

# Kvalitatiivisen aineiston tietokoneavusteisen analyysin lähtökohtia

TIMO MOILANEN

## ABSTRACT

Towards computer-assisted qualitative data analysis

This article argues that computer-assisted qualitative data analysis is not an independent method or approach. It is often assumed that the use of computers makes qualitative research more exact, more reliable and more valid. These beliefs are correct to some extent: not because software can perform the analysis, but because software makes the process of analysis easier to manage. The main analytic operations in computer-assisted qualitative data analysis are coding, retrieving, writing memos and exploring the relations between objects. However, one cannot just use the analytic operations randomly. It is crucial to use the operations in a manner which corresponds to the method one is following. For example, in *grounded theory* one applies the same analytic operations in a totally different way than in content analysis. We will examine how some Finnish studies have used the ATLAS/ti program and its features, and conclude by discussing the implications of computer-aided qualitative research.

## Tietokoneavusteisen analyysin taustaa

Ensimmäiset kvalitatiivisen aineiston analyysiin tarkoitetut ohjelmat olivat tutkijoiden itsensä laatimia pieniä sovelluksia, joilla tutkijat etsivät ratkaisua eteen tulleisiin käytännön ongelmiin.<sup>1</sup> Varsinainen tietokoneavusteinen analyysi alkoi kehittyä 1980-luvun puolivälissä, jolloin ensimmäiset varsinaiset analyysiohjelmat tulivat tutkijoiden saataville. Ohjelmat perustuivat pitkälti kahden operaation, aineiston koodauk-

sen ja koodatun aineiston hakemisen varaan. Lisäksi ohjelmissa oli tyypillisesti pari muuta toimintoa, joista yleisimpiä olivat merkkijonoperustainen etsintä sekä taustamuuttujien käyttömahdollisuus. Alan pioneerina voidaan pitää John Seidelin vuonna 1985 kehittämää Ethnograph-ohjelmaa<sup>2</sup>, minkä jälkeen markkinoille tuli muitakin ohjelmia, kuten hieman paremman käyttöliittymän omaava Textbase Alpha. Toisen sukupolven ohjelmat perustuvat edelleen pitkälti koodaukseen ja hakemiseen, mutta ne ovat vieneet operaatioiden teknisen toteutuksen aiempaa pidemmälle. Uusimpiin ohjelmiin on rakennettu toimintoja, joiden avulla voidaan kirjoittaa memoja ja tarkastella

<sup>1</sup> Tutkimusprosessin kuluessa on edelleenkin tarvetta käsitellä aineistoa, esimerkiksi lyhentää rivipituutta, mikä voidaan järkevästi automatisoida käyttämällä esimerkiksi tekstinkäsittelyohjelmien makrotoimintoja. Toimintojen avulla voidaan luoda suhteellisen monimutkaisiakin sovelluksia, kuten WP 5.1:n makrotoimintojen pohjalta kehitetty suomalainen WP<sup>index</sup>-ohjelma (ks. Sulkunen 1990; Sulkunen ja Kekäläinen 1992).

<sup>2</sup> Ks. ohjelman kehittäjän kirjoitus siitä, mitä tapahtuu, kun ohjelmaa aletaan käyttää tarkoituksiin, joihin sitä ei ole suunniteltu (Seidel 1991).

koodien tai muiden objektien välisiä suhteita. Tällaisia ohjelmia ovat mm. ATLAS/ti ja NUD•IST (ohjelmien luokittelusta ja kehityksestä tarkemmin ks. Lonkila 1993; Moilanen – Roponen 1994; Weitzman – Miles 1995).

Varsinaisten analyysiohjelmien kehityksen ajoittuminen 1980-luvulle ei ole sattumaa. Tilasto-ohjelmiahan on käytetty kvantitatiivisen aineiston analyysissä jo 1950-luvulta alkaen, ensimmäisten keskustietokoneiden kehityksen myötä. Syyt tietokoneiden eritahtiseen soveltamiseen voidaan löytää kvalitatiivisen ja kvantitatiivisen aineiston ja analyysiprosessin eroista. Ensinnäkin kvantitatiivisen aineiston keräyksessä havainnot ovat vastaajien itsensä mahdollisimman yksiselitteisesti tuottamia ja koottu havaintomatriisiin, kun taas kvalitatiivinen aineisto, kuten haastattelussa kerätty informaatio, vaatii enemmän tulkintaa, jotta saadaan tutkimuksen näkökulmasta relevantteja havainnot. Toiseksi havainnot voivat havaintomatriisin systemaattisesta järjestyksestä poiketen sijaita missä kohden haastattelua tahansa, mikä tekee aineiston hallinnasta vaikeaa. Tutkija joutuikin tulkitsemalla *tuottamaan havainnot*<sup>3</sup> jokaisesta tapauksesta erikseen, minkä vuoksi tietokonetta tarvitaan koko pitkän analyysiprosessin ajan. Kvantitatiivinen aineistohan pyritään pääsääntöisesti esikoodaamaan jo aineiston käsittelyvaiheessa, milloin analyysivaiheessa tehtävät ajot voidaan suorittaa nopeasti (tosin analyysivaiheessa voidaan joutua vielä yhdistelemään muuttujien luokkia). Kvalitatiivisessa tutkimuksessa havaintojen tuottaminen tapahtuu sen sijaan vasta aineiston analyysivaiheessa, mistä johtuen tietokonetta tarvitaan koko analyysiprosessin ajan. Keskuskoneiden aikakaudella tälle ei ollut kunnollisia teknisiä edellytyksiä, ja vasta edullisten henkilökohtaisten tietokoneiden (PC) yleistyminen 1980-luvun alussa teki kvalitatiivisen aineiston tietokoneavusteisen analyysin mahdolliseksi.

Tapahtunutta kehitystä luonnehtii ohjelmien monipuolistuminen ja ohjelmakehityksen ammattimaistuminen, ja kasvavat markkinat ovat

tehneet siitä enenevästi liiketoimintaa (esimerkiksi ATLAS/ti ja NUD•IST ovat nykyään Sage-kustantamon ohjelmistomyynnin<sup>4</sup> piirissä). Ammattimaistuminen liittyy paitsi ohjelmien toimintojen monipuolistumiseen myös käytöjärjestelmien, käyttöliittymien ja ohjelmointityökalujen muuttumiseen yhä monipuolisemmiksi ja monimutkaisemmiksi. Samaan aikaan ohjelmien kehityksen kanssa on niiden käyttö laajentunut ohjelmoijien itsensä parista tutkijoiden arkipäivän työvälineeksi. Herääkin kysymys, missä määrin ohjelmia käytetään ja mitä niillä tehdään?

### Tietokoneavusteisia menetelmiä koskeva tutkijakeskustelu

Tietokoneavusteinen analyysi on noussut aktiivisen kansainvälisen tutkijakeskustelun kohteeksi varsinaisten analyysiohjelmien kehityksen myötä. Alan perusteoksina voidaan pitää Renata Teschin kirjaa *Qualitative Research: Analysis Types and Software Tools* (1990) sekä Fieldingin ja Leen toimittamaa teosta *Using Computers in Qualitative Research* (1991), mutta nämä teokset ovat erityisesti ohjelmia koskevilta osiltaan vanhentuneita. Uusista julkaisuista tarkastelemisen arvoisia ovat Udo Kellen toimittama teos *Computer-Aided Qualitative Data Analysis: Theory, Methods and Practice* (1995) sekä Fieldingin ja Leen uusin teos *Computer Analysis and Qualitative Research* (1998). Ohjelmien ominaisuuksia ja käyttöä koskeva tieto vanhenee kuitenkin muutamassa vuodessa ja löytyy parhaiten internetistä.<sup>5</sup>

Suomessa tietokoneavusteista analyysia koskeva keskustelu on ollut melko vähäistä. Keskustelua on vaivannut ohjelmia empiirisen aineiston analyysissä soveltaneiden tutkijoiden kriittisen massan vähyys. Kun ammattitutkijat eivät juurikaan ole omaksuneet uusien analyysitekniikoiden käyttöä (ohjelmien käytöstä ks. taulukko 1), ei myöskään laajemmalle tutkija-

<sup>3</sup> Alasuutarin (1993) käyttämä ilmainen havaintojen tuottamisesta tuo tutkimusotteiden teoriasidonnaisuuden paremmin esille kuin passiivisemmat ilmaiset havainnoinnista tai reaktiivisen mielikuvan antava ilmainen havaintojen vastaanottaminen.

<sup>4</sup> Lisätietoja [www-sivulta http://www.scolari.co.uk](http://www.scolari.co.uk).

<sup>5</sup> Ohjelmaesittelyitä löytyy esimerkiksi CAQDAS-verkostoprojektin sivuilta osoitteesta <http://caqdas.soc.surrey.ac.uk>. Ohjelmien valintakriteereistä ks. Weitzman ja Miles (1995), Creswell (1998), Alexa ja Zuell (2000).

## Taulukko 1. Suomessa tehdyt empiiriset tietokoneavusteiset kvalitatiivisen aineiston analyysit.

	Pro gradu-työt	Lisens. työt	Väitökset	Muut	Yhteensä
ATLAS/ti	11	–	2	2	15
WP <sup>index</sup>	6	1	–	2	9
NUD•IST	2	–	2	5	9
Ethnograph	1	–	–	–	1
FINTWOL	–	1	–	–	1
HyperResearch	–	–	1	–	1
IZE	–	–	1	–	1
OCP	1	–	–	–	1
Textbase Alpha	1	–	–	–	1
Yhteensä	22	2	6	9	39

keskustelulle ole ollut edellytyksiä, ja julkaisut ovat usein olleet kirja- tai ohjelmaesittelyitä. Kotimaisista julkaisuista voidaan mainita Moilasan ja Roponen (1994) ATLAS/ti:tä koskeva menetelmäraportti, Eskolan (1995) toimitama teos ohjelmien käyttökokemuksista sekä Luomasen ja Räsänen (2000) erityisesti QSR NVivo -ohjelmistoa koskeva teos.<sup>6</sup>

### Ohjelmien todellinen käyttölaajuus

Kvalitatiivisen aineiston tietokoneavusteinen analyysi on Suomessa suhteellisen tuore ilmiö. Ohjelmien soveltaminen on levinnyt erityisesti 1990-luvun puolivälissä useille tieteenaloille ja eri yliopistoihin, mutta samalla myös ohjelmista saatu käyttökokemus on pirstaloitunut ja vaikeasti löydettävissä. Esimerkiksi kirjastojen käytössä ei ole mitään yhteisiä avainsanoja tai luokittelukriteereitä, joiden avulla analyysiohjelmien käyttö olisi mahdollista paikallistaa. Alla oleva taulukko perustuu keväällä 1997 lumipallo-otannalla tekemääni kartoitukseen jota on päivitetty sitä mukaa kun tietoa uudesta julkaisusta on saapunut.<sup>7</sup> Kartoituksen

tulokset ovat luonteeltaan suuntaa-antavia, koska kaikkia asianosaisia ei tavoiteta ja kaikki eivät halua tai ehdi vastata kyselyyn.

Kyselyn perusteella Suomessa tehdyt empiiriset tutkimukset näyttäisivät koostuvan pääosin pro gradu-tason tutkielmista, ja varsinaiset ammattitutkijat eivät tämän kartoituksen perusteella ole ottaneet ohjelmia omakseen. WP<sup>index</sup>-ohjelmalla on oma käyttäjäkuntansa, joka koostuu pääasiassa sosiologeista. ATLAS/ti:n käyttäjäkunta on paljon heterogeenisempi, ja ohjelmaa sovelletaan useilla eri tieteenaloilla (valtiooppi, sosiologia, käsityötiede, kansantaloustiede, sosiaalipsykologia, viestintä, psykologia). NUD•IST:in käyttäjäkunta on myös melko monipuolinen, ja edellisten lisäksi sieltä löytyy erityisesti kasvatustieteilijöitä. Muiden ohjelmien käyttö on ollut harvinaisempaa, mutta mainittakoon että Ethnographin kohderyhmään kuuluu kansatieteilijöitä, FINTWOL ja OCP on suunnattu kielitieteilijöille, HyperResearch on Macintosh-ohjelma ja Textbase Alpha vanhentunut ensimmäisen sukupolven analyysiohjelma.

Ohjelmavalinta ei kuitenkaan vielä kerro siitä, minkälaisia tutkimusotteita töissä on käytetty. Tarkastellaan lähemmin ATLAS/ti-ohjelmalla tehtyjä pro gradu -tutkielmia (tauluk-

<sup>6</sup> Poliitiikan tutkijoista ohjelmia ovat soveltaneet Moilanen (1993) ja Aarnio (1994, 1996). Moilanen käyttää ATLAS/ti-ohjelmaa empiirisen käsiteanalyysin apuna, kun taas Aarnion tarkastelu tapahtuu retoriikan ja käsitetutkimuksen näkökulmasta soveltaen kielitieteellisiä ohjelmistoja (Oxford Concordance Program, Textpack, Fintwol).

<sup>7</sup> Lähetin ohjelmien soveltamista koskevan sähköpostiviestin noin kolmellekymmenelle henkilölle, joi-

den tiesin itse joko käyttävän ohjelmia tai tuntevan toisia ohjelmien käyttäjiä. Pyysin vastaajia lähettämään tietämänsä viitetiedot minulle sekä levittämään viestiä kaikille niille, jotka ovat käyttäneet ohjelmia. Ajantaisaiset luvut ja kattavat viitetiedot on päivitetty sivulle <http://www.helsinki.fi/~tmoilane/atk>.

**Taulukko 2. Aineiston, menetelmien ja ohjelman käytön dokumentointi ATLAS/ti:ta soveltaneissa pro gradu -tutkielmissa.**

Tekijä	Haastatteluita	Koodeja	Analyysimenetelmä, taustateoria	Käytön dokumentointi
Erkkilä	11	26	vertailuteoria	jonkin verran
Kaipainen	10	~ 200	etnograf. sisällönanalyysi ja gt	runsas
Luoma	25	?	fenomenologinen	vähäinen
Moilanen	12	92	empiirinen käsiteanalyysi	runsas
Nikkonen	38	59	gt	jonkin verran
Nilson-Niemi	15	?	gt, fenomenologia	vähäinen
Salovaara	15	?	gt, kulttuurinen lähestymistapa	runsas
Saukkonen	5	544	gt, sisällönanalyysi	runsas
Vuorinen	10	?	diskurssianalyysi	vähäinen

ko 2). Ohjelmaa ei ole räätälöity yksinomaan politiikan tutkimuksen tarpeisiin (jos sellaisia metodisesti edes on), vaan sitä on sovellettu useilla eri tieteenaloilla tarkasteltaessa mm. alkoholistien selviytymisstrategioita, rintasyöpäpotilaiden elämänlaatuun vaikuttavia tekijöitä, virkamieseliitin käsityksiä politiikasta jne. Kaikissa töissä ohjelmalla oli analysoitu haastatteluaineistoa, ja useissa töissä oli kerätty myös muuta erikseen analysoitavaa aineistoa (täydentäviä kyselylomakkeita, sosiogrammeja, elämänkaarikaavioita). Käytettyjen koodien lukumäärä vaihteli suuresti 26:n ja 544:n välillä.

On ilmeistä, että lähilukuna tarkasteltavan määrällisesti rajallisemman aineiston analyysissa käytetään runsaammin koodeja kuin laajemman aineiston analyysissa. Mitä suuremmaksi aineisto käy, sitä vähemmän tutkija voi kiinnittää huomiota yksityiskohtiin (syvyys laskee). Käytetyin menetelmä oli *grounded theory* (gt), mutta sitä käytettiin lähinnä kuvaamaan koodausprosessin luonnetta ja kulkua. Teoreettiseen otantaan, saturaation ja memojen kirjoittamiseen töissä ei oltu juurikaan kiinnitetty huomiota. Raportoinnin osalta kolmessa työssä oli huomattavasti muita vähemmän tarkasteltu ohjelman käyttöä, koodien muodostamista ja analyysin konkreettista kytkemistä aineistoon. Tämä voi johtua siitä, että kyseisissä töissä käytetyt menetelmät (fenomenologia, diskurssianalyysi) eivät lähde koodauksesta liikkeelle ja niiden toteuttaminen tietokoneavusteisesti ei ole aivan yksinkertaista.

### Teoreettiset taustasitoumukset

Kvalitatiivinen aineisto on siitä erikoista, että kaikkien samassa kulttuurissa elävien on helppo ”ymmärtää” sitä. Ymmärtäminen tapahtuu julkilausumattomien arkipäiväisten tulkintäsääntöjen perusteella. Ne ovat kulttuurin jäsenille itsestään selviä ja niiden aukikirjoittaminen on loputon tehtävä<sup>8</sup>. Tämän johdosta monet kuvaavat analyysivaihetta intuitiiviseksi, välittömän tajunnan tasolla eteneväksi prosessiksi (vrt. Ehrnrooth 1990). Intuitiivinen analyysi käyttääkin hyväkseen kulttuurin tarjoamia resursseja (esiymmärrys), mutta ei ole tietoinen näistä resursseista tai ei kiinnitä niihin huomiota. Intuitiivinen tarkastelutapa havaitsee sen, mitä kulttuurissa havaitaan, mutta se ei kykene kunnolla havaitsemaan itseään tai omaa tapaansa havaita. Sikäli kun tieteellinen tutkimus problematisoi näitä arkipäivän itsestäänselvyyksiä ja tavoittelee mielenkiintoisia näkökulmia, sen täytyy kyetä ylittämään intuitiivinen tulkinta. Kriittinen tarkastelu tarvitsee tuekseen ennen kaikkea teoreettisen viitekehityksen ja siihen kytkeytyvän erottelukykyisen käsitteistön, jonka avulla aineistosta voidaan havaita aivan uusia piirteitä ja suhteuttaa se osaksi jotakin kokonaisuutta. Tämä tekee kvalitatiivisen aineiston analyysista vaativan tehtävän, koska se

<sup>8</sup> Näiden sääntöjen tutkiminen kuuluu yleisesti sosiologian, erityisesti etnometodologian tutkimuskohteeseen.

nimenomaan edellyttää tutusta aineksesta irtaantumista ja etäisyyden ottamista.

Tietokoneavusteisesta analyysistä ei ole hyötyä, mikäli tarkempi menetelmän ja substanssin hallinta puuttuu. Tietokoneavusteinen analyysi tulee selkeästi kytkeä johonkin näkökulmaan ja menetelmään, jotta yksittäiset analyysioperaatiot olisivat mielekkäitä. Ohjelmat eivät ole mitenkään neutraaleja työkaluja, vaan niiden taustalta löytyy oletuksia kvalitatiivisen tutkimuksen luonteesta. Voimakkain vaikutus ohjelmien kehitykseen on ollut *grounded theory*-metodilla, joka on yleismenetelmän maineessa (Lonkila 1995). *Grounded theoryssa* eksplikoidaan monia kvalitatiivisen tutkimuksen peukalosääntöjä, mm. aineiston otannan ja koodauksen suhteen, jotka ovat samankaltaisia myös muissa laadullisissa menetelmissä. Ohjelmat toteuttavatkin varsin hyvin *grounded theoryn* keskeisiä analyysioperaatioita. Kuitenkin tietokoneavusteisen analyysin operaatiot – koodaus, hakeminen, memot ja suhteet – ovat täysin riippuvaisia valitusta tutkimusmenetelmästä. Lähtökohtaiset erot tulevat hyvin jäljempänä esiin verrattaessa lyhyesti kahta pitkälti koodauksen perustuvaa menetelmää, *grounded theorya* ja sisällönanalyysia (sivulla 62). Erot korostuivat edelleen tarkasteltaessa menetelmiä, jotka eivät perustu pienten tekstisegmenttien koodaukseen. Esimerkiksi narratiivisessa luennassa voitaisiin tarkastella episodeja, jotka ovat kooltaan hyvin suuria ja siten vaikeita koodata. Vastaava pätee myös mm. retoriseen luentaan, jossa oleellista voi olla se, mitä ei sanota tai mikä oletetaan tiedetyksi. Tällainen lukutapa ”ylittää aineiston” siinä mielessä, että näitä kohtia ei edes voida koodata, koska niitä ei ole.

Tietokone onkin nähty vastauksena kvalitatiivisen aineiston raskaudelle, jolloin se voisi helpottaa rutiineja ja säästää aikaa. Tilanne on erityisen huolestuttava silloin, kun tutkimusprosessi on pitkällä ja työn metodinen ote on epäselvä. Kvale (1996) luonnehtii tilannetta puhumalla tuhannen sivun kysymyksestä: ”*How Shell I find a method to analyze the 1000 pages of interview transcripts I have collected*”. Kvalen mukaan koko kysymys on väärin asetettu, ja siitä voidaan löytää useita ongelmakohtia. Esimerkiksi kysymys menetelmästä esitetään aivan liian myöhään ja aineistoa on toden-

näköisesti kerätty liikaa suhteessa käytettävissä olevaan aikaan. Kvalen oivaltavasta artikkelista käy hyvin ilmi, että kvalitatiivinen tutkimus edellyttää onnistuakseen hyvää suunniteltua. Tutkija voi tosin olla herkempi aineiston suhteen ja suunnata tutkimusta enemmän kentältä tulevien vihjeiden suuntaan, mutta sekään ei voi tapahtua sattumanvaraisesti suunnittelematta. Kvale tarvitsee 17 sanaa käsittävän virkkeen perusteelliseen analyysiin 4000 sanaa. Mikäli tuhannen sivun aineistoa analysoidisiin vastaavalla tarkkuudella tarvittaisiin 235 000 sanaa, mikä tarkoittaisi käytännössä tuhannen kirjan kirjoittamista! Määrällisesti pienestä aineistosta voi saada paljon irti.

### **Tietokoneavusteinen analyysi ATLAS/ti-ohjelman avulla**

*Aineiston käsittely.* Analyysivaihetta edeltää käsittelyvaihe, jossa aineistoa valmistellaan analyysiä varten. Esimerkiksi haastattelemalla kerätty aineisto on tallennettu haastattelunauhille, jotka täytyy purkaa tekstiksi (litteroida), jotta aineistoa voitaisiin analysoida. Teknisesti haastattelunauhojen purku tapahtuu parhaiten erillisen purkulaitteen (polkimella ohjattava nauhuri) ja tekstinkäsittelyohjelman avulla. Sisällöllisesti litteraatio voidaan toteuttaa monilla eri tavoilla, joiden valinta riippuu käytettävästä tutkimusmenetelmästä. Mikäli tutkimuksessa on kiinnostuttu lähinnä faktoista, voidaan tyytyä osittaiseen, pelkästään haastatteluiden pääkohtien aukikirjoittamiseen vaikkapa ranskalaisia viivoja käyttäen. Useimmissa menetelmissä näin voimakkaasti tiivistävä litteraatio ei ole tarkoituksenmukaista, koska siinä menetetään paljon informaatiota. Esimerkiksi keskustelunalyysissa on tärkeää käsitellä aineisto siten, että aineisto sisältää mm. päällekkäin puhumiset, tauot ja niiden pituudet, äänen sävyt, painotukset jne. (ks. Koskinen ym. tässä lehdessä). Kvalitatiivisen aineiston käsittely onkin erittäin työlästä ja se vie paljon aikaa, joten on tutkimusekonomisesti tärkeää miettiä, millä tarkkuudella aineisto on tarpeen tekstuaalisoida.

ATLAS/ti for Windows ohjelmassa aineiston, analyysin ja tutkimustehtävän suhdetta jäsennetään omaperäisellä mutta hedelmällisellä tavalla. Ensinnäkin koko tutkimusprojekti (ai-

neisto ja analyysi) tallennetaan yhdeksi tiedostoksi, jota kutsutaan *hermeneuttiseksi yksiköksi*. Tutkimusaineistoja kutsutaan *primääriteksteiksi* tai dokumenteiksi, ja tutkijan analyysiä *sekundaaritekstiksi* (koodit, haut, memot, suhteet). Primääritekstit voidaan käsitellä tavallisilla tekstinkäsittelyohjelmilla (WP, Word), mutta aineisto täytyy tallentaa ATLAS/ti:n käyttämään muotoon. Tästä syystä mitään muotoilukomentoja (lihavointi, kursivointi, tyyli jne.) ei saa olla mukana, koska niiden toteutus on aina ohjelmakohtaista. Lisäksi ohjelmien tallennusmuodot muuttuvat ohjelmapäivitysten myötä lähes joka vuosi. Ratkaisu ongelmaan on tallentaa tiedostot pelkiksi ANSI (Windows) – tai ASCII (DOS) -merkistökoodausta käyttäviksi tekstitiedostoiksi. Ohjelma hyväksyy primääritekstiksi myös kuva- (esimerkiksi .jpg ja .avi) ja äänitiedostot (.wav ja .mp3), joskin näiden käyttö hakuoperaatioissa on rajoittunutta.

Tietokoneavusteinen analyysi edellyttää, että aineisto on tallennettu tiedostoiksi. Osa aineistosta kuten haastattelunauhut tulevat luontevasti käsittelyvaiheessa tekstualisoiduiksi, mutta monet muut kvalitatiiviset aineistotyytit, kuten päiväkirjat, tv- ja radio-ohjelmat, kuvat ja sanomalehtiaineistot, ovat vaikeasti tekstualisoitavissa. Tällöin on syytä punnita, onko tietokoneavusteisesta analyysistä työlle niin paljon hyötyä, että käyttäisi analyysin sijasta runsaasti aikaa aineiston käsittelyyn. Tilanne muuttuu kriittisemmäksi, jos tutkijalla ei ole aiempaa kokemusta kvalitatiivisen aineiston tietokoneavusteisesta analyysistä. Yhdistettynä aikataulupaineisiin kvalitatiivisen aineiston muutenkin vaativa analyysi voi ajautua mahdottomiin yhtälöihin, jos siihen lisätään vielä kokonaan uuden työtavan opettelu.

*Koodaus.* Koodausta ajatellaan tyypillisesti kvantitatiivisen tutkimusotteen kautta. Tällöin kyse on useimmiten aineiston esikoodauksesta (*precoding*), jossa vastaukset kvantifioidaan jatkotarkastelun helpottamiseksi. Koodaus suunnitellaan jo aineiston keruuvaiheessa ja toteutetaan käsittelyvaiheessa, kun kyselylomakkeen vastaukset merkitään havaintomatriisiin. Usein käytetään standardiksi muodostunutta Likert-asteikkoa, jossa vastaukset pisteytetään yhdestä viiteen<sup>9</sup>. Mikäli strukturoidun lomakkeen sijasta käytetään puolistrukturoitua haastatteluita, voidaan aineisto jälkeen päin koodata

(*postcoding*) kvantitatiiviseksi havaintomatriisiksi. Jälkikoodaus helpottaa aineiston käsitteilyä ja analyysia, mutta ei ole läheskään aina mielekästä, koska sen seurauksena tapahtuva aineiston reduktio (mm. aineiston narratiiviset elementit, haastateltavien omat käsitteellistämiset jne.) on liian suurta.

Myös tietokoneavusteisessa analyysissä koodaus (*coding*, myös *indexing*, *tagging*, *attributing*) on keskeisellä sijalla, mutta se tarkoittaa ja painottuu eri tavalla kuin kvantitatiivisen aineiston analyysissä. Kun kvantitatiivisen aineiston tapauksessa vastaaja antaa kyselylomakkeen ohjaamana jäsenneltyjä havaintoja itsestään, niin kvalitatiivisen aineiston kohdalla havainnot täytyy erikseen tulkinnan avulla nostaa esille. Aineistossa on tyypillisesti hyvin moniulotteista informaatiota niin haastattelijasta, haastateltavasta kuin haastattelutilanteestakin, jotka ovat keruutavasta johtuen hajautuneet haastattelun eri kohtiin. Kaikki informaatio ei ole yhden tutkimuksen näkökulmasta tärkeää, vaan tutkija joutuu valikoimaan ja tulkinnalla tuottamaan havaintoja oleellisista kohdista.

Koodauksessa runsaasta ja moneen suuntaan rönsyilevästä aineistosta nostetaan jotain *erityistä* esille. Niinpä koodauksen voi määritellä *tekstisegmentin merkitsemiseksi*, joka teknisesti koostuu (1) tekstisegmentin rajaamisesta ja sen (2) nimeämisestä koodilla. Esimerkiksi tekstin käsittelyohjelmalla tehty lihavointi tai kursivointi voidaan käsittää koodaukseksi, koska siinä on rajattu tekstialue ja merkitty se erityisellä tavalla (esimerkiksi lihavointi tarkoittaa yhteistyötä toimijoiden välillä, kursivointi riitaa jne.). Tällainen koodaus on kuitenkin kömpelöä, koska muotoilutoimintojen välityksellä käytössä oleva koodilista jää väistämättä kovin suppeaksi, minkä lisäksi muotoilutoiminnoilla merkittyä aineistoa on kömpelöä hakea ja jatkokäsittellä. Varsinaisissa analyysiohjelmissa koodaus on paljon joustavampaa, sillä koodien lukumäärä ei ole ennalta rajattu ja koodit voidaan nimetä huomattavasti monipuolisemmin (esimer-

<sup>9</sup> ”Kuinka paljon luotatte siihen, että maan hallitus ajaa oikeita asioita? Luotatteko siihen aina, useimmiten, vain joskus, vai tuskin koskaan? (1) Aina (2) Useimmiten (3) Vain joskus (4) Tuskin koskaan (8) Ei osaa sanoa” (Pesonen ja Sänkiäho 1979, 16, kyselylomakeliite).

kiksi numerolla ”601” tai käsitteellä ”nationalismi: positiivinen”). Koodausta on helppo korjata ja täydentää, ja koodien jatkokäsittely esimerkiksi hakutoimintojen avulla on teknisesti helppoa.

Koodauksen tekninen toteutus on sinänsä triviaali asia, mutta koodauksen lähtökohtia koskevat valinnat, koodien muodostaminen ja analyysiyksikön valinta, liittyvät kiinteästi työn metodiseen perustaan. *Grounded theory*-tyyppisessä tutkimuksessa koodit määritellään analyysiprosessin aikana. Toinen alkuperäisen tutkimusotteen kehittäjistä, Barney Glaser, on mennyt niin pitkälle, että hän haluaisi kieltää jopa suorien tutkimuskysymysten käytön kentätyössä (Glaser 1992, 25). Glaserin työparina toiminut Anselm Strauss lähti puolestaan kehittämään teoriaa hieman eri suuntaan, jossa huomioidaan havaintojen teoriapitoisuus ja sovelletaan aiempia teorioita ja malleja. Straussin mukaan koodaus on jatkuvasti tarkentuva prosessi, joka voidaan temaattisesti jakaa kolmeen vaiheeseen, avoimeen, aksiaaliseen ja selektiiviseen koodaukseen. Koodaus ei etene suoraan viivaisesti vaiheesta toiseen, vaan eri vaiheet tapahtuvat rinnakkain ja stimuloivat toisiaan (tarkemmin Strauss ja Corbin 1990, 61–142). Analyysiyksikön suhteen *grounded theory* ei ole tarkka, vaan koodauksen kohteena voi olla minkäläinen tekstisegmentti tahansa. Tutkimusotteeseen liittykin keskeisesti ajatus kategorioiden saturaatiosta: aineistoa kerätään lisää niin pitkään, että uuden aineiston analyysi ei tuo tutkimuskohteen kannalta uutta ainesta esille.

Sisällönanalyttisessä tutkimusotteessa koodit pyritään määrittelemään mahdollisimman kattavasti ennen analyysivaihetta, sillä koodauksen muuttuminen analyysin myötä tekisi koodista epäkoherentteja estäen luokkien väliset vertailut. Tämä ei taas *grounded theoryn* tapauksessa ole kriittistä, koska siellä ei haeta niinkään vertailtavuutta kuin saturaatiopistettä. Lisäksi vertailtavuus edellyttää yhtäläisen analyysiyksikön käyttöä, joista tyypillisimpiä ovat virke, väitelause ja kappale. Esimerkiksi vaaliohjelma voidaan jakaa väitelauseisiin, jotka kaikki koodataan erikseen. Näin voidaan vertailla vaikkapa joidenkin poliittisten ajatusten, kuten hyvinvointivaltiokeskustelun muutoksia, sekä eri puolueiden että eri ajanjaksojen välillä. Sisällönanalyysissä erkaannutaankin analyys-

sivaiheessa aineistosta koodien tasolle, ja vertailu tarkoittaa pitkälti koodifrekvenssien erojen tarkastelua tai aiheiden suhteellisten erojen tarkastelua faktorianalyysin avulla (esimerkiksi Borg 1964; Budge ym. 1987).

*Hakeminen.* Hakemista (*retrieval, decoding*) voidaan pitää koodauksen käänteistoimintona. Hakuja käytetään ensinnäkin koodauksen tarkentamiseen, jolla varmistetaan, että tulkinta on säilynyt johdonmukaisena läpi koko aineiston (samaa luokkaan kuuluvat tapaukset ovat samanlaisia). Tämä on tyypillistä *grounded theorylle*, jossa hakujen avulla muutetaan ja tarkennetaan käytettyä koodausta. Kun aineisto on mielekkäästi koodattu voidaan hakujen avulla tehdä päätelmiä koko aineistosta, esimerkiksi millaisia eroavaisuuksia haastatteluiden väliltä löytyy. Sisällönanalyysin kannalta voi olla mielekästä tarkastella, kuinka suuren osan jokin aihe vie suhteessa koko aineistoon. Ohjelmat mahdollistavat myös ehdolliset haut, mikä tekee eri osaryhmien vertailun mahdolliseksi.

ATLAS/ti-ohjelmassa on tavallisen koodatun aineiston hakemisen ohella lukuisia muita hakumahdollisuuksia. Ohjelmaa voi pyytää hakemaan koodien frekvenssien ohella myös koodien referenssit (montako riviä on koodattu). Ohjelmalla voidaan tehdä ns. semanttisia hakuja, jossa yläkoodiin kohdistettu haku tuo esille siihen kuuluvien muiden koodien alaan sijoittuvat segmentit. Esimerkiksi kategoriaan ”kansainväliset suhteet” kohdistettu haku voisi tuoda esille ”erityissuhteet” toisiin maihin, ”soitaan” ja ”rauhaan” liittyvät maininnat samoin kuin ”kansainvälisyydestä” keskustelemisen tai ”EU-suhteet”. Ohjelmalla voidaan tehdä ehdollisia hakuja, jolloin rajataan tarkastelu vaikkapa tiettyyn aikakauteen (esimerkiksi verrataan keskustelua eri vuosikymmeninä) tai tiettyjen toimijoiden (esimerkiksi hallituspuolueet, oppositiopuolueet) keskustelua. ATLAS/ti:ssä aineiston ryhmittely tapahtuu erityisten primääritekstiperheiden avulla, joissa yhteenkuuluvat dokumentit sijoitetaan samaan perheeseen.

*Memot.* Memot ovat analyysiprosessiin liittyviä muistiinpanoja. Perinteiset paperin marginaaliin tehtävät muistiinpanot ovat koodien ja memojen välimuotoja. Tietokoneavusteisessa analyysissä nämä operaatiot on selkeästi erotettu toisistaan. Myös memojen käyttö on menetelmäsidonnaista. Esimerkiksi sisällönanalyys-

sissä ei ole periaatteessa käyttöä memoille, koska koodisysteemi on määritelty jo ennen analyysivaihetta. Sen sijaan *grounded theoryssa* memojen kirjoittaminen on oleellinen osa analyysiprosessia. Memojen kirjoittaminen tapahtuu rinnan koodauksen kanssa ja sen tarkoituksena on helpottaa analyysin hallintaa ja nostaa sen abstraktiotasoa. *Grounded theoryssa* memot voidaan jakaa kahteen luokkaan, substantiaaliin ja teoreettisiin memoihin. Substantiaalisten memojen tarkoituksena on aineistoon liittyvien havaintojen spesifiointi, kun taas teoreettiset memot liittyvät aineistosta nousevan teorian spesifiointiin. Memojen tarkoituksena on yhdistää empiirisen aineiston analyysi ja tutkijan ideat, minkä seurauksena teoria kasvaa memoissa.

ATLAS/ti-ohjelmassa on hyvin monipuoliset välineet memojen käsittelyyn. Memoja voi kohdistaa yksittäisiin tai useisiin tekstisegmentteihin tai ne voi jättää ns. vapaiksi memoiksi ilman suoraa sidosta aineistoon. Tällaisia vapaita memoja voivat olla mm. analyysiprosessin etenemiseen liittyvät muistiinpanot, joissa kirjataan jo tutkimuksen analyysiosion tuloksia. Ohjelmassa on myös erillinen kommentoi-toiminto, joka on hyvin lähellä memoja. Kommentit voivat koskea niin hermeneuttista yksikköä kokonaisuutena, yksittäistä primääritekstiä, segmenttiä tai koodia. Ero memojen ja kommenttien välillä on siinä, että kommentit ovat kiinteästi sidottuja kohteeseensa (esimerkiksi koodiin kohdistuva kommentti, jossa annetaan koodille täsmällisempi sisältö), kun taas memot ovat joustavampia ja niitä voidaan tarkastella osana suhdeverkostoja.

*Suhteet.* Suhteiden tarkastelu on uusimpien ohjelmien mukanaan tuoma ominaisuus. Ensimmäisen sukupolven analyysiohjelmissa koodien välisiä suhteita ei ollut mahdollista tematisoida lainkaan, ja monissa toisen sukupolven ohjelmissa kyse on hierarkkisten verkostojen käytöstä (esimerkiksi NUD•IST-ohjelman puumalli). ATLAS/ti:n vahvuus on siinä, että se antaa työkalut tarkastella erilaisia suhteita eihierarkkisena verkostona. Niinkin itsestään selvältä vaikuttava seikka kuin tekstisegmentin koodaus pitää sisällään suhteen: segmentti kuuluu koodin alaan. Segmenttien ja koodien väliset suhteet ovat kiinteitä (eli niitä ei voi muokata) ja luonteeltaan kuuluu johonkin luokkaan-

tyyppisiä suhteita. Monet yläkoodin ja alakoodin väliset suhteet ovat vastaavanlaisia transitiivisia suhteita, esimerkiksi ”teknokratia” ja ”juristeria” ovat ”ajattelumalleja”. Haettaessa kaikki erilaiset ”ajattelumallit” tarkasteltaviksi saadaan myös ne segmentit, jotka on koodattu koodeilla ”teknokratia” ja ”juristeria”. Ohjelmassa on lukuisia muitakin suhdetyyppejä, kuten assosiaatio-suhde (symmetrinen), ominaisuus-suhde (epäsymmetrinen), vastakohtaisuus-suhde (symmetrinen), on osa jotain-suhde (transitiivinen) ja perinteinen syy-seuraus-suhde (transitiivinen). Näiden lisäksi tutkija voi itse määrittellä uusia suhdetyyppejä tarpeensa mukaan (esimerkiksi Saukkosen [1996, 44] käyttämä ”x ja y ruokkivat toisiaan”).

Paitsi että ohjelma mahdollistaa koodien välisten suhteiden määrittelyn, siinä on toimintoja erillisten verkostonäkymien muodostamiseksi ja tallentamiseksi. Yhdessä hermeneuttisessa yksikössä on helposti satoja koodeja, joten niitä ei ole mahdollista eikä mielekäästäkään tarkastella samanaikaisesti näytöllä. Tällöin voidaan käyttää ns. verkostonäkymiä (*Network view*), joihin voidaan koodien lisäksi tuoda segmenttejä, memoja ja primääritekstejä. Näin ohjelma mahdollistaa mielenkiintoisella tavalla erityisten käsittekarttojen rakentamisen.<sup>10</sup> Käsittekarttojen ideana on visualisoimalla tehdä kohde helpommin ymmärrettäväksi (käytetään myös termiä *mind-mapping*<sup>11</sup>). Kvalitatiivisen aineiston analyysissä koodien välisten suhteiden spesifiointi on aina tutkijan tehtävä. Kyse on koodien välisistä luonteeltaan sisäisistä (merkitys)suhteista, joiden merkitsevyyttä ei voida laskea, kuten kvantitatiivisen aineiston analyysissä tehdään (ulkoiset suhteet).

Verkostoilla on helppo visualisoida tutkimusta, mutta samalla niissä piilee suuria vaaroja. Koska suhteet eivät perustu ”objektiiviseen” laskemiseen vaan tutkijan systemaattiseen työskentelyyn, voidaan verkostojen avulla tuottaa helposti tutkimukselta näyttävää pseudotiedettä (tätä voi toisaalta ajatella myös siten, että verkostot tuovat tekijöiden virheet esille). Verkosto antaa helposti ylitarkan vaikutelman, koska siinä vaikeatkin käsitteet näyt-

<sup>10</sup> Käsittekarttoista tarkemmin ks. esimerkiksi <http://ksi.cpsc.ucalgary.ca/articles/ConceptMaps/>.

<sup>11</sup> Ajatuskartoista ks. Kaipainen 1996, 3.



tävät kehysten johdosta tarkan rajatuilta osana verkostonäkymää. Verkostojen arvioinnin kannalta on ongelmallista, mikäli samat koodit eivät pädekään kaikissa tapauksissa.<sup>12</sup> Tutkijan tulee olla systemaattinen ja varmistaa, ettei luo koko aineistoa koskevia verkostoja yksittäisten tapausten pohjalta. Mikäli aineistossa on tästä poikkeavia tapauksia, ne on nostettava erilleen ja selitettävä, miksi ne poikkeavat esitetystä mallista.

## Lopuksi

Perinteistä paperipohjaista työskentelyä ja tietokoneavusteista analyysia ei tule nähdä vastakkaisina. Kaikelle tietokoneen kanssa työskentelylle on ominaista, että näytöltä lukeminen on raskasta ja siltä on vaikea hahmottaa yli ruudun pituisia kokonaisuuksia. Aineistoa voidaan ja tulee tarkastella sekä paperilta että näytöltä, kunhan muistetaan päivittää tehdyt merkinnät myös ohjelmaan. Teknisesti ohjelmat ovat kehittyneet käyttäjäystävällisempään suuntaan, ja esimerkiksi ATLAS/ti:ssä tehtyjen koodausten näyttö toisessa marginaalissa parantaa ohjelman käytettävyyttä.

Tietokoneavusteinen analyysi antaa välineitä aineiston hallintaan. Tehdyt havainnot säilyvät tallessa, ja havaintojen hakeminen ja jatkotarkastelu onnistuvat ohjelmien monipuolisilla hakutoiminnoilla teknisesti ongelmitta. Tuhhat sivua aineistoa on aivan eri tavalla hallittavissa tietokoneella kuin paperilla – tosin eri asia on, tarvitaanko aina tuota tuhatta sivua. Ohjelmat näyttäisivät analyysioperaatioiden eksplikaation ja tekijän identifikaation johdosta helpottavan perinteistä kvalitatiivisen tutkimuksen arvioitavuuden ongelmaa. Tutkimusprosessi tulee läpinäkyvämmäksi, koska ohjelmien avulla aineistoon voi viitata yksinkertaisella ja teknisesti johdonmukaisella tavalla. Ohjelmat eivät välttämättä tee tutkimuksen toistettavuutta aiempaa helpommaksi, mutta niiden avulla on

helpompi seurata tutkijan päättelyä.<sup>13</sup> Periaatteessa ohjelmat antavat varsin hyvät mahdollisuudet myös usean tutkijan yhteistyöhön, mutta tätä ei ole vielä liiemmin käytetty.<sup>14</sup>

Ilmeisenä ongelmana ohjelmien soveltamisessa on uutuuden viehätysten synnyttämä harkakuva. Kuvitellaan, että ohjelma itsessään tarjoaisi ratkaisun tutkimusmenetelmällisiin kysymyksiin. Aivan liian usein käy kuitenkin niin, että työkalu alkaa muokata käyttäjänsä. Kaikki analyysiohjelmat perustuvat pitkälti koodaukseen, mutta koodaus ei sovi yhtä hyvin kaikkiin tutkimusotteisiin. Silmiin pistävää on sekin, että ohjelmien erityistoiminnot näkyvät tutkimuksissa. WP<sup>index</sup>:iä käyttävissä tutkimuksissa on ongelmasta riippumatta käytetty ratkaisuna ohjelman kontingenssianalyysiä, kun taas ATLAS/ti:tä soveltavissa tutkimuksissa on itseisarvona rakentaa verkostoja. Kvalitatiivinen aineisto tarjoaa runsaasti näkökulmia, joita tutkimusongelma ja viitekehys täsmentävät. Ongelmallista analyysi on tietysti silloin, kun tutkimuksen substanssi on epäselvä. Jos tutkija ei tiedä mitä tutkii, tulee analyysistä helposti toisistaan irrallisten teemojen luettelointia ja seilaamista.

Tietokoneiden menestys kvantitatiivisen aineiston analyysissa aiheuttaa suuria odotuksia myös kvalitatiivisen tutkimuksen suhteen. Tietokoneiden käytöstä ei kuitenkaan voi kvalitatiivisen aineiston analyysissä olla vastaavaa etua, koska aineiston analyysissä ei voida hyödyntää koneiden laskentakapasiteettia. Tilastollisen analyysin perustana olevat säännöt (esimerkiksi korrelaatiokertoimen laskenta) voidaan purkaa algoritmeiksi, joiden suorittamisessa tietokone on erittäin tehokas. Sen sijaan kvalitatiivisen aineiston analyysissä käytetyt tulkintasäännöt ovat pikemminkin periaatteita, joita on mahdoton kirjoittaa täysin eksplisiittiseen muotoon. Tarvitaan tutkijan tulkintaa, koska ainoastaan hänellä on hallussaan sellainen arkipäiväinen kontekstuaalinen tietämys, mitä il-

<sup>12</sup> Esimerkiksi Alasuutarin (1993, 23–37) kulttuurin tutkimuksellisessa laadullisessa tutkimuksessa on oleellista, että havainnot kyetään pelkistämään ja yhdistämään siten, että mahdolliset poikkeukset säännöistä kyetään erikseen selittämään (ns. absoluuttinen selittämisen malli). Kuitenkaan kaikissa kvalitatiivisen aineiston tutkimusotteissa tällaista vaatimusta ei ole.

<sup>13</sup> Esimerkiksi Saukkosen (1996, 46) työ hermeneuttisine yksiköineen on saatavissa kyseisen tutkimuslaitoksen arkistosta. Tämä ei tietenkään tietosuojasyistä ole aina mahdollista.

<sup>14</sup> Tutkijoiden välisestä yhteistyöstä ks. Ford ym. (2000).

man aineistoa ei voi ymmärtää (tietokone käsittelee aineistoa merkkijonoina, ei merkityksen omaavina semioottisina symboleina). Onkin kyseenalaista, voivatko ohjelmat helpottaa kvalitatiivisen aineiston analyysia niin paljon, että tutkimus merkittävästi nopeutuisi.

Ohjelmat tarjoavat apua analyysiprosessin tekniseen hallintaan, ei itse analyysiin. Koodit eivät synny itseksensä, vaan niiden muodostaminen vaatii hyvää perehtymistä aiempaan tutkimukseen ja tarkasteltavaan tutkimusaineistoon. Analyysiprosessin kuluessa koodit yleensä muokkautuvat siten, että niiden analyytisyyden aste kasvaa. Monesti koodaus jätetään ensimmäisellä tarkastelukerralla tarkoituksella epätarkaksi, mitä tarkennetaan, kun käydään aineisto uudelleen läpi. Esimerkiksi koodi ”hyvän hallinnon kriteerit” voi olla tarpeen eritellä useampaan alaluokkaan kuten ”hyvän hallinnon kriteerit: puolueettomuus”, ”hyvän hallinnon kriteerit: tehokkuus” jne., riippuen tietysti tutkimusongelmasta. Kvalitatiivinen aineisto antaa lukutaitoiselle tutkijalle runsaasti vinkkejä, jotka heijastuvat koodisysteemin pirstaloitumisena. Niinpä analyysissä on yleensä myös toisensuuntaista liikettä, liian analyytittimen koodaus palautetaan yleisempiin kategorioihin. Tässä mielessä analyysivaiheessa on kyse jatkuvasta tasapainotilan tavoittelusta aineiston tarjonnan ja tutkimusongelman kysynnän välillä.

On tärkeää muistaa, että kyseessä ei ole tietokoneperusteinen vaan tietokoneavusteinen analyysi. Analyysiohjelma ei analysoi tekstiä sen paremmin kuin tekstinkäsittelyohjelma kirjoittaa tekstiä. Analyysivaihe on vain yksi osa tutkimusta, ja elleivät muut vaiheet, kuten aineiston keruu, ole onnistuneet, ei analyysikään voi korjata tilannetta. Erityisesti täytyy pitää huolta siitä, että tietokoneavusteinen analyysi sopii työssä käytettyyn tutkimusotteeseen. Havaintojen tekeminen riippuu metodista, ja vasta metodi antaa ohjeet, miten analyysioperaatioita on mielekästä soveltaa. Kriittistä ajattelua ja substanssin hallintaa ei voi ostaa tietokonekaupasta.

## LÄHTEET

- Aarnio, Eeva. 1994. *Toimijakonseptiot suomalaisten puolueiden periaate- ja yleisohjelmissa. Näkökulma poliittisten kielten muutoksiin 1960-luvulta 1990-luvulle*. Lisensiaatintutkielma. Jyväskylän yliopisto, valtio-opin laitos.
- Aarnio, Eeva. 1996. Puolueohjelmat ja analyysiohjelmat. Teoksessa Kari Palonen ja Hilikka Summa (toim.), *Pelkkää retoriikkaa*. Tampere: Vastapaino.
- Alasuutari, Pertti. 1993. *Laadullinen tutkimus*. Tampere: Vastapaino.
- Alexa, Melina ja Cornelia Zuell. 2000. Text analysis Software: Commonalities, differences and limitations. The results of a Review. *Quality and Quantity* 34, 299–321.
- Borg, Olavi. 1964. *Suomen puolueideologiat. Periaateohjelmien sisältöanalyttinen vertailu sekä katsaus niiden historialliseen taustaan ja syntyprosessiin*. Vammala: WSOY.
- Budge, Ian, David Robertson ja Derek Hearl (toim.). 1987. *Ideology, Strategy and Party Change: Spatial Analyses of Post-war Election Programmes in 19 Democracies*. Cambridge: Cambridge University Press.
- Creswell, John. 1998. *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. London: Sage.
- Ehrnrooth, Jari. 1990. Intuitio ja analyysi. Teoksessa Klaus Mäkelä (toim.), *Kvalitatiivisen aineiston analyysi ja tulkinta*. Helsinki: Gaudeamus.
- Erkkilä, Tero. 2000. *Paikkatietojärjestelmät Itämeren alueen ympäristöyhteistyössä*. Pro gradu-tutkielma. Helsingin yliopisto, yleisen valtio-opin laitos.
- Eskola, Jari (toim.). 1995. *MAA KUTSUU? Kokemuksia kvalitatiivisten aineistojen mikroavusteisesta analyysistä*. Lapin yliopiston kasvatustieteellisiä julkaisuja C 9. Rovaniemi: Lapin yliopisto.
- Fielding, Nigel ja Raymond Lee (toim.). 1998. *Computer Analysis and Qualitative Research*. London: Sage.
- Fielding, Nigel ja Raymond Lee (toim.). 1991. *Using Computers in Qualitative Research*. London: Sage.
- Ford, Kate, Iddo Oberski ja Steve Higgins. 2000. Computer-Aided Qualitative Analysis of Interview Data: Some Recommendations for Collaborative Working. *The Qualitative Report* 4:3–4 [On-line serial].
- Glaser, Barney. 1992. *Basics of Grounded Theory Analysis*. Mill Valley: Sociology Press.
- Kaipainen, Minna. 1996. *Tilausompeluyrittäjä – toiminnan vahvuudet ja heikkoudet*. Pro gradu -tutkielma. Helsingin yliopisto, käsityötieteen laitos.
- Kelle, Udo. 1995. *Computer-Aided Qualitative Data Analysis: Theory, Methods and Practice*. London: Sage.
- Kvale, Steinar. 1996. The 1000 Page Question. *Qualitative Inquiry* 2:3, 275–284.
- Lonkila, Markku. 1993. *Tietokoneavusteinen kvalitatiivinen analyysi*. Lisensiaatintyö. Helsingin yliopisto, sosiologian laitos.
- Lonkila, Markku. 1995. Grounded Theory as an Emerging Paradigm for Computer-assisted Qualitative Data

- Analysis. Teoksessa Udo Kelle (toim.), *Computer-Aided Qualitative Data Analysis: Theory, Methods and Practice*. London: Sage.
- Luoma, Minna-Liisa. 1997. *Rintasyöpäpotilaan elämänlaadun keskeiset merkityssisällöt*. Pro gradu -tutkielma. Helsingin yliopisto, psylogian laitos.
- Luomanen, Jari ja Pekka Räsänen. 2000. *Tietokoneavusteinen laadullinen analyysi ja QSR NVivo -ohjelmisto*. Sosiologian tutkimuksia A 23. Turku: Turun yliopisto.
- Moilanen, Timo ja Seppo Roponen. 1994. *Kvalitatiivisen aineiston analyysi ATLAS/ti -ohjelman avulla*. Helsinki: Kuluttajatutkimuskeskus.
- Moilanen, Timo. 1993. *Virkamieseliitin käsitys politiikasta. Empiirisen käsiteanalyysin atk-avusteinen soveltaminen*. Pro gradu-tutkielma. Helsingin yliopisto, yleisen valtio-opin laitos.
- Nikkinen, Ahti. 1996. *Soitto soi ja viina virtaa. Aniskelutyöntekijöiden ja ravintolamuusikoiden näkemyksiä työtapakulttuurin muutoksista tanssiravintoloissa*. Pro gradu -tutkielma. Helsingin yliopisto, sosiaalipsykologian laitos.
- Nilson-Niemi, Eva. 1999. *Syöpäpotilaan puolison selviytyminen: fenomenologinen tarkastelu*. Pro gradu -tutkielma. Helsingin yliopisto, psykologian laitos.
- Pesonen, Pertti ja Risto Sänkiaho. 1979. *Kansalaiset ja kansanvalta. Suomalaisien käsityksiä poliittisesta toiminnasta*. Juva: WSOY.
- Salovaara, Inka. 1995. *Taikalaatikon takana. Analyysi presidenttiehdokkaiden 1994 ja poliittisten toimittajien televisiosuhteesta*. Pro gradu -tutkielma. Helsingin yliopisto, viestinnän laitos.
- Saukkonen, Kai. 1996. *Aloitetaan alusta? Selviytymisstrategiat osana alkoholismin itseään vahvistavaa prosessia*. Pro gradu -tutkielma. Helsingin yliopisto, sosiaalipsykologian laitos.
- Seidel, John. 1991. *Method and Madness in the Application of Computer Technology to Qualitative Data Analysis*. Teoksessa Nigel Fielding ja Raymond Lee (toim.), *Using Computers in Qualitative Research*. London: Sage.
- Strauss, Anselm ja Juliet Corbin, Juliet. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. London: Sage.
- Sulkunen, Pekka. 1990. *Ryhmähaastattelujen analyysi*. Teoksessa Klaus Mäkelä (toim.), *Kvalitatiivisen aineiston analyysi ja tulkinta*. Helsinki: Gaudeamus.
- Sulkunen, Pekka ja Olli Kekäläinen. 1992. *WP<sup>index 1.0</sup>. Kevyt ratkaisu raskaan työn ongelmiin*. Helsinki: Limes.
- Tesch, Renata. 1990. *Qualitative Research: Analysis Types and Software Tools*. New York: Falmer Press.
- Weitzman, Eben ja Matthew Miles. 1995. *Computer Programs for Qualitative Data Analysis: A Software Sourcebook*. London: Sage.
- Vuorinen, Tuija. 1997. *Sotilas vai nainen? Vapaaehtoista asepalvelusta suorittavien naisten sukupuoli-identiteetin diskursssianalyttinen tarkastelu*. Pro gradu -tutkielma. Helsingin yliopisto, sosiaalipsykologian laitos.