

Simo Keltanen

**SOSIAALISEN VERKON VISUAALISEN ANA-  
LYSOINNIN SELKIYTTÄMINEN VÄHENTÄMÄLLÄ  
LINKKIEN JA SOLMUJEN MÄÄRÄÄ.**



JYVÄSKYLÄN YLIOPISTO  
TIETOJENKÄSITTELYTIETEIDEN LAITOS  
2016

## TIIVISTELMÄ

Keltanen, Simo

Sosiaalisen verkon visuaalisen analysoinnin selkiyttäminen vähentämällä linkkien ja solmujen määrää

Jyväskylä: Jyväskylän yliopisto, 2016, 57 s.

Tietojärjestelmätiede, pro gradu -tutkielma

Ohjaaja: Veijalainen, Jari

Viime vuosikymmenen aikana tapahtuneen valtavan tietomassan kasvun takia verkostomaisen tiedon analysoinnista on tullut paljon vaikeampaa samalla, kun pelkät nykyiset tiedonlouhinnan menetelmät ovat käyneet riittämättömiksi. Vaikka avuksi otettaisiin ihmisen vuorovaikutustaidot, joustavuus ja luovuus, voi analysointi olla todella aikaa vievää nykyisellä linkkien ja solmujen valtavalla määrällä. Tästä johtuen, tutkielman keskeisimpänä tavoitteena oli tutkia sitä, kuinka linkkien ja solmujen määrää voidaan vähentää, ja siten selkiyttää sosiaalisten verkkojen visuaalista analysointia, käyttäen apuna graafisia elementtejä, kuten kokoa, värejä ja muotoja.

Tutkimusmenetelminä sovellettiin kirjallisuuskatsausta ja konstruktiivista toimintatutkimusta. Konstruktiona kehitettiin eräälle Suomen valtion organisaatiolle uusi visuaalinen analysointijärjestelmä ketterää sovelluskehitysmenetelmää hyödyntäen. Kehitysprojektiin osallistuivat tämän työn kirjoittaja kehittäjänä ja tutkijana sekä kahdeksan järjestelmän loppukäyttäjää. Uuteen järjestelmään liittyvä tutkimusaineisto kerättiin osallistuvalla havainnoinnilla ja puolistruktuurisin teemahaastatteluin, joihin osallistui viisi järjestelmän loppukäyttäjää. Haastateltavat valittiin heidän tehtävänkuviansa ja kokemuksensa perusteella. Haastatteluun sisältyi myös kyselylomake, jotta osa tiedosta saatiin kerättyä tiiviimmässä muodossa. Haastatteluiden litteroinnin jälkeen aineiston analysoinnissa käytettiin laskemista, asteikointia, sekä teemoittelua.

Tutkielman tärkeimpänä tuloksena voidaan pitää sektorimallin kehittämistä. Tällä uudella mallilla pystyttiin linkkien ja solmujen vähentämisen lisäksi nopeuttamaan ja helpottamaan erityisesti alustavaa analyysiä tekevien työtä enemmän, kuin muilla järjestelmään toteutetuilla toimenpiteillä. Suurin linkejä ja solmuja vähentävä vaikutus saatiin aikaiseksi visualisoimalla transaktiotieto kahden linkin ja yhden solmun sijasta vain yhdellä linkillä. Visualisointitapaa muuttamalla sekä suodattamalla ja summaamalla analysoitavaa tietoa pystyttiin visualisoitavien linkkien ja solmujen määrää vähentämään merkittävästi.

Yleisesti voidaan todeta, että linkkien ja solmujen määrää vähentämällä voidaan selkiyttää ja helpottaa visuaalista analysointia ja varsinkin yleiskuvan luontia. On kuitenkin tärkeää, että menetetty, oleellinen tietosisältö pystytään esittämään vaihtoehtoisella tavalla niin, että kaikkien käyttäjäryhmien tietotarpeet otetaan huomioon.

Asiasanat: visuaalinen analyysi, verkostomaisen tiedon visualisointi, sosiaaliset verkot, toimintatutkimus

## ABSTRACT

Keltanen, Simo

Clarifying social network visual analysis by reducing the amount of links and nodes

Jyväskylä: University of Jyväskylä, 2016, 57 p.

Information Systems Science, Master's Thesis

Supervisor: Veijalainen, Jari

The amount of data has grown exponentially in the past decade. Therefore, analyzing huge volumes of network data has become much more difficult and often inadequate using current data mining methods alone. Human involvement (i.e., interaction, flexibility and creativity) is essential, but it can be extremely time consuming given the current amount of links and nodes present in the data modeling social networks. Therefore, the main aim of this study was to investigate how to decrease the number of links and nodes which, in turn, simplifies the visualization process in social network's visual analytics, such as size, colors and shapes of the graphical elements matter.

This study consisted of a literature review of constructive action research and the development of a new visual analytics system for a Finnish State Department by utilizing an agile system development method. Participants included this author as a developer and eight employees from the above Department who acted as end-users of the system. Of these, five were selected for interviews for their long work experience. The evaluation data was collected by using participant observation and semi-structured theme interviews. Additionally, questionnaires were given to the five interviewees, which allowed for more a concise information collection. After the transcription of the interviews, the data was analyzed using counting, scaling, and thematic analysis.

The main result of this study was the development of a new sector model for visualizing social network data which reduces the number of links and nodes and also accelerates and facilitates the preliminary analysis of the data. This increases work productivity of end-users much more than any other implemented system feature. A secondary result was the reduction of the majority of the links and nodes by visualizing a transaction, opposed to the previous method of using two links and a single node with only one link. By changing the way of visualization as well as filtering and summing-up the analyzed data, it was possible to reduce the total number of links and nodes significantly.

Overall, it can be stated that by reducing the number of links and nodes it is possible to clarify and facilitate the visual analytics, especially by creating an overall picture of the analyzed social network. However, it is important to note that lost essential data content is still present in an alternative manner so that the needs of every user's information are provided for.

Keywords: visual analytics, network visualization, social networks, action research

## KUVIOT

KUVIO 1 Visuaalisen analyysiin integroituvat tieteenalat (mukautettu Keim ym., 2008).....	8
KUVIO 2 Yksinkertaista verkkoa kuvaava visualisointi (mukautettu Westphal, 2008, 71).....	14
KUVIO 3 Esimerkkejä erilaisista verkkojen tyypeistä (mukautettu Newman, 2003).....	14
KUVIO 4 Ymmärtämisen silmukka (Mukautettu Cook & Thomas, 2005, 43)....	18
KUVIO 5 Verkkomaisen tiedon visualisoinnin ontologia.....	22
KUVIO 6 Sektorimallin henkilöä kuvaavia ikoneita.....	30
KUVIO 7 Sektorimallin yhtiötä kuvaavia ikoneita.....	30
KUVIO 8 Ristiinvertailutulokset listana.....	31
KUVIO 9 Transaktion kuvaaminen ikonin avulla.....	31
KUVIO 10 Transaktion kuvaaminen yksittäisinä linkkeinä.....	32
KUVIO 11 Yhdensuuntaisten transaktioiden kuvaaminen yhtenä linkkinä osapuolten välillä.....	32
KUVIO 12 Henkilö neljällä kaimalla.....	33
KUVIO 13 Haastatteluaineiston käsittely analyysistä synteesiin (mukautettu Hirsjärvi & Hurme, 2014, 144).....	35

## TAULUKOT

TAULUKKO 1 Sektorimallin toimivuus.....	38
TAULUKKO 2 Vanhasta transaktioiden visualisointitavasta luopuminen.....	39
TAULUKKO 3 Analyysin helppous ja nopeus.....	40
TAULUKKO 4 Linkkien ja kuvakkeiden vähentymisen vaikutus.....	40
TAULUKKO 5 Graafisten elementtien toimivuus.....	42
TAULUKKO 6 Toimenpiteitä verkostomaisen, visualisoitavan tiedon tekemiseksi helpommin ihmiselle analysoitavaksi.....	49

# SISÄLLYS

TIIVISTELMÄ

ABSTRACT

KUVIOT

TAULUKOT

1	JOHDANTO.....	7
1.1	Tutkimusongelmat ja tutkimuksen rajaus .....	8
1.2	Tutkimusmenetelmät .....	9
1.2.1	Taustatiedonkeruun menetelmät.....	9
1.2.2	Visuaalisen analyysijärjestelmän kehittäminen.....	10
1.2.3	Tutkimusaineiston keruu ja analysointi .....	10
1.3	Tulokset ja niiden merkitys .....	11
1.4	Tutkielman rakenne .....	11
2	SOSIAALISEN VERKON TIEDON VISUAALINEN ANALYSOINTI.....	13
2.1	Verkostomainen tieto ja sen visualisointi.....	13
2.2	Sosiaalisten verkkojen analysointi .....	15
2.3	Visuaalinen analyysi .....	16
2.3.1	Visuaalisen analyysin prosessi.....	17
2.3.2	Visuaalisen analyysin teknisiä haasteita.....	18
2.4	Visuaaliset elementit .....	19
2.4.1	Muodot, koko ja värit .....	20
2.4.2	Visualisoitavat attribuutit .....	20
2.4.3	Solmujen väliset linkit .....	21
2.5	Yhteenveto .....	21
3	KOHDEORGANISAATIO JA MUUTOSTARPEET .....	24
3.1	Visualisoinnissa käytettävät ohjelmistot .....	24
3.2	Vanhan visualisointijärjestelmän ongelmat.....	25
4	UUDEN VISUALISOINTIJÄRJESTELMÄN LUOMINEN KONSTRUKTIIVISEN TOIMINTATUTKIMUKSEN MENETELMIN.....	26
4.1	Ratkaisun tavoitteet.....	27
4.2	Suunnittele ja toteuta.....	27
4.2.1	Tiedonhankinta.....	28
4.2.2	Sovelletut ketterät menetelmät.....	28
4.3	Uudet ominaisuudet.....	29
4.3.1	Sektorimalli .....	29
4.3.2	Transaktioita kuvaavan ikonin korvaaminen linkillä.....	31
4.3.3	Visualisoitavat attribuutit .....	32
4.4	Demonstrointi .....	33
4.5	Tiedonhankinta ja arviointi .....	33
4.5.1	Haastattelun ja kyselylomakkeen yhdistelmä .....	34

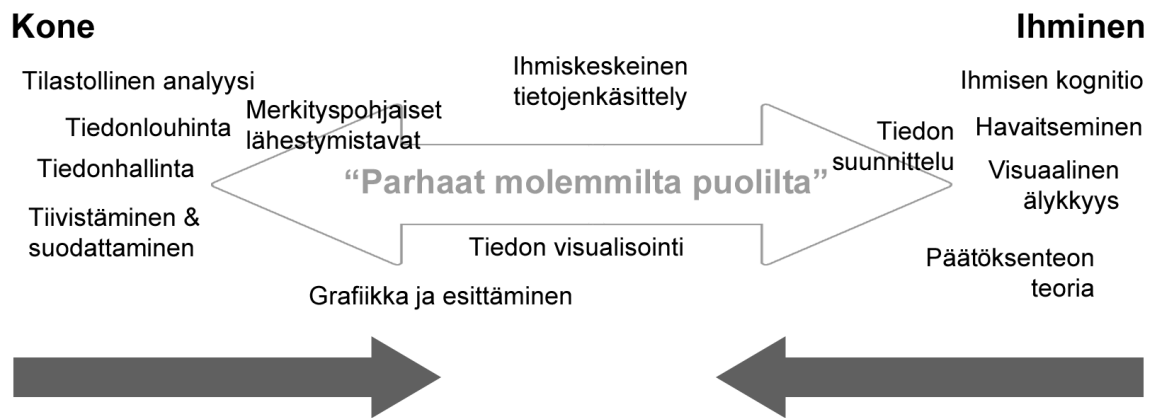
4.5.2	Havainnointi .....	34
4.5.3	Aineiston purku ja tulosten analysointi.....	35
5	TULOKSET JA JOHTOPÄÄTÖKSET .....	36
5.1	Järjestelmän käytön helppous suhteessa vanhaan.....	36
5.2	Sektorimalli.....	37
5.3	Transaktioiden esittäminen suorina linkkeinä .....	38
5.4	Analyysin helppouteen ja nopeuteen vaikuttavat ominaisuudet .....	39
5.5	Graafisten elementtien toimivuus .....	41
5.6	Oletuksena laajennettavat verkon tiedot.....	42
5.7	Tarvittavat muutokset.....	42
5.8	Johtopäätöksiä .....	43
5.8.1	Menetelmiä visuaaliseen analysointiin .....	43
5.8.2	Linkkien ja solmujen vähentäminen kadottamatta oleellista tietoa .....	43
5.8.3	Värit ja muodot visualisoinnin selkiyttäjänä .....	45
5.8.4	Ratkaisun parantaminen evaluoinnin avulla.....	46
5.8.5	Toimenpiteiden vaikutus analysoinnin helppouteen.....	47
6	POHDINTA .....	50
6.1	Tulosten arviointi.....	50
6.2	Tulosten yleistettävyys ja käytännön merkitys .....	53
6.3	Prosessin arviointi .....	53
6.4	Jatkotutkimuksen aiheita .....	54
	LÄHTEET .....	55
	LIITE 1 HAASTATTELUN MONIVALINTAKYSYMYKSET .....	57

# 1 JOHDANTO

Digitaalisen tiedon määrä maailmassa on viime vuosina lisääntynyt valtavasti, ja monet organisaatiot keräävät suuria tietovarastoja eri tietolähteistä tukemaan omaa toimintaansa, mutta eivät välttämättä osaa hyödyntää kerättyä tietoa tehokkaasti. Tietomassojen kasvaessa todella suuriksi tiedon tutkiminen ja analysointi vaikeutuu entisestään (Keim, 2002). Suurien tietomassojen louhiminen tehokkailla tietokoneilla ei pelkästään riitä, vaan etsintätyöhön tarvitaan avuksi myös yhdistelmä ihmisen yleistietoutta, joustavuutta ja luovuutta (Keim, 2002). Edellä mainittujen ominaisuuksien hyödyntämiseksi on kehitetty analysoitavan tiedon visualisointia apuna käyttävä analysointitekniikka, visuaalinen analyysi (engl. visual analytics). Visuaalisella analyysillä tarkoitetaan visuaalisesti esitetyn tiedon käsittelyä, jossa ihminen on vapaasti vuorovaikuksessa tietoon siten, että hän saa oivalluksia, voi muodostaa johtopäätöksiä sekä lopulta tehdä parempia päätöksiä saadun tiedon pohjalta (Keim, Mansmann, Schneidewind & Ziegler, 2006).

Visuaalisen analysoinnin keinoin voidaan analysoida muun muassa erilaisia sosiaalisia verkostoja (engl. social networks), rahavirtojen liikettä sekä informaation kulkua eri osapuolten välillä. Sosiaalisten verkostojen analyysistä on apua tutkittaessa esimerkiksi käyttäytymismallien samankaltaisuuksia, yhteisöjen rakenteita, sekä johtamisrakenteita (Scott, 2012, 2-3). Tärkeimpänä tietona sosiaalisten verkostojen analyysissä pidetään relaatio-tyyppistä tietoa. Muita tietotyyppisiä ovat ominaisuutta kuvaavat tiedot, sekä käsityksen muodostava tieto (engl. ideational data), jolla voidaan kuvata toiminnan merkityksiä, motiiveja, rajoituksia ja tyypityksiä (Scott, 2012, 3). Visuaalinen analyysi ei ole pelkästään tiedon visuaalista yhdistämistä, vaan siihen kuuluu muun muassa inhimillisiä tekijöitä, datan analysointia sekä päätöksentekoa (Keim ym., 2008). Visuaalisen analyysin tutkimisella edistetään työnjakoa ihmisen ja tietokoneen välillä, ja se integroi useita tieteenaloja yhteen kuvion 1 mukaisesti (Keim ym., 2008). Ideana on hyödyntää tietokonetta niissä tehtävissä, joissa se on parhaimmillaan, kuten isojen tietomassojen louhinnassa, tiivistämisessä sekä suodattamisessa ja jättää ihmiselle sellaiset tehtävät, jotka vaativat muun muas-

sa visuaalista älykkyyttä ja muita ihmiselle ominaisia taitoja, joita nykytietokoneilta ei vielä löydy (kuvio 1).



KUVIO 1 Visuaalisen analyysiin integroituvat tieteenalat (mukautettu Keim ym., 2008)

Verkko voidaan määritellä ryhmäksi solmuja ja niitä yhdistäviä linkkejä (Newman, 2003). Tietoa voidaan visualisoida erilaisten verkkomallien avulla, kuten esimerkiksi sosiaalisella verkolla, jonka solmu voi olla muun muassa henkilö tai tapahtuma, tai rahavirtoja esittävässä verkossa pankkitili tai yhtiö. Saman verkoston eri visualisoinnit ovat siis kyseisen verkoston eri malleja. Tunnistamalla keinot linkkien ja solmujen määrän vähentämiseen ilman, että menetetään oleellista tietosisältöä, visualisoinnista saadaan tehtyä selkeämpi ja helpommin analysoitava.

## 1.1 Tutkimusongelmat ja tutkimuksen raja

Tämän pro gradu -tutkielman tavoitteena oli tutkia sitä, miten laajojen, verkostojen koodaavien tietomassojen visualisoinneista voitaisiin tehdä selkeämpiä ja helpommin analysoitavia. Visuaalisen analysoinnin yhtenä suurimmista haasteista voidaan pitää kykyä havaita ja ymmärtää tietoa (Keim ym., 2006). Visualisoitavien verkkojen linkkien ja solmujen määrän kasvaessa suureksi, myös verkon kompleksisuus kasvaa, vaikeuttaen siten analyttikön työtä (Didimo, Liotta & Montecchiani, 2014).

Visuaaliseen analysointiin liittyen on tehty melko paljon tutkimusta, mutta valtaosa tutkimuksista paneutuu visualisointiohjelmistojen ominaisuuksiin sekä erilaisiin algoritmeihin, joiden avulla voidaan tuoda esiin esimerkiksi lähdeaineistosta löytyviä käyttäytymismalleja tai sosiaalisia rakenteita. Myös verkostomaisen tiedon visualisointiin käytettävistä algoritmeista löytyy paljon tutkittua tietoa, mutta graafisten elementtien (värin, koon ja muotojen) käyttäminen visualisointia selkiyttävänä sekä linkkejä ja solmuja vähentävänä muuttujana verkostomaisen tiedon visuaalisessa analysoinnissa oli aihepiiri, josta kirjoittaja ei löytänyt kovinkaan paljon tutkittua tietoa.

Tämän tutkielman tarkoituksena oli vastata seuraavaan kysymykseen:



- Millä toimenpiteillä verkostomaisen tiedon visualisoinnista voidaan tehdä helpommin ihmiselle analysoitavaa?

Varsinaisen tutkimusongelman selvittämiseksi laadittiin seuraavat apukysymykset:

1. Mitä eri menetelmiä verkostomaisen tiedon visualiseen analysointiin on olemassa kirjallisuudessa?
2. Minkä tyyppisiä linkkejä ja solmuja voidaan poistaa, ja miten menetetty oleellinen tieto voidaan kuvata muuten, ottaen huomioon kaikki erilaiset tietotarpeet?
3. Miten verkostomaisen tiedon elementtien visualisointia voidaan selkiyttää värien ja muotojen avulla?
4. Miten kehitettyä ratkaisua voidaan evaluoida ja parantaa evaluoinnin tulosten pohjalta?

Sosiaalisten verkkojen visuaalisen analyysin menetelmiin paneuduttiin tutkielman kirjallisuuskatsauksessa. Löydettyä tutkimustietoa sovellettiin kehitettävässä erään julkishallinnon tietojärjestelmän tiedon visualisointia helpommin tulkittavaksi ja analysoitavaksi. Tutkielma painottuu tiedon visualisoinnin graafiseen puoleen, eikä siinä kuvata visualisointijärjestelmän teknisempiä asioita, kuten käytettyjä visualisointialgoritmeja. Myöskään konstruktiivisen osion, kohdeorganisaation visualisointijärjestelmän taustatietokantojen rakenteita, nimiä tai muita kohdeorganisaation salassa pidettäviä, teknisiä menetelmiä ei kuvata tutkielmassa. Visualisoinnin takia tehtäviä analyysikannan muutoksia-kin on kuvattu vain yleisellä tasolla.

## 1.2 Tutkimusmenetelmät

### 1.2.1 Taustatiedonkeruun menetelmät

Tutkielmassa hyödynnetty teoretieto kerättiin kirjallisuuskatsauksen avulla. Katsauksella pyrittiin selvittämään kattavasti visuaalisen analysoinnin nykytila ja vastaamaan kysymykseen siitä, mitä eri menetelmiä verkostotyyppisen aineiston tiedon visualisointiin on olemassa kirjallisuudessa. Kirjallisuuskatsauksen tuloksena odotettiin siis löytyvän sellaisia verkostomaisen tiedon visualisointiin liittyviä menetelmiä ja ratkaisumalleja, jotka auttavat suunnittelemaan kohdeorganisaation visualisointiratkaisusta selkeämmän ja helpommin analysoitavan, sekä auttavan osaltaan vastaamaan myös varsinaiseen tutkimuskysymykseen, eli millä toimenpiteillä verkostomaisen tiedon visualisoinnista voidaan tehdä helpommin ihmiselle analysoitavaa. Aiheen monitieteisyydestä johtuen kirjallisuuskatsaukseen sisällytettiin eri tieteenalojen julkaisuja. Hakukoneena artikkelien etsimisessä käytettiin pääasiallisesti Google Scholaria. Hakusanoja olivat: "visual analytics", "visual analysis", "network visualizati-

on”, ”information visualization” sekä ”social networks”. Teknisimmät lähteet, jotka käsittelivät pääasiassa visualisointijärjestelmissä käytettäviä algoritmeja, suljettiin tutkimuksen ulkopuolelle, koska visualisoinnin tekninen toteutus on pääosin rajattu tutkielman ulkopuolelle. Lähteinä käytettiin niin konferenssijulkaisuja, aikakausjulkaisuja kuin kirjojakin. Lähteiden laatua arvioitiin niihin tehtyjen viittausten määrän, kirjoittajan sekä julkaisualustan perusteella. Relevantteimmilta vaikuttaneista lähteistä luettiin ensin tiivistelmä, jonka perusteella arvioitiin onko lähde hyödyllinen tutkimuksen kannalta ja auttaako se vastaamaan tutkimuskysymyksiin. Tarkempaan tarkasteluun otettavien lähteiden tiedot taltioitiin RefWorks -ohjelmaa apuna käyttäen. Taltioiduista 46:stä lähteestä hyödynnettiin lopulta tässä tutkielmassa 23:a lähdeä.

### **1.2.2 Visuaalisen analyysijärjestelmän kehittäminen**

Tutkielman toinen osa toteutettiin konstruktivisena toimintatutkimuksena, jonka viitekehyksenä toimi Peffersin, Tuunasen, Rothenbergin ja Chatterjeen (2007) kehittämän tieteellisen suunnittelututkimuksen metodin (engl. The design science research methodology) prosessimallin ja Seinin, Henfridssonin, Puraon, Rossin ja Lindgrenin (2011) metodiin liittyvän, jatkuvan arvioinnin periaatteen yhdistelmä. Peffersin ym. (2007) prosessimallia noudatteleva tutkielman viitekehys sisältää kuusi aktiviteettia, joita ovat: ongelman tunnistaminen ja motivointi, ratkaisun tavoitteiden määrittely, suunnittelu ja toteutus, demonstrointi, arviointi, sekä kommunikointi. Uusi visuaalinen analyysijärjestelmä toteutettiin kohdeorganisaationa toimivalle julkishallinnon yksikölle käyttäen ketteriä sovelluskehitysmenetelmiä.

### **1.2.3 Tutkimusaineiston keruu ja analysointi**

Järjestelmään liittyvä empiirinen tutkimusaineisto kerättiin pääosin osallistuvaa havainnointia sekä teemahaastattelun ja kyselylomakkeen yhdistelmää käyttäen. Toimenpiteiden linkkejä ja solmuja vähentävää vaikutusta mitattiin myös vertailemalla vanhasta ja uudesta visualisointijärjestelmästä kerättyä numeerista tietoa. Kvalitatiivisen aineiston analysoinnissa käytettiin apuna Hirsjärven ja Hurmeen (2014, 171–173) oppien mukaisesti laskemista (engl. counting), asteikointia (engl. scaling), sekä teemoittelua (engl. thematizing). Laskeminen on tässä tutkielmassa sen selvittämistä, miten useasti jokin ilmiö tai mielipide esiintyy aineistossa, ja asteikoinnilla taas kuvataan esimerkiksi sitä, miten vahvasti haastateltavat ovat jonkin asian puolella tai sitä vastaan, tai kuinka usein he tekevät jotain. Teemoittelulla tarkoitetaan analyysivaiheessa sitä, että aineistosta etsitään ja tarkastellaan piirteitä, jotka ovat nousseet esille usean haastattelun yhteydessä. (Hirsjärvi & Hurme 2014, 171–173.)

### 1.3 Tulokset ja niiden merkitys

Tutkielman tuloksena saatiin rakennettua kohdeorganisaatiolle uusi visuaalisen analyysin järjestelmä, jonka todettiin helpottavan varsinkin alustavaa analyysiä tekevien loppukäyttäjien työtä. Suurin analyysiä helpottava vaikutus saavutettiin vähentämällä järjestelmän visualisoitavien solmujen ja linkkien määrää, erityisesti analyysin alkuvaiheessa. Keskeisimpänä tutkielman saavutuksena voidaan pitää sektorimalliksi nimetyn, uuden artefaktin luomista, jonka avulla onnistuttiin vähentämään visualisoitavien solmujen määrää, selkiyttämään visualisointia ja vähentämään manuaalisen työn määrää. Sektorimallissa yhdistyy henkilöitä ja yhtiöitä kuvaavien solmujen dynaamisen ikonin välittämä tietosisältö, sekä attribuuttiin sijoitettu tiivistetty ja summattu tieto. Ikonilla, joka on solmua kuvaava visuaalinen elementti, pystytään kuvaamaan mitä tietoa kohteeseen kuuluu, ilman että verkostomaisen tiedon linkkejä ja solmuja tarvitsee laajentaa (engl. expand). Attribuuttiin sijoitetussa tiivistelmässä listataan piilotettujen linkkien ja solmujen analyysin kannalta tärkein tietosisältö. Attribuutilla tarkoitetaan tässä tutkielmassa yksinkertaistettuna solmun tai linkin ominaisuutta tai siihen liittyvää asiaa, johon pääsee käsiksi kaksoisklikkaamalla kohdetta.

Tutkielmassa onnistuttiin vastaamaan myös jokaiseen apukysymykseen ja apukysymysten vastaukset autoivat vuorostaan vastaamaan varsinaiseen tutkimuskysymykseen. Toimenpiteistä, joiden avulla verkostomaisen tiedon visualisoinnista voidaan tehdä ihmiselle helpommin analysoitavaa, koottiin yleistettävä seitsemän kohdan lista. Kyseistä listaa voidaan hyödyntää suunniteltaessa visuaalisen analyysijärjestelmän käyttöönottoa, tai kehitettäessä jo olemassa olevaa järjestelmää.

### 1.4 Tutkielman rakenne

Tämä tutkielma koostuu johdannon lisäksi viidestä muusta luvusta. Toisessa luvussa keskitytään sosiaalisen, verkostomaisen tiedon visuaaliseen analysointiin verkostomaisen tiedon visualisoinnin, sosiaalisten verkkojen analysoinnin, visuaalisen analyysin, sekä visuaalisten elementtien näkökulmasta. Kappaleen lopussa on yhteenveto kirjallisuuskatsauksessa löydetyistä asioista, jossa analysoidaan löydöksiä, sekä esitellään kirjallisuuskatsauksen pohjalta laadittu verkostomaisen tiedon viitekehys, jota käytettiin apuna tutkielman konstruktiivisessa osiossa. Kolmannessa luvussa esitellään tutkimuksen kohdeorganisaatiota ja sen muutostarpeita. Neljännessä luvussa esitellään miten ja minkälainen uusi visualisointijärjestelmä luotiin konstruktiivisen toimintatutkimuksen menetelmin, sekä kuvataan tiedonhankintaa, niin artefaktin toteuttamisen, kuin myös tutkimuksen arvioinnin näkökulmasta. Viidennessä luvussa käydään läpi tutkimuksen

tulokset. Kuudennessa luvussa arvioidaan tuloksia ja prosessia, sekä pohditaan jatkotutkimuksen aiheita.

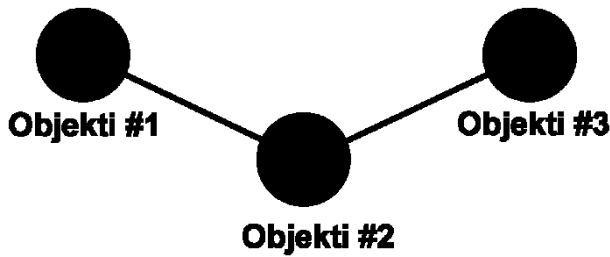
## 2 SOSIAALISEN VERKON TIEDON VISUAALINEN ANALYSOINTI

Sosiaalisen, verkostomaisen tiedon visuaalisen analysoinnin syvällisemmän ymmärtämisen saavuttamiseksi tämän kappaleen alaluvuissa tutustutaan neljään asiakokonaisuuteen. Ensimmäiseksi määritellään mitä verkostomainen tieto on ja miten sitä voidaan visualisoida. Toiseksi kuvataan mitä sosiaalinen verkko käsitteenä tarkoittaa ja miten näitä verkkoja analysoidaan. Kolmanneksi perehdytään visuaaliseen analyysiin, minkä jälkeen syvennytään visuaalisiin elementteihin ja niiden hyödyntämiseen linkkien ja solmujen visualisoinnissa.

### 2.1 Verkostomainen tieto ja sen visualisointi

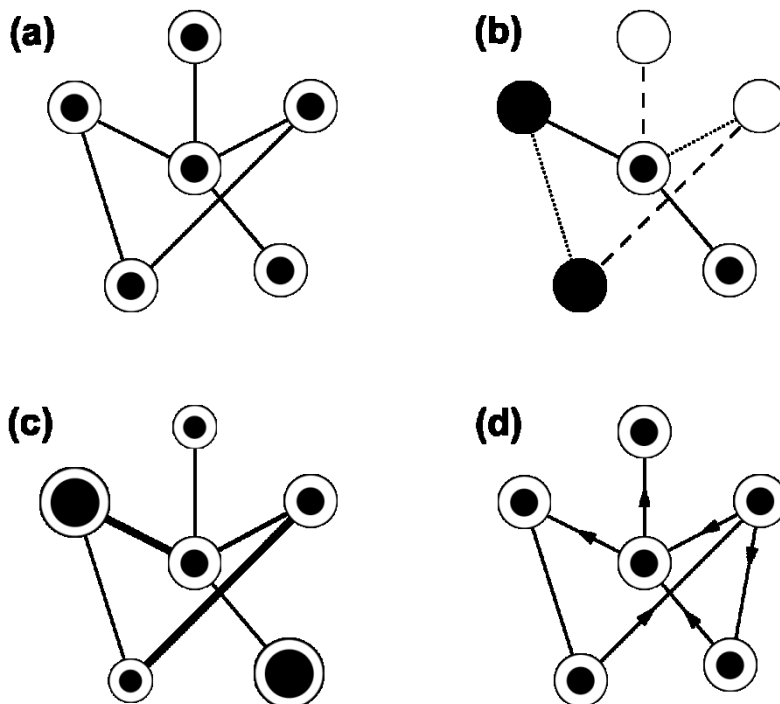
Matemaattisen verkkoteorian (engl. graph theory) mukaan yksinkertaisin verkko on tyhjä verkko (engl. null graph), joka ei sisällä yhtään solmua tai linkkiä. Saman teorian mukaan toiseksi yksinkertaisin verkko, eli triviaali verkko (engl. trivial graph) sisältää vain yhden solmun, mutta ei yhtään linkkiä. (Gross, Yellen & Zhang, 2014, 4.)

Westphalin (2008, 70) sekä Choin ja Kimin (2008) mukaan yksinkertaisimmillaan verkko on kolmen solmun ja kahden erillisen linkin muodostama kokonaisuus (kuvio 2). Yhden solmuparin muodostama, yhden tai useamman linkin sisältämä kokonaisuus ei siis vielä olisi verkko, vaan verkon muodostamiseen vaaditaan solmu, joka on yhteydessä kahteen tai useampaan muuhun solmuun (Westphal, 2008, 70). Kirjallisuudesta löytyy myös esimerkkejä, joissa kahdesta solmusta ja niiden yhdistävästä linkistä muodostuvaa kokonaisuutta pidetään yksinkertaisimpana verkkona (Hsu, 2006; Ekström, Engelson & Rydergren, 2009).



KUVIO 2 Yksinkertaista verkkoa kuvaava visualisointi (mukautettu Westphal, 2008, 71)

Newmanin (2003) määritelmän mukaan verkko on yksinkertaisimmillaan ryhmä solmuja ja niitä yhdistäviä linkkejä, mutta verkot voivat olla myös monella tavalla monimutkaisempia. Solmuja ja linkkejä voi olla useanlaisia ja molemmat voivat omata vaihtelevan kirjon erilaisia ominaisuuksia. Solmulla voidaan esimerkiksi esittää henkilöä, joka on mies tai nainen ja näiden välinen linkki voi kertoa heidän olevan ystäviä tai vihamiehiä. Linkit voivat sisältää painotuksia, joilla voidaan kertoa esimerkiksi siitä, kuinka hyvin henkilöt tuntevat toisensa. Yhteydet voivat olla myös yksisuuntaisia, jolloin linkkiä kuvataan nuolella, esimerkiksi haluttaessa ilmaista kumpi osapuoli oli puhelinkeskustelussa soittaja. (Newman, 2003.) Kuvioista 3 löytyy havainnollistava esimerkki neljästä erityyppisestä verkosta.



KUVIO 3 Esimerkkejä erilaisista verkkojen tyypeistä (mukautettu Newman, 2003)

Kuvion esittämät verkkotyypit kuvaavat seuraavanlaisia verkkoja:

- a) Yhdentyypisiä solmuja ja linkkejä sisältävä (homogeeninen verkko),
- b) erilaisia linkki- ja solmutyyppejä sisältävä (heterogeeninen verkko),
- c) vaihtelevia linkkien ja solmujen painotuksia sisältävä (painotettu verkko), sekä
- d) verkko, jossa jokaisella linkillä on oma suuntansa (suunnattu verkko) (Newman, 2003).

Verkkojen tutkimisen historia alkoi matemaattisena tutkimuksena jo 1700-luvulla. Sosiaalitieteissä verkkoja on tutkittu laajasti 1930-luvulta lähtien. Tällöin ymmärrettiin ihmisten välisten yhteyksien rakenteiden tärkeys yhteiskunnassa (Newman, 2003). Kolmekymmentäluvulta aina seitsemänkymmentäluvulle asti yhä useampi sosiaalinen antropologi ja sosiologi alkoi hyödyntää Radcliffe-Brownin sosiaalisen struktuurin (engl. social structure) käsitettä ja suhtautua vakavasti sosiaalisen elämän metaforiin: ”kudos” (engl. fabric) ja ”verkko” (engl. web). Näiden tekstiileihin perustuvien metaforien, kuten esimerkiksi ”yhteenkutoutumisen” (engl. interweaving) ymmärtämisen jälkeen etusijalle nousi sosiaalisen verkoston metafora ja tutkijat alkoivat tarkastella tutkimuksissaan sosiaalisten verkostojen tiheyttä (engl. density) ja rakennetta. (engl. texture) (Scott, 2012, 1.)

## 2.2 Sosiaalisten verkkojen analysointi

Sosiaalisten verkostojen analysoinnissa keskitytään toimijoiden (engl. actor) välisiin linkkeihin eli relaatioihin, eikä niinkään yksittäisiin toimijoihin tai heidän ominaisuuksiinsa. Sosiaalisia verkkoja kuvaavien verkkojen tietokokonaisuudet voivat sisältää tietoa useantyyppisistä sosiaalisista olioista, eli entiteeteistä, kuten henkilöistä ja organisaatioista. Näitä entiteettejä voidaan tarkastella samanaikaisesti usealla tasolla. Tutkittavat tasot voivat olla esimerkiksi yksilö, ryhmä, organisaatio tai yhteiskunta. (Hanneman & Riddle, 2005.) Sosiaalisten verkkojen analysoinnin kehittäjänä voidaan pitää antropologi Radcliffe-Brownia. Sosiaalisten verkostojen analysointi kehittyi alun perin muodoltaan suhteellisen ei-teknišenä, kohdistuen edellä mainittuihin ihmisten välisiin suhteisiin (Scott, 2012, 1).

Sosiaalisten verkkojen analysointi on hyödyllistä esimerkiksi tutkittaessa verkon rakenteen samankaltaisuuksia (engl. patterns) ja yhteisöjen rakenteita. Tärkeimmät tiedon tyypit ovat ominaisuustieto ja relaatiotieto. Ominaisuustieto kuvaa asenteita, mielipiteitä ja kohteen käytöstä, sikäli kun ne liittyvä ominaisuuksiin, kyvykkyyksiin tai ominaispiirteisiin ja kuuluvat yksilölle tai ryhmälle. Ominaisuustiedon analysointiin sopii muuttuja-analyysi, jonka avulla ominaisuuksia mitataan jonkin muuttujan arvona, kuten esimerkiksi koulutus, ammatti, tai saadut tulot. (Scott, 2012, 2-3.)

Relaatiotieto on tietoa, joka liittyy kontakteihin, sidoksiin, yhteyksiin, ryhmän suhteisiin sekä tapaamisiin ja joka yhdistää yksittäiset toimijat toisiinsa,

mutta jota ei voi muuntaa yksittäisten toimijoiden ominaisuudeksi (Scott, 2012, 3). Hannemanin ja Riddlen (2005) mukaan yhdistävät tekijät voivat liittyä esimerkiksi systeemiteorian mukaisesti materiaaliin tai informaatioon, eli ihmiset voivat liikkua organisaatioiden välillä, raha henkilöiden välillä ja kerrottu tieto siirtyy vastaanottajalle. Materiaali voi olla vain yhdessä paikassa samaan aikaan, kun taas tieto voi olla useassakin paikassa samaan aikaan (Hanneman & Riddle, 2005). Wasserman ja Faust (1994, 4) toteavat kokoavasti, että toimijoiden väliset linkit ovat kanavia aineettomien tai fyysisten resurssien siirtämiseen tai virtaamiseen (engl. flow).

Relaatiot eli suhteet eivät siis ole toimijoiden ominaisuuksia, vaan relaatiomainen systeemi rakentuu yhdistyneistä pareista, jotka ovat vuorovaikutuksessa toisiinsa. Relaatiotiedon analysointiin käytetään verkon analysoinnin metodeja, joissa toimijoiden yhteydet rinnastetaan toimijoiden välisiin linkkeihin. Relaatioiden määrää voidaan tutkia määrällisesti ja tilastollisesti, laskien esimerkiksi relaatioiden määriä tai tutkimalla relaatioiden muodostamien rakenteiden tilastollista merkitystä. Verkkoanalyysi taas muodostaa rungon kvalitatiiviselle verkon rakenteiden ja kehittämisen toimenpiteiden kuvaamiselle. (Scott, 2012, 3.)

Relaatio- ja ominaisuustieto eivät kuitenkaan ole ainoita verkostojen analysointiin sosiaalitieteissä käytettyjä tiedon tyyppisiä. Kolmatta tietotyyppiä voidaan kutsua käsityksen muodostavaksi tiedoksi, joka voi suoraan kuvastaa merkityksiä, motiiveja ja määrittämiä ja sillä voidaan myös tyypitellä toiminnassa mukana oloa. (Scott, 2012, 3.)

Sosiaaliset verkostot ovat merkityksellisiä sosiaalisen pääoman lähteitä, mutta ne ovat myös paljon muuta, kuten poliittisia konflikteja ja rahansiirtotapahtumia sisältäviä verkostoja (Scott, 2012, 8). Lainvalvontaviranomaiset käyttävät visuaalisen analyysin välineitä usein tutkiessaan rikoksia reaktiivisesti. Tällöin tutkittava entiteetti, esimerkiksi henkilö, on tiedossa ja analysointi keskittyy tämän entiteetin ympärille. Verkko muodostuu yhteyksistä esimerkiksi organisaatioihin, pankkitileihin, sijainteihin, rikostapahtumiin ja rikosteknisiin näytteisiin. Pohjimmiltaan analyysin päämäärä on laajentaa tunnettua verkostoa, jotta löydettäisiin vihjeitä tai johtolankoja, ja rikostutkija voisi tarkastella kohdetta joka näkökulmasta, ottaen selvää ketkä henkilöt liittyvät tutkittavaan kohteeseen. Yhteydet voivat löytyä esimerkiksi perheen, liiketoimien tai poliisin rekisterien kautta ja niillä pyritään näyttämään yhteyksiä ja yhteistoimintaa, joka saattaa paljastaa tärkeitä yhteyksiä muuhun rikolliseen toimintaan. Myös epäsuoria yhteyksiä, kuten linkittymistä puhelinnumeron tai osoitteen kautta voidaan hyödyntää analyysissä. (Westphal, 2008, 9.)

## 2.3 Visuaalinen analyysi

Visuaalinen analysointi on analyysin laji, jossa analyttisen päättelyn apuna käytetään interaktiivista visuaalista käyttöliittymää. Visuaalisen analyysin työkaluja ja tekniikoita käytetään tiedon yhdistämiseen, sekä suurien, dynaamisten,



epäselvien ja ristiriitaisten tietomassojen ymmärtämiseen (Thomas & Cook, 2006.). Visuaalista analyysiä ei pidä sekoittaa pelkkään tiedon visualisointiin, vaan se on ennemminkin kokonaisvaltainen lähestymistapa, johon kuuluvat visuaalinen yhdistämisen lisäksi inhimilliset tekijät, data-analyysi ja päätöksenteko. Visuaalisen analyysin tutkiminen on monitieteistä tutkimusta, joka yhdistelee asioita muun muassa visualisoinnista, tiedon louhinnasta ja kognitiotieteistä. Visuaalinen analyysi siis integroi tieteenaloja yhteen, edistään työnjakoa ihmisen ja tietokoneen välillä (kuvio 1). (Keim ym., 2008.)

Myös Cookin ja Thomasin (2005, 4) mukaan visuaalinen analysointi on monitieteinen tieteenala, johon liittyy seuraavia osa-alueita:

- Analyttiset päättelytekniikat, jotka antavat käyttäjälle mahdollisuuden saada oivalluksia, jotka tukevat arviointia, suunnittelua sekä päätöksentekoa.
- Visuaaliset esitykset ja vuorovaikutteiset tekniikat, jotka hyödyntävät ihmissilmän laajaa näkökenttää polkuna mieleen, mahdollistaen käyttäjän samanaikaisesti nähdä, tarkastella sekä ymmärtää laajoja määriä tietoja.
- Kaiken tyyppisen, myös ristiriitaisen ja dynaamisen, tietoaineiston muokkaaminen siten, että se tukee visualisointia ja analysointia.
- Tekniikat, jotka tukevat analyysin tulosten tuottamista, esittämistä ja eteenpäin välittämistä, ottaen huomioon asiansyhteyden ja tiedon vastaanottajien moninaisuuden.

Visuaalisen analysoinnin tutkimisen tavoitteena voidaan pitää tietotulvan muuttamista mahdollisuudeksi (Keim ym., 2006).

### 2.3.1 Visuaalisen analyysin prosessi

Keimin (2002) mukaan visuaalisen tiedon tutkiminen seuraa yleensä kolmen askeleen prosessia, joka tunnetaan myös Shneidermanin (1996) kehittelemänä tiedon etsimisen mantrana (engl. Information-Seeking Mantra): ”Ensin yleiskuva, zoomaa ja suodata, yksityiskohtiin tarvittaessa” (engl. Overview first, zoom and filter, then details-on-demand). Ensin käyttäjän täytyy siis luoda yleiskuva tiedosta (Keim, 2002). Yleiskuvan avulla käyttäjä löytää mielenkiintoiset tapahtumaketjut ja rakenteet, ja keskittyy niistä yhteen tai useampaan (Keim, 2002), zoomaamalla mielenkiintoisiin objekteihin ja suodattamalla pois ei-kiinnostavat objektit (Shneiderman, 1996). Analysoidakseen löydöksiä käyttäjän tulee perehtyä kohteeseen tarkemmin ja päästä käsiksi analysoitavan kohteen lisätietoihin (Keim, 2002).

Toisin kuin aikaisemmin kuvatussa tiedon etsimisen mantrassa kuvattiin, visuaalisen analyysin prosessi sisältää automaattisen analyysin menetelmiä ennen ja jälkeen vuorovaikutteisen visuaalisen esittämisen (Keim ym., 2006). Suuria ja monimutkaisia tietomassoja ei voi analysoida suoraviivaisesti, ilman etukäteen tehtyä automaattista analyysiä, jonka takia Keim ym. (2006) kehittivät edellä mainitun tiedon etsimisen mantran pohjalta seuraavan visuaalisen ana-

lyysin mantran: "Analysoi ensin – näytä tärkeys – poraudu, suodata ja analysoi pidemmälle – yksityiskohtiin tarvittaessa" (engl. Analyse First – Show the Important – Zoom, Filter and Analyse Further – Details on Demand).

Useat analyttisen pohdinnan tehtävät noudattavat kuviossa 4 kuvattua prosessia, jossa ensin kerätään tietoa, seuraavaksi kuvataan tieto paremmin analysointia tukevassa muodossa, sitten muodostetaan näkemys muokkaamalla tietoa ja lopuksi synnytetään jokin oivallukseen perustuva tuotos tai suoritetaan jokin toiminto (Cook & Thomas, 2005, 42). Tämän kaltaista kaavaa noudattavia tehtäviä voidaan kutsua ymmärtämistehtäviksi (engl. sense-making tasks) tai ymmärtämisen silmukaksi (sense-making loop), jos tehtäviä tehdään iteratiivisesti, kiertäen tehtävästä toiseen (kuvio 4) (Cook & Thomas, 2005, 42).



KUVIO 4 Ymmärtämisen silmukka (Mukautettu Cook & Thomas, 2005, 43)

### 2.3.2 Visuaalisen analyysin teknisiä haasteita

Käsiteltävän tiedon määrä on viime vuosina kasvanut valtavasti. Tiedon määrän kasvuun ovat vaikuttaneet nopeammat tietoliikenneverkot, halvempi talennuskapasiteetti, sekä tiedon nopeampi prosessointi (Keim ym., 2006). Organisaatiot pystyvät nykyään tuottamaan omista tietojärjestelmistään yhä laajemmin visuaaliseen analyysiin soveltuvaa tietoa. Suurien tietomassojen riittävä tutkimiseen ei ole kuitenkaan ollut mahdollisuuksia, minkä seurauksena tiedon keräämisen hyödyt häviävät ja tietokannoista tuleekin tiedon kaatopaikkoja. (Keim ym., 2006.)

Yleensä visualisoitavan tiedon määrä ylittää moninkertaisesti tietokoneen näytölle mahtuvan tiedon määrän, joten tiedon määrää tulee vähentää suodattamalla, summaamalla ja käyttämällä muita tiedon pelkistämisen tekniikoita. Järjestelmän kehittäjien ei tule siis ainoastaan vertailla tiedon määrän kasvua ja

laitteistojen suorituskykyä selvittääkseen ongelmasta, vaan keskittyä myös ohjelmistoihin ja algoritmeihin, tuodakseen analysoitavan tiedon oikealla tavalla tietokoneen näytölle. (Keim ym., 2006.) Yhtenä haasteena on myös tunnistaa parhaat automaattiset algoritmit käsillä olevaan ongelmaan, tunnistaa ne asiat joita ei voida automatisoida ja yhdistää parhaat analyysiä tukevat algoritmit sopivaan visualisointiin ja vuorovaikutustekniikoihin (Keim ym., 2008).

Keimin ym. (2006) mukaan tiedon merkityksen selkiyttäminen tai kyky havaita ja ymmärtää tietoa on yksi suurimmista visuaaliseen analyysiin liittyvistä haasteista. Visuaalisesti oikeanlaisen lopputuloksen tuottaminen raakatiedosta riippuu vahvasti käytetyn tiedon laadusta ja käytetyistä metodeista, sillä raakatieto voi sisältää monenlaisia laatuongelmia, kuten:

- keräysvaiheen aikana tulleet virheet,
- erilaiset häiriöt,
- ylimääräinen tieto,
- epätarkkuudet,
- puuttuvat arvot,
- kattavuusongelmat sekä
- toisteiset arvot.

Toisteisilla arvoilla tarkoitetaan sitä, että jollain yksittäisellä entiteetillä, esimerkiksi ihmisellä, on järjestelmässä useita ilmentymiä. Toisin sanoen, henkilön tiedot ovat järjestelmässä useammin kuin kerran. Lisäksi tiedon uudelleenprosessointi visuaalista analysointia varten synnyttää useita potentiaalisia laatuongelmia. Ongelmat voivat syntyä esimerkiksi tiedon migraatiossa, siivoamisessa, supistamisessa, rikastamisessa, ryhmittämisessä sekä tietojen yhdistämisessä. Tieto voi olla myös valmiiksi vanhentunutta. (Keim ym., 2006.)

Ongelmia voidaan ratkaista ja minimoida uudelleenprosessointivaiheessa mutta tämän lisäksi tulisi myös tarjota visuaalisen analyysin keinoja tiedon laatuun liittyvien ongelmien ratkaisemiseen. Tekniseltä kannalta laatuongelmat voidaan visuaalisessa analysointisovelluksessa jättää huomioimatta, mikä lisää tiedon siivoamisesta johtuvaa työtä, tai tiedon virheet ja epävarmuudet voidaan tuoda selkeästi esiin visuaalisesti, jotta analyytikko tulee tietoiseksi olemassa olevista ongelmista. (Keim ym., 2006.)

## 2.4 Visuaaliset elementit

Kuten jo edellisissä kappaleissa todettiin, verkostomaista tietoa voidaan mallintaa solmujen ja linkkien avulla. Tässä kappaleessa näitä elementtejä käsitellään visuaalisina elementteinä: värien, koon, muotojen ja visualisoitavien attribuuttien kautta.

Yksi sosiaalisen psykologian johtavista hahmoista 1930-luvulla oli Jacob Levy Moreno, jonka tärkein keksintö oli sosiogrammi (engl. sociogram) (Scott,

2012, 13–14). Sosiogrammi on graafinen esitys, jolla voidaan visualisoida ryhmien rakennetta ja sitä mihin asemaan kukin yksilö sijoittuu ryhmässä (Moreno, 1934, 432). Mielenkiintoisen sosiogrammista tekee sille määritelty notaatio, jossa linkkien ja solmujen muodon, koon ja värin avulla voidaan esittää laaja kirjo erilaisia asioita, joita käsitellään tämän kappaleen alaluvuissa.

#### **2.4.1 Muodot, koko ja värit**

Värit ja muodot voivat toimia visualisoinnissa indikaattoreina, joilla voidaan välittää erityyppistä tietoa käsiteltävistä entiteeteistä (solmuista) (Westphal, 2008, 129). Moreno (1934, 29–30) kuvasi sosiogrammissaan tyttöjä (alle 18-vuotiaita) ympyrällä, poikia (alle 18-vuotiaita) kolmioilla, sekä naisia suurella ympyrällä ja miehiä suurella kolmiolla. Tuplaympyrällä tai tuplakolmiolla ilmaistiin vuorostaan kunkin erillisen yksilön kuulumista toiseen ryhmään ja kahden yksilön välinen viiva merkitsi tunnereaktiota yksilöltä toiselle. Moreno kehitti sosiogrammille oman notaationsa värillisten ja yksiväristen linkkien esittämiseen – yksivärisessä jokainen viiva esitti viehätystä – monivärisessä jokainen eri väri esitti eri tunnereaktiota. Tavallisen viivan lisäksi hän käytti nuolta, jolla kuvattiin yksisuuntaista reaktiota, ja ristiviivaa, joka kuvasi kaksisuuntaista reaktiota. Moreno käytti myös katkoviivaa ja viivan paksuutta kuvaamaan erityisiä tunnereaktioita, kuten välinpitämättömyyttä, vihaa ja hallitsevuutta. (Moreno, 1934, 30.)

Tietyissä ohjelmissa solmuja voidaan myös visualisoida esimerkiksi piirakkakaavioilla tai jopa 3D-kuvilla. Solmun ja linkin visualisoinnin väriin, muotoon ja kokoon voidaan vaikuttaa näiden attribuuttien arvoilla. (Smoot, Ono, Ruscheinski, Wang & Ideker, 2011.)

#### **2.4.2 Visualisoitavat attribuutit**

Solmujen ja linkkien attribuuteilla voidaan määrittää solmujen ja linkkien visuaalisia ominaisuuksia, tai visualisoida ne myös tekstinä. Toisin sanoen, jos henkilön attribuuttiin on tallennettu tieto siitä, että kyseessä on nainen, voidaan solmua kuvaava ikoni visualisoida punaisena ja miesten ikoni voidaan kuvata sinisenä. Vaihtoehtoisesti sama tieto voidaan visualisoida tekstinä esimerkiksi ikonin alapuolelle. Waren (2012, 24) mukaan yleisesti attribuutiksi voidaan kutsua jonkin olion ominaisuutta jota ei voida käsitellä itsenäisenä oliona. Esimerkiksi omenan attribuutti voi olla väri, ja matkan attribuuttina voi toimia matkan kesto (Ware, 2012, 24). Attribuutin ja entiteetin erottaminen voi tosin olla joskus hankalaa (Ware, 2012, 24). Esimerkiksi työntekijän palkkaa voidaan ajatella attribuuttina tai siitä voidaan tehdä oma ”rahasumma”-entiteettinsä, joka linkitetään työntekijä-entiteettiin (Ware, 2012, 24). Palkan esittäminen omana entiteettinä mahdollistaa esimerkiksi palkan maksajien tuomisen mukaan visualisointiin, jolloin linkin avulla voidaan ilmaista kuka on palkan maksaja ja kuka saaja.

### 2.4.3 Solmujen väliset linkit

Linkki on kahden solmun välinen sidos. Linkkiä voidaan kutsua myös esimerkiksi suhteeksi, yhteydeksi tai liitokseksi (Westphal, 2008, 70). Erilaisilla linkkityypeillä tai väreillä voidaan visualisoitava tieto jakaa eri ryhmiin, (Westphal, 2008, 137) tai niillä voidaan kuvata kahden olion välistä tunnetta, kuten sosio-grammin notaatiossa on määritelty (Moreno, 1934, 30).

Viivan paksuudella voidaan kuvata tietoa siitä, että esimerkiksi henkilön ja tietyn puhelinnumeron tai osoitteen välinen yhteys löytyy lähdetiedoista useita kertoja Westphal (2008, 164). Westphal (2008) käytti esimerkkinä pankin viranomaiselle tekemiä ilmoituksia epäilyttävistä liiketoimista, joissa ilmoituksen kohteena olevan henkilön pankille ilmoittamat osoite ja puhelinnumero pysyivät samoina. Visualisoinnissa henkilön ja yhteystietojen välille oli piirretty paksumpi viiva kuvaamaan usean, samaa yhteyttä kuvaavan linkin löytymistä lähdeaineistosta. Paksumman viivan olemassaolo kertoo viranomaiselle, että kohde on toistuvasti antanut samat yhteystiedot, eikä ole esimerkiksi yrittänyt aktiivisesti peittää jälkiään muuttamalla yhteystietoja tai niiden kirjoitusasia (Westphal, 2008, 164). Edellä kuvattu esimerkki kuvaa hyvin sitä, miten yksinkertaiset visuaaliset elementit voivat auttaa visuaalisen analyysin tekemisessä.

## 2.5 Yhteenveto

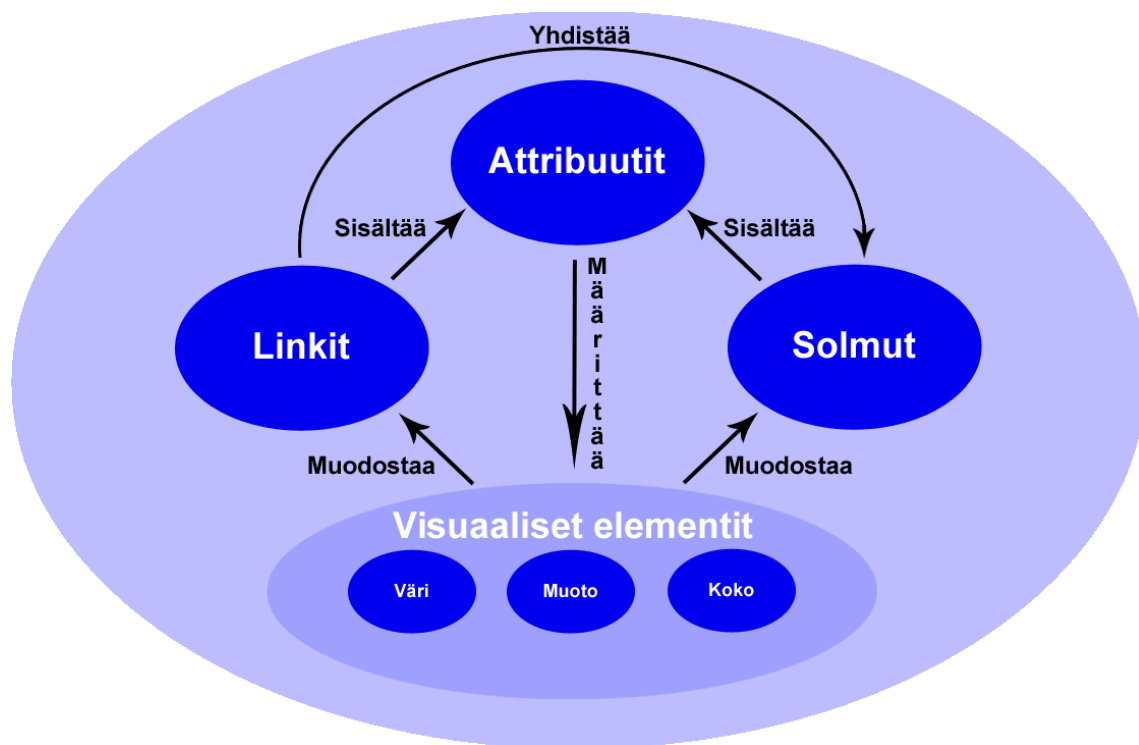
Kirjallisuuskatsauksessa käsiteltiin sosiaalisen verkon tiedon visualisointiin liittyviä asioita verkostomaisen tiedon visualisoinnin, sosiaalisten verkkojen analysoinnin, visuaalisen analyysin, sekä visuaalisten elementtien näkökulmasta. Kirjallisuuskatsauksessa käytettiin eri tieteenalojen kirjallisuuslähteitä. Lähteitä etsittiin pääasiallisesti Google Scholarilla, käyttäen seuraavia hakusanoja: "visual analytics", "visual analysis", "network visualization", "information visualization", sekä "social networks". Kirjallisuuskatsauksen tarkoituksena oli löytää niitä menetelmiä, joita käytetään verkostomaisen tiedon visuaalisessa analysoinnissa, sekä mahdollisesti löytää myös vihjeitä siihen millä toimenpiteillä verkostomaisen tiedon visualisoinnista voitaisiin tehdä helpommin analysoitavaa.

Visuaaliseen analysointiin sekä sosiaalisten verkkojen analysointiin liittyvää kirjallisuutta löytyy melko runsaasti. Sosiaalisten verkkojen analysoinnin tutkiminen on aloitettu jo 1930-luvulla, kun taas erityisesti visuaalisen analyysin tutkimukseen liittyvät varhaisimmat lähteet löytyivät 2000-luvun alulta. Verkostomaisen tiedon visualisointiin liittyvä kirjallisuus on luonteeltaan useimmiten hyvin teknistä, jonka myös Scott (2012, 1) on omassa sosiaalisen verkoston analyysiä käsittelevässä kirjassaan havainnut. Erilaisista visualisointiin liittyvistä algoritmeista löytyy paljon kirjallisuutta, mutta visualisointiin oleellisesti liittyvistä linkeistä ja solmuista, sekä niiden avulla esitettävän tiedon visualisoinnin eri tavoista, käyttäen avuksi visuaalisia elementtejä, löytyi kirjallisuutta vain suppeasti. Kattavin löytämäni visuaalisia elementtejä

selittävä ja selkeän sosiaalisen verkon visualisoinnin notaation sisältävä kirjallisuuslähde on Morenon jo vuonna 1934 kirjoittama kirja: "Who shall Survive?". Morenon teos on todella mielenkiintoinen ja sisältää suuren määrän esimerkkejä erilaisista henkilöiden ja ryhmien välisten yhteyksien visualisoinneista.

Valtaosassa kirjallisuutta visuaalisen analyysin soveltamista pidettiin tärkeänä, johtuen analysoitavan tiedon valtavasta kasvusta ja siitä, että visuaalinen analyysi on usein ainut keino analysoida monimutkaisia ja usein virheitä ja puutteita sisältäviä tapahtumaketjuja tai sosiaalisia verkostoja. Daniel Keim on onnistunut kuvaamaan niin omassa julkaisussaan kuin muiden kanssa yhdessä kirjoittamissaan artikkeleissa (Keim, 2002, Keim ym., 2006; Keim ym., 2008) todella hyvin visuaalisen analyysin prosessit ja siihen liittyvät tekniset haasteet.

Kirjallisuuskatsauksen perusteella laadittiin kuvion 5 mukainen visualisoinnin ontologia. Tämän tutkimuksen keskeisimpinä käsitteinä voidaan pitää solmuja ja niitä yhdistäviä linkkejä. Tiedon visualisoinnissa solmut ja linkit muodostuvat visuaalisista elementeistä, joita ovat väri, muoto ja koko. Solmut ja linkit sisältävät attribuutteja mutta niiden avulla voidaan myös määrittellä käytettävät visuaaliset elementit.



KUVIO 5 Verkkomaisen tiedon visualisoinnin ontologia

Kirjallisuuskatsauksella saadun teoreettisen taustan pohjalta voidaan todeta, että vaikka visuaalisen analyysin menetelmistä löytyy tutkittua tietoa, kirjallisuudesta ei kuitenkaan löydy selkeää mallia siihen miten visuaalisia elementtejä voidaan hyödyntää visualisointia selkiyttävinä ja analysointia helpottavina elementteinä. Kirjallisuuskatsauksen pohjalta voidaan myös sanoa, että yhtenä analysointia selkeästi vaikeuttavana tekijänä voidaan pitää linkkien ja solmujen

määrän kasvua. Edellisten seikkojen perusteella visuaalisen analysoinnin selkiyttämistä visuaalisia elementtejä apuna käyttäen voidaan pitää hyödyllisenä tutkimusaiheena.

### 3 KOHDEORGANISAATIO JA MUUTOSTARPEET

Tutkielman kohdeorganisaationa toimii eräs julkishallinnon yksikkö, jossa analysoidaan suuria, useista tietolähteistä saatavia, tietomassoja. Kohdeorganisaatiolla on ollut käytössään työkalut visuaalisen analyysin tekemiseen jo vuosikymmenen ajan ja järjestelmää on kehitetty pääosin yksikön omin voimavaroin. Viimeisen kahden vuoden aikana käsiteltävän tiedon määrä on kuitenkin moninkertaistunut, eikä muun työn ohella tehtävällä järjestelmän kehitystyöllä ole pystytty ratkaisemaan käsiteltävän tiedon kasvusta johtuvia ongelmia.

Kohdeorganisaatiossa käytetty visualisointiin liittyvä terminologia poikkeaa joiltain osin kirjallisuuskatsauksessa käytetystä terminologiasta. Yksittäiseen solmuun viitattaessa käytetään myös termejä ikoni tai kuvake.

#### 3.1 Visualisoinnissa käytettävät ohjelmistot

Tiedon visuaaliseen analysointiin käytetään organisaatiossa useita ohjelmistoja. Visuaalista analyysijärjestelmää rakennettaessa suurin osa työstä tehdään erilaisten tietokantojen ja rajapintojen parissa. Tutkielman sekä kohdeorganisaation loppukäyttäjien työn kannalta oleellisimpia ohjelmia ovat kuitenkin IBM:n i2 iBase, sekä Analyst's Notebook, joista jälkimmäinen toimii varsinaisena visuaalisen analyysin työkaluna ja edellinen sen lähdekantana.

Organisaatiossa käytettävän visuaalisen analysointiohjelmisto Analyst's Notebookin asetuksilla pystytään määrittämään mitä verkon linkkityyppejä ja solmuja laajennetaan oletuksena (engl. expand all), ja mitkä tiedot tulee valita erikseen laajennukseen (engl. expand with settings). Kyseinen ominaisuus on hyvin oleellinen tämän tutkielman kannalta, koska sillä voidaan rajoittaa visualisoitavan tiedon määrää, mutta tarvittaessa piilossa olevat linkit ja solmut saadaan helposti visualisoitua.



### **3.2 Vanhan visualisointijärjestelmän ongelmat**

Yhtenä vanhan visualisointijärjestelmän isoimmista ongelmista voidaan pitää suurta manuaalisen työn määrää varsinkin esianalyysivaiheessa. Johtopäätösten tekemiseen tarvittavat tiedot pystytään kyllä visualisoimaan nopeasti näytölle, mutta oleellinen tieto saattaa hukkaa tuhansien solmujen ja linkkien viidaksoon. Ajan saatossa monessa vaiheessa uudistettua vanhaa järjestelmää pidetään kehittäjien mielestä myös vaikeasti ylläpidettävänä, mikä on hankaloittanut ja siten myös osaltaan hidastanut tarvittavien muutosten tekemistä järjestelmään.

## 4 UUDEN VISUALISOINTIJÄRJESTELMÄN LUOMINEN KONSTRUKTIIVISEN TOIMINTATUTKIMUKSEN MENETELMIN

Tutkielman empiirinen osio suoritettiin konstruktivisena toimintatutkimuksena kohdeorganisaatiolle. Tutkielman viitekehystenä käytettiin Peffersin, Tuunasan, Rothenbergin ja Chatterjeen (2007) suunnittelututkimuksen metodin (engl. The design science research methodology) prosessimallin sekä Seinin, Henfridssonin, Puraon, Rossin ja Lindgrenin (2011) metodin yhdistelmää. Yhdistelmään päädyttiin, koska Peffersin ym. (2007) prosessimalli sopii hyvin konstruktivisen tutkimuksen tekemiseen, mutta koska tutkimus tehtiin konstruktivisena toimintatutkimuksena, jossa arvionti ei ole oma erillinen vaiheensa, vaan sitä tehdään jatkuvasti, lisättiin siihen Seinin ym. (2011) jatkuvan arvionnin periaate. Peffersin ym. (2007) prosessimalli sisältää seuraavat kuusi aktiviteettia:

1. Ongelman tunnistaminen ja motivointi.
2. Ratkaisun tavoitteiden määrittely.
3. Suunnittele ja toteuta.
4. Demonstrointi.
5. Arviointi.
6. Kommunikointi.

Ensimmäisen aktiviteetin tarkoitus on määritellä tutkimusongelma ja perustella ratkaisulla saavutettava arvo. Toisen aktiviteetin tarkoitus on päätellä tavoitteet tutkimusongelman määritelmästä ja siitä, mikä on mahdollista ja toteuttamiskelpoista. Kolmantena suunnitellaan ja toteutetaan artefakti. Neljäntenä demonstroidaan artefaktin käyttöä ja ratkaistaan yksi tai useampia ongelmia. Viidentenä aktiviteettina havainnoidaan ja mitataan miten hyvin tuotettu artefakti tukee ongelman ratkaisua. Kuudes aktiviteetti liittyy tutkimuksen julkistamiseen alan tutkijoille ja asiantuntijoille. Julkaisussa kuvataan ongelma ja sen tärkeys, luotu artefakti sekä sen hyödyllisyys ja uutuus. Valitun viitekehysten mukaan myös artefaktin suunnittelun

täsmällisyys ja hyödyt tulee kuvata muille tutkijoille sekä esimerkiksi kyseisen alan käytännön ammattilaisille. (Peffer ym., 2007.) Pefferin ym. (2007) prosessimalliin kuuluu myös iteraatio, jossa arviointi- ja kommunikointivaiheesta voidaan siirtyä tarvittaessa takaisin ratkaisun tavoitteiden määrittelyyn tai suunnittelu- ja toteutusvaiheeseen.

Ongelman tunnistamista ja motivointia tehtiin johdannossa sekä kappaleessa 3.2, jossa käsiteltiin vanhan järjestelmän ongelmia. Seuraavissa alakappaleissa käydään läpi sitä, miten kohdeorganisaation uuden visuaalisen analyysijärjestelmän järjestelmänkehitysprojekti suoritettiin noudattaen valittua viitekehystä.

## **4.1 Ratkaisun tavoitteet**

Tutkielman tarkoituksena oli löytää niitä keinoja, joilla laajoista, verkostomaisista tietomassoista voidaan tehdä selkeämpiä ja siten myös helpommin analysoitavia. Kohdeorganisaation kannalta ensisijaisena tavoitteena oli tehdä uudesta visuaalisen analyysin järjestelmästä edellistä järjestelmää selkeämpi ja paremmin analyysia tukeva vähentämällä muun muassa visualisoitavien linkkien ja solmujen määrää. Ratkaisusta pyrittiin tekemään yhdessä kohdeorganisaation kanssa mahdollisimman helposti ylläpidettävä ja alle kahdessa kuukaudessa toteutettava kokonaisuus. Yhtenä päätavoitteena voidaan pitää myös kohdeorganisaatiossa kehitetyn sektorimallin idean toimivuuden testaamista. Tutkimuksesta ja sen tuloksista toivotaan olevan apua myös muille visuaalista analyysia tekeville organisaatioille järjestelmien kehitysprojekteissa.

## **4.2 Suunnittele ja toteuta**

Uusi visualisointijärjestelmä suunniteltiin vanhan järjestelmän pohjalta, yhteistyössä kohdeorganisaation työntekijöiden kanssa, iteraatiivisesti ketteriä sovelluskehityksen menetelmiä käyttäen. Ennen toteuttamisvaiheen aloittamista tehty ominaisuuksien suunnittelu pidettiin melko yleisellä tasolla, koska uusia ominaisuuksia suunniteltiin myös vaiheittain toteutusjaksojen välissä. Poikkeuksena voidaan kuitenkin pitää sektorimallia, jonka suunnittelu oli aloitettu yhdessä kohdeorganisaation kanssa jo ennen tutkielman teon alkua ja jonka ensimmäinen versio saatiin valmiiksi, juuri ennen varsinaista ketterien menetelmien vaiheen alkua.

Suunnittelua tehtiin pääasiallisesti pienryhmäpalavereissa ja tietokantoihin sekä visualisointisovellukseen tehtävät muutokset toteutettiin pääsääntöisesti itsenäisenä työskentelynä. Toteutus tehtiin käyttäen kehitysmenetelmää, joka perustui nopeisiin iteraatioihin ja iteraatioiden välissä oleviin suunnittelupalavereihin, joissa edellisen version tai työvaiheen tulokset

testattiin ja analysoitiin kehittäjien yhteistyönä sekä kirjattiin muistiin vaadittavat muutos- ja korjaustoimenpiteet. Loppukäyttäjät olivat suunnittelupalavereissa pääroolissa. Suurin osa suunniteltujen ominaisuuksien toteutuksesta liittyi tietokantoihin tehtävien uusien näkymien ja taulujen luontiin, sekä muihin tietokantoihin liittyviin töihin. Tämä vaihe eroaa Peffersin ym. (2007) metodista siten, että mukaan on otettu Seinin ym. (2011) metodin mukaisesti toimintatutkimukseen kuuluvat toistuvat testaamisen ja arvionnin vaiheet.

#### **4.2.1 Tiedonhankinta**

Organisaatissa työskenteleviä järjestelmän loppukäyttäjää pyydettiin noin kuukautta ennen toteutusvaihetta keräämään tietoja ongelmakohtista ja puutteista, joita nykyinen visuaalisen analyysin järjestelmä sisältää, ja antamaan mahdollisia korjausehdotuksia ongelmiin. Ikävä kyllä loppukäyttäjät eivät olleet koonneet pyydettyä listaa, vaan ongelmakohtia ja vanhasta järjestelmästä löytyviä puutteita jouduttiin keräämään kehitystyön aikana keskustelemalla loppukäyttäjien kanssa ja järjestämällä ylimääräisiä työpalavereja, joissa asioita pohdittiin ryhmässä.

Ennen varsinaisen järjestelmäkehityksen aloittamista pidettiin myös yksi pidempi palaveri, jossa esiteltiin tutkimuksen tavoitteita ja tutkimusmetodia sekä keskusteltiin niistä asioista, joilla voitaisiin vähentää linkkien ja solmujen määrää. Palaveriin osallistui tutkijan lisäksi useita vanhan visuaalisen analyysijärjestelmän aktiivisia loppukäyttäjää.

Toteutusvaiheessa jokaisen iteraation alussa kerättiin tietoa edellisen vaiheen onnistumisesta ja siitä, miten toteutettua järjestelmän osaa tulisi vielä kehittää, jotta se palvelisi mahdollisimman hyvin tarkoitustaan.

#### **4.2.2 Sovelletut ketterät menetelmät**

Ensimmäisen kahden viikon aikana rakennettiin uuden visualisointijärjestelmän pohja käyttäen apuna vanhan järjestelmän rakennetta niin, että sitä täydennettiin jo tässä vaiheessa jonkin verran loppukäyttäjiltä kerättyjen korjausehdotusten pohjalta. Ensimmäisen kahden viikon aikana suunniteltiin ja toteutettiin myös ensimmäinen versio sektorimallista. Seuraavina neljänä viikkona järjestelmää kehitettiin inkrementaalisesti viikon iteraatioissa siten, että maanantaina pidetyssä tunnin palaverissa tutkija esitteli edellisen viikon aikaansaannokset. Mukana olevat kahdeksan järjestelmän loppukäyttäjää hyväksyivät, hylkäsivät, tai pyysivät korjauksia ominaisuuksiin. Erimielisyyksien syntyessä turvauduttiin äänestykseen. Seuraavana vaiheena palaverissa valittiin ne uudet ominaisuudet, jotka oli tarkoitus toteuttaa kuluvan viikon aikana. Kokouksen lopuksi valittiin loppukäyttäjien joukosta kokenut visuaalisen analyysijärjestelmän käyttäjä beta-testaajaksi. Aikaa seuraavan järjestelmän version toteuttamiseen sovituin ominaisuuksin sekä alfa-testauksen suorittamiseen oli tutkijalla maanantain iltapäivästä kyseisen viikon torstai-iltaan. Viikoittain beta-testaukseen valitulle yhdelle

loppukäyttäjälle annettiin aikaa perjantaihin klo 13.00 asti testata alkuviikon aikaansaannokset. Tutkija teki tiivistelmän viikottaisista beta-testausraporteista, joita käytettiin seuraavan maanantain palaverissa esiteltäessä edellisen viikon tuotoksia. Testausraportit auttoivat myös kehityskohteiden löytämisessä kyseisellä viikolla tuotettuihin uusiin ominaisuuksiin.

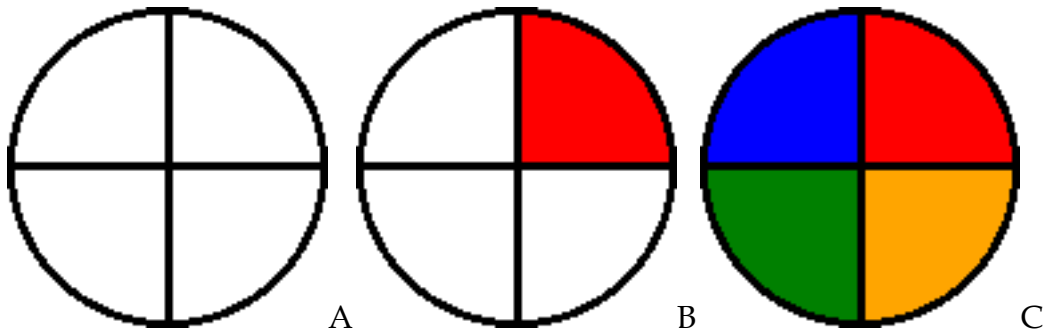
### 4.3 Uudet ominaisuudet

Tutkielmassa kohdeorganisaation visuaaliseen analyysijärjestelmään toteutettiin useita uusia ominaisuuksia. Osa järjestelmään luoduista uusista ominaisuuksista toteutettiin, jotta järjestelmä olisi helpommin ylläpidettävä ja päivitettävä, mutta loppukäyttäjille tärkeimpiä ovat visualisoinnissa näkyvät uudet ominaisuudet.

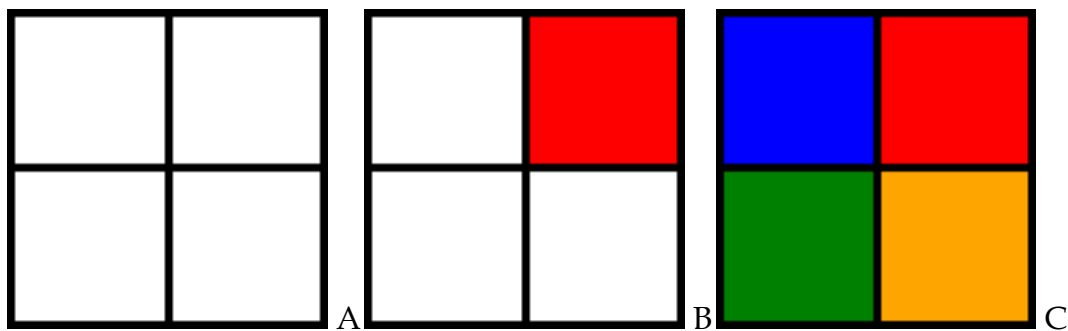
Tämän luvun alaluvuissa esitellään toteutetuista visuaalisista ominaisuuksista tärkeimmät sekä kuvataan niillä tavoiteltuja hyötyjä. Tekstiä selkiyttävät esimerkkikuviot on lainattu uudesta visualisointijärjestelmästä, mutta kuvioiden sisältämä tekstitieto on anonymisoitu ja muokattu joiltakin osin kuvaavampaan muotoon.

#### 4.3.1 Sektorimalli

Näkyvin muutos aikaisempaan visualisointiratkaisuun oli sektorimallin käyttöönotto henkilöitä ja yhtiöitä kuvaavissa ikoneissa. Aikaisemmassa visualisointitavassa kaikkia henkilöitä kuvattiin samanlaisella ikonilla. Sektorimallissa henkilöä kuvaava pyöreä ikoni (kuvio 6) on jaettu neljään samankokoiseen lohkoon, joista kukin lohko kuvaa tiettyä ristiinvertailurekisteriä siten, että jos kyseinen lohko on väriltään jokin muu kuin valkoinen, löytyy kohteelle jotain tietoa kyseisestä rekisteristä. Ristiinvertailurekisterillä tarkoitetaan siis rekisteriä, jonka tiettyjä tietoja on vertailtu ristiin kohderekisterin tietojen kanssa. Neljää lohkoa käytettäessä erilaisia ikoneja voidaan muodostaa maksimissaan kuusitoista kappaletta. Yhtiöiden osalta sektorimalli toimii samalla periaatteella, sillä erotuksella, ikoni on neliön muotoinen (kuvio 7). Mallin värit on valittu väriympyrää hyväksi apuna käyttäen niin, että vastakkaisissa lohkoissa olevat värit ovat toistensa vastavärejä. Valitsemalla vastavärit pyrittiin kontrastin avulla erottamaan sektorit selkeästi toisistaan (Stone, 2006). Lohkot eroteltiin toisistaan lisäksi selkeällä mustalla ristikolla, jotta kokonaan valkoiset mallia noudattavat ikonit erottuisivat tavallisista, visualisoinnissa mahdollisesti käytettävistä, pallon ja neliön muotoisista ikoneista.



KUVIO 6 Sektorimallin henkilöä kuvaavia ikoneita



KUVIO 7 Sektorimallin yhtiötä kuvaavia ikoneita

Kuvioiden 6 ja 7 esittämät ikonit kuvaavat seuraavanlaisia henkilöitä ja yhtiöitä:

- a) Kohteesta ei löydy tietoja tarkasteltavista ristiinvertailurekistereistä,
- b) kohteelle löytyy yksi tai useampi merkintä yhdestä tarkasteltavasta ristiinvertailurekisteristä ja
- c) kohteelle löytyy tietoa kaikista tarkasteltavista ristiinvertailureistereistä.

Sektorimallin ideana on se, että pystytään toteamaan nopeasti, yhdellä vilkaisulla, minkä tyyppistä tietoa henkilöön tai yhtiöön liittyy. Usein halutaan myös tietää tarkemmin minkälaista ja kuinka vanhaa tietoa kohteena olevat rekisterit pitävät sisällään. Tätä tarkoitusta varten jokaiselle henkilölle ja yhtiölle lisättiin ”rekisteritarkistukset”-niminen kenttä, joka sisältää taustarekistereistä löydetyt tiedot tiivistelmänä (kuvio 8). Kyseisen tiedon saa näkyviin kaksoisklikkaamalla kohteen ikonia. Tiivistelmässä lasketaan muun muassa rekisterikohtaisten osumien määrät, sekä esitetään yksittäisten osumien tärkeimmät tiedot jokaisen rekisterin osalta listana siten, että oletettavasti tärkeimmät osumat näkyvät listan alussa. Tiivistelmän on tarkoitus auttaa koneellisesti kohteen profiloinnissa sekä nopeuttaa analyytikon työskentelyä. Ennen vastaava lista koottiin usein manuaalisesti alustavan analyysin vaiheessa. Sama taustarekisteritieto löytyy tarvittaessa myös kohteeseen linkitettyinä tietona.

**RISTIINAJON TULOKSET:**

Rekisteri 1: 1kpl.  
 Rekisteri 2: 3kpl.  
 Rekisteri 3: 2kpl.  
 Rekisteri 4: 4kpl., (1 xx).

**Rekisteri 1 osumat:**

2015/...

**Rekisteri 2 osumat:**

2014/...

2013/...

2013/...

**Rekisteri 3 osumat:**

2015/...

2010/...

**Rekisteri 4 osumat:**

2015/... xx...

2015/...

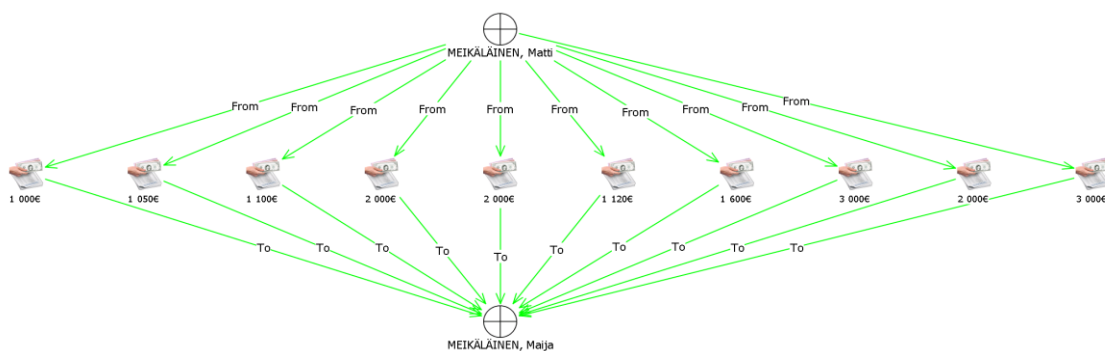
2010...

2007...

KUVIO 8 Ristiinvertailutulokset listana

#### 4.3.2 Transaktioita kuvaavan ikonin korvaaminen linkillä

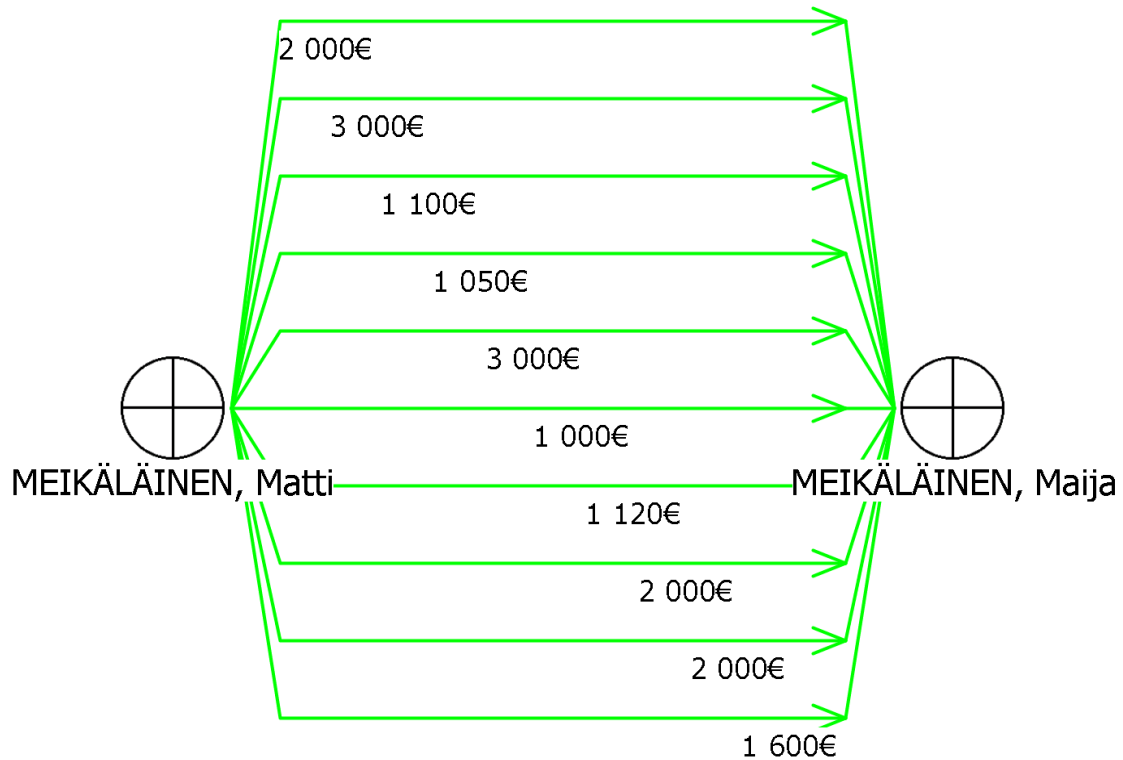
Transaktio-tyyppinen tieto, joka kuvaa esimerkiksi rahan liikettä pankkitililtä toiselle pankkitilille kuvattiin aikaisemmin yhden ikonin ja kahden linkin avulla. Jo pelkästään kahden henkilön välinen, kymmenen transaktiota sisältävä, rahaliikenteen visualisointi toi kuvaan suuren määrän linkkejä ja ikoneja (kuvio 9).



KUVIO 9 Transaktion kuvaaminen ikonin avulla

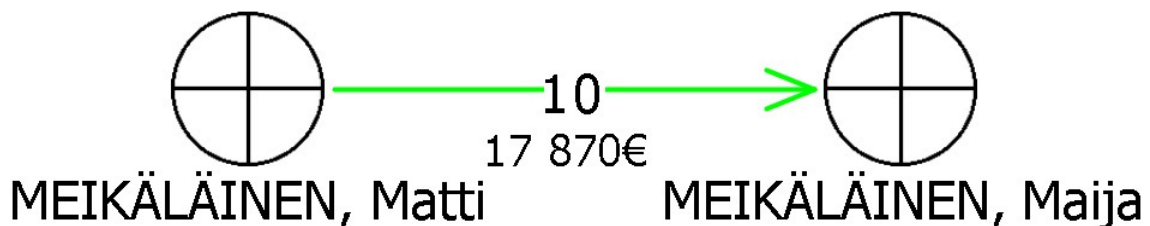
Suurta linkkien ja ikonien määrää päätettiin lähteä vähentämään muuttamalla transaktiotiedon esitystapaa niin, että ikonit korvattiin osapuolten välisillä linkeillä ja nuolen suunnalla kuvataan rahavirran suuntaa. Visualisoinnissa käytetty i2 iBase:n ja Analyst's Notebookin yhdistelmä mahdollistaa transaktio-

tyyppisten linkkien esittämisen niin, että jokainen linkki visualisoidaan omana nuoli-tyyppisenä linkkinään (kuvio 10).



KUVIO 10 Transaktion kuvaaminen yksittäisinä linkkeinä

Visualisointi voidaan määritellä analyytikon toimesta myös siten, että kaikki yhteen suuntaan liikkuvat rahavirrat lasketaan yhteen, ja transaktioiden lukumäärä kuvataan rahasumman yläpuolella lukuna (kuvio 11).



KUVIO 11 Yhdensuuntaisten transaktioiden kuvaaminen yhtenä linkkinä osapuolten välillä

### 4.3.3 Visualisoitavat attribuutit

Uuteen visualisointiratkaisuun lisättiin joitakin visualisoitavia attribuutteja. Attribuuttien tarkoituksena on tuoda jokin tärkeä tieto, kuten henkilön nimi (kuvio 12), helposti ja nopeasti havaittavaksi ilman, että sitä tarvitsee hakea kohteeseen linkitetystä tiedosta tai kohteen piilossa olevista attribuuteista, joihin pääsee käsiksi kaksoisklikkaamalla ikonia tai linkkiä.



Järjestelmässä on esimerkiksi useita tapauksia, joissa lähdetietokannasta löytyy useita samoja entiteettejä, esimerkiksi henkilöä tai yhtiötä, kuvaavia toisteisia ilmentymiä. Toisin sanoen sama yhtiö tai henkilö löytyy tietojärjestelmästä useamman kuin yhden kerran. Projektin aikana havaittiin, että tätä ongelmaa ei aina muisteta, mikä saattaa johtaa siihen, että kaikkea analyysin kannalta tärkeää tietoa ei ole mukana analyysissä.

Jotta toisteiset arvot eivät unohtuisi, lisättiin jokaiselle henkilölle ja yhtiölle niin sanottu ”kaimatieto”. Kaimatieto tuotiin esille visualisoitavana attribuutina, joka kertoo kaimojen lukumäärän. Kuviosta 12 löytyy uudesta järjestelmästä otettu esimerkki henkilöstä, jolle löytyy neljä kappaletta nimikaimoja. Ominaisuutta toteutettaessa arveltiin, että kaimatietoa voitaisiin käyttää myös apuna siivottaessa toisteisia arvoja päärekisteristä.



KUVIO 12 Henkilö neljällä kaimalla

#### 4.4 Demonstrointi

Demonstrointi voidaan toteuttaa Peffersin ym. (2007) mukaan esimerkiksi suorittamalla kokeita, tapaustutkimuksella tai muulla sopivalla toiminnalla. Tässä tutkielmassa uuden tyyppistä visualisointitapaa demonstroitiin ja testattiin loppukäyttäjien toimesta tuotantodatalla siten, että he hyödynsivät järjestelmää normaalissa päivittäisessä työssään. Koekäyttäjille annettiin aluksi yhteinen koulutus järjestelmän käyttöön. Uusista ominaisuuksista koostettiin pikaopas ja varsinkin alkuvaiheessa käyttäjiä opastettiin myös henkilökohtaisesti. Henkilökohtaisen opastuksen aikana havaitut asiat kirjattiin ylös joko heti tai välittömästi opastuksen päätyttyä. Myös loppukäyttäjää pyydettiin kirjaamaan ylös testivaiheessa havaittuja asioita.

#### 4.5 Tiedonhankinta ja arviointi

Arvioinnissa testattiin uuden artefaktin tehokkuutta ja vaikuttavuutta toimintaympäristöön ja käyttäjiin (March & Smith, 1995) sekä yleisesti sitä miten artefakti tukee asetettuja tavoitteita (Peffers ym., 2007). Vanhaa ja uutta

visualisointitapaa vertailtiin linkkien ja solmujen määrien osalta kvantitatiivisesti, käyttäen kummassakin visualisointitavassa samaa lähdeaineistoa. Uuden visualisoinnin arviointi perustui käyttöympäristössä tehtyihin havaintoihin sekä käyttökokemuksiin, jotka kerättiin haastattelemalla loppukäyttäjii. Loppukäyttäjien työtä ei haluttu häiritä liikaa, joten esimerkiksi suoritusajkoja ei mitattu.

#### **4.5.1 Haastattelun ja kyselylomakkeen yhdistelmä**

Haastattelu toteutettiin teemahaastatteluna, jolla pyrittiin keräämään loppukäyttäjien ajatuksia, kokemuksia ja uskomuksia (Järvinen & Järvinen, 2012, 185) liittyen yksikön visualisointijärjestelmään. Haastatteluun valittiin mukaan viisi kahdeksasta uuden visuaalisen analyysijärjestelmän version kehittämiseen osallistuneesta henkilöstä. Jokaisella valitulla henkilöllä oli usean vuoden käyttökokemus aiemmasta versiosta. Ennen haastattelua haastateltavat olivat ehtineet käyttää uutta järjestelmän versiota kahdesta kuuteen viikkoa. Haastatelluista loppukäyttäjistä kaksi oli niin sanottua esianalyysiä tekeviä, toiset kaksi varsinaista syvempää analyysiä tekeviä ja viidenneksi haastateltavaksi valittiin yksi analyytikkojen esimiehistä, jonka toimenkuvaan kuuluvat vahvasti myös tiedon analysointiin liittyvät tehtävät. Haastattelut kestivät puolesta tunnista vajaaseen puoleentoista tuntiin.

Haastattelut toteutettiin puolistrukturisena teemahaastatteluna. Kerätty data oli pääsääntöisesti kvalitatiivista, mutta toteutetun artefaktin onnistumisen tasoa mitattiin myös haastattelun loppuun sijoitetuilla kyselylomakkeen kysymyksillä. Kysymyslistan tarkoituksena oli saada kerättyä oleellinen tieto myös tiivistetyssä muodossa (Hirsjärvi & Hurme, 2014, 66). Kahdella rinnakkaisella menetelmällä kerättyä aineistoa käytettiin myös reliabiliteetin todentamiseen siten, että jos molempien tutkimusmentelmien tulokset ovat samat, voidaan aineistoa pitää luotettavampana kuin vain yhtä tutkimusmenetelmää käytettäessä (Hirsjärvi & Hurme, 2014, 186). Vastausvaihtoehtoja annettiin 5-portaisella Likert-asteikolla (Likert, 1932, 14) 1:stä (täysin eri mieltä) 5:een (täysin samaa mieltä), johon lisättiin vielä kuudes vaihtoehto: "en osaa sanoa", jotta voitaisiin erottaa neutraali suhtautuminen ja osaamattomuus vastata kysymykseen. Haastateltavilla oli myös mahdollisuus kommentoida lomakekysymyksiin valitsemiaan vastausvaihtoehtoja. Haastatteluiden aineistoa käsiteltiin luottamuksellisena tietona ja julkaistavat johtopäätökset, lainaukset ja tulokset tarkistettiin kohdeorganisaation taholta, ennen tutkimuksen julkaisua.

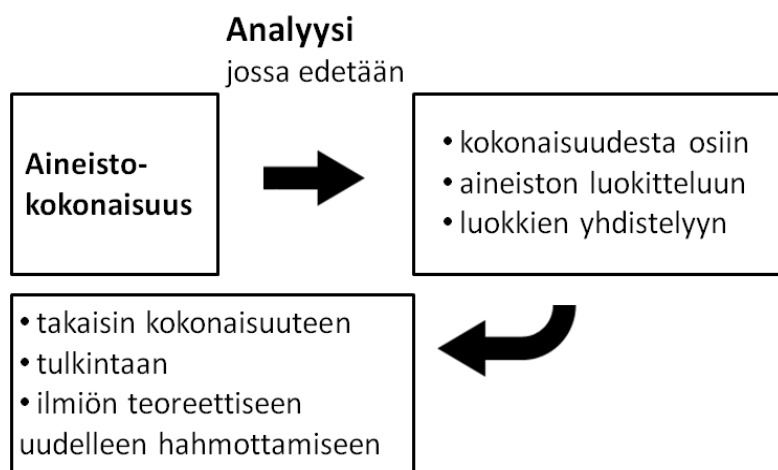
#### **4.5.2 Havainnointi**

Tutkimuksessa käytettiin myös osallistuvaa havainnointia, jossa tutkija osallistui tutkittavaan toimintaan ja jossa tutkimukseen osallistuneille ilmoitettiin havaintojen tekemisestä heti tutkimuksen alettua (Järvinen & Järvinen, 2012, 185). Havainnointia käytettiin, koska pelkät haastattelut eivät kerro mitä todella tapahtuu (Hirsjärvi, Remes & Sajavaara, 1997), mutta

havainnoimalla pystytään selvittämään sitä mitä ihmiset tekevät ja miten he toimivat eri julkisissa tilanteissa (Järvinen & Järvinen, 2012, 185). Havainnot kirjoitettiin muistiin ja niitä verrattiin haastatteluissa kerrottuihin asioihin. Näin voitiin esimerkiksi tarkistaa tekevätkö ja ajattelevatko haastateltavat samoin kuin haastatteluissa sanoivat (Järvinen & Järvinen, 2012, 151).

### 4.5.3 Aineiston purku ja tulosten analysointi

Haastattelut litteroitiin tietokoneella niiltä osin kuin ne käsittelivät valittuja teemoja (Hirsjärvi & Hurme, 2014, 141). Seuraavaksi haastattelut käytiin läpi yksitellen, ja aineisto purettiin vielä teemoittain yhteen Excel-tilaukseen, jokainen teema omalle välilehdelle. Haastatteluaineiston analysoinnissa käytettiin Hirsjärven ja Hurmeen mallia, jossa ensin eritellään ja luokitellaan aineistoa, jonka jälkeen synteessin avulla luodaan kokonaiskuva (kuvio 13).



KUVIO 13 Haastatteluaineiston käsittely analyysistä synteessiin (mukautettu Hirsjärvi & Hurme, 2014, 144)

Aineiston analysointitapoina käytettiin pääasiassa laskemista, asteikointia, sekä teemoittelua (Hirsjärvi & Hurme, 2014, 171–173). Analysoinnissa pyrittiin tarkastelemaan myös yhteyksiä esiin nousevien seikkojen osalta. Yhteyksiä tarkasteltiin tyypittelemällä vastauksia ja jakamalla haastateltavia niiden perusteella ryhmiin. Syntyneitä ryhmiä verrattiin esimerkiksi haastateltavien työtehtäviin.

Haastattelun lopuksi jokaisen haastateltavan kanssa täytetyn kyselylomakkeen vastauksia vertailtiin haastatteluissa saatuihin vastauksiin. Jos haastattelun vastauksessa ja kyselylomakkeen vastauksessa oli ristiriitaisuuksia, pyrittiin ristiriitaisuuteen löytämään syy esimerkiksi haastateltavan lomakkeen täytön yhteydessä antamista kommentteista. Varsinaisilla teemahaastatteluilla kerättyä aineistoa jouduttiin täydentämään analyysivaiheessa kahden henkilön puhelinhaastatteluilla.

## 5 TULOKSET JA JOHTOPÄÄTÖKSET

Tässä kappaleessa tarkastellaan sitä miten hyvin kohdeorganisaation uusi visuaalisen analyysijärjestelmän versio tukee sille asetettuja tavoitteita. Analysoitava aineisto on pääosin peräisin haastatteluista, mutta analysoinnissa on käytetty apuna myös järjestelmänkehitysprojektin aikana tehtyjä havaintoja sekä järjestelmästä kerättyjä tietoja linkkien ja solmujen määristä. Erityisesti aineistosta pyrittiin löytämään tutkimukseen osallistuneiden vastauksia siihen, mitkä toimenpiteet tekivät verkostomaisen tiedon analysoinnista helpompaa. Haasteluiden lainauksista poisjätetyt kohdat on merkitty kahdella ajatusviivalla, joiden välissä on välilyönti, eli " - " ja lisäykset ja muutokset on merkitty hakasulkeiden sisään, esimerkiksi [taustarekisterin nimi].

### 5.1 Järjestelmän käytön helppous suhteessa vanhaan

Kun haastateltavia pyydettiin kertomaan miltä uutta järjestelmää tuntui käyttää, tuli kolmen haastateltavan kommentteista selkeästi esiin vanhoista toimintatavoista pois oppimisen vaikeus. Haasteltavien kommentteja olivat esimerkiksi: "Mun mielestä tämä on vähän sekavan oloinen" ja "Sanotaan, että kun tulee tällainen iso juttu, niin se on vähän semmoista tottumista". Toisaalta kaksi haastateltavaa vastasi seuraavasti: "Helpottaa käyttöä" ja "Siinä sitten näkee paljon helpommin sen kokonaisuuden ja sen ketkä on niitä, joita kannattaa läheteä tarkemmin katsomaan. Kyllä se on paljon selkeämpi siinä mielessä".

Alustavaa analyysiä tekevät haastateltavat kokivat uuden järjestelmän muita haastateltavia helppokäyttöisempänä. Kolme viidestä mainitsi myös sektorimallin käyttöä helpottavana ominaisuutena, vaikka kaksi heistä jäikin hiukan kaipaamaan vanhasta järjestelmästä sitä, että ennen ikonin värillä pystyttiin osoittamaan henkilön sukupuoli.

## 5.2 Sektorimalli

Kuten jo sektorimallia esiteltäessä mainittiin, sen perimmäisenä ideana on, että analyytikko pystyy näkemään yhdellä vilkaisulla sen, mistä valituista neljästä taustarekisteristä kyseiseen henkilöön tai yhtiöön löytyy tietoa. Malliin kuuluu myös rekisteritarkistukset-kenttä, jonka tarkoituksena on antaa tiivistettyä tietoa siitä, kuinka paljon, kuinka vanhaa, missä tilassa olevaa ja minkä tyyppistä tietoa kukin rekisteri kohteesta sisältää.

Kun haastateltavilta kysyttiin kuinka usein he joutuvat laajentamaan sektorimallin avulla piilotetut linkit näkyviksi, erottuivat alustavaa analyysiä tekevien vastaukset varsinaista analyysiä tekevien vastauksista siinä, että heille vaikutti olevan selkeämpää missä tapauksessa heidän tarvitsi laajentaa verkkoa taustarekisterin tiedoilla ja milloin ei. Useissa tapauksissa sektorimallin sisältämää tietoa pidettiin riittävänä alustavien johtopäätösten tekemiseen. Vaikka taustarekisterien tietoja joudutaankin vielä tuomaan mukaan visualisointiin linkkeinä ja ikoneina, pidettiin sektorimallin linkkejä vähentävää vaikutusta hyvänä:

[– –]ensimmäinen expandaus ei koskaan enää räjähdä, siinä sitten näkee paljon helpommin sen kokonaisuuden ja sen ketkä on niitä, joita kannattaa lähteä tarkemmin kattomaan. Kyl se on niinku paljon selkeämpi siinä mielessä.

Sektorimalliin kuuluvan, taustarekistereistä kootun tiivistelmätiedon sisältöön oltiin yleisesti tyytyväisiä. Eniten siitä katsottiin olevan hyötyä alustavaa analyysiä tekeville loppukäyttäjille, jotka kopioivat tiivistelmätiedon ja täydentävät sitä tapauksesta riippuen erilaisilla lisätiedoilla, ennen kuin siirtävät tiedon varsinaista analyysiä tekevien hyödynnettäväksi toiseen järjestelmään. Sektorimalliin valittujen neljän rekisterin rekisteritietojen lisäksi yksi haastateltava kaipasi tiivistelmätietoon mukaan myös yksikön päärekisterin tietoja listana, kommentoiden tiivistelmätietoa seuraavasti: ”Se on täydellinen, paitsi, että siitä puuttuu ne [rekisterin nimi]-historiat.”

Kuten taulukosta 1 voidaan havaita, haastattelun lomakeosiossa kaikki haastateltavat olivat yksimielisiä siitä, että sektorimalli nopeuttaa alustavien johtopäätösten tekemistä. Sektorimallin hyödyllisyyden tasosta koko analyysiprosessin ajan ei oltu yksimielisiä. Kaksi viidestä oli täysin samaa mieltä, että sektorimallista on hyötyä koko analyysiprosessin ajan, toiset kaksi olivat väitämästä jokseenkin samaa mieltä ja yksi valitsi neutraalin kannan. Huomioitavaa oli kuitenkin se, että täysin samaa mieltä olevista toinen edusti varsinaista analyysiä, ja toinen alustavaa analyysiä tekevää loppukäyttäjää. (taulukko 1.)

TAULUKKO 1 Sektorimallin toimivuus

Väittäjä	Vastaukset
Sektorimalli nopeuttaa alustavien johtopäätösten tekemistä.	5 x täysin samaa mieltä.
Sektorimallista on hyötyä koko analyysiprosessin ajan.	2 x täysin samaa mieltä, 2 x jokseenkin samaa mieltä, 1 x neutraali kanta.
Sektorimalli nopeuttaa kokonaisuuden hahmottamista.	2 x täysin samaa mieltä, 3 x jokseenkin samaa mieltä.

Tehtyjen laskelmien perusteella sektorimallin avulla pystyttiin piilottamaan visualisoinnista yhteensä noin 10 prosenttia järjestelmän kaikista linkeistä ja 13,6 prosenttia järjestelmän kaikista solmuista. Jos vanhanmalliset transaktiotapahtumat, jotka jätettiin myös uuteen visualisointiin, päätetään poistaa tarpeettomina, sektorimallin avulla piilotettavien tapahtumien osuus kasvaisi 32,3 prosenttiin ja linkkien määrä 17,7 prosenttiin suhteessa kaikkiin solmuihin. Sektorimallin vaikutus linkkien ja solmujen määrään kuitenkin vaihtelee laajasti analyysin kohteesta ja tehtävän analyysin tyypistä johtuen, joten yksittäisissä tapauksissa sektorimallin linkejä ja solmuja piilottava vaikutus voi vaihdella hyvin suuresta lähes olemattomaan.

### 5.3 Transaktioiden esittäminen suorina linkeinä

Projektissa transaktiityyppisen tiedon visualisointi muutettiin yhden ikonin ja kahden linkin (kuvio 9) mallista vain yhden, osapuolten välisen, linkin malliin (kuviot 10 & 11). Uudenlainen visualisointitapa oli aiheuttanut hämmennystä kahdelle haastatellulle, eikä ominaisuutta ollut osattu käyttää oikein. Osaltaan asiaa oli sotkenut vielä se, että vanhanmallisen visualisointitavan linkit ja ikonit oli jätetty järjestelmään käyttäjien toivomuksesta, vaikka ne oli otettu oletuslaajennusasetuksista pois.

Vaikka uuden visualisointitavan käyttöönotossa oli ollut ongelmia, pitivät alustavaa analyysiä tekevät loppukäyttäjät uutta tapaa analyysiä helpottavana tekijänä: "Se helpottaa kyllä, varsinkin esikäsittelyvaiheessa se helpottaa sitä". Kun haastateltavilta kysyttiin onko heillä vielä tilanteita, joissa he haluaisivat visualisoida transaktiotapahtumat vanhan mallin mukaan, vastasi yksi viidestä törmänneensä kerran tällaiseen tilanteeseen. Hän koki, että vanhalla mallilla tapahtumakokonaisuuteen liittyvien yksittäisten transaktioiden tapahtuma-ajankohdat pystyi visualisoimaan selkeämmin taulukkomuodossa. Myös uudessa analyysijärjestelmässä tapahtuma-ajankohdat pystytään visualisoimaan kuvaan sekä esittämään taulukkomuodossa, mutta ongelmia tuotti se, että taulukkomuotoinen esitystapa tuotetaan uudessa järjestelmässä eri tavalla.

Kyselylomakkeella haluttiin lisäksi kartoittaa haasteltavien mielipidettä siihen, voidaanko vanhanmallinen transaktioiden visualisointi poistaa kokonaan järjestelmästä. Kysymys jakoi mielipiteet selkeästi kahtia, eikä erilaisilla

työtehtävillä ollut vaikutusta vastauksiin. Kaksi viidestä oli täysin sitä mieltä, että vanhanmallisesta transaktioiden visualisoinnista voitaisiin luopua, kun taas kaksi viidestä oli jokseenkin eri mieltä. Asiaa kommentoitiin esimerkiksi seuraavanlaisesti: ”En mä kyllä niitä poiskaan haluaisi, voi niitä joskus tarvita” Yksi viidestä ei osannut sanoa kantaansa asiaan. (taulukko 2.)

TAULUKKO 2 Vanhasta transaktioiden visualisointitavasta luopuminen

Väittäjä	Vastaukset
Transaktiotapahtumien kuvaamiseen riittävät suorat linkit osallisten välillä ja vanhanmalliset transaktiotapahtumat voidaan poistaa.	2 x täysin samaa mieltä, 2 x jokseenkin eri mieltä, 1 x en osaa sanoa.

Uudessa järjestelmässä tehtyjen laskelmien perusteella vanhanmallisessa visualisointiratkaisussa transaktiotyyppiset tapahtumat edustavat noin 58 prosenttia kaikista järjestelmän solmuista ja niihin liittyvät linkit noin 57 prosenttia kaikista järjestelmän linkeistä. Esittämällä transaktiot suorilla linkeillä osallisten välillä pystytään visualisoinnista jättämään pois selvästi yli puolet kaikista järjestelmän solmuista. Linkkien osalta vähennys on puolet vanhanmallisen visualisointitavan linkkien määrästä, eli noin 28,5 prosenttia kaikista järjestelmän linkeistä. Jos samansuuntaiset linkit visualisoidaan yksittäisten linkkien sijasta yhtenä linkkinä (kuvio 11), saadaan visualisoitavien linkkien määrää vielä selkeästi pienemmäksi.

#### 5.4 Analyysin helppouteen ja nopeuteen vaikuttavat ominaisuudet

Neljä viidestä haastateltavasta koki sektorimallin analyysiä helpottavana ja nopeuttavana tekijänä, kun taas yksi viidestä totesi työskentelytapojen muuttuneen, mutta työn tehostumista tai helpottumista oli hänen mielestään vaikea arvioida. Sektorimallin hyötyinä pidettiin kuvakkeiden vähentymistä ja sitä että nähdään heti mistä eri rekistereistä kohdehenkilöistä löytyy tietoa. Molemmat alustavaa analyysiä tekevistä haastateltavista kokivat myös rekisteritarkistusten tuloksista muodostetun listan nopeuttavan heidän työtään ja vastasivat kysymykseen: ”Mitkä ominaisuudet nopeuttavat tai helpottavat analyysin tekemistä?” seuraavanlaisesti: ”No siis tämä ristiinvertailujuttu ihan ehdottomasti, että sä näet heti sitä henkilöä klikkaamalla, että sillä on rekisterihistoriaa, onko avoimia, onko – -. Se on ihan sairaan hyvä. Mä oon tykännyt siitä aivan älyttömän paljon ” ja ”Ristiinvertailutiedot, eli sehän on aivan loistavaa, kunhan vaan [tietokanta-ajot] on kondiksessa”. Haastattelunsa lopussa edellisen vastauksen antanut haastateltava totesi myös seuraavasti rekisteritarkistuslistaan viitaten: ”Meillä säästyy tosi paljon aikaa siinä, että meidän ei tarvitse noita rekistereitä käydä yksitellen läpi”. Toinen varsinaista analyysiä tekevistä haastateltavista

toi myös esille, että ilman sektorimallin mukaista kuvaketta hän ei olisi eräässä tutkittavassa kokonaisuudessa ymmärtänyt hakea heti alkuvaiheessa lisätietoa eräästä rekisteristä ja löytänyt tutkittavan kohteen kannalta hyödyllistä tietoa. Yksi viidestä haastateltavasta mainitsi myös transaktiotyypisten tapahtumien esittämisen suorina linkkeinä osallisten välillä analyysiä nopeuttavana ja helpottavana tekijänä.

Kyselylomakkeen vastaukset olivat samansuuntaisia haastattelun vastausten kanssa. Uuden järjestelmän koettiin melko vahvasti nopeuttavan analysointia ja ilmoitusten esikäsittelyä, kun taas analysointia helpottavasta vaikutuksesta oltiin ainoastaan jokseenkin samaa mieltä. Manuaalisen työn määrän vähentymisestä olivat täysin samaa mieltä molemmat alustavaa analyysiä tekevät loppukäyttäjät, kuten haastateluistakin kävi esille. Yksi haastateltavista oli asiasta jokseenkin samaa mieltä ja kaksi haastateltavaa ei nähnyt uuden visualisointitavan vähentävän manuaalista työtä. Jokseenkin eri mieltä ollut haastateltava kommentoi väittämään seuraavasti: ”Manuaalinen työ tulee lähinnä esikäsittelystä – ei se mulla ainakaan ole vähentänyt”. (kuviokuva 16.)

TAULUKKO 3 Analyysin helppous ja nopeus

Väittäjä	Vastaukset
Uusi visualisointitapa nopeuttaa analysointia.	3 x täysin samaa mieltä, 2 x jokseenkin samaa mieltä.
Uusi visualisointitapa helpottaa analysointia.	5 x jokseenkin samaa mieltä.
Uusi visualisointitapa nopeuttaa ilmoitusten esikäsittelyä.	2 x täysin samaa mieltä, 2 x jokseenkin samaa mieltä, 1 x en osaa sanoa.
Uusi visualisointitapa vähentää manuaalista työtä.	2 x täysin samaa mieltä, 1 x jokseenkin samaa mieltä, 1 x neutraali kanta, 1 x jokseenkin eri mieltä.

Haastateltavista neljä viidestä oli täysin samaa mieltä siitä, että linkkien määrän vähentyminen helpottaa analysointia, ja yksi viidestä oli asiasta jokseenkin samaa mieltä (taulukko 4). Väitteeseen ”Kuvakkeiden määrän vähentyminen helpottaa analysointia.” kaikki haastateltavat vastasivat olevansa täysin samaa mieltä (taulukko 4).

TAULUKKO 4 Linkkien ja kuvakkeiden vähentymisen vaikutus

Väittäjä	Vastaukset
Linkkien määrän vähentyminen helpottaa analysointia.	4 x täysin samaa mieltä, 1 x jokseenkin samaa mieltä.
Kuvakkeiden määrän vähentyminen helpottaa analysointia.	5 x täysin samaa mieltä.



## 5.5 Graafisten elementtien toimivuus

Yhtenä haastattelun teemana oli graafisten elementtien toimivuus uudessa järjestelmässä. Graafisista elementeistä keskityttiin sektorimallin ikoneihin, yleisesti linkkeihin sekä visualisoitaviin attribuutteihin. Käsiteltäviä asioita olivat esimerkiksi värit ja muodot.

Haastateltavat pitivät sektorimallin värejä toisistaan erottuvina, mutta kaikilla oli jonkin tasoisia ongelmia muistaa mitä ristiinvertailurekisteriä mikäkin väri, tai sektori edustaa, yleisimmin esiintyvää väriä lukuun ottamatta:

Pystyy ne erottamaan, mut itelle on tullu tutuksi vasta yks pala ja mahdollisesti toinen pala. Ne mitkä on yleisiä, niin ne sä muistat, mitä ne tarkoittaa - -. Mulla ei ole ohjekirjaa vieressä, että voisin tarkistaa, mitä ne muut mahdollisesti oli, koska pääosa niistä tulee [rekisterin nimi]-tiedoista.

Syynä siihen, miksi kaikkien värien tai niiden sijaintien merkitystä ei vielä muistettu, pidettiin käyttökokemuksen vähäisyyttä ja sitä, että tiettyjä värejä eli tiettyjen rekisterien tietoja esiintyy toisia harvemmin. Asia kommentoitiin esimerkiksi: "Sit kun mä käytän sitä vähän enemmän, niin opin ne värit väkisin, oli ne minkä värisiä tahansa. Erotumisessa ei ole ollut mitään ongelmaa". Henkilöt, joiden kuvakkeet sisälsivät paljon eri värejä, koettiin erityisesti esianalyysiä tekevien keskuudessa kiinnostavimpina. Haastateltavat sanoivat, että "Kun mä näen, että sillä on paljon värejäni, niin me katsotaan mitä se on syönyt. Et siinä mielessä - paljon värejä - tässä on jotain." ja "Ainahan se heti herättää, jos siinä on useampia niitä värejä ja se on heti niinku hei, tässä on joku."

Visualisoitavien attribuuttien määrää pidettiin sopivana ja niiden koettiin antavan lisäarvoa: "Niitä ei ole liikaa ja ne mitä siellä nyt on, ne palvelee". Yksi haasteltavista piti kaimatiedosta kertovaa attribuuttia (kuvio 12) erityisen hyvänä ideana sanoen: "Erityisen hyvänä kuitenkin voin mainita sen, voit antaa vaikka vuoden [yksikön nimi]-palkinnon, sitten kun saat projektisi valmiiksi, oli se mikä tuli ihan puskista, oli tämä kaimojen lukumäärätieto. Se on perkuleen kiva juttu. Se on tosi kiva juttu." Kuten on jo aiemmin mainittu, kaimatieto lisättiin muistuttamaan toisteisista arvoista, jotta kaikki tärkeä tieto muistettaisiin hakea mukaan analyysiin. Kaimatieto-attribuuttia kehiteltäessä ja kokeillessa eräs projektiin osallistuvista loppukäyttäjistä sanoi, että hän ei tiennyt, että kyseinen ominaisuus olisi edes mahdollista toteuttaa järjestelmään.

Linkkityyppien väreihin kaikki haastateltavat eivät osanneet ottaa selkeää kantaa. Kaksi viidestä mainitsi erikseen, että transaktioiden vihreä väri on hyvä ja visualisoinnissa selkeästi erottuva. Yksi viidestä haastatellusta halusi yhden linkkityyppin väriä muutettavan, koska kyseiset linkit voidaan sekoittaa toiseen, samalla värillä visualisoitavaan linkkityyppiin. Haastatteluiden ja havaintojen perusteella voidaan todeta, että loppukäyttäjät vaativat lisää koulutusta linkkityyppien eri värien hyödyntämisessä työssään.

Kaksi viidestä haastateltavasta oli täysin samaa mieltä ja kolme viidestä jokseenkin samaa mieltä siitä, että uusi visualisointitapa tekee visualisoinnista selkeämpää. Uuden visualisointitavan graafisten elementtien selkeydestä täysin

samaa mieltä oli kolme viidestä ja jokseenkin samaa mieltä kaksi viidestä. Graafisten elementtien värien onnistumisesta neljä viidestä oli täysin samaa mieltä, vain yhden viidestä ollessa jokseenkin samaa mieltä. (taulukko 5.)

TAULUKKO 5 Graafisten elementtien toimivuus

Väittäjä	Vastaukset
Uusi visualisointitapa tekee visualisoinnista selkeämpää.	2 x täysin samaa mieltä, 3 x jokseenkin samaa mieltä.
Uuden visualisointitavan graafiset elementit, kuten kuvakkeet ja linkit, ovat selkeitä.	3 x täysin samaa mieltä, 2 x jokseenkin samaa mieltä.
Uuden visualisointitavan graafisten elementtien värien valinnassa on onnistuttu.	4 x täysin samaa mieltä, 1 x jokseenkin samaa mieltä.

## 5.6 Oletuksena laajennettavat verkon tiedot

Kolme viidestä haastateltavasta oli sitä mieltä, että verkon laajentamisen oletusasetuksia voisi muuttaa niin, että vähintään kohteiden yhteystiedot, kuten osoitteet ja puhelinnumerot, jätettäisiin oletuksena olevista laajennusasetuksista pois. Yksi viidestä kertoi, että ei käytä ollenkaan oletusasetuksia laajentaessaan verkon tietoja, vaan valitsee visualisointiin tuotavat kohteet ja linkkityypit aina tilannekohtaisesti.

Haastatteluissa kävi ilmi, että analyysin tekemisen alkuvaiheessa halutaan saada tutkittavasta kohteesta yleiskuva, johon tarvitaan suhteellisen pieni osa koko järjestelmän tiedoista. Tässä vaiheessa esimerkiksi kohteiden yhteystiedoilla ei ole merkitystä, tai ne ovat pikemminkin visualisointia häiritseviä tekijöitä ja voivat viedä alkuvaiheessa huomion epäolennaisiin asioihin. Vasta varsinaisen syvemmän analyysin vaiheessa, yleiskuvan saamisen jälkeen, on aika hakea kohteiden välisiä yhteyksiä ja pohtia niiden merkityksiä. Yhteystiedot edustavat noin neljää prosenttia kaikista järjestelmän solmuista.

## 5.7 Tarvittavat muutokset

Haastateltavilta kysyttiin myös yleisesti mitä asioita heidän mielestään järjestelmässä pitäisi vielä parantaa tai mitä muutoksia sinne tulisi tehdä. Yksi ehdotus koski muun muassa erään tietotyypin purkamista osiin niin, että vain osa tiedosta olisi valittuna mukaan tiedon laajennuksen oletusasetuksiin. Muut ehdotukset koskivat tiedon linkittämiseen ja tiettyjen, muita kuin henkilöitä ja yhtiöitä koskevien, toisteisten tietojen yhdistämiseen visualisoinnissa. Tarkka lista kehitys- ja muutosehdotuksista tullaan luovuttamaan tutkielman valmistuttua kohdeorganisaation ja järjestelmän jatkokehitysprojektin käyttöön.

## 5.8 Johtopäätöksiä

Visuaalisen analyysijärjestelmän visualisoinnin suunnittelu ja toteutus niin, että se täyttää erilaisten käyttäjien ja käyttäjäryhmien tarpeet, on ollut haastava mutta antoisa kokemus. Tämän kappaleen alakappaleissa käydään ensin läpi sitä miten tutkimustulokset vastaavat tutkimusongelmaan liittyviin apukysymyksiin, minkä jälkeen näiden vastausten avulla vastataan varsinaiseen tutkimuskysymykseen: ”Millä toimenpiteillä verkostomaisen tiedon visualisoinnista voidaan tehdä helpommin ihmiselle analysoitavaa?”

### 5.8.1 Menetelmiä visuaaliseen analysointiin

Visuaalisen analyysin menetelmiä käsiteltiin kirjallisuuskatsauksessa, jossa tämän tutkielman apukysymykseen: ”Mitä eri menetelmiä verkostomaisen tiedon visuaaliseen analysointiin on olemassa kirjallisuudessa?” saadaan vastaukseksi esimerkiksi kirjallisuuskatsauksessa käsitelty tiedon etsimisen mantra: ”Ensin yleiskuva, zoomaa ja suodata, yksityiskohtiin tarvittaessa” (Shneidermanin, 1996) sekä siitä erityisesti visuaalista analysointia varten muokattu visuaalisen analyysin mantra: ”Analysoi ensin – näytä tärkeys – zoomaa, suodata ja analysoi pidemmälle – yksityiskohtiin tarvittaessa” (Keim ym., 2006).

Visuaalisessa analysoinnissa on tärkeää, että ne tehtävät, jotka ovat nopeita ja helppoja toteuttaa tietokoneen avustamana, kuten laskeminen ja tietojen yhdistely tietokannassa, annetaan tietokoneen tehtäväksi. Myös kirjallisuuskatsauksessa kuvailtua, iteratiivista ymmärtämisen luuppia (kuvio 4) (Cook & Thomas, 2005, 42) voidaan käyttää menetelmänä verkostomaisen tiedon visuaalisessa analysoinnissa. Varsinkin ensimmäisessä ymmärtämisen luupin iteraatiossa tiedon kuvaaminen uudessa, paremmin analyysiä tukevassa muodossa voidaan toteuttaa automaattisesti tietokoneella, jolloin analyytikko voi keskittyä näkemyksen muodostamiseen ja ratkaisun synnyttämiseen.

### 5.8.2 Linkkien ja solmujen vähentäminen kadottamatta oleellista tietoa

Uuden visualisointijärjestelmän yhtenä tavoitteena oli vähentää verkon linkkien ja solmujen määrää, jotta analysoinnista saataisiin tehtyä helpompaa. Linkkien ja solmujen vähentyminen ei kuitenkaan saa johtaa yhdenkään käyttäjäryhmän kannalta oleellisen tiedon katoamiseen. Tässä kappaleessa tullaan vastaamaan seuraavaan, aiheeseen liittyvään tutkielman apukysymykseen:

- Minkä tyyppisiä linkkejä ja solmuja voidaan poistaa, ja miten menetetty oleellinen tieto voidaan kuvata muuten, ottaen huomioon kaikki erilaiset tietotarpeet?

Uusilla visualisointimenetelmillä pystyttiin vähentämään oletuksena visualisoinnissa laajennettavien linkkien ja solmujen määrää huomattavasti. Suurin linkkejä vähentävä vaikutus saavutettiin sillä, että transaktiotieto kuvattiin

kahden linkin ja yhden solmun sijasta yhdellä transaktion osallisten välisellä linkillä. Tällä toimenpiteellä pystyttiin vähentämään visualisoitavien solmujen määrä alle puoleen, sekä pääsemään eroon yli 28 prosentista kaikista järjestelmän linkeistä. Myös sektorimallin käyttöönotto vähensi visualisoitavien linkkien ja solmujen määrää huomattavasti.

Erityisesti sektorimallia suunniteltaessa jouduttiin pohtimaan sitä, minkä tyyppisiä linkkejä ja solmuja voidaan poistaa, ja miten menetetty oleellinen tieto kuvataan muuten. Organisaatiossa tehtyjen havaintojen ja haastattelujen perusteella sektorimalliin kuuluvien taustarekisteritietojen linkit ja solmut menettivät suurelta osin merkityksensä sektorimalliin siirryttäessä, erityisesti alustavaa analyysiä tehdessä. Kohdetta alustavasti tarkasteltaessa voi riittää jo pelkästään ikonin perusteella nähtävä tieto siitä, että kohteelle löytää jotain tietoa yhdestä tai useammasta ristiinvertailurekisteristä. Aina näin ei kuitenkaan ole. Jotta oleelliseksi katsottua ja yleisimmin käytettyä tietoa ei menetettäisi, se tallennettiin listan muodossa kohteen yhdeksi attribuutiksi (kuvio 8). Listan ja ikonin antamaa tietoa pidettiin loppukäyttäjien haastatteluiden perusteella riittävänä varsinkin alustavien johtopäätösten tekemisessä.

Vaikka sektorimalli toimiikin useissa tapauksissa linkkejä ja solmuja vähentävänä mallina, löytyy erityisesti varsinaisen analyysin puolelta vielä tapauksia, joissa verkostomainen esitystapa myös sektorimalliin sisältyvien taustarekisteritietojen osalta on hyödyllistä. Esimerkiksi eri kohteiden välisiä, verkosta löytyviä yhteyksiä sektorimallilla ei pystytä esittämään tehokkaasti, minkä johdosta taustarekisteritietoja esittäviä linkkejä ja solmuja ei poistettu järjestelmästä, vaan ne poistettiin ainoastaan oletuslaajennusasetuksista. Tällä tavoin visualisoinnista saadaan selkeämpi ja myös taustarekisteritietojen linkkeihin ja solmuihin pääsee tarvittaessa käsiksi laajennusasetuksia muuttamalla. Käsiteltyyn kysymykseen voidaankin sektorimalliin osalta vastata seuraavasti:

- Visualisoinnista voidaan poistaa kohteisiin liittyvien, sektorimalliin kuuluvien taustarekisterien linkkejä ja solmuja.
- Linkkitietoa ei kuitenkaan tule poistaa kokonaan järjestelmästä, vaan ainoastaan oletuslaajennusasetuksista.
- Menetettyä tietoa voidaan kuvata sektorimallia apuna käyttäen, ikonilla ja tiivistettyä tietoa sisältävällä attribuutilla. Ikonilla kuvataan sitä, mitä tietoja kohteeseen kuuluu, ja attribuutti antaa tarkempaa tietoa taustarekisterien sisältämistä tiedoista.

Transaktiotiedon esitystavalla laskettiin olevan suuri vaikutus visualisoitavien linkkien ja solmujen määrään. Sama asia pystyttiin esittämään kahden linkin ja yhden solmun sijasta vain yhdellä osallisten välisellä linkillä (kuvio 10), tai jopa tiiviimmin, kuvaamalla kaikki samansuuntaiset transaktiot yhdellä linkillä (kuvio 11). Vanhanmallisen transaktiotiedon esittämistavassa transaktion tarkempi tietosisältö, kuten esimerkiksi tapahtuma-aika kuvattiin solmun attribuuteissa. Uudessa visualisointitavassa sama tietosisältö tuodaan kunkin linkin attribuuttiin, joka mahdollistaa edelleen samaan tietoon pääsyn. Käsiteltyyn kysymykseen voidaankin transaktiotiedon osalta vastata seuraavasti:

- Visualisoinnista voidaan poistaa ylimääräiset linkit ja solmut, joiden tietosisältö pystytään esittämään suoralla, kahden osallisen välisellä linkillä.
- Menetetty tieto voidaan siirtää poistetuista linkeistä ja solmusta, yhteen uuteen linkkiin.

Jotkin ennen ainoastaan linkkien kautta verkosta saatavat tiedot päätettiin tuoda selkeästi näkyviin visualisoitavana attribuuttina (kuvio 12). Visualisoituja attribuutteja käytettiin useissa eri solmutyypeissä, ja niitä pidettiin haastattelussa hyödyllisenä visualisoinnin osana. Osa attribuuttina visualisoitavasta tiedosta ei linkity muihin analysoinnin kannalta oleellisiin tietoihin, minkä johdosta aikaisemmin kyseistä tietoa kuvanneet linkit ja solmut käyvät tarpeettomiksi. Käsiteltyyn kysymykseen voidaankin visualisoitavien attribuuttien osalta vastata seuraavasti:

- Visualisoinnista voidaan poistaa sellaiset linkit ja solmut, jotka eivät linkity muihin analysoinnin kannalta tärkeisiin tietoihin ja jotka voidaan kuvata visualisoitavalla attribuutilla.

Niin alustavaa analyysiä kuin myös varsinaista analyysiä tekevien käyttäjien mukaan ensivaiheessa verkon sisällöstä tulisi laajentaa vain välttämätön tieto. Haastatteluiden perusteella oletuslaajennusasetuksista voidaan jättää pois esimerkiksi kohteiden yhteystiedot ja tiettyjä muita kohteisiin liittyviä, vähemmän tärkeitä tietoja.

Tiivistettynä voidaankin sanoa, että linkkejä ja solmuja vähennettäessä verkon oleellinen tieto saadaan säilytettyä, erilaiset tietotarpeet huomioiden, seuraavin keinoin:

1. Kuvataan tieto vaihtoehtoisella tavalla.
2. Ilmaistaan ikonilla, minkä tyyppistä tietoa kohteisiin liittyy.
3. Kootaan visualisoinnista poistuneesta tiedosta tiivistetty ja jäsenelty listaus, joka sisältää kaiken oleellisen tiedon.
4. Tallennetaan tietoja poistettujen solmujen sijasta linkkeihin.
5. Käytetään visualisoitavia attribuutteja kuvaamaan tärkeimpiä tietoja.
6. Ei poisteta muulla tavalla visualisoitavaa tietoa järjestelmästä, vaan poistetaan se ainoastaan oletuksena laajennettavien tietojen listalta.

### 5.8.3 Värät ja muodot visualisoinnin selkiyttäjinä

Väreillä ja muodoilla voidaan visualisoinnissa viestittää monenlaisia asioita. Tämän johdosta tutkielman kolmannen apukysymyksen avulla haluttiin tutkia, miten verkostomaisen tiedon visualisointia voidaan selkiyttää värien ja muotojen avulla.

Selkeimmin värien ja muotojen merkitys tuli esille sektorimallissa, joissa molemmilla elementeillä oli onnistumisen kannalta ratkaiseva merkitys. Muotoja käytettiin erottamaan yhtiöitä ja henkilöitä kuvaavat solmut toisistaan ja vä-

reillä eroteltiin kohteisiin liittyvät rekisteritiedot toisistaan. Sektorimallissa värien ja muotojen avulla pystyttiin kertomaan minkä tyyppistä tietoa kohteista löytyy ja siten ne myös auttoivat vähentämään visualisoitavien linkkien ja solmujen määrää.

Järjestelmässä käytettiin kahdeksaa erilaista linkkityyppiä, jotka oli määritelty pääsääntöisesti erivärisiksi. Vähemmän tärkeäksi katsottujen linkkien väreiksi valittiin huomaamattomampia värejä, kuten harmaa, ja tärkeimpien linkkien väreinä käytettiin esimerkiksi punaista väriä. Yhtenä tavoitteena oli, että jo pelkästään linkkien värejä vilkaisemalla havaittaisiin minkä tyyppistä tietoa milloinkin käsitellään ja että kauemmas kuvaa zoomattaessa vähemmän tärkeät tiedot häviäisivät taustalle, tärkeimpien jäädessä vielä näkyviin.

Linkkien värien lisäksi myös niiden muodoilla oli tärkeä rooli visualisoinnissa. Esimerkiksi transaktioiden linkit esitettiin nuolina, jos molemmat osapuolet olivat tiedossa. Osa linkeistä esitettiin katkoviivalla, jolla kuvattiin epävarmaa tai varmistamattomaa tietoa. Myös linkin paksuutta olisi haluttu käyttää Westphalin (2008, 164) kuvaamalla tavalla kertomaan jonkin yhteyden olevan järjestelmässä useita kertoja, mutta tutkielmassa käytetty ohjelmisto ei tukenut kyseistä toimintoa.

Tiettyjen visualisoitavien attribuuttien edessä käytettiin erilaisia värikkäitä kuvioita, herättämään analyytikkojen huomiota. Kuvioiden avulla visuaaliset attribuutit pystyttiin erottamaan myös visualisointia vähän kauempaa tarkasteltaessa. Kysymykseen: "Miten verkostomaisen tiedon elementtien visualisointia voidaan selkiyttää värien ja muotojen avulla?" voidaankin värien ja muotojen osalta vastata seuraavasti:

- Solmujen väreillä ja muodoilla voidaan korostaa jotain tiettyä tietoa.
- Visualisoitavia attribuutteja voidaan korostaa tuomalla tekstin eteen huomiota herättävä, värikäs kuvio.
- Värien ja muotojen avulla voidaan kertoa mitä tietoa kohteeseen liittyy, mikä vähentää kyseisten tietojen visualisoinnin tarvetta.
- Linkkien värien ja muotojen avulla voidaan auttaa tuomaan esiin sitä, minkä tyyppistä tietoa visualisointi esittää.
- Linkkien esittäminen nuolella helpottaa esimerkiksi transaktion suunnan havaitsemista.
- Linkin esittäminen katkoviivalla auttaa erottamaan varmistettua tai luotettavaa tietoa varmistamattomasta tai epäluotettavasta tiedosta.

#### **5.8.4 Ratkaisun parantaminen evaluoinnin avulla**

Seuraavaksi pohditaan tutkielman apukysymystä: Miten kehitettyä ratkaisua voidaan evaluoida ja parantaa evaluoinnin tulosten pohjalta? Tutkielmassa käytettiin Seinin ym. (2011) jatkuvan arvioinnin periaatetta. Uudet ominaisuudet testattiin erikseen, mutta kuitenkin osana visuaalisen analyysijärjestelmän kokonaisuutta. Sovellettu ketterän kehityksen menetelmä mahdollisti sen, että kaikki loppukäyttäjät voivat kertoa mielipiteensä ja aidosti

vaikuttaa visuaalisen analyysijärjestelmän ratkaisuihin. Loppukäyttäjien aktiivinen osallistuminen toteutusvaiheessa järjestelmän beta-testaamiseen ja järjestelmän arvointiin mahdollisti ongelmatilanteisiin puuttumisen hyvin nopeasti. Loppukäyttäjät olivat koko ajan tavoitettavissa, jos heidän mielipidettään ja osaamistaan tarvittiin jonkin ongelman ratkaisemiseksi tai toteutusidean toimivuuden arviointiin.

Elokuun ensimmäisenä päivänä 2015 tuotantokäyttöön luovutetun järjestelmän arviointi toteutettiin organisaatiossa tehtyjen havaintojen sekä syyskuun puolivälin jälkeen toteutettujen loppukäyttäjien teemahaastatteluiden ja kyselylomakkeen yhdistelmän avulla saatujen vastausten perusteella. Tämän lisäksi uusia visualisointiratkaisuja vertailtiin vanhan järjestelmän vastaaviin ratkaisuihin linkkien ja solmujen määrän osalta. Arviointi suoritettiin Peffersin ym. (2007) viitekehyksen mukaisesti havainnoimalla ja mittaamalla toteutetun artefaktin kykyä ratkaista tunnistettu ongelma, sekä Marchin ja Smithin (1995) mukaisesti testaamalla artefaktin tehokkuutta ja vaikuttavuutta loppukäyttäjiiin.

Vaikka kehitettyä ratkaisua voidaankin pitää tulosten perusteella onnistuneena, tuli esille myös monta kohtaa, joissa ratkaisua voidaan ja tulee kehittää jatkossakin. Kohdeorganisaatiossa päätettiin, että visualisointiratkaisun jatkokehitys aloitetaan välittömästi tämän tutkielman valmistuttua. Jatkokehityksessä tullaan käyttämään samanlaista ketterän kehityksen menetelmää kuin tässä tutkimuksessa esiteltiin aiemmin ja lopullinen tehtävälista kehityskohteista tullaan laatimaan sekä tämän tutkimuksen arvioinnin tuloksista että erikseen vain kohdeorganisaation käyttöön kerätystä aineistosta.

Tiivistetysti voidaan sanoa, että kehitettyä ratkaisua on hyvä ensin evaluoida kehitysvaiheessa pienissä paloissa, osana kokonaisuutta. Valmiin ratkaisun evaluointi pitää perustaa loppukäyttäjien todellisiin käyttökokemuksiin riittävän pitkältä aikaväliltä. Saatuja evaluoinnin tuloksia tulee käyttää uuden järjestelmäversion kehittämisessä.

### 5.8.5 Toimenpiteiden vaikutus analysoinnin helppouteen

Kirjallisuuskatsauksen pohjalta voidaan todeta, että visuaalisen analysointiratkaisun kehittäminen vaatii kohdeympäristön hyvää tuntemista, sekä osaamista tunnistaa koneiden ja inhimillisten kykyjen parhaat puolet ja soveltaa niitä oikeassa suhteessa. Tämän ymmärtäminen antaa hyvän pohjan vastata tutkimuksen pääkysymykseen:

- Millä toimenpiteillä verkostomaisen tiedon visualisoinnista voidaan tehdä helpommin ihmiselle analysoitavaa?

Käytännössä analysoitava tieto tulee esianalysoida koneellisesti, ennen ihmisen tekemää analyysiä, käyttäen apuna muun muassa tiedonlouhinnan ja tilastollisten menetelmien keinoja. Analysoitavasta tiedosta voidaan tehdä helpommin analysoitavaa pelkistämällä sitä esimerkiksi summaamalla tai suodattamalla.

Analysoitavasta kokonaisuudesta tulee aluksi pystyä luomaan yleiskuva ja siirtä yksityiskohtiin vasta tarvittaessa, kuten Shneidermanin (1996) tiedon etsimisen mantrassakin todetaan.

Kuten myös Keim ym. (2006) tähdensi, visualisoinnissa käytetyn raakatiedon laadulla on iso merkitys visuaalisen analyysin onnistumiselle. Toteutetussa visualisointijärjestelmässä raakatiedon epätarkkuuksia sekä esimerkiksi toisteisten arvojen olemassa olo pyrittiin tuomaan Keimin ym. (2006) oppien mukaisesti visuaalisesti selvästi esille, jotta analyytikot huomaisivat ottaa kyseiset seikat huomioon omassa työssään. Uudessa visualisointijärjestelmässä käytettiin esimerkiksi katkoviivaa kuvaamaan epävarmaa tietoa ja visualisoitua attribuuttia kuvaamaan toisteisten arvojen olemassaoloa. Erityisesti toisteisista arvoista kertovaa, nimen samantyyppisyyden perustuvaa kaimatietoa pidettiin hyvänä uudistuksena, jonka avulla pystyttiin helposti havaitsemaan, että kohteesta saattaa löytyä vielä jotain analyysin kannalta oleellista tietoa.

Tutkielman tulosten perusteella linkkien ja kuvakkeiden vähentymistä pidettiin selkeästi analyysiä helpottavana tekijänä. Tulos vahvistaa myös Didi-mon, Liottan ja Montecchianin (2014) tutkimuksen tuloksia siitä, että linkkien ja solmujen määrän kasvaessa suureksi myös verkon kompleksisuus kasvaa, mikä taas vuorostaan vaikeuttaa analyytikon työtä. Eniten linkkien määrää pystyttiin vähentämään visualisoimalla kahden osapuolen välinen transaktiotieto kahden linkin ja yhden solmun sijasta vain yhdellä linkillä. Myös sektorimallin käyttöönotolla pystyttiin vähentämään linkkien ja solmujen määrää huomattavasti. Jotta järjestelmä tukisi kaikkia organisaatiossa tarvittavia analyysin tarpeita, tuodaan visuaaliseen järjestelmään suuri määrä myös sellaisia linkkejä ja solmuja, joita käytetään vain joissakin analyysin vaiheissa tai vain tietyissä tehtävissä. Edellä mainittuja tietoja, kuten kohteiden osoitteita, ei tarvitse ottaa mukaan oletuslaajennusasetuksiin, mikä vähentää visualisoitavien tietojen määrää ja helpottaa siten osaltaan analysointia.

Sektorimalliin liittyvistä rekisteritarkistuksista tehtyä tiivistettyä listaa pidettiin erityisesti alustavaa analyysiä tekevien loppukäyttäjien toimesta analyysiä helpottavana ja nopeuttavana ominaisuutena. Kohdeorganisaatiossa työskentelyn aikana havaittiin myös, että kaikkien analyytikkojen käytössä olevilla kahdella 30-tuumaisella sekä isoon resoluutioon kykenevillä tietokoneen näytöillä oli selkeä analyysin tekemistä helpottava vaikutus.

Kirjallisuuskatsauksessa ja tapaustutkimuksen tulosten perusteella voidaan tuottaa taulukon 6 mukainen lista niistä toimenpiteistä, jotka vastaavat tutkielman varsinaiseen kysymykseen niistä toimenpiteistä, joilla verkostomaisen tiedon visualisoinnista voidaan tehdä helpommin ihmiselle analysoitavaa.



TAULUKKO 6 Toimenpiteitä verkostomaisen, visualisoitavan tiedon tekemiseksi helpommin ihmiselle analysoitavaksi

<b>Toimenpide</b>
Tutustu kohdeympäristöön ja järjestelmän käyttötarkoitukseen hyvin ennen minkään toimenpiteen suorittamista.
Hyödynnä tietokoneen kykyä esimerkiksi suurien tietomassojen louhinnassa, tietojen pelkistämässä ja yhdistämisessä.
Mahdollista alussa tarvittavan yleiskuvan luonti analysoitavasta kohteesta ilman ylimääräisiä, häiritseviä linkkejä ja solmuja.
Tuo visualisoitavan tiedon epätarkkuudet, kuten toisteiset arvot ja epävarmat tiedot, selvästi analyytikon nähtäville.
Poista ylimääräiset linkit ja solmut sekä kuvaa kadotettu, tarpeellinen tieto vaihtoehdoisella tavalla, käyttäen esimerkiksi visuaalisia elementtejä tai tiivistettyä tekstimuotoista esitystapaa.
Evaluoivat tehtyjä ratkaisuja yhteistyössä loppukäyttäjien kanssa ja jatkokehittä visualisointia uuden tiedon valossa.
Käytä tiedon visuaalisessa analyysissä riittävän isoa näyttöä, isolla resoluutiolla.

## 6 POHDINTA

Viime vuosien aikana tapahtunut digitaalisen tiedon valtava kasvu on vaikeuttanut tiedon analysointia. Vaikka verkostomaisen tiedon visuaalisen analysoinnin apuna käytettäisiin tiedonlouhinnan ja tilastollisen analyysin menetelmiä, tulee tietokoneella jalostettu tieto vielä pystyä visualisoimaan mahdollisimman ymmärrettävässä muodossa. Tämän tutkielman tavoitteena oli löytää toimenpiteitä, joilla verkostomaisen tiedon visualisoinnista voidaan tehdä helpommin ihmiselle analysoitavaa. Tiivistetty lista kyseisistä toimenpiteistä löytyy taulukosta 6.

Aihetta lähestyttiin etsimällä aikaisemmista tutkimuksista tietoa visuaalisesta analyysistä, verkostomaisen tiedon visualisoinnista ja sosiaalisista verkoista. Yhtenä kirjallisuuskatsauksen päämäärinä oli kartoittaa esimerkiksi sitä, mitä erilaisia menetelmiä verkostomaisen tiedon visuaaliseen analysointiin löytyy. Aikaisemmista tutkimuksista toivottiin myös löydettävän ratkaisumalleja, joita voitaisiin hyödyntää tutkielman konstruktiivisessa toimintatutkimuksessa, jossa pyrittiin kehittämään, ketteriä järjestelmänkehitysmenetelmiä hyväksi käyttäen, erään organisaation visuaalisesta analyysiratkaisusta helpommin analysoitavan. Seuraavissa alakappaleissa arvioidaan tutkimustuloksia ja prosessia, sekä pohditaan lopuksi jatkotutkimuksen aiheita.

### 6.1 Tulosten arviointi

Tässä tutkielmassa haluttiin ensisijaisesti vastata kysymykseen siitä, millä toimenpiteillä verkostomaisen tiedon visualisoinnista voidaan tehdä helpommin ihmiselle analysoitavaa. Ennen kuin kysymykseen pystyttiin vastaamaan kattavasti, tuli tutkielmassa pystyä vastaamaan seuraaviin apukysymyksiin:

1. Mitä eri menetelmiä verkostomaisen tiedon visuaaliseen analysointiin on olemassa kirjallisuudessa?

2. Minkä tyyppisiä linkkejä ja solmuja voidaan poistaa, ja miten menetetty oleellinen tieto voidaan kuvata muuten, ottaen huomioon kaikki erilaiset tietotarpeet?
3. Miten verkostomaisen tiedon elementtien visualisointia voidaan selkiyttää värien ja muotojen avulla?
4. Miten kehitettyä ratkaisua voidaan evaluoida ja parantaa evaluoinnin tulosten pohjalta?

Ensimmäisen apukysymyksen tarkoituksena oli löytää menetelmiä verkostomaisen tiedon visuaalisen analyysin suorittamiseen. Mielestäni parhaiten kirjallisuudesta löytämistäni menetelmistä verkostomaisen tiedon visuaaliseen analyysin tekemiseen sopii Keim ym. (2006) visuaalisen analyysin mantra: ”Analysoi ensin – näytä tärkeys – zoomaa, suodata ja analysoi pidemmälle – yksityiskohtiin tarvittaessa”, joka eroaa tiedon etsimisen mantrasta (Shneiderman, 1996) selkeimmin siinä, että se ottaa huomioon myös visuaalisessa analysoinnissa oleellisen tiedon koneellisen analysoinnin. Visuaalisen analyysin mantran kohta: ”yksityiskohtiin tarvittaessa” on mielestäni erityisen tärkeä, sillä alussa yleiskuvaa luotaessa visualisoinnissa ei saa olla ylimääräisiä, häiritseviä tekijöitä. Sama asia tuli selkeästi esille myös loppukäyttäjien haastatteluissa.

Toisella apukysymyksellä pyrittiin löytämään ne ”ylimääräiset” linkit ja solmut, jotka voitaisiin poistaa visualisoinnista. Poistettavien linkkien tietosisältöä ei kuitenkaan saa kadottaa, jos se on jollekin käyttäjäryhmälle visuaalisen analyysin kannalta oleellinen tieto, joten lisäksi jouduttiin pohtimaan sitä, miten menetetty tieto kuvattaisiin muuten. Kohdeorganisaation uutta järjestelmää kehitettäessä todettiin, että transaktiotieto voitiin esittää puolella alkuperäisestä linkkimäärästä ja kokonaan ilman solmuja. Ennen kahdessa linkissä ja yhdessä solmussa sijainnut tieto pysyttiin siirtämään jäljelle jääneen linkin tietosisällöksi. Yleistäen voidaan todeta, että visualisoinnissa voidaan poistaa sellaisia linkkejä ja solmuja, jotka voidaan korvata yhdellä kahden osapuolen välisellä linkillä. On kuitenkin huomattava, että ratkaisu voidaan yleistää vain niihin tilanteisiin, joissa solmua kuvanneeseen tapahtumaan tai ilmiöön kuuluu vain kaksi osallista, sillä linkkiin linkittäminen ei ole yhdessäkään tuntemassani visualisointijärjestelmässä mahdollista.

Myös tutkielman kohdeorganisaation ideasta lähtenyt sektorimallin idea muovautui tutkimuksen aikana toimivaksi malliksi poistaa ristiinvertailurekistereihin liittyviä linkkejä ja solmuja erityisesti analyysin alkuvaiheessa. Mallin dynaamisen ikonin lisäksi siihen liittyy oleellisena osana tiivistetty lista niistä tiedoista, joita ennen kuvattiin linkkien ja solmujen avulla. Linkitettyä ristiinvertailutietoa ei kuitenkaan poistettu kokonaan järjestelmästä, vaan ainoastaan oletuslaajennusasetuksista. Sektorimalli noudattaa hyvin Keimin ym. (2006) visuaalisen analyysin mantran periaatteita analysoimalla tietoa alussa koneellisesti. Mallissa muodostetaan ohjelmallisesti erilaisia ikoneja, jotka kertovat mitä tietoja kohteisiin kuuluu. Seuraavassa vaiheessa analyttikko voi tarvittaessa katsoa tietoja yksityiskohtaisemmin listasta, joka on myös luotu ohjelmallisesti tiivistämällä ja suodattamalla alkuperäistä tietoa. Jos tämäkään ei riitä, hän voi laajentaa oletuksena piilotetut ristiinvertailurekisterien linkit ja solmut. Sekto-

rimalli ei siis kaikissa tapauksissa auta poistamaan linkkejä ja solmuja kokonaan järjestelmästä, mutta voi hyvinkin tehokkaasti vähentää niiden visualisoinnin tarvetta, varsinkin alustavia johtopäätöksiä tehtäessä.

Kolmannella apukysymyksellä haettiin vastausta siihen, miten verkostomaisen tiedon visualisointia voidaan selkiyttää värien ja muotojen avulla. Useat uuden järjestelmän uudistuksista liittyivät värien ja muotojen käyttöön, ja niitä käytettiin Westphalin (2008, 129) oppien mukaisesti indikaattoreina, välittämään tietoa visualisoitavista kohteista. Värit ja muodot ovat erityisen tärkeitä sektorimallissa, jossa värit ja niiden sijainti kertovat mitä tietoja kohteeseen liittyy, ja ikonin muoto erottaa kaksi mallissa käytettyä solmutyyppiä, eli henkilöt ja yhtiöt toisistaan. Sektorimallin värien osalta tutkielmassa onnistuttiin mielestäni hyvin, mutta kauempaa tarkasteltaessa henkilöä kuvaava ympyrä ja yhtiötä kuvaava neliö eivät erotu riittävän hyvin toistaan. Tähän ongelmaan voi mahdollisesti auttaa erään haastateltavan ehdotus visualisoida analysoinnin kannalta yleensä henkilöihin verrattuna vähemmän tärkeitä yhtiöt pienemmillä ikoneilla.

Neljännessä apukysymyksessä pohdittiin sitä, miten kehitettyä ratkaisua voidaan evaluoida ja parantaa saatujen tulosten pohjalta. Uuden järjestelmän kehityksessä pidettiin tärkeänä noudattaa Seinin ym. (2011) jatkuvan arvioinnin periaatetta. Jokainen uusi ominaisuus testattiin ja arvioitiin niiden valmistuttua tutkijan sekä yhden loppukäyttäjän toimesta. Testaukset suoritettiin käyttäen tuotantodataa, ja lopullinen ominaisuuksien arviointi ja hyväksyntä suoritettiin viikottaisessa kokouksessa, jossa oli mukana eri käyttäjäryhmiä edustavia loppukäyttäjiä. Tuotantokäyttöön otetun uuden visuaalisen analyysijärjestelmän kokonaisuuden evaluointi tehtiin tämän tutkielman tuloksia ja johtopäätöksiä käsittelevissä, sekä tässä nimenomaisessa luvussa. Menettely noudatti tutkielman viitekehikseksi valittua Peffersin ym. (2007) konstruktivisen tutkimuksen prosessimallia, jonka viidentenä aktiiviteettina havainnoidaan ja mitataan toteutetun artefaktin onnistumista. Ratkaisun arviointi perustui havaintoihin, haastatteluihin, sekä haastatteluiden ohessa käytetyn kyselylomakkeen tuloksiin. Haastatteluiden avulla koottiin myös lista korjattavista ja kehitettävistä asioista, sekä kerättiin kokemuksia käytetystä ketterästä järjestelmänkehitysmenetelmästä organisaation seuraavia projekteja silmällä pitäen.

On tärkeää, että evaluointia ei jätetä vain loppuun, vaan sitä suoritetaan jatkuvasti. Osassa haastateltavista oli havaittavissa, että he eivät olleet käyttäneet järjestelmää riittävän kauan, jotta olisivat voineet arvioida hyvin kaikkien uusien ominaisuuksien hyötyjä.

Tutkimuksen apukysymyksiin vastaamalla pystyttiin löytämään vastauksia myös varsinaiseen tutkimuskysymykseen, jossa etsittiin niitä toimenpiteitä, joilla visualisoinnista voidaan tehdä ihmiselle helpommin analysoitavaa. Taulukon 6 listalla tiivistin tärkeimmät ja vaikuttavimmat kyseisistä toimenpiteistä mahdollisimman yleistettävässä muodossa.

## 6.2 Tulosten yleistettävyyden ja käytännön merkitys

Selkeimmin tämä tutkielma tuo uutta tietoa alalle esittelemällä ja kehittämällä sektorimallin sekä arvioimalla sen vaikutusta analysointia selkiyttävänä tekijänä. Tutkielman havaintojen ja haastatteluiden perusteella tuli hyvin selväksi, että joissakin tapauksissa linkitetyn tiedon kokonaan järjestelmästä poistamisen sijaan analyttikoille tulee mieluummin tarjota vaihtoehtoisia tapoja visualisoida tietoa niin, että vain kyseisen työtehtävän kannalta oleellinen tieto saadaan helposti tietokoneen näytölle. Esimerkiksi sektorimalliin sisältyvien solmujen ei tarvitse tulla näkyviin verkkoa laajennettaessa oletusasetuksin, mikä tuo kuvaan selkeyttä, mutta kyseisiin solmuihin ja linkkeihin tulee päästä laajennusasetuksia muuttamalla.

Toinen tutkielmassa kehitetty uusi ominaisuus, josta ei löytynyt mainintaa kirjallisuudesta, oli kaimatiedon, eli toisteisten arvojen esittäminen visualisoitavana attribuuttina. Kaimatietoa pidettiin hyvänä uudistuksena, joka auttaa analyttikkoja havaitsemaan, että analysoitavaan kohteeseen saattaa löytyä järjestelmästä lisätietoa.

Analyysiprosessit ja tietojärjestelmien arkkitehtuurit ovat hyvin samankaltaisia, pyrittiin sitten havaitsemaan rahanpesua, sisäpiirikauppoja, vakuutuspehtoksia tai muunlaista rikollista toimintaa (Westphal, 2008, xii), joten tuloksia voitaneen hyödyntää monenlaisten organisaatioiden laajojen, verkostomaisien tietomassojen visuaalisten analyysijärjestelmien kehittämisessä. Tuloksista eniten voivat hyötyä eri esitutkintaviranomaiset sekä vakuutusyhtiöt ja pankit, joissa analysoitavan tiedon määrä on suurta ja joiden toimenkuvaan kuuluu epäilyttävien tapahtumien tunnistaminen ja analysointi, niin oman liiketoiminnan kuin lainsäädännönkin näkökulmasta.

Suurin käytännön merkitys tutkimuksella on luonnollisesti kohdeorganisaatiolle, jonka visuaalisen analyysin järjestelmästä saatiin tehtyä helpommin käytettävä, varsinkin alustavaa analyysiä tekeville loppukäyttäjille. Järjestelmästä saatiin tehtyä myös edellistä versiota helpommin ylläpidettävä ja nopeammin päivitettävä. Kohdeorganisaatiossa tullaan myös jatkamaan uuden version kehitystä tämän tutkielman tuloksiin perustuen sekä lanseeraamaan uusi visuaalinen analyysijärjestelmä myös niiden analyttikkojen tuotantokäyttöön, jotka eivät osallistuneet tutkielmaan liittyvään järjestelmänkehitysprojektiin.

## 6.3 Prosessin arviointi

Tutkielman avulla pystyttiin löytämään keinoja, jotka auttoivat osaltaan selkiyttämään visualisointia sekä helpottamaan sosiaalisen verkoston analysointia. Tutkielma valmistui tavoitellussa aikataulussa. Kehitettävääkin tosin löytyi; esimerkiksi projektiin osallistuneille järjestelmän loppukäyttäjille olisi voitu antaa pidempi aika käyttää järjestelmää ennen haastatteluiden aloittamista, jotta he olisivat osanneet ottaa vahvemmin kantaa kaikkiin haastatteluissa käsitel-

tyihin teemoihin. Aikataulun venyttäminen olisi tosin viivästyttänyt koko tutkielman valmistumista suunnitellussa aikataulussa. Myös haastateltavien määrä oli suhteellisen pieni, mikä omalta osaltaan voi haitata tulosten yleistettävyyttä. Toisaalta haastateltavat edustivat noin kolmannesta kaikista järjestelmän todellisista loppukäyttäjistä, eikä projektiin osallistuneita, kaikkia haastateluun osallistumisen vaatimuksia täyttäneitä loppukäyttäjiä ollut kyseistä viittä loppukäyttäjää enempää.

Tutkielmassa esitettyjen linkkien ja solmujen vähentymistä kuvaavat prosenttiarvot voivat olla hankalasti hahmotettavia eri tilanteissa, koska järjestelmään haluttiin jättää transaktioita kuvaavien tapahtumien osalta myös vanhan visualisointitavan linkit ja solmut. Järjestelmän loppukäyttäjät pystyivät valitsemaan transaktiotiedon visualisoinnin kahdesta eri vaihtoehdosta, uuden visualisointitavan ollessa oletusarvona. Edellisestä ominaisuudesta johtuen järjestelmässä oli valtavasti ”ylimääräisiä” solmuja ja linkkejä, jotka vaikeuttivat tarkkojen prosenttiarvojen laskemista.

## 6.4 Jatkotutkimuksen aiheita

Hyvänä jatkotutkimuksen aiheena tälle tutkielmalle olisi selvittää sitä, miten sosiaalisen verkon visuaaliset elementit tulisi suunnitella, jotta niiden avulla pystyttäisiin esittämään sosiaalisten verkkojen analyysissä käytettäviä solmujen keskeisyyttä (engl. centrality) kuvaavia lukuja mahdollisimman selkeässä muodossa. Freeman (1979) esitteli artikkelissaan kolme tällaista lukua: 1) asteluku (engl. degree), joka kuvaa paikallista keskeisyyttä verkossa, 2) läheisyysluku (engl. closeness), joka kuvaa laajemmin keskeisyyttä, ottaen huomioon myös solmujen väliset etäisyydet, sekä 3) välisyyttä kuvaava luku (engl. betweenness), joka mittaa sitä, miten keskeinen kyseinen solmu on muiden solmujen ja niiden ryhmien välillä. Tutkielman laatimisen aikana kokeiltiin visualisointiohjelmistosta löytyviä keskeisyyttä kuvaavia ominaisuuksia ja erityisesti asteluvun käyttämistä niin, että linkkien lukumäärät korvattiin transaktioiden molemmille suunnille lasketuilla yhteissummilla. Ominaisuutta ei ehditty tutkia ja testata laajemmin, mutta sosiaalisten verkkojen analysoinnissa käytetyissä menetelmissä vaikuttaisi olevan paljon potentiaalia transaktiotiedon visuaalisessa analysoinnissa.

## LÄHTEET

- Choi, T. Y., & Kim, Y. (2008). Structural embeddedness and supplier management: A network perspective\*. *Journal of Supply Chain Management*, 44(4), 5-13.
- Cook, K. A., & Thomas, J. J. (2005). Illuminating the path: The research and development agenda for visual analytics (No. PNNL-SA-45230). *Pacific Northwest National Laboratory (PNNL)*, Richland, WA (US).
- Didimo, W., Liotta, G., & Montecchiani, F. (2014). Network visualization for financial crime detection. *Journal of Visual Languages & Computing*, 25(4), 433-451.
- Ekström, J., Engelson, L., & Rydergren, C. (2009). Heuristic algorithms for a second-best congestion pricing problem. *NETNOMICS: Economic Research and Electronic Networking*, 10(1), 85-102.
- Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social networks*, 1(3), 215-239.
- Gross, J. L., Yellen, J., & Zhang, P. (2014). *Handbook of graph theory* (2. painos). CRC press.
- Hanneman, R. A., & Riddle, M. (2005). Introduction to social network methods. Riverside, CA: University of California, Riverside. Haettu 7.4.2015 osoitteesta [http://www.researchgate.net/profile/Robert\\_Hanneman/publication/235737492\\_Introduction\\_to\\_social\\_network\\_methods/links/0deec52261e1577e6c000000.pdf](http://www.researchgate.net/profile/Robert_Hanneman/publication/235737492_Introduction_to_social_network_methods/links/0deec52261e1577e6c000000.pdf)
- Hirsjärvi, S., & Hurme, H. (2014). *Tutkimushaastattelu: teemahaastattelun teoria ja käytäntö*. Gaudeamus Helsinki University Press.
- Hirsjärvi, S., Remes, P., & Sajavaara, P. (1997). *Tutki ja kirjoita*. Tammi.
- Hsu, C. C. (2006). Generalizing self-organizing map for categorical data. *Neural Networks, IEEE Transactions*, 17(2), 294-304.
- Keim, D. A. (2002). Information visualization and visual data mining. *Visualization and Computer Graphics, IEEE Transactions on*, 8(1), 1-8.
- Keim, D. A., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). *Visual analytics: Definition, process, and challenges* (s. 154-175). Springer Berlin Heidelberg.
- Keim, D. A., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006, July). Challenges in visual data analysis. In *Information Visualization, 2006. IV 2006. Tenth International Conference on* (s. 9-16). IEEE.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Moreno, J. L. (1934). *Who shall survive?: A new approach to the problem of human interrelations*. Nervous and Mental Disease Publishing Co.
- Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2), 167-256.

- Peppers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
- Scott, J. (2012). *Social network analysis* (3. painos). Sage.
- Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly*, 35(1), 37-56.
- Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on* (s. 336-343). IEEE.
- Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27(3), 431-432.
- Stone, M. (2006). Choosing colors for data visualization. *Business Intelligence Network*.
- Thomas, J. J., & Cook, K. A. (2006). A visual analytics agenda. *Computer Graphics and Applications, IEEE*, 26(1), 10-13.
- Ware, C. (2012). *Information visualization: perception for design*. Elsevier.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.
- Westphal, C. (2008). *Data Mining for Intelligence, Fraud & Criminal Detection: Advanced Analytics & Information Sharing Technologies*. CRC Press.



## LIITE 1 HAASTATTELUN MONIVALINTAKYSYMYKSET

Uusi visualisointitapa nopeuttaa analysointia.

1 | 2 | 3 | 4 | 5 | 6

Uusi visualisointitapa helpottaa analysointia.

1 | 2 | 3 | 4 | 5 | 6

Uusi visualisointitapa nopeuttaa ilmoitusten esikäsittelyä.

1 | 2 | 3 | 4 | 5 | 6

Uusi visualisointitapa vähentää manuaalista työtä.

1 | 2 | 3 | 4 | 5 | 6

Uusi visualisointitapa tekee visualisoinnista selkeämpää.

1 | 2 | 3 | 4 | 5 | 6

Uuden visualisointitavan takia ei menetetä analysoinnin kannalta oleellista tietoa.

1 | 2 | 3 | 4 | 5 | 6

Uuden visualisointitavan graafiset elementit, kuten kuvakkeet ja linkit, ovat selkeitä.

1 | 2 | 3 | 4 | 5 | 6

Uuden visualisointitavan graafisten elementtien värien valinnassa on onnistuttu.

1 | 2 | 3 | 4 | 5 | 6

Sektorimalli nopeuttaa alustavien johtopäätösten tekemistä.

1 | 2 | 3 | 4 | 5 | 6

Sektorimallista on hyötyä koko analyysiprosessin ajan.

1 | 2 | 3 | 4 | 5 | 6

Sektorimalli nopeuttaa kokonaisuuden hahmottamista.

1 | 2 | 3 | 4 | 5 | 6

Linkkien määrän vähentyminen helpottaa analysointia.

1 | 2 | 3 | 4 | 5 | 6

Kuvakkeiden määrän vähentyminen helpottaa analysointia.

1 | 2 | 3 | 4 | 5 | 6

Transaktiotapahtumien kuvaamiseen riittävät suorat linkit osallisten välillä ja vanhan malliset transaktiotapahtumat voidaan poistaa.

1 | 2 | 3 | 4 | 5 | 6

Uusi visualisointijärjestelmä voidaan ottaa koko yksikön käyttöön nykyisellään.

1 | 2 | 3 | 4 | 5 | 6

Vastausvaihtoehdot:

1. täysin eri mieltä
2. jokseenkin eri mieltä
3. neutraali kanta
4. jokseenkin samaa mieltä
5. täysin samaa mieltä
6. en osaa sanoa