnäytetyön innoittajana, vaikka tutkimuksen kohteena olevan arkiston kuvat ovatkin Penttisen tutkimuksessa printti- ja itselläni digitaalisessa muodossa.

Kuvien sisällönkuvailuun on kehitetty erilaisia lukuohjeita sekä analyysikehikoita. Lukuohje onkin ollut kuvien sisällönkuvailun keskeinen tutkimuksen kohde (esim. Shatford Layne 1994, Karhula 1991). Näitä malleja on sovellettu erilaisissa kuvakokoelmissa. Eri aloilla on olemassa erilaisia kuvakokoelmia. Jotain kuvakokoelmien moniulotteisuudesta kertoo se, että esimerkiksi taiteen alalla löytyy kokoelmia taideteoksista itsestään, tallenteita taideteoksista, taidejäljennöksiä ja erilaisia taideteoksiin liittyviä luon-

noksia ja tekstejä. Lisäksi kuvat ovat eri muodoissa, kuten dioina, valokuvapaperilla tai digitaalisessa muodossa. (Shatford Layne 1994.)

Shatfordin (1994, 583–584) ideaan tukeutuen Peter Enser on ehdottanut neljää ominaisuutta (attributes), jotka pitää tunnistaa kuvaa kuvailtaessa:

- Biografiset (*biographical*) teoksen syntyhetkeen liittyvät ominaisuudet kuten tiedoston sijainti, omistaja sekä tekijät.
- Aihe (*subject*): kohde, aika, paikka, tila, toiminta tai tapahtuma.
- Ilmiasu (*exemplified*): kuvan fyysinen muoto.
- Suhteet (*relationship*): looginen yhteys muihin kuviin/teksteihin, esimerkiksi valmis maalaus ja luonnos.

Shatford (1986) esittää myös, että valokuva voidaan kuvailla myös kysymällä kysymykset kuka, mitä, missä ja milloin.

Taulukko 1. Shatfordin lukuohje (1986)

	Nimetty yksilökohde	Yleinen kohdeluokka	Symboliset viittaukset
<b>Kuka?</b>	nimetty henkilö, eläin, esine	henkilö-, eläin-, esinetyyppi	myyttiset olennot
<b>Mitä?</b>	nimetty tapahtuma	tapahtuma- tai tilannetyyppi	symbolinen paikka
<b>Missä?</b>	nimetty paikka	maantieteellinen tai arkkitehtoninen paikkatyyppi	symbolinen paikka
<b>Milloin?</b>	pvm, kellonaika	vuoden- tai vuorokaudenaika	ajan symboloima tunne, abstraktio

Shatford Layne (1994) on todennut, että kuvien sisällönkuvailun tulisi tarjota pääsy kuviin mieluiten kaikkien niiden sisältämien ominaisuuksien kautta. Kuvan sisällönkuvailun kompleksisuutta kuvastaa sekin, että mitä kuvasta poimitaan, voi olla hyödyllistä vain joko sen yleiselle tai erityiselle identiteetille. Hakijalle esimerkiksi kuva Tampereen Näsinneulasta voi olla hyödyllinen siksi, että se esittää yleisesti ottaen tornia, tai toiselle hakijalle siksi, että se esittää juuri Näsinneulaa. Ideaalitulanteessa pitäisi siis mahdollistaa pääsy kaikkien mahdollisten yleisten ja erityisten identiteettien mukaan. Lisäksi Shatford Layne (1994, 585) huomauttaa, että tulisi tarjota pääsy hyvin ryhmiteltyihin kuviin, ei vain yksittäisiin kuviin. Ryhmiteltyjen kuvien idea on mielestäni kaukonäköinen, sillä se muistuttaa tavallaan luvussa 4 kuvailtuja, niin kutsuttuja folksonomioita.

Karhula (1991) on kritisoinut Shatfordia siitä, ettei hän ole kyennyt osoittamaan, mihin ja miten lukuohjetta voitaisiin soveltaa, Hänen mielestään Shatfordin teoreettinen pohja on puutteellinen. Karhula on huomauttanut, että valokuvan sisällöstä ei ole oleellista puhua aiheena, vaan pikemminkin sen keskeisestä sisällöstä. Käyttäjätutkimusten mukaan hakupyynnöiden termit eivät kuvaa vain valokuvien elementtejä, vaan sisältöä, tunnelmaa ja merkitystä. (Karhula 1991, 176, 80–200.)

Sahavirran & Sormusen (2001) mukaan kuvien sisällönkuvailua voitaisiin parantaa tutkimalla kuvan eri käsitetasoja. Olennaisia kehittämiskohteita ovat sellaiset kuvailumenetelmät, jotka mahdollistavat vaihtoehtoisten selailunäkymien luomisen kuva-aineistoon. Merkitykset ja representaation tasot, jotka ovat ongelmallisia kuvailuissa ja kyselyissä, ovat usein kuvaa silmäiltäessä helposti tulkittavissa. Kuvien subjektiivisen tulkinnan aiheuttamaan ongelmaan on tarjottu edellä kuvattuja hahmopohjaisia indeksointi- ja kuvailumenetelmiä, mutta niiden heikkoutena on kuvien merkityksen varioin-

ti. Haetut merkitykset saattavat kadota jos ne sanallistetaan tai pakotetaan jäykkiin hahmomalleihin.

Karhula (1991) huomauttaa, ettei visuaalisuuden lisääminen välttämättä poista verbaalikielisen kuvailun tarvetta. Mikäli tiedonhakukomentoja aiotaan visualisoida, visualisoitavat hakukomennot täytyy kuitenkin ensin tuntea verbaalikielisinä käsitteinä. Näin varmistetaan, että käytettävät visuaaliset piirteet vastaavat mahdollisimman hyvin tarkoitettua käsitettä. (Karhula 1991, 17.) Kuitenkin tietojärjestelmien kapasiteetin lisääntyessä on mahdollista tehdä esimerkiksi selailukuvien (*thumbnails*) silmäily yhä joustavammaksi. Erilaisten yhdistelmämallien kehittäminen, joissa käytettäisiin sekä kielellistä että kuvallista selailumahdollisuutta, vaikuttavat potentiaalisimmalta vaihtoehdolta.

Uusia vaatimuksia digitaalisten kuva-arkistojen kehittämiseksi asettaa ainakin digitaalisessa muodossa olevien kuvien määrän lisääntyminen. Käyttäjäystävälliset sovellukset, kuten digitaalisten kameroiden ja kuvankäsittelyohjelmien kehittyminen, sekä sosiaalisen median palvelut ovat helpottaneet kuvien tuottamista ja jakamista. Yhä useampi ihminen voi tuottaa sekä jakaa kuvia yhä nopeammin. Uusia vaatimuksia asettaa myös lisääntyvä käyttäjien suora pääsy digitaalisiin kuvakokoelmiin, kuten esimerkiksi museoiden digitaalisiin kuvakokoelmiin (esim. Markkula & Sormunen 2000). Kuvien sisälönkuvailun tutkimuksen ja kuvahallinnan (*image management*) kehittämisen tarve on lisääntynyt huomattavasti, sillä digitaalisten kuva-arkistojen pitäisi kyetä hallitsemaan kuvatulvaa sekä parantamaan löydettävyyttä.

## 2.6 Yhteisöllinen sisällönkuvailu verkossa

Tiedon tuottamisen ja jakamisen tavat ovat muuttuneet ns. sosiaalisen median myötä. Sosiaalinen media on käsitteenä kritisoitu, eikä sitä ole tarkasti määritelty. Se on Web 2.0 -käsitteen myötä ilmaantunut muoti-ilmaus, josta alettiin puhua vuosituhannen vaihteissa. Sosiaalinen media muodostuu käytännöistä, joissa viestintä suuntautuu yhdeltä monelle tai muutamalle niin, että ihmiset jakavat tekemäänsä tai muiden tuottamaa aineistoa omaksi ilokseen. (Lietsala & Sirkkunen 2010, 264–267.) Tässä tutkielmassa alaluvussa 2.6.3 käytän esimerkkinä *Flickr* -kuvapalvelua, koska siellä käyttäjät voivat jakaa ja avainsanoittaa valokuvia.

Valokuvien intellektuaalisen kuvailun uusi ja kasvava osa-alue on käyttäjien internetin kuvapalveluissa tekemät avainsanoitukset kuville. Valokuvien sisällönkuvailun kannalta erilaiset sisältöjen jakeluun tarkoitettut sivustot ovat kiinnostavia, sillä niissä käyttäjät voivat tallentaa valokuvia ja kuvailla niitä avainsanoilla (*tags*). Avointa avainsanoitusta käytetään esimerkiksi nettisivujen, kuvien, musiikin, kirjojen ja akateemisten julkaisujen organisoimiseen. Avoimen avainsanoituksen edelläkävijöistä verkossa tunnetuimpia on *Flickr*-kuvapalvelu. (Mesnage 2009, 1; Tuominen 2006, 171.) Myöhemmin *Flickr:n* ovat suosiossa syrjäyttäneet muut sovellukset, kuten *Instagram*. Käytän kuitenkin tässä kirjallisuuskatsauksessa esimerkkinä *Flickr* -palvelua, koska sitä on tutkittu paljon.

Enser (2008, 7) huomauttaa, etteivät internetin kuvapalvelut ole varsinaisesti kokoelmia, ainakaan siinä mielessä kuin kokoelma ymmärretään ammattilaisten parissa. Internetin kuvapalveluilta puuttuu ammatillisesti järjestettyjen kokoelmien perusominaisuuksia, kuten ammattimainen kokoelman hallinta. Kuvapalveluiden valtava kapasiteetti niin kuvien kuin avointen avainsanoitusten tekijöiden suhteen on kuitenkin lopullisesti muuttanut käsityksemme kuvakokoelmista. (Enser 2008, 7–8.)



## 2.6.1 Käyttäjät avainsanoittajina valokuvapalveluissa

Sisältöjen jakeluun tarkoitettut sivustot internetissä, kuten verkon valokuvapalvelut, ovat yleistyneet vuodesta 2004 lähtien. Niiden toimintaidea perustuu siihen, että käyttäjät voivat itse luokitella verkosta löytämiään sivustoja ja dokumentteja. Näin syntyy yhteisöllisesti käsityönä tuotettua ja sanalistoina (*tags*) esitettyä metadattaa. Se on dokumenttien kuvailutietoa, joka on linkitetty verkkoresursseihin. Avainsanat ovat vapaasti muo-  
toiltuja dokumentteja kuvaavia termejä. (Tuominen 2006, 169–170.)

Oleellista avoimissa avainsanoituksissa on, että niiden varaan rakentuvat sanastot eivät ole rakenteeltaan hierarkkisia. Avoimissa avainsanoituksissa ei siis ole täsmällisesti määriteltyjä ylä- ja alakäsitteitä. Sen sijaan niiden välille luodaan automaattisesti assosiatiivisia viittauksia merkittyjen sisältöjen perusteella. Lähestymistapa avoimissa avainsanoituksissa sanastojen rakentumiseen on siis täysin päinvastainen verrattuna taksonomiseen luokitteluun: sen sijaan, että luokittelusta pyrittäisiin tekemään mahdollisimman täsmällinen, sen annetaan rakentua käyttäjien määräämällä tavalla. (Tuominen 2006, 171–173.)

Käyttäjien määrän lisääntyessä suosituimmista otsikoista muodostuu sanastoja tai luokitteluja, joita kutsutaan folksonomioiksi (*folksonomies*) *folk* ja *taksonomy* -sanoja yhdistäen (Tuominen 2006). Folksonomiat ovat Saartin (2009) määritelmän mukaan käyttäjien verkkoyhteisöissä luomia taksonomioita eli avoimia luokittelujärjestelmiä, jotka muodostuvat vapaasti määritellyistä avainsanoista (*tags*). Tuomisen (2006, 172) mukaan folksonomiat ovat vaatimattomia rakennelmia verrattuna kontrolloituihin sanastoihin ja luokituskaavoihin. Folksonomiat ovat epähierarkkisia luokituksia, ennemminkin assosiaatioita ihmisten, avainsanojen ja dokumenttien välillä (Mesnage 2009, 1).

Avointa avainsanoitusta tukevat kuvapalvelut mahdollistavat sen, että vaikka kuvatulva on valtava, on myös olemassa valtava määrä käyttäjiä tekemässä työtä kuvien löydettävyyden lisäämiseksi. Kuvapalvelut suosittelevat esimerkiksi uutta kirjanmerkkiä lisäville muiden jo käyttämiä termejä. Tämä lisää folksonomian yhdenmukaisuutta, käyttäjät hyödyntävät tiettyä käsitettä tarkoittaessaan samoja sanoja. (Tuominen 2006, 171–172.) Tuomisen mukaan tällöin yksilölliset ja yhteisölliset suodatusmekanismit limittyvät toisiinsa. Käyttäjä työskentelee itsensä sekä yhteisönsä hyväksi, vaikka hänen motiivinsa olisikin vain omien avainsanoitusten jäsentäminen. Parviälly mahdollistaa kollektiivisen folksonomian tuottamisen pienellä vaivalla. (Tuominen 2006, 174)

## **2.6.2 Avoin avainsanoitus kuvapalveluiden kuvatulvan suodattajana**

Valokuvista on mahdoton poimia automaattisesti avainsanoja, sillä kuva ei niitä sisällä. Myös intellektuaalisestikin kuvailtaessa avainsanojen liittämien kuvaan on hidasta. Fokker ym. (2006) esittävätkin ratkaisuksi artikkelissaan *Tagging in peer-to-peer wikipedia. A Method to induce cooperation* käyttäjien vapaaehtoisen avoimen avainsanoituksen. Avoin avainsanoitus auttaa kuva-aineiston järjestelemisessä sekä parantamaan hakukoneiden tehokkuutta, koska sisältö on käyttäjien luokittelemaa. Kun käyttäjät itse kuvailevat kuvien sisältöä, sisältö on luokiteltu käyttäjille tutulla ja yhteisesti jaetulla sanastolla. Kuvien kuvailussa ja haussa käytetyt termit ovat samanlaisia, koska käyttäjät itse kuvailevat kuvia. Käyttäjien kuvanhakutapojen tunteminen onkin tärkeää, kehitettäessä uusia kuvanhakumenetelmiä. (Fokker ym. 2006.)

Käyttäjät voivat avainsanoittaa valokuvan miten oudolla termillä tahansa. Jos käyttäjä yrittää etsiä esimerkiksi tiettyä elokuvaa, mutta ei muista elokuvan nimeä, tekijöitä tai näyttelijöiden nimiä, hän voi käyttää hakuterminä vaikkapa elokuvassa esiintynyttä si-

nistä *Impala*-merkkistä autoa. Koska muut käyttäjät ovat vapaasti voineet avainsanoitaneet elokuvan, suurella todennäköisyydellä joku on voinut avainsanoittaa elokuvan samalla termillä. Käyttäjien vapaudesta on hyötyä erityisesti epätavallisten termien ollessa kyseessä. (Fokker ym. 2006.) Niin kutsutun parviällyn voimalla voidaan siis laajentaa valokuvan kuvailua, joka parantaa erilaisten merkitysten näkymistä avainsanoissa (Sigurbjörsson & van Zwol 2008). Seuraavassa osiossa esittelen *Flickr*-kuvapalvelun esimerkkinä palvelusta, jossa käyttäjät voivat tehdä avainsanoitusta kuville.

### 2.6.3 Käyttäjien avainsanoitus Flickr -kuvapalvelussa

*Flickr* -kuvapalvelu on internetissä sisältöjen jakeluun tarkoitettu sivusto, joka hyödyntää tageja eli käyttäjien tekemää avainsanoitusta. *Flickr*-palvelussa käyttäjät voivat tallentaa ottamiaan valokuvia tai videoita palveluun, ja otsikoivat niitä avainsanoilla. *Flickr* -kuvapalvelulla on jo yli 8.5 miljoonaa rekisteröitynyttä käyttäjää, ja se sisältää satoja miljoonia kuvia sekä videoita ladattuna, avainsanoitettuna ja organisoituna. Käyttäjien tuottaessa ja linkittäessä jatkuvasti uutta valokuvamateriaalia kuvien määrä internetissä kasvaa eksponentiaalista vauhtia. (Sigurbjörsson & van Zwol 2008; Tuominen 2006, 169–171.)

*Flickr*-kuvapalvelun kuvien määrä on lisääntynyt viime vuosina valtavasti. Kun Tuominen (2006, 171) haki marraskuussa 2005 yhdellä suosituimmista avainsanapilven tagilla eli *iPod*-termillä, hän löysi 16 858 kuvaa. Marraskuussa 2014 löysin *iPod*-termillä jo 259 201 kuvaa, vaikka se ei teknologian kehityksen valossa ymmärrettävästikään ollut enää suosituimpien termien avainsanapilvessä (<http://www.flickr.com/photos/tags/>). Vastaava termi avainsanapilvessä oli *iPhone*, jolla löytyi jo 5 507 651 kuvaa. Marraskuussa 2014 suurimmilla sanoilla avainsanapilvessä oli ladattu jo kymmeniä miljoonia

kuvia kuten esimerkiksi *instagramapp* (56 222 175 latausta). Muita kaikkien aikojen suosituimpia avainsanoja olivat marraskuussa 2014 mm. *wedding*, *Japan*, *travel*, *canon*, *family*, *nature* ja *nikon*.

Flickr:n ja muiden kuvapalveluiden suosio osoittaa, että käyttäjät ovat halukkaita tarjoamaan kuvailutietoja intellektuaalisen kuvailun keinoin saadakseen kuvansa paremmin muiden saataville. (Sigurbjörnsson & van Zwol 2008.) Näiden avointa avainsanoitusta hyödyntävien verkkopalveluiden mahdollistamat suositukset, arvioinnit ja käyttötilastot lisäävät huomattavasti niiden suodatuskapasiteettia. (Tuominen 2006, 169–171.)

Suosituksia kuvapalveluissa helpottavat siis relevantin aineiston löytymisen kuvatulvas-ta. Käyttäjä voi tutkia *Flickr*-kuvapalvelussa, millaista aineistoa muut käyttäjät indeksoineet esimerkiksi hakusanalla *reggae*, ja tehdä tämän perusteella sanahakuja. Kuvapalvelut voivat tarjota käyttäjälle hakusanaan tilastollisesti liittyviä muita avainsanoja. Muiden käyttäjien laatimista avainsanoista hakeminen on usein tehokkaampaa kuin hakukoneen, esim. Googlen käyttäminen, koska hälyä on vähemmän. Annettujen linkitysten määrä kertoo jo jotain dokumentin relevanssista, vaikka sosiaalisen median kuvapalveluissa ei olisikaan yhteisesti sovittua aineistonvalintapolitiikkaa. (Tuominen 2006, 171.)

#### **2.6.4 Avainsanapilvet, tagipelit ja avainsanasuosituksset**

Käyttäjien tekemä sisällönkuvailu on kuitenkin yhä hallitsematonta ja kontrolloimatonta. Avainsanoituksen sekä parantamiseksi on kehitetty erilaisia parannuskeinoja. Tutkimuksessa on kiinnitetty huomiota käyttäjien tekemän avainsanoituksen parantamiseen kehittämällä erilaisia keinoja tukemaan avainsanoitusta sekä lisäämään innostusta

avainsanoitukseen. Myös avainsanojen navigointia ja käytettävyyttä on tutkittu (ks. Mesnage 2009).

Avainsanoitus on hyödyllisimmillään, kun tagit kootaan samaan yhteyteen isoksi parveksi. Esimerkiksi *Flickr*-kuvapalvelussa avainsanapilvien avulla voi etsiä kuvia (ks. kuva 1). Tällaista avainsanapilveä (*tag cloud*) voidaan ajatella jonkinlaisena folksonomian visualisoituna ilmentymänä. (Lemare 2008.) Avainsanapilvi on Fokker:n ym. (2006) määritelmän mukaan visuaalinen esitys avainsanojen (*tags*) tilastollisista suhteista pohjautuen niiden esiintymistiheyteen. (Fokker ym. 2006.) Avainsanapilvi näyttää linkit suosituimpiin tageihin, fontin koko riippuu tagien suosion tasosta. Parvet ja niihin liittyvät tagit mahdollistavat tagipohjaisen navigoinnin (Mesnage 2009, 1).



Kuva1. Avainsanapilvi Flickr-kuvapalvelussa (www.flickr.com)

Fokker ym. (2006) esittelevät työkalun, jolla voi navigoida esimerkiksi wikipedian multimediaversiossa avainsanapilviä hyväksi käyttäen. Käyttäjä voi valita erilaisia näkymiä avainsanapilvessä, kuten viimeisimmät avainsanoitukset, henkilökohtaiset- tai kaverien avainsanoitukset sekä suosituimmat taggaukset, jotka näkyvät kun käyttäjä liikuttaa kursoria avainsanan yli. (Fokker ym. 2006.) Mesnage (2009) kuitenkin huomauttaa, että nykyiset folksonomioiden navigointimahdollisuudet ovat rajoittuneita, useimmissa kuvahakukoneissa etsi-toiminto on pääasiallinen toiminto joka hyödyntää avainsanoja. Tägeja selaillessa useimmat järjestelmät antavat vain muutaman avainsanaan liittyvän avainsanan sitä klikatessa, mutta eivät anna tagiin liittyviä, samansukuisia avainsanoja (*related tags*) useisiin tageihin samaan aikaan.

Käyttäjien avoimen avainsanoituksen hyödyllisyys riippuu käyttäjien motivaatiosta (esim. Sigurbjörnsson & van Zwol 2008.) Käyttäjien avainsanoituksen parantamiseksi on kehitelty niin kutsuttuja tagipelejä tai avainsanapelejä, jotka ovat työkaluja joilla voidaan lisätä relevantteja avainsanoja kuville tietokonepelin avulla. Käyttäjät pelaavat peliä ehdottamalla avainsanoja kuville jotka ilmestyvät ruudulle ja saavat pisteitä, jos he ehdottavat samoja tageja kuin muut pelaajat (Sigurbjörnsson & van Zwol 2008, 328.)

Sigurbjörnsson ja van Zwol (2008) ovat tutkineet artikkelissaan *Flickr tag recommendation based on collective knowledge* Flickr-kuvapalvelun käyttäjien avainsanoituskäytäntöjä. He ovat analysoineet kuinka käyttäjät avainsanoittavat kuvia ja millaisia avainsanoja (*tags*) he tuottavat. Sigurbjörnsson ja van Zwol tutkivat millaista informaatiota Flickr-kuvapalvelun avainsanoitus sisältää. He saivat luokiteltua 52% kokoelman avainsanoista suosituimpiin kategorioihin. Kaikista suosituin kategoria oli paikka (28%), seuraavaksi suosituin kategoria olivat esineet tai kohteet (16%), sitten ihmiset tai ryhmät (13%) sekä tapahtuma tai toiminta (9%) ja ajankohta (7%). (Sigurbjörnsson & van Zwol 2008, 329.)

Sigurbjörnsson ja van Zwol (2008, 329) havaitsivat että käyttäjien kuville antamien avainsanojen määrä vaihteli. Joillain kuvilla oli vain yksi tagi, kun taas toiset kuvat olivat saaneet suorastaan uuvuttavan määrän tageja. Sigurbjörnsson ja van Zwol huomasiivat, että heidän otoksessaan (yli 52 000 000 kuvaa) avainsanojen määrä kuvaa kohti noudatti *Flickr*:ssä niin kutsuttuja 80/20 sääntöä eli *Pareton* periaatetta (*Power law*). Eniten tageja eli kuusi tai enemmän oli saanut n. 7 000 000 kuvaa, kun taas suurin osa kuvista eli noin 15 000 000 oli saanut vain yhden avainsanan. (Sigurbjörnsson & van Zwol 2008, 334.)

Sigurbjörnsson ja van Zwol (2008) ovat esittäneet suosituksia edellä mainitun tutkimuksen pohjalta käyttäjien avainsanoitukseen. Yleisimmin käytetyt avainsanat ovat liian yleisluontoisia, kuten viisi yleisintä avainsanaa *Flickr*:ssä (2008): *2006*, *2005*, *wedding*, *party*, ja *2004*, jotka eivät identifioi kuvaa tarpeeksi tarkasti. Vain kerran esiintyviä avainsanoja oli myös valtava määrä, ne taas olivat usein joko väärin kirjoitettuja tai liian erityisiä, jotta ne parantaisivat löydettävyyttä. Esiintymistiheysjakauman keskivälissä olevat avainsanat ovatkin Sigurbjörnssonin ja van Zwolin (2008, 329) mukaan hyödyllisimpiä avainsanasuosittelusten tekemiseen.

### **3 TUTKIMUSTEHTÄVÄ JA TUTKIMUSMENETELMÄ**

Tässä luvussa esittelen tutkimuksen lähestymistavan, keskeiset tutkimuskysymykset sekä aineistonkeruumenetelmän. Ensimmäisessä luvussa 3.1 esittelen tutkimuskysymykset. Luvussa 3.2 kerron tutkimuksen lähestymistavasta ja perustelen lähestymistavan valinnan. Luvussa 3.3 esittelen ja kuvailen tutkimusprosessin kulun.

Tutkimus on laadullinen ja aineistolähtöinen. Olen kerännyt aineiston itse havainnoinnin, haastatteluiden ja keskusteluiden avulla. Tutkimukseen osallistuivat Tampereen yliopiston viestintäyksikön työntekijät. Lisäksi olen dokumentoinut tutkimuskohdetta valokuvin ja saanut sähköpostitiedonantoja.

#### **3.1 Tutkimuskysymykset**

Tutkimuskysymykset ovat muotoutuneet kehityshankkeen lähtökohdista ja ensimmäisellä tutustumiskierroksella syntyneen keskustelun pohjalta. Alustavien keskusteluiden ja sähköpostiviestien pohjalta nousi heti esille joukko kehitystarpeita. Kehitystarpeiksi mainittiin mm. yhtenäinen arkistointikäytäntö, ohjeet ja arkiston kuvaus. Lisäksi toivottiin nykyistä helpompaa hakua sekä vaivatonta kuvien jakamista yliopiston ulkopuolelle. Kehitystarpeita oli paljon ja niitä nousi esille myös haastatteluissa. Tässä tutkimuksessa keskityn kuvakokoelman tarkasteluun ja esitän myös aineiston pohjalta nousseita kehitysehdotuksia. Tutkimuskysymykset ovat siis tarkentuneet tutkimusprosessin edetessä ja aineistonkeruun ohessa.

Tässä tutkimuksessa haetaan vastauksia seuraaviin kysymyksiin:



- Mitkä ovat Tampereen yliopiston viestintäyksikön digitaalisen valokuva-arkiston käyttötarpeet?
- Minkälaisista kuvista (kuvatyytit) arkiston sisältö koostuu ja millainen on kuvan elinkaari arkistossa?
- Millaisia ovat kuvan kuvailun käytännöt valokuva-arkistossa?
- Millaisia kehitysnäkökulmia internetin kuvanhallintasovellukset voivat tarjota arkistolle?

Näihin tutkimuskysymyksiin 1-2 vastaan luvussa 4 Kuva-arkiston nykytilanne. Tutkimuskysymykseen 3 vastaan luvussa 5 Kuvien kuvailun käytännöt arkistossa. Tutkimuskysymykseen 4 vastaan luvussa 8 Kuvailun kehittäminen nettisovellusten tapaan.

### **3.2 Tutkimuksen lähtökohdat ja tutkimuksen toteutus**

Tutkimuskohteena on Tampereen yliopiston viestintäyksikön digitaalisessa muodossa oleva valokuvakokoelma. Tutkimus on laadullinen tapaustutkimus, jossa kuvaillaan yhden kuva-arkiston piirteitä. Tutkimuksen tarkoituksena on selvittää viestintäyksikön digitaalisen kuvakokoelman kuvien valinta-, järjestämis- ja seulontaperiaatteita. Tutkimuksessa pyritään esimerkkitapauksen avulla selvittämään ja erittelemään digitaalisessa muodossa oleviin kuviin ja niiden kuvailuun liittyviä haasteita.

Tutkimus lähti liikkeelle vuonna 2010 lukiessani sähköpostiviestin Tampereen yliopiston viestintäyksikön tarjoamasta projektista joka kohdistui kuva-arkiston kehittämiseen. Näin siis tutkimukselle oli valmista tilausta organisaatiopuolelta. Aihe vastasi suunnitte-

lemaani graduaihetta. Tutkimus on toteutettu osana Tampereen yliopiston viestintäyksikön kuva-arkiston kehittämishanketta.

Tutkimukseni voi määritellä laadulliseksi tapaustutkimukseksi, sillä olen kerännyt aineistoni kvalitatiivisin menetelmin. Laadullisessa tutkimuksessa tutkimuskohteeseen perehdytään usein kokonaisvaltaisesti. Laadullisessa analyysissä ei käytetä koetilanteita, vaan tilanteet ovat usein luonnollisia. Aineisto voidaan kerätä itse, mutta myös olemassa olevia aineistoja voidaan käyttää (esim. Alasuutari, 1994). Kaikki laadullinen tutkimus ei ole tapausten tutkimusta. Tapaustutkimus saa myös vaikutteita myös muista laadullisen tutkimuksen suuntauksista. Tapaustutkimuksen voi ymmärtää ennemminkin lähestymistavaksi tai tutkimukselliseksi strategiaksi kuin aineiston keruu- tai analyysimenetelmäksi. Kuitenkin tapaustutkimukselle on löydettävissä määrittävä yhteinen tekijä: tapaustutkimuksessa käsitellään yhtä tai useampaa tapausta (*case*) joiden määrittely, analysointi ja ratkaisu tapaustutkimuksen keskeinen tavoite. (Eriksson & Koistinen 2005, 1-4.)

Tapaustutkimuksessa tapaus voi olla erilaisissa organisaatioissa toteutettava rajattu projekti ja kehityshanke. Tutkimuksessani lähtökohtana on viestintäyksikön kuva-arkiston kehityshanke, jossa pyritään parantamaan kuvien kuvailun prosessia jotta kuvat löytyisivät paremmin. Tutkimuksessani tapauksena laajemmassa mielessä on siis viestintäyksikön digitaalinen kuvakokoelma. Olen lähestynyt kohdetta valitsemalla joukon työntekijöitä tapauksiksi, tarkastelemalla heidän työtehtäviään. Tapaukset valitaan usein niiden ainutlaatuisuuden, erityisyyden tai teoreettisen mielenkiintoisuuden vuoksi (Eriksson & Koistinen 2005, 23). Tapaustutkimus valitaan usein lähtökohdaksi kun tutkijalla on vähän kontrollia tapahtumiin, aiheesta tehty vain vähän empiiristä tutkimusta tai tutkimuskohteena jokin tämän ajan elävässä elämässä oleva ilmiö. Tapaustutkimukselle

ominaisia kysymyksiä ovat mitä, miten ja miksi- kysymykset (Eriksson ja Koistinen 2005, 4-5.)

### **3.3 Empiirisen aineiston keruu**

Olen koonnut empiirisen aineiston kahdessa osassa. Keväällä 2011 keräsin aineistoa seminaarityötäni varten ja keväällä 2014 tekemälläni haastattelulla täydensin sekä päivitin tietoja valokuva-arkiston nykytilanteesta ja kehityksestä. Vaikka osa aineistosta on jo pari vuotta vanha, on se edelleen suhteellisen tuore. Aihe on edelleen ajankohtainen. Keväällä 2014 varmistin viestintäyksiköstä ettei aineisto ole olennaisilta osin muuttunut (Venetjoki, R. 2014.)

Aineistonkeruumenetelmänä tutkimuksessa on käytetty havainnointia, keskustelua ja haastateltu viestintäyksikön työntekijöitä. Kävin viestintäyksikössä useita kertoja. Ensimmäisen, kartoittavan tapaamisen jälkeen laadin haastattelukysymykset. Tapauksiksi valittiin aluksi 6 työntekijää ja heidän tehtävänsä. Tutkittava kohde on kuvakokoelman kuvailuprosessi, josta voi saada tietoa vain toimijoiden kautta. Tässä tutkimuksessa viestintäyksikön työntekijät toimivat informanteina. Kysymyksillä pyrittiin kartoittamaan jokaisen työntekijän näkökulmasta kuviin liittyviä työvaiheita. Eri henkilöillä prosessit tietysti vaihtelivat. Haastattelukierroksella pyrin kartoittamaan kuinka kuvien tallennuksen, kuvailun ja haun prosessi kokonaisuudessaan toimii.

Täydentävässä aineistonkeruussa haastattelin ja havainnoin kuvaajaa uudelleen, sillä kuvaaja oli tällä välin vaihtunut ja näin tarjoutui myös mahdollisuus rikastaa havaintoja ja päivittää kokoelman tietoja. Informanteina oli siis yhteensä 7 työntekijää, joista kuutta haastateltiin ensimmäisillä haastattelukierroksilla 2011 ja yhtä 2014.

Haastatteluun ja keskusteluun 2011 osallistui haastatteluissa kuusi henkilöä, joista havainnoitavana oli kolme henkilöä. Havainnoinnissa, keskustelussa ja haastatteluissa 2011 käytettiin vastausten tallennusmenetelmänä käsin kirjoittamista koska se soveltui työvaiheiden kuvaamiseen parhaiten. Lisäksi otin ruudunkaappauskuvia kuva-arkiston kansiorakenteesta.

Haastattelut toteutettiin useammalla käyntikerralla työaikatauluista johtuen. Kävin viestintäyksikössä kolme kertaa: 3.12.2010, 15.4.2011 ja 6.5.2011. Ensimmäisellä kerralla tapasin Riikka Venetjoen (tiedottaja, www-sivuston kokonaisuus, intran ylläpito ja kehitys), Marjo-Rita Valtosen ja Pirjo Achtén (tiedottaja, lehdistötiedotteet). Heidän kanssaan keskustelimme kuvien järjestämisestä. Kirjasin tapaamisen pohjalta kysymyksiä ja ongelmia joita keskustelussa oli ilmennyt, joiden pohjalta laadin seuraavan tapaamisen haastattelukysymykset.

Toisella kerralla tein haastattelukierroksen. Informantteina toimivat Pirjo Achté (tiedottaja, lehdistötiedotteet), Aikalaisen kuvaaja Teemu Launis, Heikki Laurinolli (tiedotomittaja Aikalainen -lehden päätoimittaja), Tiina Wesslin (tiedottaja, tapahtumatiedotus, Aikalaisen toimittaja) ja Sinikka Jokinen (taittaja). Työntekijät esittelivät minulle työvaiheitaan. Keskustelimme samalla aiheesta, ja vapaamuotoinen kysely suoritettiin tekijän työpisteellä, jossa hän pystyi tarkistamaan prosessin eri vaiheita.

Pyrin hahmottamaan kuvan elinkaaren eri vaiheet Viestintäyksikössä. Haastattelu toteutettiin pyytämällä informantteja kuvaamaan piirtämällä ja kirjoittamalla kuvan kuvailun prosessin, jota varten oli laadittu valmiiksi täytettävä paperi. Prosessikaavion täyttämiseen kului noin 20-30min. Prosessikaavio löytyy liitteenä (Liite 1).

Kävin vielä kolmannen kerran viestintäyksikössä haastattelemassa Riikka Venetjokea ja havainnoimassa arkistoa. Otin myös ruudunkaappauskuvia kuva-arkiston kansiorakenteesta.

2011 aineistossa kartoitin kuvan elinkaarta laajemmin ja 2014 haastattelussa keskityin tarkemmin kuvan kuvailuvaiheeseen. Viestintäyksikön työnjaossa kuvaajalla on olennainen rooli kuvan kuvailussa, sillä hän tallentaa uudet kuvat kameralta palvelimelle. Viestintäyksikössä muutkin työntekijät osallistuvat kuvailuun, mutta koen aiheelliseksi keskittyä kuvaajan prosessiin. Kuvaajalla on ratkaiseva rooli tuoreiden, uusien kuvien kuvailussa. Kuvaajan K1 haastattelurunko löytyy liitteenä (Liite 2). Lisäksi täydensin aineistoa myös käyttäjien kuvien hakemisen tapojen osalta. Kevään 2014 haastattelussa äänitin kuvaajan K2 haastattelun. Haastattelurunko löytyy liitteenä (Liite 3). Haastattelun tarkoitus oli täydentää 2011 tehdyssä haastattelussa saatuja tietoja. Lisäksi aineistoa on täydennetty henkilökohtaisilla tiedonannoilla sähköpostitse.

### **3.4 Aineiston analyysi**

Aineiston sisällönanalyysi aloitettiin tunnistamalla haastattelu- ja havaintoaineiston pohjalta informanttien työtehtäviä. Haastattelussa kartoitettiin laajemmin koko kuvan elinkaareen liittyviä työvaiheita ja työtehtäviä sekä tarkennettiin sitten kysymyksiä koskemaan itse kuvien kuvailua. Kokonaisuuden kartoittamiseksi ja prosessin hahmottamiseen oli tehtävä havaintomatriisi, jota varten tunnistettiin aineistosta taustaksi informanttien työtehtäviä. Havaintomatriisin avulla pyrittiin ryhmittelemään kuvan kuvailun kannalta olennaiset tehtävät sekä löytämään päällekkäisyyksiä tehtävissä kuten myös tehtäviä, joita suorittivat vain yksi tai muutama työntekijä.

Näiden vaiheiden hahmottamisen ja erittelyn jälkeen saatoinkin alkaa järjestämään haastattelu- ja havainnointiaineistoa analyysia varten. Aineiston koodaaminen aloitettiin jaotellamalla haastattelu- ja havaintoaineistoa tutkimuskysymysten 1-3 mukaisiin temaattisiin kokonaisuuksiin. Aineistosta on koodattu lauseita ja kappaleita, jotka kohdistuvat tutkimuskysymysten mukaisiin teemoihin.

Aineiston järjestämisen jälkeen seurasi varsinainen analyysi. Kävin haastatteluaineistoa lukemalla läpi tunnistuen aineistosta koodattujen teemojen mukaan järjestettyjä lauseita ja kappaleita. Vertailin informanttien käsityksiä toisiinsa ja näin aineistosta alkoi hahmottua esiin tutkimuksen kannalta kiinnostavia seikkoja. Haastatteluaineiston erittelyä täsmensin omien havaintojen ja muistiinpanojen pohjalta. Näistä havainnoista ja muistiinpanoista oli hyötyä analyysivaiheessa, sillä informantit näyttivät minulle työvaiheitaan jolloin sain visuaalista informaatiota joka ei luonnollisestikaan tallennu äänitiedostoon. Kävimme lähinnä kuvaajien kanssa työvaiheet läpi yksittäisien kuvatiedostoesimerkkien avulla. Tällöin informantit tekevät myös asioita joita he eivät sano ääneen, ja ilman havainnointia prosessi olisi jäänyt vaikeaksi ymmärtää. Vaikka videokuva olisi sinällään ollut hyvä tallennusmuoto työtehtävien kuvaamiseen, hylkäsin ajatuksen, koska halusin että informantit voivat rauhassa esitellä työtään. Lisäksi oletin saavani rehellisempiä mielipiteitä kahdenkeskisessä keskustelussa.

Analyysivaiheessa huomasin miten tärkeää on ymmärtää arkiston kansiorakenne ja kuvan kuvailun prosessin eri vaiheet sekä kuvan elinkaari Tampereen yliopiston viestintäyksikön kuva-arkistossa. Kuvan kuvailuun liittyviä työvaiheita ja kuva-arkiston kansiorakennetta olen kuvannut sanallisesti kansiorakenteesta otettujen ruudunkaappauskuvien kansionäkymien avulla, joita tutkimalla arkiston kansiorakenne on kokonaisuudessaan helpompi hahmottaa. Myös informanttien laatimat prosessikuvaukset ovat helpottaneet eri vaiheiden tunnistamista.

Aineiston sisällönanalyysissä nostettiin esille tehtävät jotka liittyvät kuvien kuvailuun. Työntekijöiden tehtäväkenttää kuvaavan havaintomatriisin pohjalta nousi esille että kuvaaja on keskeinen henkilö kuvien kuvailun prosessissa. Tutkimuksen fokus painottuu kuvaajan tehtäviin, jotka kohdistuvat kuvien kuvailuun. Aineistoa koodatessa on otettu huomioon kaikki haastattelut, mutta koska kuvaajalla on erityinen asema kuvien kuvailuprosessissa, erotan kuvaajien suorat lainaukset muista työntekijöistä lyhenteillä k1 ja k2. Muiden työntekijöiden suhteen käytän suorissa lainauksissa lyhennettä h1, h2, h3 jne. Kuvattaessa kuvan elinkaarta ja työntekijöiden tehtäviä, sekä haastateltujen esittelyssä voivat nimet sekä toimenkuvat (esim. tiedottaja) tulla esille selkeämmän kokonaiskuvan saamiseksi.

## **4 KUVA-ARKISTON NYKYTILANNE**

Tässä luvussa vastaan kahteen ensimmäiseen tutkimuskysymykseen. Luvussa 4.1 vastaan kysymykseen mitkä ovat Tampereen yliopiston viestintäyksikön digitaalisen valokuva-arkiston käyttötarpeet? Luvuissa 4.2, 4.3 ja 4.4 vastaan kysymykseen minkälaisista kuvista (kuvatyytit) arkiston sisältö koostuu ja millainen on kuvan elinkaari arkistossa?

Tässä luvussa kartoitetaan Tampereen yliopiston viestintäyksikön kuva-arkiston nykytilannetta. Tässä luvussa kuvaillaan myös arkiston sisältöä, kartoitetaan arkiston käyttötarpeita ja esitellään kuvan elinkaari viestintäyksikön arkistossa. Pohjana kartoitukselle ovat olleet vapaamuotoiset tiedustelut, työtehtävien havainnointi sekä informanttien haastattelut. Lisäksi on havainnoitu ja dokumentoitu digitaalista valokuva-arkistoa. Luvussa 4.1 esitellään Tampereen yliopiston viestintäyksikön digitaalisen valokuva-arkiston keskeiset käyttötarpeet. Luvussa 4.2 esittelen kuva-arkiston sisällön ja luvussa 4.3 käsitellään kuvien päätymistä arkistoon elinkaarimallin avulla. Luvussa 4.4 esittelen kuvatyytit viestintäyksikön arkistossa.

### **4.1 Viestintäyksikön digitaalisen arkiston käyttötarpeet**

Tampereen yliopiston viestintäyksikkö huolehtii yliopiston keskitetystä sisäisestä ja ulkoisesta viestinnästä ja tarjoaa valokuvapalveluita. Kuvia etsitään arkistosta eri tarkoituksiin. Kuvapyyntöjä tulee niin yliopiston sisältä kuin ulkopuolisiltakin asiakkailta. Kuvia pyydetään esimerkiksi markkinointi- ja esitemateriaaliksi, kuten esimerkiksi opinto-oppaisiin. Arkisto tarjoaa valokuvia yliopiston ja yhteistyökumppanien käyttöön sekä mediakäyttöön. Ulkoiset käyttäjät voivat pyytää viestintäyksikön kautta arkistossa



olevia kuvia käyttöönsä. Lisäksi Viestintäyksikön kuva-arkiston kuvia käytetään yliopiston verkkojulkaisujen kuvitukseen.

Arkiston keskeiset käyttötarpeet ovat:

1. *Aikalainen* -lehden juttukuvitus ja kuvien arkistointi
2. Yliopiston muiden (verkko) julkaisujen kuvitus
3. Yliopiston erilaisten esitemateriaalien, verkkosivujen ym. kuvitus
4. Kuvapyynnöt yliopiston ulkopuolelta

(Venetjoki, R. 2010.)

Viestintäyksikön arkiston kuvilla on myös erilaisia statuksia niiden käyttöoikeuksien mukaan. Kuvia on talonsisäiseen käyttöön, kaikille käyttöön, ja kansainväliseen käyttöön. Eri kuvilla on eri käyttöoikeudet, joka vaikuttaa myös kuvien järjestämiseen arkistossa.

Tässä tutkielmassa keskitytään etenkin *Aikalaisen* painetun sekä verkkoversion kuvituskuviin. Arkiston pääasiallinen käyttö painottuu *Aikalainen*-lehden juttukuvitukseen ja kuvien arkistointiin. *Aikalainen* on tieteen, opetuksen ja kulttuurin sanomalehti, joka ilmestyy yhdeksän kertaa vuodessa. Painetun version lisäksi *Aikalainen* ilmestyy myös verkkolehtenä, johon käytetään myös arkiston kuvia. Kierto kuvan ottamishetkestä arkistoon on siis nopea. Yksinkertaistetusti kuva otetaan, siirretään koneelle digitaalisessa muodossa olevaksi tiedostoksi ja julkaistaan jonka jälkeen se päättyy arkistoon, josta se voidaan hakea uudelleen käytettäväksi. Kuvan elinkaari ei kuitenkaan ole näin yksinkertainen. Tarkastelen yksityiskohtaisemmin kuvan elinkaarta luvussa 4.3 Kuvan elinkaari viestintäyksikön arkistossa.

## 4.2 Arkiston sisältö

Viestintäyksikön kuva-arkiston voi jaotella sekä digitaaliseen että paperiseen (fyysiseen) arkistoon. Viestintäyksikön fyysisen arkiston paperikuvat sijaitsevat kellariarkistossa, jossa paperikuvat, diat ja negatiivit on arkistoitu riippukansioihin ja mappeihin osa teeman, osa kuvausajankohdan mukaan. Jonkin verran sekä viestintäyksikön kellariarkiston että yliopiston arkiston paperikuvia on skannattu digitaaliseen muotoon omaan kansioonsa levypalvelimelle, mutta historiallisten kuvien käyttöä tapahtuu harvoin. Arkiston tuoreimmat ja samalla käytetyimmät kuvat sijaitsevat levypalvelimen digitaalisen arkiston kansioissa.

Tuoreimmat ja käytetyimmät kuvat löytyvät digitaalisessa muodossa levypalvelimella. Levypalvelimen kuva-arkisto sisältää myös jonkin verran skannattuja vanhoja kuvia yliopiston arkistosta. Kuvat on arkistoitu osin teeman mukaisiin kansioihin, osin kuvaajan mukaan. Kuvien tunnistetietojen määrä ja laatu vaihtelevat. Tunnistetiedot ovat lähinnä kuvatiedostojen nimissä, josta tarkempi kuvaus edempänä. Erillisellä palvelimella ovat lisäksi *Aikalainen* -verkkolehden kuvituskuvat, jotka ovat usein käsiteltyjä versioita myös painetussa lehdessä käytetyistä kuvista. (Venetjoki 2010.)

Viestintäyksikön digitaalisen kuva-arkiston kuvat voi karkeasti jaotella nykykuviin ja arkistokuvaan. Käytön jälkeen nykykuva muuttuu arkistokuvaksi, joka voidaan ottaa uudelleen käyttöön.

Viestintäyksikön kuvakokoelma jakautuu kahdelle verkkolevylle, joista ensimmäisellä verkkolevyllä oli vuonna 2014 yhteensä 41 725 kpl kuvaa ja toisella 41 749 kpl. Toisen verkkolevyn kuvista kuitenkin suurin osa kuvista on kopioita. Lisäksi kokoelman kuvia on myös alakansioissa joita ei tähän laskelmaan saatu kartoitettua. Kuvakokoelmasta löytyy henkilökuvia 12 492 kpl (tilanne 25.10.2014). Jos toisen verkkolevyn kuvakopi-

oita ei lasketa arkiston kuvamäärään, voidaan arvioida että henkilökuvien osuus arkiston kuvista on 29,9 %. Arkiston henkilökuvien määrän näkee suhteellisen helposti, koska ne ovat kaikki omassa ”henkilökuvat”-kansiossaan (Renvall, J. 2014.)

Henkilökuvat -kansio onkin järjestämisen suhteen koko kuva-arkiston selkein kansio. Kansio sisältää vain henkilökuvia jotka on arkistoitu henkilön sukunimellä yhdenmukaisesti. Kansiorakenne on hierarkkinen niin, että L:yhteinen –aseman alla sijaitseva X-aseman pääarkisto-kansio kattaa kaikki alla olevat eri teeman mukaan nimetyt kansiot, kuten henkilökuvat -kansion, eri kuvaajien omat kansiot, historiakuvat –kansion, ryhmäkuvat –kansion sekä *Aikalaisen* verkkokuvat –kansion. Lisäksi X –asemalta löytyy paljon muita kansioita. Kaikki kansiot eivät ole pääarkistokansiossa, sillä *Aikalainen* –kansio löytyy L:yhteinen aseman alta.

### **4.3 Kuvan elinkaari: miten kuva päätyy arkistoon?**

Kuvan elinkaarella tarkoitetaan tässä kuvan vaiheita sen syntyhetkestä hävittämiseen tai siirtymiseen pitkäaikaiseen säilytykseen. Digitaalisen valokuvan elinkaari eroaa perinteisen valokuvan elinkaaresta arkistointiin liittyvien käsitteiden sekä seulonnan osalta. Digitaalisia kuvia voidaan poistaa kameran muistista heti kuvaushetken jälkeen, joten seulonta voi tapahtua jo elinkaaren alkumetreillä. Digitaalisten valokuvien kohdalla puhutaankin pitkäaikaisesta ja lyhytaikaisesta säilytyksestä pysyvän ja määräaikaisen säilytyksen sijaan. (Penttinen 2007, 46–47.)

Seuraavassa esitellään valokuvan elinkaari yleisellä tasolla, jonka jälkeen tarkastellaan valokuvan elinkaarta viestintäyksikön digitaalisessa kuva-arkistossa. Valokuvan elinkaaren voi yleisesti sanoa alkavan kuvausimpulssista, joka johtaa valokuvaustapahtu-

maan. Kuvausimpulssi tarkoittaa kuvatarpeen syntymistä. Kuvausimpulssi on syy joka aiheuttaa kuvan ottamisen, esimerkiksi tarve ottaa kuvituskuva ratsastusaiheiseen lehti-juttuun.

Seuraavassa vaiheessa on itse valokuvaustapahtuma, jossa kuva tallentuu digitaaliseen muotoon kameralle. Tämän jälkeen, mikäli valokuva päätetään säilyttää, siihen liitetään oleelliset kontekstitiedot. Kontekstitiedot voivat olla kamerassa automaattisena asetuk-sena, kuten päivämäärä, tai ne voidaan lisätä erikseen esimerkiksi tiedostonimeen kun kuva tallennetaan tietokoneelle. Seuraavaksi valokuvasta tehdään käyttökuva, käsitte-lemällä kuva käsitellään sellaiseen muotoon jossa sitä käytetään ja lopulta kuva arkis-toidaan. Valokuvia voidaan seuloa ja järjestää myös jälkeinpäin. Digitaalisesta valoku-vasta voidaan myös tehdä eri versioita, sitä voidaan muokata, kopioida ja tallentaa eri muodoissa. Digitaalisen valokuvan elinkaaren vaiheet ovat siis lyhyesti: kuvausimpuls-si, valokuvaustapahtuma, digitaalinen kuva, hävittäminen tai ensimmäinen seulonta, kontekstitietojen liittäminen, kuva-aineiston hallinta eli kuvien käsittely, arkisto tai käyttö. (Penttinen 2007, 47-48).

Seuraavassa esitellään kuvan elinkaari viestintäyksikössä. Elinkaaren pohjana ovat in-formanttien piirtämät prosessikaaviot, joissa he ovat kuvailleet kuvan vaiheita omassa työskentelyssään. Tässä luvussa käytän esimerkkinä *Aikalaisen* digitaalisessa muodossa olevan kuvituskuvan elinkaarta. Uuden kuvan elinkaari kulkee *Aikalaisen* toimituksessa käsi kädessä lehtijutun kanssa. Päätoimittaja ja kuvaaja käyvät toimituspalavereita, jois-sa pohditaan kuvitusta. Lähtökohtaisesti pyritään aina saamaan uusi ja tuore kuva, jol-loin näkökulmakin aiheeseen on tuore. Kiireisin vaihe onkin juuri ennen lehden ilmes-tymistä, jolloin prosessi on kehämäinen: kuvaaja, päätoimittaja sekä taittaja käyvät läpi eri vaihtoehtoja. Tässä prosessissa myös kuvakoko muuttuu ja kuvaa tallennetaan eri tiedostomuotoihin. Mikäli tuoretta kuvaa ei voida käyttää, kuva haetaan kuva-arkistosta.

Viestintäyksikössä valokuvan elinkaaren katsotaan alkavan tapahtumasta tai juttuideasta, joka toimii edellä mainittuna kuvausimpulssina. Seuraavassa vaiheessa on itse valokuvaustapahtuma, jossa kuva tallentuu digitaaliseen muotoon kameralle. Viestintäyksikön arkistossa kuva esikäsitellään ja siirretään keskeneräisten kuvien kansioon. Seuraavaksi taittaja hakee kuvan kansioista, suunnittelee taiton, tallentaa kuvan käsiteltävien kuvien kansioon ja ilmoittaa ohjeet kuvan käsittelystä kuvaajalle. Kuvaaja hakee kuvan kansioista sekä tekee lopulliset käsittelyt (esim. rajaus, värimääritys) ja lopulta siirtää kuvan valmiit kuvat kansioon. Kun tämä kiireisin vaihe on saatu valmiiksi, kuva arkistoidaan omalla tiedostonimellään. Kuva arkistoidaan eri kansioihin, esimerkiksi kuvaajalla on omat kansionsa. Lisäksi kuva siirretään teeman mukaiseen kansioon, josta se voidaan taas hakea käyttöön.

#### **4.4 Kuvatyyppit Tay:n viestintäyksikön kuva-arkistossa**

Digitaalisen kuvan elinkaari on edellisessä luvussa kuvattu erittäin yksinkertaisena. Tässä yhteydessä on vielä syytä esitellä tarkemmin erilaiset kuvan lähteet, kuvatyyppit ja käyttötarkoitukset, ts. mitä erilaisia versioita kuvista arkistossa on. Tällä pyrin selvittämään syitä siihen miksi kuvien kuvailuprosessi ja arkiston kansiorakenne eivät ole suoraviivaisempia.

Voidaan todeta, että digitaalinen kuva ja samalla digitaalinen kuva-arkisto ovat paljon kompleksisempia käsitteitä kuin paperikuva. Kuvaa voidaan monistaa ja siitä tehdään eri versioita. Digitaalisessa ympäristössä käsite originaali muuttuu problemaattiseksi. Voidaankin pohtia, mikä kuvan lukuisista versioista sitten on se arkistoitava, säilyttämisen arvoinen. Viestintäyksikössä arkistoidaan kuva useammassa muodossa useaan kansioon.

Kuva arkistoidaan eli tallennetaan arkiston kansioon heti kun se siirretään kamerasta palvelimelle. Lisäksi kuva arkistoidaan siinä muodossa, joka on lehdessä painettu. Muodolla tarkoitetaan ilmiä, esimerkiksi värimäärittelyjä sekä rajausta, en tiedostomuotoa. Painotaloa varten kuva muutetaan painokelpoiseen tiedostomuotoon, tekemällä siitä uusi versio. Kaikki *Aikalaisessa* julkaistut kuvat arkistoidaan Arkisto-kansioon sekä *Aikalainen* -kansioon.

Haastatelluista informantti H5 jaotteli arkiston kuvat kolmeen eri kuvan lähde - kategoriaan, jotka ovat:

- Aikalaisen kuva
- arkistokuva
- uusi kuva

*Aikalaisen kuvan* elinkaari tuli jo esiteltyä edellä kappaleessa, mutta lyhyesti kuva päätyy arkistoon näin: 1) jutun valinta, 2) kuvapyyntö kuvaajalle, 3) kuvaaja lähettää kuvan (verkkolevy-[www-palvelin](#)), 4) [L:/yhteinen/www-arkisto/etusivun](#) valokuva (kuva saa uuden nimen), 5) kopiointi [www-palvelimelle](#) [www.uta.fi/ajankohtaista/kuvat](#) [vp-julkaisujärjestelmällä](#), 6) kuva julkaistaan joko *Aikalainen* printtilehdessä tai verkkoversiossa.

Arkistokuva tarkoittaa tässä kuvaa joka haetaan arkistosta joko sisäiseen tai ulkoiseen käyttöön. Sisäisessä käytössä prosessi etenee näin: 1) jutun valinta, 2) teemakuvan valinta kuva-arkistosta: muistinvarainen/kysely huomioitava julkaisu-oikeudet, 3) kuvan käsittely ja prosessi jatkuu samalla lailla kuin *Aikalaisen kuvan* elinkaaren tapauksessa.

Ulkoisella käytöllä tarkoitetaan tässä ulkoa tulevaa kuvapyyntöä: 1) kuvapyyntö ulkoa: kartoitetaan tema ja tarve, 2) arkiston selailu ja kuvien valinta, 3) kuvien siirto [www-](#)

palvelimelle ”katseluvälimuistiin”, 4) asiakas valitsee, 5) valittujen kuvien lähetys sähköposti/cd, 6) kuvien poisto palvelimelta.

Uusi kuva eroaa Aikalaisen kuvan käytöstä siinä, ettei sitä oteta tiettyä juttua varten vaan se kuvataan suoraan arkistoa varten. Kuvatarve voidaan ikään kuin ennakoida. Tällaisia kuvia ovat hyvin usein esim. yleiset kuvat yliopistosta tai opiskelijoista. Uusi kuva päättyy arkistoon näin: 1) kuvasuunnittelu kuvaajan kanssa, 2) kuvaaja kuvaa, 3) kuvan valinta (kuvaaja), 4) käsittely (kuvaaja), 5) tallennus x: www- ja esitekuvia, 6) www-palvelimelle kopiointi vp-järjestelmään, 7) julkaisu.

Työntekijät tallentavat kukin osaltaan kuvan arkistoon kun lopullinen käsittely on tehty. Kuva voi olla siis täysin uusi, erityistä tarkoitusta varten kuvattu, uudesta kuvasta muokattu, vanha arkistokuva, vanhasta kuvasta muokattu tai jopa ulkoisista kuvapankeista hankittu ja esimerkiksi Aikalainen-lehdessä julkaistu kuva. Kuvat voivat päättyä erilaisiin käyttötarkoituksiin, kuten Aikalaisen kuva (esim. pääuutiskuva), www-kuva (esim. kuvapalkki tai banneri). Kuvat päättyvät myös medialle ja erialaisiin sisäisiin ja ulkoisiin julkaisuihin kuten vaikkapa tiedotteisiin ja opinto-oppaisiin.

## 5 KUVIEN KUVAILUKÄYTÄNNÖT ARKISTOSSA

Tässä luvussa vastaan kolmanteen tutkimuskysymykseen eli millaisia ovat kuvan kuvailukäytännöt valokuva-arkistossa? Aluksi luvussa 5.1 käsitellään arkiston työntekijöiden työnjakoa ja yhteistyöstä kuvien kuvailun suhteen. Luvussa 5.2 eritellään kuvien kuvailuelementtejä ja niiden esiintymistä arkistossa informanttien ja havaintoaineiston pohjalta.

### 5.1 Kuvien kuvailun työnjako

Tampereen yliopiston viestintäyksikön arkistossa ei ole erillistä arkistonhoitajaa, joten työntekijät hoitavat kuvailun ja kuvien tallentamisen arkistoon oman työnsä ohessa. Tosin toiset työntekijät ovat enemmän tekemisissä kuvien kanssa päivittäisessä työssä, kuten tietysti kuvaaja, joka tuottaa uutta kuvamateriaalia. Haastattelemani kuvaaja luonnehtiikin osuvasti sitä miten työntekijät hoitavat arkistoon tallentamisen omien töidensä ohessa:

K2: Mää en ole arkistonhoitaja vaan mä oon valokuvaaja jonka pitää tuottaa lisää sinne että muutenhan tää olis se että mä hengaan päivät pitkät täs koneella ja katon eikä uutta kuva tuu ollenkaan (...) tässä tulee se että siitä niinkun puuttuu yks henkilö, mun pitäis vielä olla sellanen tyyppi joka osais niinku pitää sitä loogisessa järjestyksessä.

Kuvaajan mainitsema tilanne lienee hyvin yleinen eri organisaatioissa; kuvien kuvailu hoidetaan muun työn ohessa eikä arkiston systemaattiseen järjestämiseen ole usein palkattu erikseen arkistonhoitajaa. Haastatteluissa usein mainittu ongelma onkin kiire. Kuvan kuvailun tekevät Viestintäyksikön henkilökunta eli lähinnä kuvaajat ja toimittajat.



Kuvien haku perustuu muutaman avainsanan ja muistin varaan. Toimittajien pitää muistaa kansiorakenteet, mitä missäkin kansiossa on, tai kuka kuvaaja on kuvannut minkäkin tapahtuman. Kansioiden nimissä ja kuvien tiedostonimissä on paljon implisiittistä tietoa, joten kuvan tiedostonimi ei kerro välttämättä kontekstia tuntemattomalle mitään.

Työntekijä, joka on ollut pitkään talossa, osaa esimerkiksi etsiä valokuvan jostain erityisestä tapahtumasta, jos hän muistaa kuka oli kuvaamassa tapahtumaa. Näin hän voi haakea kuvan kuvaajan kansioista. Kuvia haetaankin usein arkistosta muistinvaraisesti, muutaman vihjeen perusteella. Kuvaaja antaa usein konkreettisen esimerkin tällaisesta tilanteesta:

K2: ...näähän on tällaisia että (...) ”muistaks sä kun sä olit silloin ja silloin kuvaamassa, niin siellä oli se ja se henkilö, mutta että hänestä ei löydy nyt kuvaa...

H1: ensin kysyn H4 tai H5 onko halutusta aiheesta kuvaa, koska he tuntevat arkiston sisällön paremmin (...) jonkun verran selailen kansioita, mutta materiaalia on niin paljon, että kansioiden nimestä en löydä etsimääni ja materiaali on aika vanhaa...

## **5.2 Kuvien kuvailun elementit**

Kuvailu ei ole suoraviivainen prosessi. Arkistossa on erilaisia kuvia eri käyttötarkoitusta varten, kuten luvussa 4.4 osoitettiin. Kuvasta on lisäksi monta eri versiota. Toisaalta informantit olivat hyvin perillä miten kuvailu pitää tehdä (jotta kuvat löytyisivät) eli millainen tiedostonimi kuvalle annetaan kun se tallennetaan arkistoon. Malli tiedostonimien muotoilemiseksi näkyy toki aiemmin tallennettujen tiedostojen tiedostonimistä.

Lisäksi kuvan kuvailuun ja kuvan hakuun liittyy olennaisesti työntekijöiden yhteistyö: työntekijän pitää tietää keneltä voi kysyä jos etsitty kuva ei löydy. Pidempään talossa olleet muistavat paremmin mitä eri kansioissa on. Kuvan kuvailuun liittyykin paljon työntekijöiden muistinvaraista tietoa. Informanttien käsitykset siitä, pitäisikö kuvia kuvailla tarkemmin, erosivat jonkin verran. Lisäksi haastatteluissa ilmeni eriäviä käsityksissä arkiston toimivuudesta. Tätä kysymystä tarkastellaan yksityiskohtaisemmin luvussa 5.3.

Havaintoaineiston (esim. ruudunkaappauskuvat arkistosta) perusteella voi jo sanoa, että kuvien kuvailu ei ole täysin yhdenmukaista. Vaikka kansiorakenne on hierarkkinen, samalla tavalla nimettyjä kansioita löytyy eri paikoista eivätkä tiedostonimet ole yhdenmukaisia.

### **5.2.1 Esimerkki kuvan kuvailusta arkistossa**

Esitän tässä esimerkkinä yhden kuvatiedoston nimeämisen. Valitsin tapaukseksi henkilökuvan tallentamisen, koska informanttien mukaan ja omien havaintojen perusteella henkilökuvat oli selkeimmin järjestetty arkiston kansio. Henkilökuvat -kansio sijaitsee Viestintäyksikön L:yhteinen aseman X:pääarkisto -kansiossa. Henkilökuvat -kansion sisällä kansiot ovat nimetty kuvatun henkilön mukaan niin, että sukunimi on ennen etunimeä. Lisäksi kansiot ovat aakkosjärjestyksessä sukunimen mukaan. Esimerkiksi kansio voi olla nimeltään Pirkkala Pekka. Suurin osa kansioista on nimetty näin yhdenmukaisesti, mutta löytyy myös kansioita joiden nimissä on muitakin attribuutteja kuin nimi, esimerkiksi ”aamupala ja Ogutu”.

Pyysin informantteja antamaan konkreettisen esimerkin, miten henkilökuva tallennetaan arkistoon. Halusin nimenomaan tietää tarkasti tiedostonimen, joka kuvalle annetaan kun

se tallennetaan henkilökuvat-kansioon. Informanttien puhuessa kuvien kuvailusta he tarjosivat yhdenmukaisen kuvauksen siitä, mitä tietoja kuvatiedoston tulee sisältää kun se tallennetaan arkistoon. Toisaalta kun he kertoivat konkreettisen esimerkin tiedostonimestä, kaikki antavat erilaisen kuvauksen. Eroja on siis kuvaus elementtien järjestyksessä ja muodossa, ei kuitenkaan sisällössä.

Kuvaajilla oli samat elementit mutta ne olivat eri järjestyksessä. Informantit käyttivät esimerkeissään eri nimiä, mutta yhdenmukaisuuden vuoksi annan esimerkin samalla nimellä. Kuvaajat muistivat mainita kuvaajan nimikirjaimet joka on tekijänoikeuksien kannalta olennainen.

K1: pvm\_sn\_en\_kn.jpg

esim: 1.1.2011\_Holli\_Kaija\_K1.jpg

K2

sn\_pvm\_kn.jpg

esim: Holli270314K2

Sen sijaan muut informantit eivät maininneet kuvaajaa tiedostonimi –esimerkissään: Kuvaajan nimi mainittiin kyllä sanallisesti kerrottaessa. Toisaalta voi olla etteivät kaikki vain muistaneet mainita asiaa.

H5

vkkpv tunniste

20110506Holli

Viestintäyksikön arkistossa keskeisiä tietoja ovat: kuvaajan tiedot, kuvausaika sekä mitä tai kuka kuvassa on. Kuvatiedosto on voitu nimetä esimerkiksi: tre-talo\_11-04-08.jpg. Tiedostonimi voi olla myös olla muotoa: tynkkynen\_mantyla\_jr.jpg, jossa jr ovat kuvaajan nimikirjaimet. Tiedostonimen lisäksi kuvan tiedoissa (*details*) näkyy kuvan tiedostomuoto, kuvanottohetken päivämäärä ja kellonaika sekä viimeisin päivämäärä ja kellonaika, jolloin kuvaa on muokattu. Kuvan koko sekä mitat näkyvät myös. Jos kursoria liikuttaa tiedostonimen päälle, saa näkyviin selailukuvan (*thumbnail*) joka on sormenpään kokoinen pienoiversio itse kuvasta.

### **5.2.2 Henkilö**

Shattfordin attribuuteista kuka esiintyy paljon viestintäyksikössä, erityisesti henkilö kuvien osalta. Henkilö vastaa tavallaan myös mitä-attribuuttiin. Arkistossa henkilökuvia löytyy henkilökuvat –kansioista ja ryhmäkuvat kansioista. toki henkilön/henkilöiden nimiä käytetään kuvailutiedoissa myös muista kansioista löytyvissä kuvissa, kuten esimerkiksi historiakuvat –kansiossa.

Henkilökuvat oli ainut selkeästi järjestetty kansio, jossa tiedostonimessä esiintyi henkilön nimi. Kansion selkeys ja helppo haettavuus selittyvät kuvan luonteella. Henkilokuva noudattaa tarkkaa formaattia ja aihe on lähes yksiselitteinen, kuvan hakija tietää millä nimellä kuva on kuvailtu.

### **5.2.3 Päivämäärä**

Haastatteluissa kaikki informantit mainitsivat päivämäärän tärkeänä metatietona. Päivämäärä tiedostonimessä sisältää kontekstitietoja, joita voisi olla hidasta kuvailla. Haas-

tatteluissa mainittiin, että päivämäärä kertoo kuvassa olevan vuodenajan, miljööön (esimerkiksi minkä näköisiä ihmisiä kuvassa on, mitä heillä on päällään) ja mikä tilaisuus kuvassa mahdollisesti on meneillään (esimerkiksi valmistujaiset keväällä). Pelkästään tiedostonimestä voi jo päätellä paljon kuvan sisällöstä ilman että itse kuvaa tarvitsee katsoa.

#### **5.2.4 Tiedostonimi: kuka, mitä, missä ja milloin?**

Viestintäyksikön arkistossa kuva arkistoidaan heti kun se siirretään koneelle. Kuvan tiedostonimeen kirjataan kuvan keskeiset tiedot. Keskeisiä ovat tiedot, jotka mahdollistavat kuvan löytymisen tarvittaessa. Yhden tiedoston tiedostonimeen ei mahdu paljon tietoa.

K2: ...se on ihan maalaisjärjelläkin ajateltuna että mitä tai kuka, missä ja milloin mutta en mä tiedä ei se missä oikeestaan...kuka henkilö tai mikä tapahtuma ja koska ja sit kuka on kuvannut sen on tärkeimmät, koska se missä kysymys on vähän sit sellanen extra, jota voi sit myöhemmin kuvailla lisää että ja sen yleensä sit kuvaaja itse viimestään niin mä nään kyllä ja muistan että missä toi on otettu...

Arkistossa tiedostonimissä esiintyi siis seuraavia attribuutteja: kuka, mitä, missä, milloin. Arkistossa kuka –attribuutti on järjestelmällisesti edustettuna nimenomaan henkilökuissa ja ryhmäkuissa, yleensäkin kuvissa joissa esiintyy ihminen tai ihmisiä. Mitä attribuutti voi olla jokin esine, kuten esimerkiksi rakennus. Missä kertoo paikan ja milloin kertoo kuvanottamisajankohdan. Tärkeimmäksi attribuutiksi mainitaankin milloin.

Muitakin tietoja voitaisiin haastateltavien mukaan kuviin liittää:

K2: Ne on aika silleen aihepiireittäin mä oon luonu kansioita. Aika paljon se kuvio menee niin et mä oon nyt kuvannut jonkun keikan, niin mä tuun tähän koneelle ja hoidan kuvat silleen omaan kansioonsa, että siel on selkeesti päivämäärä ja mitä ja sit kuka on kuvannu (...) tärkeintä on vaan mainita just tää, että mitä siinä kuvassa on, kuka siinä on ja koska se on otettu ja sit että se on kenen kuva se on. Kuka sen kuvan on ottanu. Niin nää on ne ehdottomasti tärkeimmät ne pitää aina muistaa mainita. Ja sit siitä eteenpäin on aika semmosta bonusta mutta voi olla jopa tärkeetäki bonustietoa se metatietojen syöttäminen...

### **5.2.5 Kameran metatiedot**

Kameran metatietoja pidetään kehittämisen arvoisena kuvien kuvailutietojen suhteen. Metatieto tarkoittaa tietoa tiedosta, eli dokumentista tai dokumentin sisällöstä. Digitaalinen kuva saa metatietoja heti kuvanottohetkellä, sillä digitaalisessa kamerassa kuvaan tallentuu automaattisesti tietoja. Haastatteluista ilmeni, että etenkin päivämäärä on hyödyllinen kamerasta tallentuva tieto. Kameraan on myös mahdollista lisätä metatietoja jotka jäävät myös ikään kuin kuvan sisälle, tiedoston itsensä muistiin.

#### **Jääkö kamerasta jotain tietoja?**

K2: Sitä pitää jotenki itsestäänselvytenä niin sieltä jää siis se päivämäärä olettaen että toi kamera on ajan tasalla (...) se jättää tonne noin tiedon mulle siitä että koska se on kuvattu (date created) (...) kellonaikaki näkyy jopa (...) Nää tulee automaattisesti et mun ei tarvii niistä huolehtia.

Haastatteluissa tuli ilmi, että arkistossa käytetyn kameran yhteensopimattomuus kuvankäsittelyohjelman kanssa on aiemmin tuottanut ongelmia. Näin jo hyvin aikaisessa vai-

heessa saatetaan kadottaa tietoja, eikä muita tietoja kuin kuvanottoaika säilyy siirrettävässä kuva koneelle. Haastatteluissa ilmeni, että yhteensopimattomuusongelmat kameran ja kuvankäsittelyohjelman kanssa ovat aiheuttaneet myös kellonajan heittoja, ja siten tuottaneet virheellistä tietoa kuvan ottamisen ajankohdasta. Arkistossa onkin sisällytetty päivämäärä tiedostonimeen mahdollisimman pian, heti tallennusvaiheessa.

Arkistossa on käytössä *Adobe Bridge* -kuvankäsittelyohjelma. Haastattelujen välillä versio on päivitetty, ja nykyinen versio on nopeuttanut kuvien purkua ja kuviin saa automaattisesti sisäistä metatietoa.

Kameran ajantasaisuusongelmista huolimatta kameran metatietoja pidetään arkistossa hyödyllisinä. Arkistossa hyödynnetään kameran metatiedoista lähinnä päivämäärää, joka kertoo milloin kuva on otettu.

Kameran valmiina asetuksina olevista metatiedoista löytyy jo muitakin tietoja. Sinne on myös mahdollista lisätä tapauskohtaisesti enemmänkin kuvaan liittyviä metatietoja.

K2: On tässä muutaki työtilaa mulla mut kuten huomaat tää *Bridge* on sydän tässä työkentelyssä. Hirvee määrä kaikkea tietoja löytyy sieltä mutta sitte nää on melkein tärkeimmät et mä nään tosta sen että koska tää on tehty. Sitte täältä näkyy kaikki nää tekniset tiedot nää suljinajat sun muut.

### **5.3 Kuvan kuvailun kehittäminen käyttäjien silmin**

Tutkimuksen alkuvaiheessa kehitystoiveiksi mainittiin esimerkiksi yhtenäinen arkistointikäytäntö, arkistointiohjeet sekä arkiston kuvaus, nykyistä helpompi haku, vaivaton

kuvien jakaminen ulkopuolelle sekä mahdollisesti arkistohallintaohjelma (Venetjoki 2010).

Useimmat haastateltavista toivoivat kuviin lisää metatietoja, jotta kuvia voitaisiin hakea mahdollisimman monenlaisten attribuuttien mukaan. Lisäksi muutamalla tutkimukseen osallistuneella oli jo näkemyksiä millainen tulevan käyttöjärjestelmän pitäisi olla. Jotkut haastateltavista olivat jo pohtineet millaisia kuvailumahdollisuuksia tulevassa järjestelmässä tulisi olla.

K2: ...tää on tämmönen nopee systeemi jolla me nyt ollaan menty, mutta kun me ruvetaan kehittään tota kuva-arkistoa, kuvapankkia eteenpäin sillonhan se tulee elintärkeeks (...) heittää samantien kaikki asiasanat mitkä liittyy siihen (kuvaan) mut se taas riippuu siitä millanen se meidän softa tai systeemi tai whatever pilvipalvelu tulee olemaan. Mut siinä vaiheessa kun me ruvetaan siirtään niitä kuvia siihen meidän systeemiin (...) niin siinä vaiheessa siinä on iso homma sit käydä läpi kaikki kuvat ja laittaa sinne ne kaikki metatiedot ja asiasanat. Tarkastaa se et siel on varmasti kuvaajan nimi, copywritet ja kaikki.

K1: Arkistointijärjestelmä voisi pakottaa kategorisoimaan kuvan tietyillä tiedoilla. Tämä helpottaisi kuvan loogista arkistointia (...) kuvan sisälle metatietoihin voisi lisätä kuvailua.

Haastateltujen esittämissä kehitysehdotuksissa tulivat esille kameran metatiedot ja parempi pääsy arkistoon. Kuvan kuvailujärjestelmän pitäisi mahdollistaa samalla vapaamuotoinen kuin luokiteltukin kuvailu. Monet toivoivat pakollista metatietokenttää. Myös kuvien purkamisen haluttiin olevan nopeampaa. Kuvan hakemisen arkistosta toivottiin helpottuvan hakusanojen avulla.



H1: Työskentelyä varten luodut kansiot ja kuvien siirrot niiden välillä ovat selkeät. Itse kuva-arkisto ei ole toimiva, ei ole hakusanoja.

Toivottiin hakutoimintoa ja että arkiston sisältöä ja kansiojakoja selkiytettäisiin. Tällä hetkellä jokainen kuvaaja laittaa kuvansa omaan kansioon, kaikki kuvat pitäisi yhdistää. Lisäksi kuvien kuvailuun toivottiin tarkempi kuvaus kuvien käyttöoikeuksista. Kuvasta ei aina nykyisin kuvailutiedoin selviä, mihin tarkoitukseen sitä on lupa käyttää.

## 6 TULOSTEN YHTEENVETO JA TULKINTA

Tässä luvussa tarkastelen tutkimuksen tuloksia ja arvioin niitä aiemman tutkimuksen valossa. Tarkastelen tutkimuksessa ilmenneitä viestintäyksikön digitaalisen kuva-arkiston ongelmia. Esitän myös kehitysehdotuksia arkiston työntekijöiden haastatteluis- sa ilmenneiden ongelmien sekä kehitystoiveiden pohjalta.

Viestintäyksikössä työskentelyä varten luodut kansiot ja kuvien siirtoprosessit niiden välillä koettiin selkeiksi. Itse kuva-arkistoa ei kuitenkaan koettu ole toimivaksi. Työntekijöiden käsitykset arkiston toimivuudesta erosivat hieman. Työnkuvasta riippuen osalle kuvien haku oli hankalampaa kuin toisille.

Kuva-arkistossa kuvien haku- ja löytäminen perustui paljon muistiin ja tottumukseen. Järjestelmän toimivuuden kannalta muistinvaraisuus voi olla riski, sillä työntekijöiden vaihtuessa katoaa paljon hiljaista tietoa, ja kansiojärjestelmän ja tiedostonimien ymmärtäminen voi viedä aikaa. Haastatteluis- sa ilmeni, että kuva-arkiston keskeiset ongelmat ja kehittämistarpeet liittyvät kuvien löytymiseen ja kuvailutietoihin.

Shatford (1986) jakaa kuvan tasot siihen mitä kuvassa konkreettisesti on (*of*), ja mistä se kertoo (*about*). *About*-tasolla tehdään jo kuvan sisällönanalyysiä. *About*-taso käsittää siis kuvan ydinsanomana eli keskeisen sisällön. (Shatford 1986.)

Tiedostonimissä ei ole paljon informaatiota, joten kuvailutiedot ovat verrattain niukkoja. Viestintäyksikön kuva-arkistossa kuvien kuvailuun ei ole tietokantaa, jossa olisi kenttä sisällönkuvailuun. Kuvankäsittelyohjelmassa kuitenkin on mahdollista lisätä metatietoja avoimeen tekstikenttään. Viestintäyksikön arkistossa kuvat kuvaillaankin vain Shatfordin (1986) edellä määrittelemällä *of*-tasolla. Mitä kuvassa on, määritellään tie-

dostonimessä kuitenkin vain yhdellä sanalla tilanpuutteen vuoksi. Arkistossa annetaan kuvalle usein vain yksi avainsana.

Tutkimukseni yhtenä oletuksena oli se, että kuvien kuvailussa tärkeimpiä attribuutteja ovat Shatfordin esittämät elementit kuka, mitä, missä ja milloin (kts.luku 2.5). Viestintäyksikön kuva-arkistossa kuva kuvaillaan useimmiten ja kuvan luonteesta tai käyttötarkoituksesta riippuen vastaamalla näihin kysymyksiin. Viestintäyksikön arkistossa keskeisiä tietoja ovat: kuvaajan tiedot, kuvausaika sekä mitä tai kuka kuvassa on.

Arkistossa henkilökuvat-kansiossa on selkein järjestys. Tiedostonimessä esiintyy henkilön nimi, joka vastaa suoran kysymyksen kuka. Kansion selkeys ja helppo haettavuus selittyvät kuvan luonteella. Henkilokuva noudattaa tarkkaa formaattia ja aihe on lähes yksiselitteinen, kuvan hakija tietää millä nimellä kuva on kuvailtu.

Viestintäyksikön Aikalaisen kuvat kuvaillaan siis hyvin samoin tiedoin kuin lehtikuvia yleensä. Käytännössä esimerkiksi lehtikuvia kuvailtaessa esitetään kuvista perustietoina esimerkiksi kuvaaja, milloin kuva on otettu, kuvauspaikka sekä julkaisutiedot. Sisällönkuvailussa käytetään kuvaan liitettyä kuvatekstiä, sekä luetellaan kuvassa esiintyvät henkilöt, paikat sekä tapahtuma, jos sellainen voidaan nimetä. Lisäksi voidaan käyttää erilaisia avain- ja asiasanoja sekä luokituksia.

Viestintäyksikön informanttien haastattelussa tuli ilmi, että Shatfordin kuvailukysymys milloin eli kuvan ottamishetki (kellonaika/päivämäärä) sisältää etenkin viestintäyksikön arkistossa paljon tietoa, josta on hyötyä kuvaa hakiessa. Päivämäärä tiedostonimessä kertoo esimerkiksi kuvassa olevan vuodenajan, miljöö (esimerkiksi minkä näköisiä ihmisiä kuvassa on, mitä heillä on päällään) ja mikä tilaisuus kuvassa mahdollisesti on meneillään (esimerkiksi valmistujaiset keväällä). Pelkästään tiedostonimestä voi jo päätellä paljon kuvan sisällöstä ilman että itse kuvaa tarvitsee katsoa. Ajanpuut-

teen ja kuvailutilan puutteen vuoksi pyritään tiivistämään johonkin elementtiin tietoa mahdollisimman paljon.

Karhulan (1991) mukaan valokuvan olemus perustuu kohteisiin. Karhula tarkoittaa kohteella valokuvan ulkopuolista, havaittavaa objektiä, ominaisuutta, ilmiötä tai toimintaa. Karhula määrittelee valokuvan elementeiksi kaikki visuaaliset olennot, joita voidaan nimetä, kuten hahmot, hahmon osat tai hahmoja muodostamattomat elementit. Viestintäyksikön arkistossa kuvailussa käytetään vähän mitä-attribuuttia. Näin kuvailutiedoista jäävät usein pois sanat, jotka kertoisivat kuvan kohteen.

Kehittämistarpeita nähtiin kameran metatiedoissa. Hakutoiminto ja hakusanat kuvaarkistoon nopeuttaisivat hakemista. Kansiojärjestelmä ja kuvailut tehdään pitkälti aiemmin toimituksessa työskennelleen kuvaajan järjestyksen pohjalta. Ongelmia aiheutti haastateltujen mukaan myös kuvaajan vaihtuminen. Kuvaajan työsuhteet ovat lyhyitä, joka vaikuttaa arkistointijärjestelmän kehittämiseen.

## 7 ARKISTON ONGELMIA JA KEHITYSNÄKÖKULMIA

Haastateltujen mukaan tiedostonimen lisäksi kuvien sisälle pitäisi saada tiedot kuvauspaikasta ja tilaisuudesta, jossa kuva on otettu. Usein nämä tiedot jäävät pois, koska tiedostonimessä ei ole tilaa. Ongelmaksi koettiin oleellisten tietojen katoaminen. Toimittajat joutuvat useasti tekemään salapoliisityötä yrittäessään selvittää itse näitä tietoja jälkikäteen. Lisäksi ongelmaksi koettiin kansiojärjestelmän hajanaisuus. Tämä johtuu siitä että jokainen kuvaaja tallentaa kuvansa omaan kansioon. Informanttien mukaan kaikki kuvat pitäisi sijoittaa samaan paikkaan. Tietojen säilymiseksi ehdotettiin arkistointijärjestelmää, joka kuvaa tallennettaessa pakottaisi täyttämään oleelliset tiedot metatietokenttiin. Lisäksi www-tiedostoille toivottiin yhteistä arkistointikäytäntöä.

Haastatteluissa tuli myös ilmi että kuva-arkistossa oli kokeiltu kuvatietokantaohjelmaa. Ohjelmaa ei kuitenkaan koettu toimivaksi ja siitä luovuttiin. Kuvien luokittelu ei ollut toiminut, sillä se oli aikaa vievää ja monimutkaista. Osassa haastatteluissa tulikin esille varautunut suhtautuminen tietokantaohjelmaan. Nykyiseen kansiojärjestelmään oltiin tyytyväisiä, mutta siihen toivottiin hakusanakenttää ja hakutoimintoa. Vaikutti siltä, että pidempään talossa olleet ja kansiojärjestelmää enemmän käyttäneet työntekijät olivat hieman tyytyväisempiä kansiojärjestelmään ja kuvien löydettävyyteen. Suhtautumiseen vaikuttivat myös työtehtävät. Kuvan hakemiseen vaikuttaa mitä haetaan ja mihin tarkoitukseen, ovatko hakutoiveet tarkasti rajattuja vai löyhästi määriteltyjä.

Keskeisiä kehitystoiveita olivat myös yhtenäinen arkistointikäytäntö, arkistointiohjeet sekä arkiston kuvaus, nykyistä helpompi haku, vaivaton kuvien jakaminen yliopiston ulkopuolelle sekä mahdollisesti arkistonhallintaohjelma (Venetjoki 2010).

Yhteenvetona arkiston keskeisiä ongelmia olivat:

- hakusanakentän ja hakusanatoimintojen puute
- kuviin liitettävien kuvailutietojen vähyys
- kansiojärjestelmän hajanaisuus
- puuttuvat käytännöt ja välineet, joilla kuvia voitaisiin jakaa vaivattomasti yliopiston ulkopuolelle.

Viestintäyksikön arkistossa hakuja voidaan tehdä vain yksinkertaisina merkkijonohakuina tiedostonimien perusteella. Kuvakokoelmaa ei ole organisoitu tiedonhakujärjestelmän alle, joka tukisi hakujen tekemistä esimerkiksi hakusanayhdistelmien avulla. Kuvailu on siis hyvin yksinkertaista ja rajoittunutta, koska olennaisen tiedon, jonka avulla dokumentin voisi hakea, tulee mahtua hyvin pieneen tiedostonimikenttään.

Tähän ongelmaan liittyen kehityksen arvoisena kohteena voi pitää kameran metatietoja. Kuvailun ollessa yksinkertaista ja suppeaa, kameran metatietoihin voitaisiin lisätä kuvailua, ja näin jo suhteellisen helpolla keinolla lisätä kuvailutietojen määrää. Kameran muistiin ja kuvan tietoihin jää jo nyt viestintäyksikössä kellonaika ja päivämäärä. Lisäksi sinne jää teknistä tietoa, kuten esimerkiksi kuvanottohetkellä käytetty suljinaika.

Kameraan on myös mahdollista lisätä omin sanoin metatietoja jotka jäävät myös ikään kuin kuvan sisälle, tiedoston itsensä muistiin. Tällöin jo automaattisesti tulevien teknisten tietojen lisäksi voitaisiin asettaa joitain tietoja valmiiksi kameran muistiin. Jos esimerkiksi kuvattava kohde on tiedossa, voitaisiin ennen kuvauskeikkaa asettaa kameraan valmiita kuvailutietoja, kuten henkilökuvassa esiintyvän nimi tai kuvaajan nimi. Nykyisiin kameroihin näitä tietoja on mahdollista tallentaa.

Kameran metatiedot näkyvät kuvan tiedoissa, kun kuvat on siirretty kuvankäsittelyohjelmaan. Kuvankäsittelyohjelmassakin on mahdollista lisätä joitain kuvailutietoja metatietokenttiin. Tosin hakujärjestelmä ei viestintäyksikössä tue hakemista näistä kuvan sisäisistä tiedoista. Tiedot, eli merkkijonot, joiden perusteella voidaan hakea järjestelmästä, lisätään siis tiedostonimeen vasta kun kuvat tallennetaan kansioihin.

Toinen kehittämiskohde, joka liittyy olennaisesti myös kameran metatietoihin, on hakemisto. Ensinnäkin hakemistorakenneratkaisun ja käyttöjärjestelmän tulisi sopia kameran järjestelmään. Lisäksi hakemistorakenneratkaisun tulisi tukea kameran metatiedoista hakemista. Haastateltujen vuonna 2011 mainitsema kameran ja ohjelmiston yhteensopimattomuus lienee täysin ratkaistavissa oleva ongelma, kun käytössä on kameran järjestelmää tukeva järjestelmä.

Kun hakujärjestelmä ottaisi huomioon kuvan jo kamerasta saamat kuvailutiedot, säästettäisiin paljon aikaa ja vaivaa. Lisäksi ei tietoa ei hukkuisi matkalla. Kuvan löydettävyyttä paranisi olennaisesti, kun mukana olisivat nämä tiedot.

Näillä teknisillä sovelluksilla päästäisiin viestintäyksikön arkistossa jo eteenpäin. Tallennusvaiheessa kuluu paljon aikaa kansioihin tallentamiseen, arkistonimien miettimiseen ja yhdenmukaisten kuvailukäytäntöjen pohtimiseen. Monia mutkia ja vaiheita voitaisiin suoristaa, kun merkkijonohakua lisättäisiin. Lisäksi kuvailutiedot säilyisivät kiinnittyneenä kuvaan.

Teknisten tietojen lisäksi kuvailutietojen pitäisi kertoa myös sisällöstä. Nykyisin monet kuvatiedostonhallinnan sovellukset, kuten esimerkiksi *Photoshop* ja *Adobe Lightroom* tukevat täydentävien metadatojen tallennusta ja hakemista. Ilman erillisiä tietokantaohjelmiaakin kuviin voisi liittää kuvailua. Ainakin kokeileminen on vaivatonta, sillä lataaminen on maksutonta.

Käyttöönottoa lykätään usein organisaatioissa joissa aika on kortilla. Tällaisten ohjelmistojen käyttöönottojen ajatellaan usein olevan aikaa vieviä ja kalliita. On totta, että kun aikaisempia käytäntöjä muutetaan, se on aina jonkinlainen ponnistus. Mielellään tehdään niin kuin on ennenkin tehty. Usein ajatellaan, että tarvitaan erikseen aikaa ja rahaa vievää ohjelmistokoulutusta. Näin ei kuitenkaan ole, jos miettii vaikka arkielämäämme. Useat meistä ottavat kuukausittain tai jopa päivittäin käyttöön uusia sovelluksia vapaa ajalla internetissä käyttäessämme pelkästään käyttämällä jotain uutta sivustoa. Ainakin suuremmissa organisaatioissa, kuten yliopistot, löytyy lisäksi erillinen ATK-tuki tai tukihenkilö, joka voi ratkoa järjestelmien käyttöönottoon liittyviä ongelmia.

Jos kuvanhallinta olisi tiedostojen kuvailutietohakuja tukevaa, arkistoon ei tarvittaisi erillistä arkistonhoitajaa tai tallentajaa. Useimmissa organisaatioissa työntekijät tekevät kuvailua muun työnsä ohessa. Kuvaaja voisi liittää tiedostoon kuvailutietoja kameran kuvailutietojen lisäksi jo tallennusvaiheessa. Hakeminen helpottuisi, jos haun tuloksia voitaisiin rajata hakemalla mahdollisimman monella sanalla. Joitain metatietokenttiä voitaisiin rajata myös pakollisiksi. Esimerkiksi tekijänoikeudet -kenttä pitäisi täyttää ennen kuin kuvan voi tallentaa, jolloin tieto jäisi varmasti kuvan sisälle.

Tiedostohaun lisäksi säilyisi tietysti selailumahdollisuus. Kuvien selailu on tärkeä hakumenetelmä vaikkapa hakiessa kuvitusideaa Aikalaisen juttuun tai etsiessä jotain yleisluontoista kuvaa. Esimerkiksi aiheeseen yleisesti sopivaa kuvaa etsitään usein silmäillen sormenpääkuvia. Tässäkin tapauksessa hakemistoratkaisussa olisi hyvä, jos tulosjoukkoa voi rajata.

Kun haku on monipuolista ja helppoa, kuvien jakaminen ulkopuolellekin on vaivattomampaa. Jotta kuvapankkiajatus voi toteutua, tulevaisuudessa kuvat tulisi kuvailla niin,



että ne löytyisivät myös muille käyttäjille kuin arkiston työntekijöille. Kuvien jakamisesta arkiston ulkopuolelle ja avoimesta arkistosta jatkan seuraavassa luvussa 8.

## 8 KUVAILUN KEHITTÄMINEN NETTISOVELLUSTEN TAPAAN

Luvussa 7 esittelin konkreettisia vaihtoehtoja arkiston hallintaan. Tässä luvussa vertailen viestintäyksikön empiiristä aineistoa netin kuvailusovelluksien käytäntöihin ja tutkimukseen, joita käsittelin kirjallisuuskatsauksessa luvussa 2.6. Luvussa 7 esitin, että ilman erillistä tietokantaohjelmaakin voidaan käytännön hakemistorakenneratkaisulla parantaa kuvien hallintaa. Tässä luvussa esitän tulevaisuuden kehitysnäkökulmia: haluan ottaa huomioon digitaalisten aineistojen avoimuuden ja helpon pääsyn sekä jaettavuuden näkökulmat, jotka ovat varsin pinnalla tämän päivän julkisessa keskustelussa. Mielestäni nämä näkökulmat on hyvä ottaa huomioon jo varhaisessa vaiheessa suunnitteluprosessia.

Viestintäyksikön kuva-arkiston rakennetta voi luonnehtia suljetuksi mutta vapaamuotoiseksi. Suljetulla tarkoitan sitä, että pääsy arkistoon on rajattu. Vapaamuotoisuudella tarkoitan, että arkisto on rakentunut käyttäjien muovaamana samanaikaisesti käytössä ollessaan, tai sitä mukaa kun kuvia on tullut. Aluksi paperikuvat olivat kansioissa kellarissa, digitaalisten kuvien tullessa käyttöön niitä on yritetty järjestää parhaan mukaan työnteon ohessa, kuvavirran koko ajan kasvaessa. Arkiston työntekijät eli käyttäjät ovat itse kuvailleet kuvia niin, että ne löytyisivät parhaiten.

Digitaalisten kuvien määrä on kasvanut räjähdysmäisesti. Viestintäyksikön arkistossa on aiemmin yritetty kuvien hallintaa kokeilemalla kuvatietokantaohjelmaa, jonka ohessa kokeiltiin sanaston luomista. Kokeilu ei kuitenkaan onnistunut (kts. luku 7). Ratkaisuna yhä kasvavalle kuvien määrälle ja viestintäyksikön kaltaiselle arkistolle luokituskielten luominen olisi aikaa vievää ja kallista, eikä kovin tehokasta. Kuvien kierto on niin nopeaa, että kuvien käyttö- ja jakamistapa ovat menneet näiden vaiheiden ohi.

Viestintäyksikön kuva-arkistoa on vaikea tyhjentävästi nimittää joko kuvakokoelmaksi tai kuva-arkistoksi. Tässä mielessä sen rakentuminen muistuttaa internetin kuvapalveluita. Enser (2008, 7-8) muistuttaa, etteivät internetin kuvapalvelut ole varsinaisesti kokoelmia, ainakaan siinä mielessä kuin kokoelma ymmärretään ammattilaisten parissa. Internetin kuvapalveluilta puuttuu ammatillisesti järjestettyjen kokoelmien perusominaisuuksia, kuten ammattimainen kokoelman hallinta tai luokittelujärjestelmä. Kuvapalveluiden valtava kapasiteetti niin kuvien kuin avointen avainsanoitusten tekijöiden suhteen on kuitenkin lopullisesti muuttanut käsityksemme kuvakokoelmista.

Luvussa 7 esitin, että ilman erillistä tietokantaohjelmaakin voidaan käytännön hakemistorakenneratkaisulla parantaa kuvien hallintaa. Haastatteluissa tuli ilmi, että viestintäyksikön arkistossa on pohdittu mahdollisesti tulevaisuudessa toteutettavaa ratkaisua, josta esimerkkinä mainittiin internetin kuvapankit. Kuvapankin sisältö voisi olla mahdollisesti avoin käyttäjille ja yhteinen muiden yliopistojen kanssa.

Erilaiset kuva-arkistot, kuten yliopistojenkin arkistot, ovat hajautuneet eri paikkoihin. Arkistot ovat suljettuja käyttäjiltä. Mahdollisilla kuvan ostajilla, esim. medially ei ole pääsyä selailemaan arkistoa. Myöskään arkiston työntekijät eivät pääse aina verkon kautta katselemaan kuvia. Arkiston kansiot sijaitsevat siis eri paikkoihin hajautuneena. Esimerkiksi tästä syystä arkistossa on turhia kopioita kuvista. Näistä syistä olisi käytännön kannalta suotavaa saada ainakin osa kuvista internettiin.

Esitänkin, että mallia tällaisen kuvapankin suunnitteluun voisi löytyä internetin nettisovelluksista. Nettisovelluksissa, kuten *Flickr* (kts luku 2.6.3), on ratkaistu ainakin kaksi kuvahallintaan liittyvää ongelmaa. Ensinnäkin kuvien valtava määrä vie säilytystilaa. Omilla palvelimilla säilyttämisen sijaan erilaisissa pilvipalveluissa on enemmän kapasiteettia säilytykseen. Lisäksi avoimuus käyttäjille lisää käyttöä ja jakamista. Samalla se

lisää kuvan kuvailua, kuten luvussa 2.6.3 kerrotaan. Ainakin yliopistojen yhteinen kuvapankki, jossa kuvailua yhdistettäisiin, olisi mahdollinen. Haastatteluissa ilmeni myös toive että arkisto olisi avoin. Näin kuvat olisivat selailtavissa myös ulkopuolisille tahoille jotka ostavat kuvia. Tässä voisi olla myös pohdinnan aihetta myös pohtiessa rahoitusta yliopiston arkiston toimintaan.

Yksi kehitysnäkökulma tällaisessa kuvapalvelussa voisi olla avoin avainsanoitus. Viestintäyksikön arkistossa kuvailussa ei käytetä täsmällisesti määriteltyjä ylä- ja alakäsitteitä. Tässä mielessä kuvailu muistuttaa avointa avainsanoitusta. Tuominen (2006, 171–173) toteaa, että olennaista avoimissa avainsanoituksissa on, että niiden varaan rakentuvat sanastot eivät ole rakenteeltaan hierarkkisia. Sen sijaan niiden välille luodaan automaattisesti assosiatiivisia viittauksia merkittyjen sisältöjen perusteella. Lähestymistapa avoimissa avainsanoituksissa sanastojen rakentumiseen on siis täysin päinvastainen verrattuna taksonomiseen luokitteluun: sen sijaan, että luokittelusta pyrittäisiin tekemään mahdollisimman täsmällinen, sen annetaan rakentua käyttäjien määräämällä tavalla.

Yliopistojen kesken jaetun kuvapalvelun etuja voisivat olla ainakin yhteinen työskentely. Kuten yhteisölliselle tekemiselle on ominaista, tekemisen jakaminen nopeuttaa ja jakaa vastuuta. Tuominen (2006, 171–172) antaa tästä esimerkin: internetin kuvapalvelut suosittelevat esimerkiksi uutta kirjanmerkkiä lisääville muiden jo käyttämiä termejä. Yhteisöllisen tekemisen etuna on muun muassa joukkoälyn idea: kuvat pitäisi oikeastaan kuvailla niin kuin suurin osa ihmisiä ne hakisi.

Kun käyttäjät itse kuvailevat kuvien sisältöä, sisältö on luokiteltu käyttäjille tutulla ja yhteisesti jaetulla sanastolla. Fokker ym. (2006) toteaa, että kuvien kuvailussa ja haussa käytetyt termit ovat samanlaisia, koska käyttäjät itse kuvailevat kuvia. Lisäksi Tuominen (2006, 171–172) huomauttaa, että avoimissa avainsanoituksissa sanastojen raken-

tuminen on suorastaan päinvastainen verrattuna taksonomiseen luokitteluun, koska sen annetaan rakentua käyttäjien määrämällä tavalla. Näin on myös viestintäyksikön arkistossa, kuvat on järjestetty kansioissa ja kuvailtu niin että työntekijät ne mahdollisimman hyvin löytäisivät.

Karhula (1991, 48–50) totesi kaukonäköisesti jo 1990-luvun alussa, että kuvan olennainen tietosisältö jäsentyy paremminkin sen käyttökontekstista kuin kuvan itsensä ominaisuuksista. Kuvan käyttökonteksti ratkaisee siis olennaisen tietosisällön. Kuvan tietosisältö voi siis muunnella eri aikoina eri paikoissa eri käyttäjillä. Tässä tavoitetaan mielestäni jotain olennaista kuvan kuvailusta. Ajatus on myös hyvin nykyaikainen käyttäjälähtöisyydessään.

Muitakin hyviä ideoita kuvien kuvailuun on esitetty jo 1990-luvulla. Shatford Layne (1994, 585) huomauttaa, että arkiston tulisi tarjota pääsy hyvin ryhmiteltyihin kuviin, ei vain yksittäisiin kuviin. Ryhmiteltyjen kuvien idea on mielestäni kaukonäköinen, sillä se muistuttaa tavallaan luvussa 2.6.4 kuvailtuja, niin kutsuttuja folksonomioita. Saartin (2009) mukaan folksonomiat ovat käyttäjien verkkoyhteisöissä luomia taksonomioita eli avoimia luokittelujärjestelmiä, jotka muodostuvat vapaasti määritellyistä avainsanoista (tags). Tässä tapauksessa taksonomia ei ole ammattilaisen luoma, vaan käyttäjien kuvailusta muodostunut luokittelujärjestelmä. Avainsanoitus on hyödyllisimmillään, kun tagit kootaan samaan yhteyteen isoksi parveksi.

Lemare (2008) toteaa, että avainsanapilveä voidaan ajatella visuaalisena ilmentymänä folksonomiasta. *Flickr*-kuvapalvelun avainsanapilvi (ks. kuva 1) on hyvä esimerkki folksonomiaa hyödyntävästä luokittelujärjestelmästä, jossa kootaan avainsanat pilveen tai parveen. Avainsanapilvi on Fokker:n ym. (2006) määritelmän mukaan visuaalinen esitys avainsanojen (*tags*) tilastollisista suhteista pohjautuen niiden esiintymistiheyteen.

Mesnage (2009, 1) toteaa, että avainsanapilvi näyttää linkit suosituimpiin tageihin ja avainsanan fontin koko riippuu tagien suosion tasosta. Parvet ja niihin liittyvät tagit mahdollistavat tagipohjaisen navigoinnin.

Sigurbjörnsson ja van Zwol (2008, 329) tutkivat millaista informaatiota *Flickr*-kuvapalvelun avainsanoitus sisältää. Suosituin kategoria oli paikka (28%), seuraavaksi suosituin kategoria olivat esineet tai kohteet (16%), sitten *ihmiset* tai ryhmät (13%) sekä tapahtuma tai toiminta (9%) ja ajankohta (7%). Nämä kategoriat muistuttavat Shatfordin (1986) lukuohjeen attribuutteja. Koska suosituimmat avainsanat kuuluvat esineet tai kohteet-kategoriaan, käyttäjät luultavasti hakevat eniten tähän kategoriaan kuuluvilla termeillä. Tästä syystä viestintäyksikönkin arkistossa olisi hyvä lisätä esineet tai kohteet-kategorian kuvailua.

Kuva-arkiston luokittelusysteemi sijoittuu tavallaan hierarkkisten luokittelusysteemien ja folksonomia-mallin välimaastoon. Kumpikaan malli ei suoraan päälle liimattuna sovi kuva-arkiston luonteeseen. Ammattimaisen käsittehierarkioita hyödyntävän arkistonhallinnan ylläpitäminen on liian työlästä ja hidasta kuvien nopean kierron takia. Toisaalta arkistossa on luotu jonkinlaista järjestystä kuvien juuri käsitteiden avulla, kuten luvussa 5.2. esitellään. Folksonomiat tai täysin avoimesti rakentuva avainsanoitus taas on liian hallitsematon, kuvia pitäisi pystyä hallitsemaan myös perinteisessä mielessä. Folksonomia malli perustuu liikaa kuvien suosituimmuusjärjestykseen, se elää ajassa. Yliopistot pyrkivät säilyttämään historiaansa, joten kuvien järjestämisperiaatteiden tulisi olla selkeitä ja läpinäkyviä.

Molemmista mallien ääripäistä voisi siis saada hyviä käytäntöjä ja ideoita arkiston kehittämiseen. Pilvipalvelut ja jakaminen helpottavat säilytystä ja mahdollistavat vaivattoman pääsyn, toisaalta kuvailun käsittehierarkioita on tutkittu paljon, joka on hyvä ottaa

huomioon palveluiden suunnittelussa. Avainsanoituksessa käyttäjien itse hauissaan käyttämät sanat tallentuisivat kuvan metadataksi, joka taas mahdollistaisi kuvaan pääsyn monella eri tavalla. Haastateltavat toivoivat kuviin lisää kuvailutietoja, jotta voitaisiin hakea mahdollisimman monenlaisten attribuuttien mukaan. Kuvan käyttöarvo paranee, kun se voisi olla avoimella pääsillä ja löydettävissä mahdollisimman monella eri sanalla.

## 9 YHTEENVETO JA JOHTOPÄÄTÖKSET

Tutkimuksen tavoitteena oli kartoittaa ja kuvata Tampereen yliopiston viestintäyksikön kuva-arkiston sisältö, käyttötarpeet ja kuvailukäytännöt sekä vertailla työntekijöiden tehtäviä, kuvailukäytäntöjä ja käsityksiä arkiston toimivuudesta. Tavoitteena oli selvittää viestintäyksikön digitaalisen kuvakokoelman sisältöä ja kuvien käyttötarpeita sekä kuvien kuvailun käytäntöjä. Tutkimuksessa pyritään esimerkkitapauksen avulla selvittämään ja erittelemään digitaalisessa muodossa oleviin kuviin ja niiden kuvailuun liittyviä haasteita.

Tutkimuksessa haettiin vastauksia seuraaviin kysymyksiin:

- Mitkä ovat Tampereen yliopiston viestintäyksikön digitaalisen valokuva-arkiston käyttötarpeet?
- Minkälaisista kuvista (kuvatyyppit) arkiston sisältö koostuu ja millainen on kuvan elinkaari arkistossa?
- Millaisia ovat kuvan kuvailukäytännöt valokuva-arkistossa?
- Millaisia kehitysnäkökulmia internetin kuvanhallintasovellukset voivat tarjota arkistolle?

Vastauksia näihin tutkimuskysymyksiin haettiin laadullisen tutkimuksen keinoin. Halusin nimenomaan käyttäjän näkökulman omin sanoin kerrottuna. Arkiston kanssa päivittäin tekemisissä olevat työntekijät osaavat kertoa parhaiten arkistosta. Edellä mainituista syistä valitsin tutkimusmenetelmäksi haastattelun, keskustelun ja havainnoinnin. Menetelmänä tutkimuksessa käytettiin haastatteluja ja havainnointia. Empiirinen aineisto ke-



rättiin kahdessa osassa, täydensin ja päivitin keväällä 2014 tekemälläni haastattelulla keväällä 2011 kerättyä aineistoa.

Tutkimus on toteutettu osana Tampereen yliopiston viestintäyksikön kuvakokoelman kehittämishanketta. Jo tutkimuksen alkuvaiheessa kehitystarpeita ilmeni paljon ja niitä nousi esille myös haastatteluissa. Kehitystarpeiksi mainittiin mm. yhtenäinen arkistointikäytäntö, ohjeet ja arkiston kuvaus. Lisäksi toivottiin nykyistä helpompaa hakua sekä vaivatonta kuvien jakamista yliopiston ulkopuolelle.

Kartoituksen ja työntekijöiden haastatteluiden pohjalta haluttiin löytää ja nostaa esille kuva-arkiston yleisiä ja erityisiä piirteitä, sekä esittää kehitysehdotuksia kuva-arkiston jatkoa ajatellen. Tulosten perusteella voikin sanoa, että menetelmä toi esille sellaisia arkiston piirteitä, joita ei olisi löytynyt muilla menetelmillä. Ainakin voi sanoa, että tutkimuksessa tuli kuvailtua itse arkisto. Lisäksi kuvien kuvailun prosessi saatiin kirjoitettua ylös ja selvennettyä, josta voi olla hyötyä arkistoa kehitettäessä.

Tutkimukseni yhtenä oletuksena oli se, että kuvien kuvailussa tärkeimpiä attribuutteja ovat Shatfordin (1986) esittämät elementit kuka, mitä, missä ja milloin (kts.luku 2.5). Viestintäyksikön kuva-arkistossa kuva kuvaillaan useimmiten ja kuvan luonteesta tai käyttötarkoituksesta riippuen vastaamalla näihin kysymyksiin. Viestintäyksikön arkistossa keskeisiä tietoja ovat: kuvaajan tiedot, kuvausaika sekä mitä tai kuka kuvassa on. Yksi tärkeimmistä kuvailun attribuuteista vastaa kysymykseen milloin. Päivämäärä on arkistossa keskeinen, kuvasta paljon kertova tieto.

Tutkimuksessa ilmeni, että kuvan kuvailutiedot ovat tiedostonimessä, joten siihen ei mahdu paljon kuvailua. Käytäntö on kuitenkin tulevaisuuden kannalta riittämätön. Olisi toivottavaa, että kuvailutiedoista ilmenisi, mitä kuvassa on. Päivämäärä kuvailutietona ei kerro ulkopuolisille paljonkaan. Tulevia käyttäjiä ajatellen olisi keskeistä saada ku-

vaan liitettyä enemmän mitä ja kuka kysymyksiin vastaavia tietoja. Arkiston keskeiset ongelmat ja kehittämistarpeet liittyivät kuvien löytymiseen ja kuvailutietoihin.

Tutkimuksessa ilmeni, että kameran metatiedot on yksi keskeinen kehityskohde. Kameran metatietoihin voidaan liittää kuvaa koskevaa tietoa, joka jää ikään kuin kuvan sisään kuvailutiedoksi. Näistä tiedoista kuvaa voidaan hakea jos hakemistoratkaisu tukee kameran metatiedoista hakemista. Hakutoiminto ja hakusanat kuva-arkistoon nopeuttaisivat hakemista. Toinen kehittämiskohde olisikin kameran metatietohakuja tukevan järjestelmän käyttöönotto.

Kehitysehdotuksia miettiessäni yhdistin kaksi hieman erilaista näkökulmaa kuvien hallintaan ja kuvailuun. Koska digitaalisten kuvien tuottaminen on niin nopeaa, niiden määrä kasvaa isojen organisaatioiden arkistoissa. Näin ollen tarvitaan lisää säilytystilaa, johon pilvipalvelut tarjoavatkin ratkaisuja. Lisäksi kuvien löytyminen kuvamassasta vaatii uusia ratkaisuja kuvailun ongelmaan. Yksittäinen ihminen ei siihen enää kykene. Tässä tutkielmassa näkökulma viestintäyksikköön oli nykyhetken ja tulevaisuuden rajapinnassa.

Kuva-arkiston luokittelusysteemi sijoittuu tavallaan hierarkkisten luokittelusysteemien ja folksonomia-mallin välimaastoon. Kumpikaan malli ei suoraan päälle liimattuna sovi kuva-arkiston luonteeseen. Molemmista käytännöistä voi kuitenkin saada hyviä ideoita arkiston kehittämiseen. Pilvipalvelut ja jakaminen ovat tätä päivää ja tulevaisuutta, toisaalta kuvailun käsitehierarkioita on tutkittu paljon, joka on hyvä ottaa huomioon palveluiden suunnittelussa.

Viestintäyksikön arkistossakin voitaisiin toteuttaa jatkotutkimusta. Ilmaiseksi ladattavan kuvanhallintaohjelman kokeilu (esim. *Adobe Lightroom*) voisi olla mielenkiintoinen projekti esimerkiksi opinnäytetyötä tekeväälle opiskelijalle. Kokeilussa voitaisiin tutkia

kameran metatietojen yhteensopivuutta ohjelman kanssa ja miten tiedostohaku toimii. Myös jo arkistossa olevien kuvien siirtäminen ohjelmaan, sekä täydentävien kuvailutietojen liittäminen niihin, olisi mielestäni kokeilun arvoista. Näin voitaisiin selvittää, miten aikaa vievää kuvien kuvaileminen on.

Mitä tulee aiempaan kuvien kuvailun tutkimukseen, monet jo 1980-luvulla luodut mallit pitävät yhä paikkansa kuvia kuvailtaessa. Shatfordin (1986) lukuohje osoittautui käyttökelpoiseksi kuvailuelementtien kartoitukseen. Tutkimuskirjallisuudesta löytyi internetiä ennakoivia ajatuksia kuvan kuvailusta, kuten esimerkiksi Shatford Laynen (1994, 585) ryhmiteltyjen kuvien idea saa nykyaikaisen vastineen folksonomioista ja avainsanapilvistä.

Karhulan (1991, 48–50) näkemykset kuvien kuvailun käyttäjälähtöisyydestä ovat relevantteja. Kuva-arkistoissa on koettu isoin murros digitaalisten kuvien ja internetin yleistyttyä. Voisi todeta, että pääsääntöisesti kuva-arkistot ovat muuttumassa suljetuista ja eri paikkoihin hajautetuista avoimiksi ja jaetuiksi. Tulevaisuudessa internetin kuvapalvelujen ja käyttäjien tekemän avoimen avainsanoituksen tutkiminen tuskin vähenee.

## Lähteet

- Alasuutari P. 1994. Laadullinen tutkimus. 2. uudistettu painos. Tampere: Vastapaino.
- Chen, H. & Rasmussen, E. 1999. Intellectual access to images. *Library Trends*. Saatavilla findarticles-palvelusta (<http://findarticles.com/>) (käytetty 15.3.2010).
- Enser, P. G. B. 1995. Pictorial information retrieval. *Journal of Documentation*, (51) 2, 126–170.
- Enser, P. G. B. 2008. Visual image retrieval. *Annual Review of Information Science & Technology*, (42) 1, 3–42.
- Eriksson, P. & Koistinen, K. 2005. Monenlainen tapaustutkimus. Kuluttajatutkimuskeskuksen julkaisuja 4:2005. Helsinki: Kuluttajatutkimuskeskus.
- Fokker, J., Buntine, W. & Pouwelse, Johan. 2006. Tagging in peer-to-peer Wikipedia. A method to induce cooperation. (<http://www.emse.fr/OSIR06/2006-osir-p39-fokker.pdf>) (käytetty 6.3.2011).
- Hirsjärvi, S., Remes, P. & Sajavaara, P. 2004. Tutki ja kirjoita. 10. osin uudistettu painos. Helsinki: Tammi.
- Järvelin, K. & Sormunen, E. Dokumentit kateissa? 1999. Tiedon tallennus ja haku avuksi. Teoksessa Mäkinen (toim.) Tiedon tie. Johdatus informaatiotutkimukseen. Helsinki: Btj Kustannus, 110–142.
- Jørgensen, C. 2003. Image retrieval: Theory and research. Lanham (Md.): Scarecrow Press.

Kaakinen, A. 2008. Sisältöpohjainen kuvatiedonhaku ja kuvahaun keinot. Pro gradu - tutkielma. Tampere: Tampereen yliopisto. Saatavilla Tampereen yliopiston tutkielmatietokannasta (<http://tutkielmat.uta.fi>) (käytetty 12.3.2010).

Karhula, P. 1991. Sanomalehtivalokuvien sisällönkuvailun kehittäminen. Pro gradu-tutkielma. Tampere: Tampereen yliopisto.

Klijn, E. & de Lusenet, Y. 2004. Sepiades. Cataloguing photographic collections. Amsterdam: European Commission on Preservation and Access (ECPA). Saatavilla (<http://www.knaw.nl/sepia.pdf>) (käytetty 2.10.2014).

Lemare, P. 2008. Social tagging and music information retrieval. Journal of new music research, (37) 2, 101–114. Saatavilla Ebscohost-palvelusta <<http://search.ebscohost.com>> (käytetty 7.4.2010).

Leary, W.H.1985.The Archival appraisal of photographs: a RAMP study with guidelines. Paris:Unesco.

Lybeck, J. et al. 2006. Arkistot yhteiskunnan toimiva muisti. Asiakirjahallinnon ja arkistotoimen oppikirja. Helsinki: Arkistolaitos.

Mesnage, C. & Carman, M. 2009. Tag navigation. Saatavilla acm-portaalista <<http://portal.acm.org>> (käytetty 8.4.2010). 29–32. (Käytetty 5.4.2010).

Markkula, M. & Sormunen, E. 2000. End-user searching challenges indexing practices in the digital newspaper photo archive. Information Retrieval, 1, 259–285.

Penttinen, L. 2007. Yrityksen valokuva-arkisto ja arkistointisuunnitelma: tapaus Rautaruukki Oyj. Pro gradu-tutkielma. Informaatiotutkimuksen laitos. Tampere: Tampereen yliopisto.

Sahavirta, H. & Sormunen, E. 2001. Kuvahaun perusongelmat. Polku Panofskysta digitaaliseen kuva-arkistoon. *Informaatiotutkimus* 20 (3), 60–70.

Sepiades. Recommendations for cataloguing photographic collections. 2003. Klijn, E. (toim.) Amsterdam: European Commission on Preservation and Access (ECPA). <<http://www.knaw.nl/ecpa/sepia/workinggroups/wp5/sepiadestool/sepiadesdef.pdf>>

Shatford Layne, S. 1994. Some issues in the indexing of images. *Journal of the American Society for Information Science*, (45) 8, 583–588.

Shatford, S. 1986. Analyzing the subject of a picture: A Theoretical approach. *Cataloging and Classification Quarterly*, 6:39-62.

Sigurbjörsson, B. & van Zwol, R. 2008. Flickr tag recommendation based on collective knowledge. *Proceedings of the 17th International World Wide Web Conference (WWW2008) Beijing*. 327–336.

Stvilia, B. & Jörgensen, C. 2009. User-generated collection-level metadata in an online photosharing system. *Library & Information Science Research*, 31(1), 54–65.

Suominen, V. 2009. Bibliografisen valvonta: johdatus luetteloinnin ja sisällönkuvailun menetelmiin. Helsinki: Btj Kustannus.

Vainio, M. 1996. Digitaalisen kuva-arkiston tekniikka ja sen mahdollistamat kuvainformaation tallennuksen ja -haun mallit. *Laudaturtyö*. Tampere: Tampereen yliopisto.

Vakkari, P. 1999. Tiedonhankinnan tukeminen ja informaatiotutkimus. Teoksessa Mäkinen, I. (toim.) Tiedon tie. Johdatus informaatiotutkimukseen. Helsinki: Btj Kustannus, 9–31.

### **Verkkolähteet:**

<http://www.uta.fi/hallinto/yliopistopalvelut/viestinta/avuksi/valokuvapalvelut.html> (käytetty 14.4.2011)

[http://www.narc.fi/Arkistolaitos/oppikirja/asiakirjahallinnon\\_oppikirja.pdf](http://www.narc.fi/Arkistolaitos/oppikirja/asiakirjahallinnon_oppikirja.pdf) (käytetty 08.10.2014)

Neugebauer, T (2005) Image indexing. Photography Media Journal (2005). <http://www.photographymedia.com>. Viitattu 15.3.2010.

### **Tiedustelut:**

Renvall, J. 2014. Viestintäyksikön kuva-arkisto. Henkilökohtainen tiedonanto. Sähköposti johanna.nevalainen@uta.fi 25.10.2014

Venetjoki, R. 2014. Viestintäyksikön kuva-arkisto. Henkilökohtainen tiedonanto. Sähköposti johanna.nevalainen@uta.fi 27.03.2014.

Venetjoki, R. 2011. Viestintäyksikön kuva-arkisto. Henkilökohtainen tiedonanto. Sähköposti johanna.nevalainen@uta.fi 22.11.2010.

## **LIITE 1. Prosessikaaviot**

Hahmottele tähän kuvan elinkaari prosessina Viestintäyksikön arkistossa:

Hahmottele tähän millainen olisi mielestäsi ideaalinen prosessi:

Hahmottele tähän kuvan tyypillinen (useimmiten esiintyvä) elinkaari prosessina viestintäyksikön kuva-arkistossa?

Hahmottele tähän, miten tyypillinen (useimmiten esiintyvä) kuvan kuvailuprosessi viestintäyksikön kuva-arkistossa menee?



## **LIITE 2. Haastattelukysymykset kuvaajalle K1**

Mitä kameran kuvatietoja jää kuvatiedostoon pysyvästi?

Mitä tietoja kuvatiedostoihin liitetään tallennusvaiheessa?

Mitä muita tietoja kuvatiedostoihin olisi mielestäsi hyvä liittää?

### **LIITE 3. Haastattelukysymykset kuvaajalle K2**

Millainen koulutus sinulla on?

Miten olet hakeutunut tai päätenyt tälle alalle?

Työkokemus: Millaisia työpaikkoja/tehtäviä sinulla on ollut aiemmin?

Millaisia ovat nykyiset työtehtävät, miten ne painottuvat eri asioihin?

Millainen käsitys sinulla on nykyisestä työnkuvasta?

Onko työssä rutiinin omaisia piirteitä, voiko toteuttaa luovia ideoita jne..?

Miten työnjako toimii työyhteisössä?

Montako kuvaajaa viestintäyksikössä on?

Ovatko kaikki työntekijät samanaikaisesti työsuhteessa?

Mikä rooli kuvaajalla on kuvien kuvailussa?

Tekeekö joku muukin kuvailua?

Jääkö kameran muistista metatietoja/kuvatietoja kuvaan?

Mitä kameran kuvatietoja jää kuvatiedostoon pysyvästi?

Mitä tietoja kuvatiedostoihin liitetään tallennusvaiheessa?

Miten kuvia etsitään arkistossa?

Onko kuva-arkistossa hakutoimintoa?