

Tekijä Anni Tolvanen		
Työn nimi Silminnähten neutraali? – Teoreettinen katsaus feministisemmän datavisualisaation tekoprosessiin		
Laitos Median laitos		
Koulutusohjelma Visuaalisen viestinnän muotoilun koulutusohjelma		
Vuosi 2019	Sivumäärä 78	Kieli suomi

Tiivistelmä

Informaatiomuotoilu on tiedon jäsentämistä ihmisen käyttöön sopivaan muotoon, ja sen rooli on korostunut informaatiohälyn lisääntymisen ja massiivisten, jatkuvasti kasvavien datamäärien myötä. Erityisesti datavisualisaatioilla on keskeinen rooli niin tieteellisten tutkimusaineistojen jäsentelyssä, tutkivassa journalismissa kuin markkinoinnissakin. Nämä tietojenkäsitteilyä ja visuaalisen viestinnän muotoilua yhdistellen valtavista aineistomääristä tuotetut esitykset ovat myös varsin vakuuttavia.

Intersektionaalisen feminismin mukaan inhimillinen toiminta on useiden historiallisesti kerrostuneiden, risteävien valtarakenteiden vyyhti. Niinkutsuttu datafeminismi pyrkii analysoimaan, tuomaan esiin ja purkamaan näitä rakenteita juuri datan keräämisen, käsittelyn ja visualisoinnin parissa. Ongelmana on ennen kaikkea, että nykyisin datan parissa työskentelee verrattain pieni, yksipuolinen ja etuoikeutettu joukko ihmisiä: on siis tehtävä huomattavasti töitä, jotta alan näkökulmat monipuolistuisivat eikä oman aikamme vinoumia muurattaisi osaksi tulevaisuuden informaatioinfrastruktuuria.

Kun valtava määrä dataa saatetaan siistiin ja selkeään muotoon, syntyy helposti viettelevä harha kokonaisuuden hahmottamisesta ja kaikkitietävyydestä. Kaikki tieto on kuitenkin sidoksissa sen sen tuottaneeseen näkökulmaan ja kontekstiin, joten datafeministit haluavat ravistella tätä visualisaatioiden neutraalia ja objektiivista auraa. He myös vaativat lisää tapoja ilmaista visuaalisesti aineiston ja esityksen osittaisuutta, siihen sisältyviä rajauksia sekä näkökulmasidonnaisuutta.

Tässä tutkielmassa käydään läpi datavisualisaation työprosessin vaihe vaiheelta pohtien ja esitellen, miten projektiin kuuluvia valintoja ja rajoituksia voitaisiin tehdä feministisemmin – siis läpinäkyvämmiin, kriittisempiin sekä aidosti monenlaisia näkökulmia huomioonottavasti. Koska fokus on visuaalisen viestinnän muotoilijan roolissa, keskitytään erityisesti visuaaliseen jäsennykseen: kontekstin ilmentämiseen, esitystapaan, graafiseen rehellisyyteen, epävarmuuden ja aukkojen kuvaamiseen – siis neutraalin vaikutelman purkamiseen rakenteellisesta selkeydestä tinkimättä.

Selvityksen perusteella vaikuttaa siltä, että kriittisestä otteestaan huolimatta intersektionaalisella feminismillä on paljon positiivista annettavaa tiedon visualisoinnille: se tarjoaa työkaluja valtarakenteiden ja vinoumien huomioimiseen sekä inspiraatiota niiden ilmentämiseen ja purkuun. Aiempaa moninaisempien näkökulmien esiintuonti visualisaatioissa on luontevaa, sillä alan klassikkokirjoittajatkin korostavat, että hyvä visualisointi yhdistelee paljon tietoa useista eri lähteistä. Visualisaation tärkein päämäärä, eli pyrkimys antaa mahdollisimman totuudenmukainen ja selkeä kuva käsiteltävästä aiheesta tarkentuu feminismin myötä: totuudenmukaisuuden vaatimus koskee myös esitettävän tiedon rajallisuuden ja kontekstisidonnaisuuden myöntämistä ja esittämistä. Näin saadaan aikaiseksi aivan kaikkien kannalta parempia visualisaatioita.

ANNI TOLVONEN
SILMINNÄHDEN
NEUTRAALI?
TEOREETTINEN
KATSAUS
FEMINISTISEN
DATAVISUALISAATION
TEKOPROSESSIIN

Taiteen kandidatin opinnäytetyö
Median laitos
Visuaalisen viestinnän muotoilun koulutusohjelma
Taiteiden ja suunnittelun korkeakoulu
Aalto-yliopisto
Kevät 2019

PHSöL-

1 Johdanto	3
2 Lähtökohdat	8
2.1 Datavisualisaatiot	8
2.2 Datafeminismi	13
3 Datavisualisaation työprosessi	25
3.1 Tarve ja tavoite	29
3.2 Konteksti ja kohderyhmä	31
3.3 Rajausta ja asiayhteys	33
3.4 Aineisto, tutkimus ja järjestäminen	37
3.5 Luonnostelu ja kokeilu	42
3.6 Graafinen rehellisyys	46
3.7 Epämääräisyyden esittäminen	48
3.8 Tuotanto ja viimeistely	56
3.9 Selkeys ja uskottavuus	61
3.10 Yhteenveto	66
4 Reflektio	68
Lähteet	74

1 TOHTO

Taipaleeni visuaalisen viestinnän muotoilun parissa alkoi aikoinaan halustani jäsentää ja muotoilla tietoa mahdollisimman ymmärrettävään ja vaikuttavaan muotoon, jotta voisin auttaa edistämään eettisesti ja ekologisesti kestävämmän yhteiskunnan kehitystä. Tiedon visualisointi onkin graafisen suunnittelun keskeisiä osa-alueita: sen kenttä ulottuu laajasti ymmärrettynä taitoista ja muista käyttöliittymistä pedagogisiin kuvituksiin.

Algoritmien ja esineiden internetin tuottamien valtavien datamäärien aikakauden sarastaessa erityisesti koneellisesti suurista aineistomääristä luotujen datavisualisaatioiden rooli on kohonnut keskeiseksi tieteessä ja sen popularisoinnissa, mutta myös journalismin ja markkinoinnin saralla. Datavisualisaatiot ovat paitsi voimakas tiedon jäsentelyn, myös vaikuttamisen keino: suurten aineistojen analysointi paljastaa rakenteellisia epäkohtia ja

syy-seuraussuhteita sekä voi tuoda jämäkästi esiin jonkin ongelman koko laajuuden.

Datan keruu, käsittely ja visualisointi ovat, kuten kaikki muukin inhimillinen toiminta, useiden risteävien valtarakenteiden ja dynamiikkojen vyyhti. Monen muun alan kehitystä seuranneen tavoin olen perin huolestunut siitä, että tulevaisuuden tietoteknistä infrastruktuuria juuri tälläkin hetkellä visioi, muotoilee, koodaa ja rakentaa verrattain pieni ja yksipuolinen ihmisjoukko – tämän voi päätellä paitsi IT-firmojen henkilöstöä kuvaavista tilastoista, myös heidän työnsä tuloksista, jotka tuon tuostakin ilmentävät ajattelemattomuutta ja suoranaisia laiminlyöntejä. Viimeistään nyt tarvitaan tietotekniseen ja viestinnälliseen osaamiseen linkkittynyttä empatiakykyä, kriittisyyttä ja vakaata sitoutumista moninaisten näkökulmien huomioimiseen muutenkin kuin markkinointitemppuna.

Haluankin tutkielmassani perehtyä siihen, mikä rooli valppaalla visuaalisen viestinnän muotoilijalla on voi olla datan parissa työskentelyssä: intersektionaalisena feministinä minulle on tärkeää tunnistaa ja mahdollisuuksieni mukaan purkaa niin omaan toimintaani kuin ympäröiviin systeemeihin sisältyviä haitallisia valtarakenteita. Tutkinuskysymyksenäni onkin, miten tehdä feministisestä näkökulmas-

ta parempia datavisualisaatioita. Eettisen agendan lisäksi mukana on käytännön uteliaisuus datan visualisoinnin työprosessia kohtaan: minua kiehtoo mahdollisuus oppia suunnittelemaan pitkälle automatisoituja, muuntuvien systeemejä, joilla datavisualisaatioita tehdään. Tutkimustyön ohessa yritinkin harjoitella koodaustaitojeni ja visuaalisen viestinnän mutoilijan asiantuntemukseni yhdistämistä, vaikkakin melko vaatimattomalla menestyksellä.

Datafeministien mielestä keskeinen ongelma on – IT-kuplan työntekijöiden homogeenisyyden lisäksi – että pyrittäessä rakenteelliseen selkeyteen datavisualisaatioista karsitaan paljon tiedon kontekstiin asettamisessa auttavia yksityiskohtia, huomautuksia ja epämääräisyyksiä. Modernismin ihanteista kumpuava eleettömän estetiikan avulla visuaalisatioissa pyritään esittämään puolueeton, yleispätevän objektiivinen kokonaiskuva, jonka yksityiskohtia tutkailemalla vastaanottaja voi tehdä omat johtopäätöksensä. Tieto ei kuitenkaan synny tyhjiössä, ja se esitetään aina jostakin näkökulmasta ja arvovalinnoista käsin; jotkut näkökulmat vain ovat normalisoidumpia ja siten näkymättömämpiä kuin toiset.

Datafeministit kehottavatkin tiedostamaan aiempaa kriittisemmin visualisointiprosessiin – aina aiheen ja datan valinnasta aineiston analysoinnin

kautta lopullisen visuaalisen esityksen laadintaan – sisältyvät valinnat, rajaukset ja materiaaliset puitteet. He myös kannustavat kehittämään lisää keinoja näiden rajausten ilmentämiseen ja vastaanottajien osallistamiseen. Näin saadaan aikaan kriittiseen ajatteluun herätteleviä, aivan kaikkien kannalta parempia ja totuudenmukaisempia visualisaatioita.

Käyn siis tässä tutkielmassa läpi datavisualisaation laatimisen periaatteita ja käytännön prosessia. Selvennän käyttämiäni käsitteitä kuvin sekä havainnollistan mainitsemiäni muotoiluratkaisuja esittelemällä ja analysoimalla niin inspiroivan onnistuneita kuin varoittaviakin esimerkkejä. Kaikki tutkielmassa esiintyvät kuvat on muokattu karkeiksi bittikarttakuviksi visuaalisen yhtenäisyyden luomiseksi. Olen myös halunnut tällä kuvankäsittelytekniikalla ilmentää datan keruussa, käsittelyssä ja visualisoinnissa tapahtuvaa pelkistymistä: jokainen esitys on aina vain tietystä näkökulmasta laadittu matalaresoluutioinen heijastuma kuvatusta asiasta, ja jokainen versiointi pelkistää sitä entisestään.

Datavisualisaation laatimisen työprosessia läpikäydessäni pyrin vaihe vaiheelta tutkimaan prosessin varrella tehtäviä valintoja ja rajoituksia sekä pohtimaan, millä tavoin nämä päätökset voitaisiin tehdä

entistä vastuullisemmin ja läpinäkyvämmiin, siis feministisemmin. Visualisaation laatimisen työprosessin läpikäytyäni tarkastelen vielä erikseen onnistuneen visualisaation arviointikriteereitä. Ennen tätä avaan kuitenkin seuraavassa luvussa vielä lisää datavisualisaation käsitettä sekä selvennän entisestään datafeministisiä lähtökohtiani.

2 LÖÖFFTÖ- BOFF- DATA

2.1 DATA- VISUALISAAATIO

Informaatiomuotoilu on tiedon jäsentämistä sel-
laiseen muotoon, joka soveltuu mahdollisimman
hyvin ihmisen käytettäväksi – näin sen ainakin
määrittelevät tutkija Robert E. Hornin hengessä suo-
malaisen tiedon visualisoinnin pioneerit Juuso Ko-
popnen, Jonatan Hildén ja Tapio Vapaasalo (2016,
19) perusoppaassaan *Tieto Näkyväksi*. Monet visuaa-
lisen viestinnän sovellusalueet, esimerkiksi taitto tai
muut käyttöliittymät, voidaan nähdä informaation
muotoiluna.

Suppeammin ymmärrettynä informaatiomuotoilusta puhuttaessa viitataan tietoa välittävät kuviin, ja muun muassa Koposen, Hildénin ja Vapaasalon (2016, 27–40) ja Alberto Cairon (2012, xv–xvi) tavoin jaottelen ne edelleen tässä työssä karkeasti kahteen ryhmään: infografiikoihin ja visualisointeihin. Infografiikka on usein käsin laadittu, mahdollisesti tietyn tarinan kertova viestinnällinen, selittävä grafiikka. Visualisaatio ei puolestaan perinteisen määritelmän mukaan välitä mitään ennalta määrättyä viestiä, vaan paljastaa visuaalisella jäsentelyllään uusia piirteitä aineistosta. Visuaalisaatio siis avaa siihen syventyvälle vastaanottajalle mahdollisuuden tutkia ja oivaltaa, vetää omat johtopäätöksensä. (Koponen, Hildén & Vapaasalo 2016, 27–40; Alberto Cairo 2012, xv–xvi) Keskityn tässä tutkielmassa ennen kaikkea visualisointeihin, sillä tämän määritelmän hengessä niitä pidetään lähtökohtaisesti neutraalimpina kuin tarinan kertovat, elämykselliset infografiikat.

Tehostuneiden ja automatisoitujen tiedonkeruumenetelmien ansiosta aineistot ovat nykyisin valtavia, ja niiden pohjalta tuotetut visualisatiotkin ovat varsin usein nimenomaan koneellisesti

tuotettuja. Tällaisessa datavisualisaatiossa suuri määrä tiettyyn tietorakenteeseen – taulukkoon tai luokittelupuuhun – kerättyä raakaa aineistoa jäsennellään näkyvään muotoon tarkoitukseen laaditun ohjelman tai muun systeemin avulla. Tämä esitysrakenne eli skeema pystyy käsittelemään monia rakenteellisesti samankaltaisia aineistoja ja tuottaa niiden pohjalta tiettyjen periaatteiden mukaan karkean visualisaation, jota datavisualisti tai jopa datan käsittelyyn vihkiytymätön graafikko voi sittemmin muotoilla eteenpäin. (Koponen, Hildén & Vapaasalo 2017, 21–22)

Valtavien datamäärien roolin korostuminen on viime vuosikymmeninä tasaisesti kasvattanut kiinnostusta datavisualisaatioita kohtaan. Jo kuluneen mantran mukaan elämme algoritmien kerääminen suuraineistojen (*big data*) ja somen kyllästämän informaatioähkyn aikakautta, jota vaatimatomat savanniapinan aivomme eivät pysty käsittelemään.

Big Data

Big datalla viitataan epämääräisesti niihin valta-
viin datamassoihin, joita esimerkiksi sosiaalisen
median mainosalgoritmit, navigointipalvelut, in-
ternetsivustojen logit tai vaikkapa kauppaketjujen
kanta-asiakasprofiilit keräävät käyttäjä-kuluttaji-
en toiminnasta. Lähitulevaisuudessa muun muas-
sa esineiden internet ja itseajavat autot tulevat
kartuttamaan valtavia datamassoja entisestään.
Big data kertyy ja muuttuu jatkuvasti, sillä sen ke-
ruu on automatisoitu. Valtavien, epämääräisten da-
tamäärien ylläpitoon, tallennukseen ja analysoin-
ti liittyy useita vielä ratkaisemattomia haasteita.
(Gandomi & Haider 2015)

Lentävän lause *"data is the new oil"* ja siihen liit-
tyvä intohimoinen keskustelu, jossa lauahdus mil-
loin julistetaan oikeaan osuvaksi, milloin kaukaa-
haetuksi liioitteluksi, kuvaa osuvasti suuren datan
ympärillä käyvää kuohuntaa.

Koska kuitenkin näköaistimme on verrattaen hyvä hahmottamaan säännönmukaisuuksia, tiedon saattaminen visuaaliseen muotoon auttaa meitä sen ymmärtämisessä ja sisäistämisessä: visualisointi auttaa vähentämään ähkyä tekemällä informaatiosta tehokkaasti selkeämpää, yksinkertaisempaa ja intuitiivisesti hahmotettavaa, jolloin datahälyssä operoivien kuluttaja-kansalaisten on helpompi keskittyä olennaiseen.

Ensinnäkin datavisualisaatiot voivat siis tarjota tutkimustietoa popularisoivia oivaltamisen mahdollisuuksia uutisgrafiikoita kämmentietokoneiltaan sormeileville valveutuneille kansalaisille – kunninahimoisimmat visualisaatiot tehdään tällä hetkellä *Washington Postin* ja *The New York Timesin* kaltaisten suurten uutismedioiden grafiikkatoimituksissa. Toisaalta visualisaatiot ovat keskeinen työkalu myös tieteellisen perustutkimusdatan käsittelyssä: joskus aineiston säännönmukaisuudet tai selkeät mittausvirheet oivalletaan vasta, kun niistä on tuotettu mielekäs visualisaatio.

Lisäksi datavisualisaatioilla on olennainen rooli markkinoinnin sovellutuksissa: tästä mainittakoon ilmeisimmänä esimerkkinä verkkosivujen liikennettä tilastoiva ja visualisoiva *Google Analytics* -palvelu ja dominoivan hakukoneen globaaliin käyttöön

liittyviä tilastoja visualisoiva *Google Trends*, joiden tarjoaman tiedon avulla markkinoijat kautta maailman ennakoivat tulevia kysyntäpiikkejä ja optimoivat sivustojensa liikennettä. Lisäksi pian osaksi arkeamme vakiintuva esineiden internet tulee tuottamaan massiivisia määriä dataa, jonka käsittelyyn tarvitaan yhä hienostuneempia ja intuitiivisempia keinoja: datavisualisaatiot tulevat todennäköisesti olemaan keskeinen työkalu niin laitteiden kehittäjille ja niiden keräämää dataa eri tavoin hyödyntäville, kuin myös laitteiden kanssa vuorovaikuttaville käyttäjille osana palautetta antavaa käyttöliittymää.

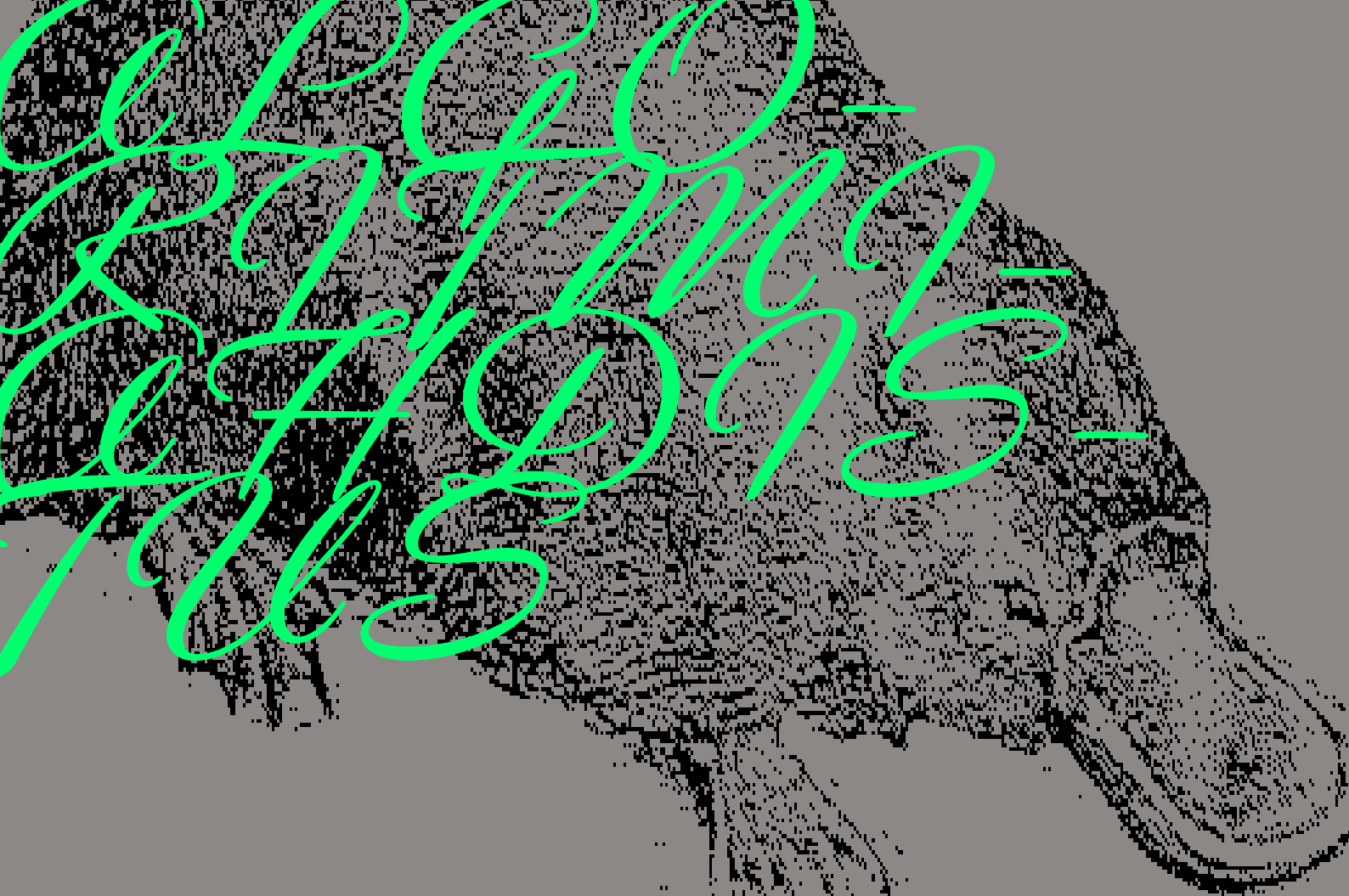
2.2 *DATAFEMINISMI*

Samaan aikaan kun julistetaan avoimen datan ja sen nokkelan visualisoimisen viitoittavan tietä ymmärrettävämpään, läpinäkyvämpään, optimoidumpaan sekä älyllisesti vastuullisempaan ja kaikin puolin valaistuneempaan tulevaisuuteen, on muistettava, että data on vain juuri sitä mitä se on: käsittelemättömiä signaaleja, ärsykyksiä, numeroita. Jotta siitä saataisiin informaatiota ja lopulta sisäistettyä, ymmärrettyä tietoa, on se prosessoitava ihmisen ymmärrettävään muotoon ja asetettava kontekstiin (lisää informaation, datan ja tiedon eroista,

katso Adriaans 2012/2018). Kriittisempää ja tiedostavampaa otetta datan käsittelyyn ja visualisointiin peräänkuuluttaa esimerkiksi niin kutsuttu datafeministinen näkökulma. Catherine D'Ignazion ja Lauren Klein (2018) alleviivaavat toimittamassaan *Data Feminism* -kirjassa¹ datan tuotannon ja keruun kehollisten ja materiaalisten puitteiden sekä vallitsevien valta- ja hyötysuhteiden kriittisempää ja läpinäkyvämpää tiedostamista ja – tässä kohtaa juuri datan visualisoijien rooli korostuu – niiden ilmentämistä.

Datan keräys, käsittely ja visualisointi ovat, kuten kaikki muukin inhimillinen toiminta, useiden risteävien valtarakenteiden ja dynamiikkojen vyyhti. Marginalisoitujen ihmisten haavoittuvat vähemmistöasemat ovat useiden esimerkiksi sukupuoleen, seksuaaliseen suuntautumiseen, henkiseen ja fyysiseen terveyteen sekä luokka- ja etniseen taustaan liittyvien tekijöiden monimutkaisia yhdistelmiä. Asian tarkastelu vain yhdestä näkökulmasta, esimerkiksi toisen aallon (1960–80-lukujen) feministille ominaisen mies-nais-erottelun perspektiivistä pelkistää liikaa. Se, että hyvätuloiset valkoiset naiset saadaan suuryritysten ja valtioiden johtoon ei vielä riitä, sillä kukaan ei ole vapaa ennen kuin *kaikki* ovat vapaita.

2 Kirjan luonnos vuodelta 2018 on tällä hetkellä verkossa kommentoitavana, ja se julkaistaan virallisesti syksyllä 2019.



KUVA 1. Surullisenkuuluisana esimerkkinä Pöytäkirjan IT-kuplan luokittelujen suppeudesta käyvät vaikkapa useimmat kasvojen-tunnistusalgoritmit. Niiden opettamiseen on epähuomiossa käytetty liian suppeaa (toisinsanoen valkoista) kuvavalikoimaa ihmiskasvoista, jolloin niillä on vaikeuksia tunnistaa rodullistettuja käyttäjiä. (Pater 2016, 185)

Tuoreempaan esimerkkinä algoritmeihin valetuista vinoumistakin mainittakoon Facebookin ja Instagramin pornografiantorjunta-algoritmit. Niiden käsitys kyseenalaisesta seksuaalisesta materiaalista on paremman tiedon puutteessa koodattu naurettavan suppeaksi. Niinpä saan ainakin itse omasta sosiaalisen median syötteestäni jatkuvasti lukea julkilausumia siitä, kuinka jälleen uusia queer-piirien seksiposiitivisia julkaisuja tai käyttäjätilejä on asetettu jäähylle solmuun menneiden algoritmien takia. Näin cybertodellisuuden rakenteet edelleen hankaloittavat moninaisuutta koskevan tiedon leviämistä vähemmistöääniä vaientaen.

Kuvassa kaikkien lokeroitien ulkopuolelle pudotettujen totemieläin, vesinokkaeläin (*Ornithorhynchus anatinus*). Laji elää edelleen harmoniassa luonnon kanssa autuaan tietämättömyydestä siitä kauhistuksesta, jonka niiden luonnottomana pidetty sovimattomuus länsimaiseen luokittelujärjestelmään aikoinaan aiheutti Eurooppalaisille.

Juuri tästä risteämisen ja leikkaamisen ajatuksesta (engl. *intersect*)² nimensä saanut intersektionaalinen feminismi on teoreettinen näkökulma, joka pyrkii nostamaan esiin, analysoimaan, kyseenalaistamaan ja purkamaan näitä rakennevyyhtejä. Se on ominaista erityisesti 90-00-lukujen kolmannen ja 2010-luvulla alkaneelle neljännen aallon feminismille. Kun puhun tässä työssä feminismistä, viittaan juuri intersektionaalisuutta korostavaan neljännen aallon feminismiin, jota itse vahvasti kannatan. Myös D’Ignazion ja Kleinin (2018) tovereineen ajama datafeministinen projekti nojaa intersektionaalisuuteen. Sen sijaan, että oman aikamme vinoumat valetaan osaksi tulevaisuuden informaatioinfrastruktuuria, on tärkeää, että niitä määrätietoisesti puretaan.

Sitä paitsi data ja sen visualisointi ovat liian hyviä työkaluja jätettäväksi huomiotta. Valtavien aineistojen keruu ja visualisointi ovat omiaan paljastamaan laajamittaisia rakenteellisia epäkohtia ja systemaattisia vinoumia. Suuri aineisto saattaa kytkeä erilliset kokemukset osaksi suurempaa vyyhtiä ja paljastaa, että kyse ei ole yksilön vaan koko järjes-

2 Käistettä käytti ensimmäisen kerran feministitutkija Kimberlé Williams Crenshaw vuonna 1989.

telmän ongelmasta. Naseva datavisualisointi myös toimii vastaansanomattomana todistusaineistona. Esimerkiksi sukupuolen huomioonottaminen ihmisiä kuvaavissa tilastoissa alkoi Suomessakin kunolla vasta pitkäjänteisen tasa-arvoaktivismin myötä 1970–1980-luvuilla: tätä ennen naisia koskevaa tietoa ei juuri katsottu tarpeelliseksi tuottaa. Vasta erottelujen tekeminen antoi välineet tasa-arvokehityksen seuraamiseen ja todisteita niin nais- kuin mieserityisten ongelmien tarkasteluun. (Veikkola 2005, 85) Sinänsä silkka mies-nais-erottelu on tietenkin liian kapea, ennen kaikkea siinä mielessä, että sukupuolia on enemmän kuin kaksi. Lisäksi intersektionaalisuuden hengessä vielä enemmän selitysvoimaa on datalla, joissa ihmisten tilanteita on eroteltu myös muilla mittareilla.

Datan keruuseen ja käsittelyyn liittyy paljon valtarakenteita ja piileviä vinoumia: datan kanssa työskentelevät ihmiset ja instituutiot tuovat prosessiin välttämättä mukanaan omat ennakko-oletuksensa ja sokeat pisteensä. Keskeinen ongelma on, että yleisesti ajatellaan, että data on 'raakaa syötettä' (*raw input*), höyryävän tuore palanen todellisuutta, faktoja. Aineistoon päätyneet luvut ovat kuitenkin erehtyväisten ihmisten usein kiireessä tekemän, lukuisia tiedostettuja ja tiedostamattomia kompro-

misseja ja rajanvetoja sisältäneen polveilevan prosessin tulostetta (*output*). (D'Ignazio & Klein 2019) Käytetyt menetelmät ohjaavat valintoja, ja historiallisen kerrostumisen seurakuksena päädytään usein kierrättämään, uusintamaan ja lujittamaan jo olemassaolevia ratkaisuja ilman, että niiden sopivuutta ja mielekkyyttä käsillä olevaan tehtävään tarkasteltaisiin kriittisesti. Ei ole esimerkiksi lainkaan merkityksetöntä, että esimerkiksi valtaosa maailman ihmisiä koskevasta datasta kerätään tällä hetkellä markkinointitarkoituksiin eikä esimerkiksi tieteellistä perustutkimusta varten. Data ja sen visualisointi, kuten kaikki muukin ihmisten tekemä, ei siis koskaan ole neutraalia tai objektiivista.

Data ja siihen liittyvät prosessit kuitenkin *näyttävät* neutraaleina, koska kylmät luvut häivyttävät niiden keräämiseen osallistuneet ihmiset taka-alalle. Ongelma kiteytyy nähdäkseni myös vakiintuneessa visualisaation määritelmässä, naiivissa uskomuksessa siihen, että visualisaatio vain neutraalisti jäsentää objektiivisia havaintoja, joista kukin voi vetää omat johtopäätöksensä. Myös visualisaatioiden estetiikka korostaa tätä suhtautumista: paljon tietoa koodaavissa visualisoinneissa tähdätään selkeyden ja luettavuuden takaamiseksi melkolailla yhtenäisiin, yksinkertaisiin ja pelkistettyihin ratkaisuihin.

NÖÖKE- MÄMEN

kuva 2. Näkeminen tapahtuu etäältä, ilman selkeää kehollista vuorovaikutusta: sen avulla on mahdollista hankkia informaatiota turvallisen välimatkan päästä. Näön ylikorostuminen on Harawayn (1988, esimerkiksi 281-582) mukaan johtanut muiden aistien, kehollisuuden ja paikkasidonnaisuuden ylenkatsumiseen ja unohtamiseen. Hybristä lisää entisestään se, että tieteen ja teknologian avulla pystymme nykyisin "katsomaan" – tai Harawayn (ibid.) retorikkaa mukaillen pikemminkin kytymättömästi tarkastelemaan – lähes kaikkea, missä ja milloin vain.

Oheinen hahtikuussa 2019 julkaistu ensimmäinen kuva mustasta aukosta on sekä massiivinen datavisualisaatio, joka on koostettu Kabe Doumanin nimissä johtaman työryhmän laatiman algoritmin avulla kahdeksan radioteleskoopin datasta.

Tämä kaikesta tiettynä historiallisena ajanjaksona tietyssä kulttuurissa – massa- ja markkinointiviestinnän ammattilaisten keskuudessa 1900-luvun Euroopassa ja Yhdysvalloissa – turhana pidetystä karsittu modernistinen estetiikka on saanut etuoikeuden olla objektiivisuuden merkki (Wang, 2018).

MacLuhanin (1964/1994) ”The medium is the message”-klassikkolausahduksen hengessä voidaan siis sanoa, että yksinkertaistukset, yleistykset, kaksikulotteiset projektiot, hillityt värit ja litteät geometriset muodot luovat mielikuvan siitä, että aihetta tarkastellaan viileän rationaalisesti, turvallisen välimatkan päästä. Feministifilosofi Donna Haraway (1988, 581–590) puhuu ’jumaluustempusta’ (*the god trick*): data esitetään usein kaikennäkevän jumalan näkökulmasta.

Harawayn (1988; Rolin 2005, 100–101) ja kumppaneiden muotoileman tiedon sosiaalisen rakentamisen ja feministisen standpoint-teorian mukaan kaikki tieto on kuitenkin aina osittaista, näkökulma- ja kontekstisidonnaista: tutkimuksen kohteena olleet ja dataa keränneet, prosessoineet ja muotoilleet kehot historioineen, olosuhteineen ja kykyineen vaikuttavat siihen, mitä on saatu selville. Jotakin jää tai jätetään aina ulkopuolelle, joko epähuomiossa tai tarkoituksellisesti.



Standpoint- eli tutkimuksen näkökulmasidonnaisuuden teoria ajautuu vahvoissa muodoissaan relativismiin: totuus on suhteellista, kulttuurista ja kontekstista toiseen vaihtelevaa sopimuksenvaraista ja valtapeleissä määräytyvää. Nykyisin akateemisessa feministisessä tieto-opista ja tieteenfilosofiasta käytävässä keskustelussa standpoint-teoriaa pidetään postmodernistisuudessaan turhan jyrkkänä, mutta sen esittämiä näkemyksiä on myös huolella jalostettu eteenpäin (Rolin 2005, 97–111). Yksinkertaistettu standpoint-teoria sopii kuitenkin tämän tutkielman tarpeisiin problematisoimaan datan visualisoinnissa piilevää vallankäyttöä.

Järkevyyttä ja objektiivisuutta estetiikallaan viestivä visualisaatio siis näyttäytyy neutraalina eli arvovapaana 'näkymänä ei-mistään' (Haraway 1988, 581; D'Ignazio & Klein 2018, "Chapter One: Bring Back The Bodies"). Esitys on kuitenkin aina jonkun tekemä, ja tämän tekijän näkökulman vaikutus on aina välttämättä läsnä: valintoja pidetään niin itsestäänselvinä, että ne ovat jääneet niin tekijältä kuin lukijaltakin huomaamatta. Visualisaatio, jonka rajoja ei ole tiedostettu ja tuotu esiin, tarjoaa vain *oletusnäkymän* kertoen yhtä paljon siitä, mitä visualisaation tekijät pitävät normaalina kuin itse esitettävästä ilmiöstä.

Ilman selitystä ja kontekstia ei saada vastauksia kysymyksen siitä, miksi data on juuri sellaista kuin se on (D'Ignazio & Klein 2018, "Chapter Five: Numbers Don't Speak For Themselves"). Hajanaisten faktojen tarkastelu altistaa kaikenlaisille päättelyvirheille, kun heuristiikat – maalaisjärki, käytännön elämässä hioutuneet peukalosäännöt ja stereotyyppiat – astuvat kuvaan. Näin on erityisesti tapauksissa, joissa data kuvaa marginalisoidun ryhmän tilannetta: johtuuko nuorison alhainen äänestysprosentti heidän välinpitämättömyydestään vai turhautumisestaan demokratiaan? Entä onko yliopistolla raportoitujen ahdistelutapausten lisääntyminen seuraus-

ta varsinaisen ahdistelun lisääntymisestä vai uhreja syyllistävän ilmapiirin hälventymisestä? (D'Ignazio & Klein *ibid.*)

Tarkasteltavan ilmiön ulkopuolelle astumisen ja sen objektiivisen, kokonaisvaltaisen tarkastelun illuusiosta kumpuavaan hybrikseen on siis suhtauduttava varoen. Visualisaatioon liittyvät tulkinnanvaraisuudet ja epävarmuustekijät hämärtyvät viettelevän vakuuttavan esitystavan alle, ja valppainkaan vastaanottaja ei aina muista pohtia ensisilmäyksellä syntyneen tulkintansa taustalla vaikuttavia oletuksia ja valtasuhteita. D'Ignazion (2015) mielestä valvutuneen tiedon visualisoijan tehtävänä onkin pyrkiä herättelemään vastaanottajia visualisaatioidensa kriittisempään tarkasteluun. On keksittävä uusia tapoja ilmentää epävarmuutta, ulkopuolisuutta, puuttuvaa dataa ja metodologisia ongelmia. On tutkittava keinoja ilmentää datan takana olevia materiaalisia puitteita. On kehitettävä mahdollisuuksia olla visualisaation kanssa vuorovaikutuksessa, osallistua sen luomiseen, käydä sen kanssa keskustelua ja esittää eriäviä näkemyksiä.

Tiiviisti ilmaistuna informaatiomuotoilijan on siis kannettava osansa vastuusta siinä, mitä ja miten näytetään ja ketä se palvelee. (D'Ignazio 2015). Tätä

vastuuta alleviivaa entisestään se, että nykyisin visualisaation tekijät ovat usein monialaisia, tutkivia graafisia journalisteja, jotka osallistuvat aktiivisesti sisällön toimittamiseen ja datan käsittelyyn muutenkin kuin visuaalisessa mielessä (Koponen, Hildén & Vapaasalo 2016, 318–321). Kaiken kaikkiaan datafeminismi tarjoaa siis aiempaa laajaemmän ja syvemmän näkökulman tiedon visualisoijan eettisiin ohjenuoriin.

3
D O C T O R
V I S I T
C L I N I C
S E C R E T
L I O N
P R O
S E S S I O N

Tässä osuudessa käyn läpi datavisualisaation laatimisen prosessia, ja sen eri vaiheissa tapahtuvia valintoja. Pohdin samalla, millä tavoin nämä valinnat voi tehdä datafeminismin näkökulmasta kestävämmän eli inklusiivisemmän, vastuullisemmän ja läpinäkyvämmän. Käytän prosessin jäsentelyyn ensinnäkin Vesa Kuuselan (2000, 190–201) yleiskatsausta onnistuneen tilastokuvino tekemisestä. Samalla sovellan Koposen, Hildénin ja Vapaasalon (2016, 308–315) viisivaiheista visualisoinnin tuottamisen työvaiheiden jäsentystä, johon on yhdistelty alan aiempia malleja (muun muassa Ben Frylta, Moritz Stfanderilta, Frederik Ruysilta ja Tamara Munzneilta) ja kirjoittajien omasta praktiikastaan tekemiä havaintoja.

Lisäksi mukana on Jessica Hullmanin ja Nicholas Diakopouloksen (2011, 2233) jako visualisaation toimituksellisiin kerroksiin: visualisaatioista on erotettavissa datan, kuvallisen esittämisen, kielellisten selitteiden sekä interaktiivisuuden taso. Kaikki mainitut jäsentelyt täydentävät toisiaan.

Visualisaation laatiminen on laaja projekti, jota tehdessä toimitaan ainakin kolmessa eri roolissa: *sisältövastaavana*, joka grafiikan käsittelemän aiheen asiantuntijuutensa tai muun perehtyneisyytensä

avulla huolehtii grafiikan tiedollisesta sisällöstä, *informaatiomuotoilijana*, joka hallitsee tiedon mielekkään visuaalisen jäsentämisen niin viestinnällisellä kuin tekniselläkin tasolla, sekä *toteuttajana*, joka tuottaa julkaisukelpoisen ulkoasun käyttöliittymineen (Koponen, Hildén & Vapaasalo 2016, 302–307).

Renessanssineroutta hipova graafinen journalisti kykenee huolehtimaan näistä kaikista rooleista yksinkin, mutta on myös varsin realistista ja tavalista, että visualisaatio laaditaan monialaisessa työryhmässä. Mielekkääseen kokoonpanoon kuuluu esimerkiksi sisällön asiantuntija, tietojenkäsittelyn ammattilainen ja visuaalisen viestinnän muotoilija: eri asiantuntemusten kohtaamisesta syntyy monipuolisia ja tarkoituksenmukaisia ratkaisuja. (Koponen, Hildén & Vapaasalo 2016, 302–307)

Lopullinen visuaalisen ulkoasun suunnittelu on vain projektin viimeinen silaus. Aiemmat työvaiheet vaikuttavat huomattavasti enemmän lopputuloksen onnistumiseen. Itseäni juuri datavisualisatioissa kiehtoo se, että ne rakennetaan dynaamisina systeemeinä, joiden avulla varsinaisen käsityön ja siten myös inhimillisten virheiden määrä pyritään minimoimaan. Tietenkin on huomioitava, että mitä enemmän työtä ohjelmille säilytetään, sitä olennai-

sempaa on, että ne on alunperinkin tehty kyllin kriittisellä mielellä, jottei vinoumia automatisoida.

Kaikesta sisältö- ja tietojenkäsittelytieteellisen asiantuntijuuden korostamisesta huolimatta myös ammattitaitoisella visuaalisen viestinnän muotoilijalla on tärkeä panos annettavana jokaiseen suunnitteluprosessin vaiheeseen. Ensinnäkin hän tuo projektiin paljonpuhuttua (esimerkiksi Joutsela & Hovinen 2018) design-ajattelua: luovan prosessin tuntemusta, ideomistekniikoita ja luovaa rohkeutta. Lisäksi Koponen, Hildén ja Vapaasalo (2016, 313–314) korostavat erityisesti viestintäosaamista, joka pääsee oikeuksiinsa varsinkin kohderyhmän määrittelyssä ja sen tarpeiden huomioimisessa.

Voitaneen siis sanoa, että visuaalisen viestinnän muotoilijan tehtävä on olla ”yleisön edunvalvoja” (Isotype-ryhmän hengessä Koponen, Hildén & Vapaasalo 2016, 305), tai datafeministien sanoin tuoda lihaa ja verta olevat ihmiset huomioiva näkökulma mukaan muutoin korostuneen kylmiä numeroita käsittelevään prosessiin. Tästä näkökulmasta juuri designerille on siis myös luontevaa olla myös kriittinen feminististen näkökulmien esiintuoja visualisaatiota laadittaessa. Luovan ajattelun ammattilaisena hän voi kysyä kiperiä kysymyksiä: onko työryhmässä tarpeeksi erilaisia näkökulmia? Pitäisi-

kö mukaan kenties ottaa joku kohderyhmän tai tutkittavan ihmisjoukon edustaja? Entä miksi käytössä oleva data on juuri sellaista kuin se on?

3.1 TARVE JA TAVOITE

Visualisointiprojekti voi lähteä liikkeelle joko viestinnällisistä tavoitteista tai tietystä aineistosta. Joka tapauksessa projektin alussa on pohdittava, mihin kysymyksiin visualisaation halutaan vastaavan (Koponen, Hilden Vapaasalo 2016, 309 & 315). Kuusela (2000, 193) aloittaa vielä perustavanlaatuisemmalla kysymyksellä: tarvitaanko kuvaajaa ylipäättään? Kuten Koponen, Hildén ja Vapaasalo (2016, 325) kiteyttävät, ”kuvalla on valtava voima”: parhaimmillaan se selventää, mahdollistaa oivalluksen, kiinnittää huomion, saattaa jopa tunnetasolla liikuttaa sekä muuttaa katsojan toimintaa.

Kuitenkin, kuten edellisessä luvussakin todettiin, visualisaatio johtaa herkästi myös harhaan. Tietoisesti vilpillisesti laadittuja esityksiä on harvassa, mutta on aivan liian helppoa päätyä tilanteeseen, jossa tekijät epähuomiossa siirtävät visualisointiin omat vinoumansa. Tämä on kohtalokasta, koska grafiikan lukeminen on niin intuitiivista, että sen antamaa väärää käsitystä on liki mahdoton korjata

jälkikäteen¹. Näin ollen kuvan tarve ja tarkoitusperät on harkittava erityisen tarkkaan: Kuusela (2000, 193) toteaa jopa, että ”puuttuva kuvio on parempi kuin epäonnistunut kuvio”.

Visualisaation perimmäinen ja välttämättömin tehtävä on välittää oikeaa tietoa selkeimmällä mahdollisella tavalla (Kuusela, 2000, 193; Koponen, Hildén & Vapaasalo 2016, 32 & 326). Tämä funktio ei saa vaarantua missään projektin vaiheessa. Muut tavoitteet, kuten huomion kiinnittäminen, toimintaan kehottaminen tai viihdyttäminen ovat aina alisteisia tälle ensisijaiselle päämäärälle. Lähteiden on oltava kunnossa, ja kaikkien johtopäätöksiä, yhteyksien, rinnastusten ja vertailujen, jotka kuviossa on löydettävissä, on oltava aineistolla perusteltavissa (Tufte 2006, 79).

Näin ollen myös edellä esitetty kriittinen vaatimus kuvion tai aineiston rajallisuuden, epätarkkuuden tai piilomotiivien esittämisestä on varsin perusteltu: feminisittisen standpoint-teorian perusteella ”oikeaa tietoa” saadaan vain toisiaan täydentäviä näkökulmia yhdistelemällä ja kunkin tiettyyn ai-

1 Koponen, Hildén ja Vapaasalo (2016, 325) mainitsevat, että tämä ilmiö tunnetaan kognitiivisen ja havaintopsykologian parissa nimellä kuvan ylivoimaefekti.

kaan ja paikkaan sidotun näkökulman rajat tiedostamalla. Tuften (2006, 79) vaatimusten vaikutus- aluetta siis laajennetaan koskemaan myös aineiston kontekstia ja siihen liittyviä valtarakenteita. Tekijän on tarkastettava lähteittensä lisäksi myös omat ja sidosryhmiensä etuoikeudet ja vinoumat tehdäkseen kaikkensa niiden kerrannaisvaikutusten minimoimiseksi.

3.2 KONTEKSTI JA KOHDERYHMÄ

Kuten visuaalisessa viestinnässä ylipäättään, designin kohderyhmä on yksi keskeinen suunnittelutyötä ohjaavia tekijöitä. Yleisön arvioitu kuvanlukutaito määrää monet visuaaliset ratkaisut, esimerkiksi sen, kuinka erikoistunutta tai abstraktia kuvaajaslangia voidaan käyttää. Lukutaitoon liittyy myös suoraan kyky suhtautua kriittisesti kuvaajan esittämiin väitteisiin; näin ollen laajemmalle yleisölle tehdyn grafiikan pitäisi olla vielä läpinäkyvämpi kuin asiantuntijakäyttöön laaditun.

Valpas visualisti on toki myös tietoinen yleisönsään mahdollisesti kukoistavista yleisistä ennakkoluuloista sekä muista kognitiivisista ja sosiaalisista vinoumista. Täytyykö jonkin tiedon esittämiseen kiinnittää erityistä huomiota siksi, että se on vastaanottajille todennäköisesti yllättävää? Entä mikä on kohderyhmälle itsestään selvää – kenties niin itsestään selvää, että päättelymalli säteilee sellaisiin-kin tilanteisiin, joihin se ei oikeasti sovellu? Vaikka ihmiset eivät ole tyhmiä ja laiskoja, he ovat väsyneitä ja kiireisiä (Tufte 1990, 34) – nykyisin polarisointuneessa sosiaalisessa mediassa tosin yleensä myös peloissaan tai vihaisia – jolloin niin informaatioähkyn aiheuttava kaoottisuus kuin liian vähäisen selittävän kontekstin tarjoileminen altistavat päättelyvirheille (D’Ignazio ja Klein 2018, ”Chapter Three: ‘What Gets Counted Counts’”).

3.3 RAJAT JA ASIAYHTEYS

Kuvio on myös osattava liittää oikeaan asiayhteyteen (Kuusela 2000, 194). Esityksen mahdollistamat vertailut sekä aineiston oikea rajaaminen ovat tässä tärkeitä tekijöitä. Ihmiset etsivät merkitysyhteyksiä toisiinsa liittymättömienkin asioiden väliltä, kun

ne rinnastetaan – esimerkiksi elokuvan ja sarjaku-
van kerronta perustuvat pitkälti tähän. Tiedon visu-
alisoinnissakin rinnastusta voi käyttää paitsi uusien
oivallusten synnyttämiseen, myös totaaliseen har-
haanjohtamiseen.

Klassinen esimerkki epärehellisyydestä on rin-
nastaa samaan kuvaajaan keskenään korreloivia²
asioita, jolloin annetaan samalla ymmärtää, että nii-
den välillä on myös kausaalinen eli syy-seuraus-suh-
de. (Pater 2016, 165) Rajaamisen suhteen esimerkik-
si esitettävä aikaikkuna on valittava rehellisesti: jos
kuvataan jotakin dramaattista muutosta, sitä edel-
tänyt ja sen jälkeinen aika on myös syytä sisällyttää
mukaan esitykseen.

Tapausesimerkkinä juuri rajaamiseen liittyväs-
tä vastuusta toimikoon suomalaisen Lucifyn (2018)
laatima datavisualisaatio Syyrian sisällissotaa Eu-
rooppaan paenneista turvapaikanhakijoista ja
muusta Eurooppaan vuoden 2012 jälkeen suuntau-
tuneesta muuttoliikkeestä . Tyylikäällä, interaktii-
visella kartalla 20–25 turvapaikanhakijan joukkoja
edustavat pisteet vaeltavat vääjäämättömänä vir-

2 Erillisten muuttujien välinen tilastollinen riippu-
vuussuhde, josta ei kuitenkaan voida vielä pää-
tellä syy-seuraus-suhdetta.

tana Lähi-Idästä ja Afrikasta kohti Eurooppaa, antaen kylmäävän käsityksen humanitaarisen kriisin laajuudesta. Visualisointi saikin huomattavaa näkyvyyttä ja kiitosta kansainvälisessä lehdistössä aina The Guardiania myöten.

Datavisualisointienkin poliittisuutta tutkinut Ruben Pater (2016, 162–163) tosin huomauttaa, että visualisaation data oli rajattu koskemaan vain Eurooppaan, eikä se esitä lainkaan sitä valtaosaa pakolaisista, jotka saavat turvapaikan huomattavasti lähempää lähtömaitaan – Lähi-Idän maat ottavat nimittäin vastaan yli 5 kertaa enemmän pakolaisia kuin Eurooppa.

Toki Lucify on tehnyt tämän Eurooppaan keskittyvän rajauksen selväksi jo otsikkotasoa myöten. Laajempaa kontekstia myös käsitellään visualisaation oheisartikkelissa muun muassa Eurooppaan ja Lähi-Itään suunnanneiden pakolaisten määrää jalopallokenttien avulla havainnollistamalla. Patel (2016, 163) haluaa kuitenkin nostaa esiin, että herkkä aihe olisi ollut tärkeää esittää vähemmän eurooppakeskeisestä näkökulmasta jo kartalla, joka lähti leviämään viraalisti ilman selitteitään. Kuvan viesti olisi ollut aivan erilainen, jos Euroopan tapahtumat olisi nähty myös suhteutettuna kaikkiin Lähi-Idän maihin päätyneisiin pakolaisiin.

Esitys voidaan toki pyrkiä kytkemään tekstin avulla laajempaan kontekstiin tai kiinnittää huomioon kuviosta puuttuvaan dataan. Toisaalta Lucyfy-esimerkki osoittaa myös sen, että kuva on vaikutavuudeltaan huomattavasti voimakkaampi kuin sen ohessa oleva teksti: harva kiireinen sosiaalisen median selaaja tuskin edes vierittää sivustoa oheistekstiin asti. Sen sijaan jokaisen mieleen piirtyy kuva vain ja ainoastaan Eurooppaan vyöryvästä tieheästä pistevirrasta. Koponen, Hildén ja Vapaasalo (2016, 325) korostavatkin, että grafiikan lukeminen on kovin intuitiivista, ja lukutapaa on lopultakin hyvin vaikeaa muuttaa kirjallisilla tulkintaohjeilla.

3.4 AINEISTO, TUTKIMUS JA JÄRJESTÄMINEN

Niputan tähän kohtaan kaksi erillistä työvaihetta: aineiston hankkimisen sekä sen tutkimisen ja järjestämisen. Työvaiheet ovat perin teknisiä sekä niin visuaalisen viestijän kuin niissä tehtävien eettis-poliittisten valintojenkin näkökulmasta toisiinsa limittyneitä.

Onnistunut visualisointi pohjaa asianmukaisesti kerättyyn tai valittuun ja sopivaan muotoon

siistittyyn aineistoon – nykyisin usein (tilasto)da-
taan (Koponen, Hildén & Vapaasalo 2016, 311; Kuu-
sela 2000, 193). Aineiston valikointi ja työstö on
niin abstraktia, että sen yhteydessä tapahtuvat to-
dennäköisimmin esityksen laadinnan pahimmat
laiminlyönnit. Nämä huomiotta jättämiset ovat
usein myös tiedostamattomia, joten niitä ei myös-
kään kommunikoida esityksessä (Hulman & Dia-
kopoulos, 2011, 2234). Jos aineisto kerätään itse, sen
keräysprosessi saattaa olla vinoutunut esimerkik-
si epäedustavan otoksen tai johdattelevien kysymys-
ten kautta. Valmiina saatu aineisto puolestaan saa-
tetaan hyväksyä sellaisenaan, ilman kunnollista
kriittistä pohdintaa sen alkuperästä, kategorisoin-
neista tai rajauksesta. Saatetaan esimerkiksi ottaa
annettuna aineiston jako miehiin ja naisiin: ehkäpä
esimerkiksi vastaajien ikien luokittelu on esitettävän
aiheen kannalta paljon kiinnostavampaa, ja pako-
tetun sukupuolibinäärin visualisointi vain ylimää-
räistä visuaalista melua.

Oma osansa prosessia – surullisenkuuluisan
vitsin mukaan noin 80 % datatieteilijän työajas-
ta – on datan siistiminen visualisointeihin sopi-
vaan muotoon. D’Ignazio ja Klein (2018, ”Chapter
Four: Unicorns, Janitors, Ninjas, Wizards, and Rock
Stars”) argumentoivat, että kaikki data on synty-

jään, määritelmällisesti ja välttämättä sotkuista: epämääräisyydet rakenteessa ja ilmaisussa ovat niitä tahroja tai jälkiä, joita datan keräämisen suunnitelleet ja toteuttaneet ihmiset sekä keräämisen prosessi ja ympäristö ovat siihen jättäneet. Datan aukkoja ja sotkuisuutta olisi siis syytä arvostaa eräänlaisena metadatana, jonka voi myös tuoda jotenkin visualisatiossa ilmi.

Valitettavan usein tämä sotkuisuus on myös ainoa konteksti, jota datalle on saatavissa. Internetissä nimittäin rehoittaa täysin juuretonta, vain epämääräisen otsikoinnin selittämää avointa dataa ja lisää tulee jatkuvasti. Kenties meidän kaikkien pitäisi alkaa olla vaativampia datamme laadun suhteen?

Search here	
Type	
	comma2k19
	BRATS2013 Tumor NoTumor Dataset (TT-1)
	COCO 2017
	30M Factoid Question-Answer Corpus (FQAQA)
	Indiana University - Chest X-Rays (IWX) View
	Stanford Cars
	SPM_val_targz
	PASCAL Visual Object Classes (VOC)
	Amazon reviews - Polarity
	Labeled Optical Coherence Tomography (OCT)
	SPM_train_targz
	Invasive Ductal Carcinoma (IDC) Histology Image Dataset
	MNIST
	Oxford 102 Flowers
	Food-101
	DeepLadon (10,094 CT scans with lesions)
	Downsampled Open Images V4 Dataset
	Yelp reviews - Polarity
	Wildcraft-103
	Common Crawl corpus - training-parallel-commoncrawl-gpt2 View Download View Download

KUVAT 5 & 6. Academic Torrents on vielä melko hyvin dokumentoitua ja selitettyä avointa dataa, vaikka aihepiirit vaihtelevatkin melko kaaottisesti luokitelluista siitepölyhiukkasista ravintola-arvioihin.

Kaikesta kontekstistaan puhdistettu data on kuin rihmastostaan erotettu, formaldehyldiin säilöty sieni: olennaisinta siinänsä kuitenkin on juuri sen maanalainen rihmasto ja sen toiminta osana metsän ekosysteemiä.

Hankittua aineistoa tutkitaan, järjestellään ja rikastetaan erilaisin tietojenkäsittelyllisin ja tilastollisin menetelmin. Siitä suodatetaan pois käsiteltävä asian kannalta epäolennainen tieto ja korostetaan mielenkiintoisia löydöksiä – juuri tällaisten rajausten yhteydessä saatetaan epähuomiossa tosin hukata jokin olennainen näkökulma.

Onkin tärkeää pohtia intersektionaalisen kriittisesti, mitkä aineiston ominaisuudet viestinnällisten tavoitteiden kannalta kiinnostavia ja tarkoituksenmukaisia, eli onnistutaanko syntyneellä jäsentelyllä todella ilmentämään asiaa, jota sen avulla yritetään kuvata (Koponen, Hildén & Vapaasalo 2016, 315) Onko kyseessä siis osuva tai riittävän kattava metonymia käsiteltävälle aiheelle (Hullman & Diakopoulos 2011, 2234)? Karkeana esimerkkinä epäonnistuneesta metonymiasta mainittakoon bruttokansantuote asukasta kohden -tieto, jota yritetään edelleenkin silloin tällöin käyttää ilmentämään kansalaisten hyvinvointia – vertaus ei ole onnistunut, sillä bruttokansantuote kasvaa myös esimerkiksi luonnonkatastrofien ja sotien seurauksena, eikä kerro mitään varallisuuden jakautumisesta kansalaisten kesken.

3.5 LUONNOSTELU JA KOKEILU

Luonnostelun ja kokeilun myötä pyritään päättämään aineistolle sopiva visuaalinen esitystapa. Tämä on siis yksi prosessin visuaalisimmista vaiheista, joten keskityn erityisellä mielenkiinnolla muun muassa esitystavan hahmottelemiseen. Esitystavalla tarkoitetaan piirakkadiagrammin tai pylväskuvaajan kaltaista visuaalista ratkaisua, jonka avulla data lopulta esitetään.

Luonnostelu on, kuten kaikessa suunnittelussa (Koponen, Hildén & Vapaasalo 2016, 312) mukaan erittäin olennainen osa prosessia. Aineiston visuaalinen jäsentämienn esimerkiksi lyijykynäluonnoksin on feministisestäkin näkökulmasta relevanttia ensinnäkin siksi, että se kehollisuudessaan vapauttaa ajattelua tietokoneiden ja muiden työvälineiden asettamista rajoitteista. Käsien luonnostelu on mahdollista myös ei-ammattilaisille: näin ollen niin sisältöasiantuntijoita (Koponen, Hildén & Vapaasalo 2016, 313), visualisaation aiheena olevia ihmisiä kuin lopullisen työn kohderyhmääkin voidaan osallistaa suunnitteluprosessiin.

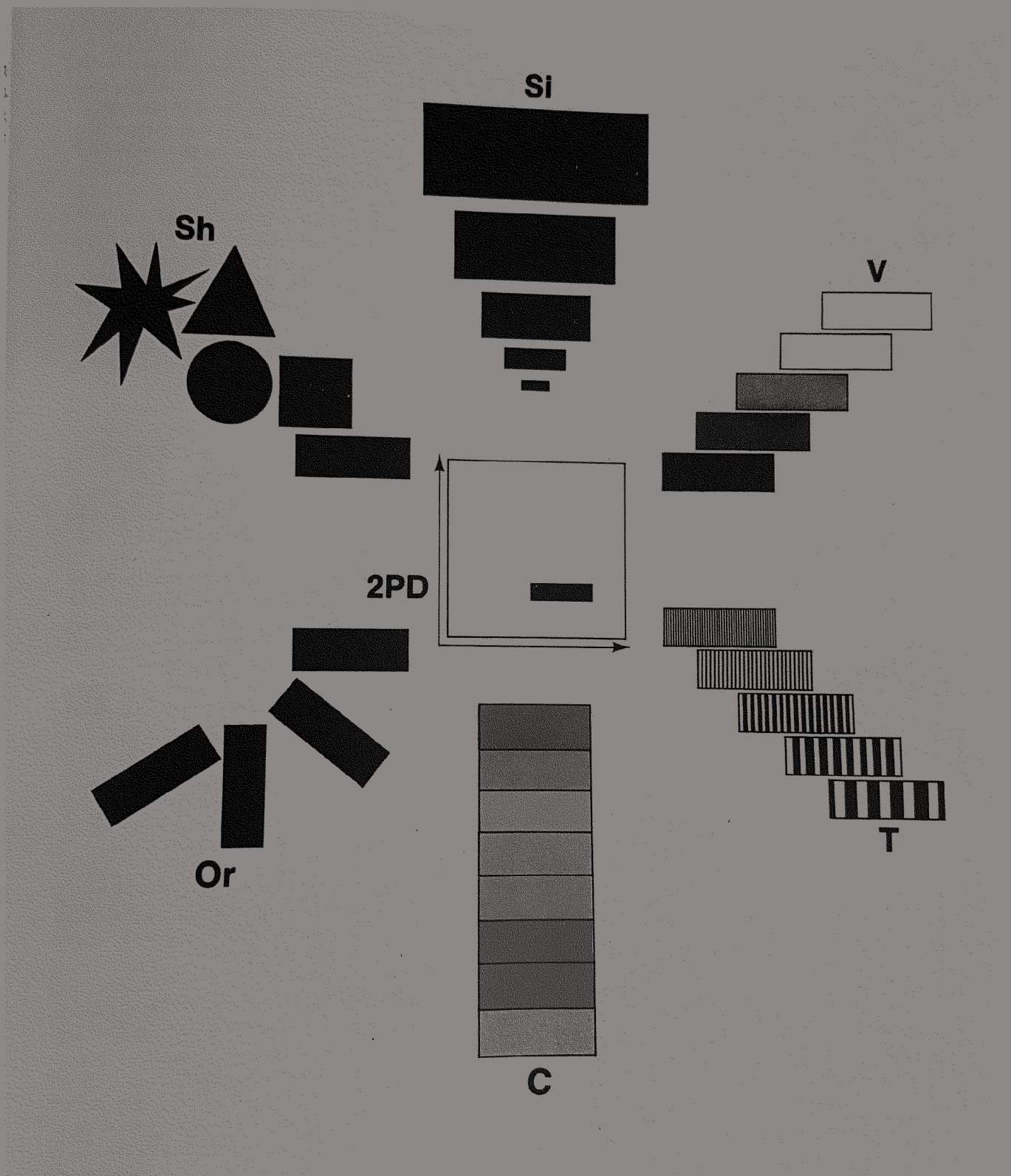
Toisaalta myös tietokoneavusteinen oikealla datalla luonnostelukin on perusteltua: lopullisen ai-

neiston ominaisuudet eivät saa tulla yllätyksenä vasta loppumetreillä. Varsinaisen datan kanssa luonnostelu johtaakin usein myös uusiin aineistoa koskeviin havaintoihin ja ohjaa näin prosessia. Tässä vaiheessa voidaan myös huomata, että jokin olennainen näkökulma jää tyystin puuttumaan. Onneksi visualisaation laadinta on ketterä, syklinen prosessi (Koponen, Hildén ja Vapaasalo 2016, 308–315): työn edetessä tehtyjen uusien havaintojen ja näkökulmien muutosten myötä on rohkeasti vain palattava aiempiin työvaiheisiin. Tiedon visualisointi mahdollistaa myös useiden lähteiden hyödyntämisen ja tiedon kerrosten yhtäaikaista esittämisen (Tufta 2006, 78–79). Näin ollen visualisti voi valppaalla kokonaisesityksellä tehdä oikeutta aiemmin liian suppeasti käsitellylle aiheelle.

Luonnostelun myötä data siis pyritään koodaamaan visuaalisiksi muuttujiksi, kuten pisteiden sijainniksi kuvapinnalla: aineistolle yritetään löytää sen ominaisuudet parhaiten esiin tuova esitystapa. Saman aineiston voi periaatteessa visualisoida jopa sadoin eri tavoin (Bertin 1983/2010), ja on tekijän ammattitaidon varassa laatia aineistoon parhaiten, toisin sanoen vähiten harhaanojohtavasti sopiva esitystapa. Visuaalista ilmaisuja ohjaavat ainakin aineiston

tilastotieteelliset, matemaattiset ja tietojenkäsittelylliset ominaisuudet sekä aiemmin mainitut haluttu viesti sekä kohderyhmä (Salmelin 2003, 1763).

Lisäksi olennaisesti vaikuttavat havaintopsykologiset realiteetit, kuten värioppi, Gestalt-lait sekä muu tieto visuaalisten vihjeiden tulkintakyvystä. Esimerkiksi esityksen värivalinnat on tehtävä havaintopsykologisin, ei esteettisin eikä etenkin luutuneisiin totunnaisuuksiin perustuen: jos esimerkiksi merkitään tympeän kliseisesti poikia sinisellä ja tyttöjä vaaleanpunaisella, väreistä tummempi eli sininen muodostaa valkoisen paperin kanssa suuremman kontrastin, korostaen poikadatan kuvan huomionarvoisimmaksi elementiksi.



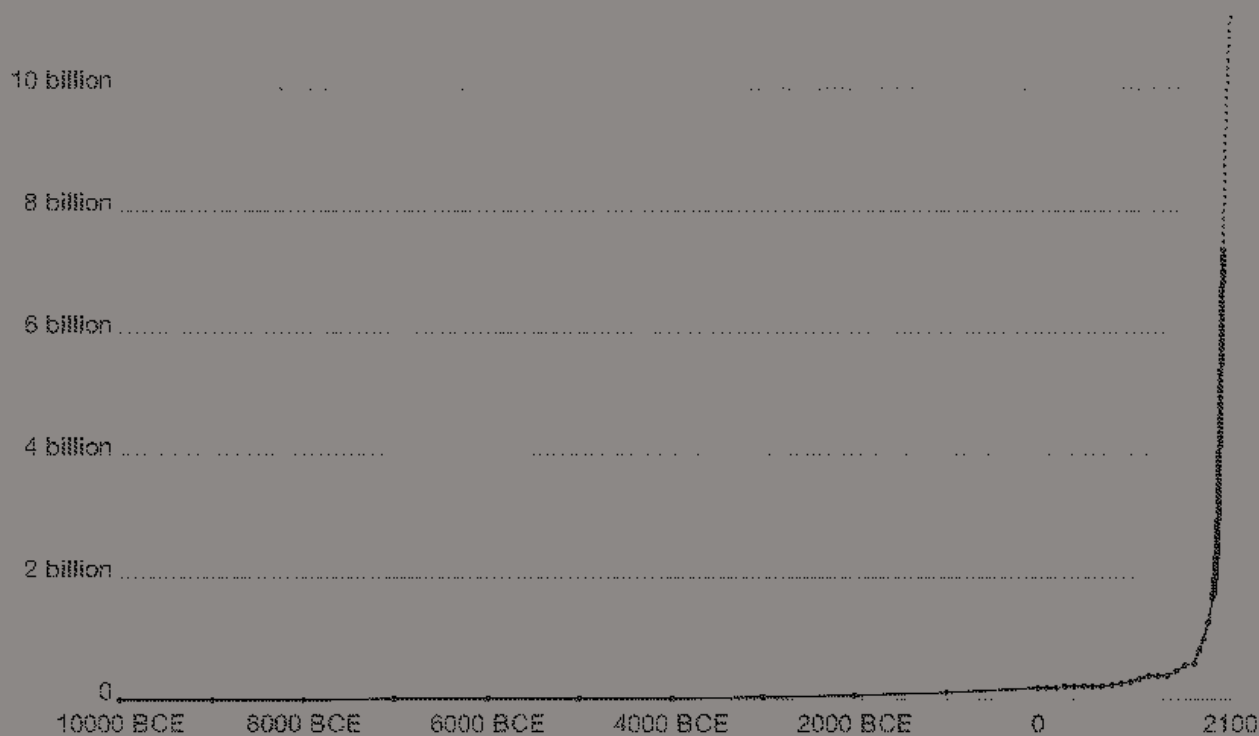
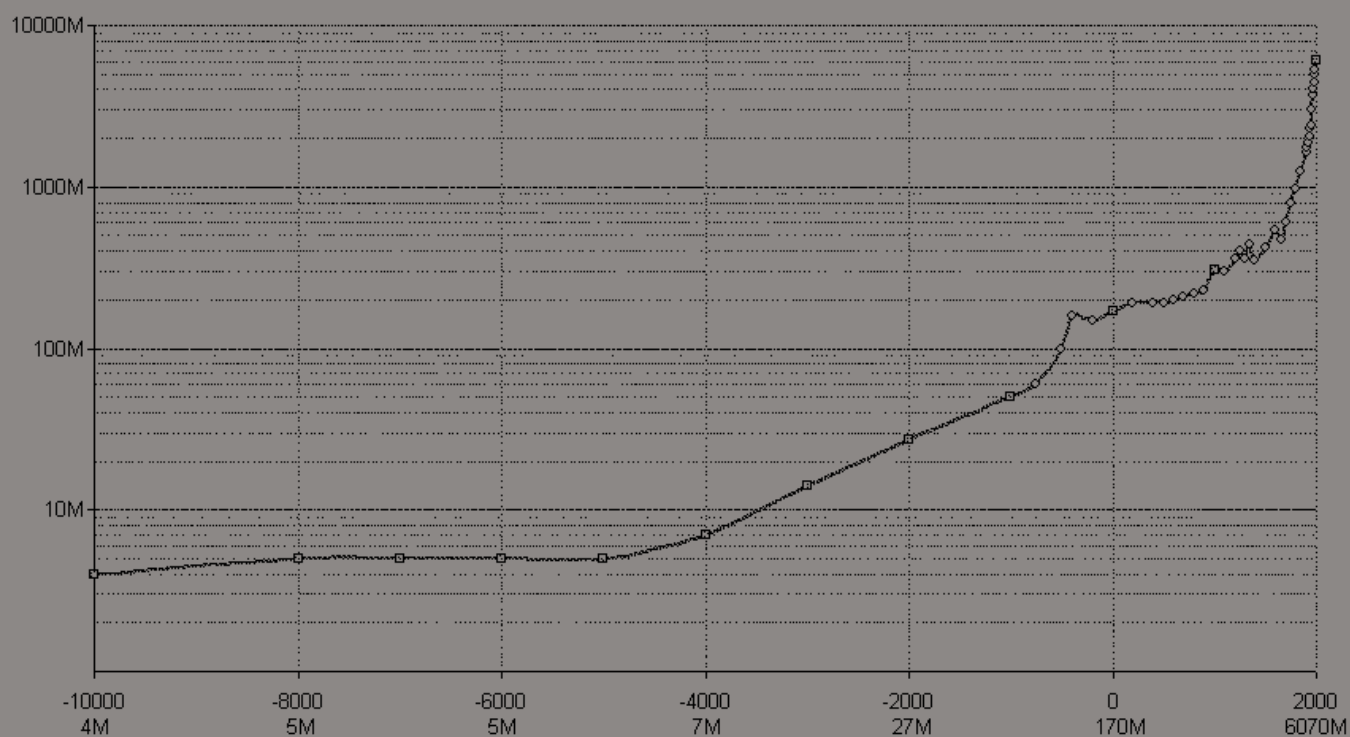
KUVA 7. Informaatiomuotoilun kieliopista on tarjolla ansiokas kokoava esitys esimerkiksi Koposen, Hildénin ja Vapaasalon *Tieto näkyväksi* -oppaassa (2016). Kuvassa on Bertinin (1983, 42) lähes maaginen, kokoava esitys visuaalisista muuttujista: keskeltä ylhäältä alaken myötäpäivään kuvassa ovat koko, tummuus, gradientti, väri, kulma, muoto ja keskellä sijoittuminen kaksiulotteisella pinnalla.

3.6 GRAAFINEN REHELLISYYS

Visuaalista esitystä laadittaessa on tärkeää panostaa on graafiseen rehellisyyteen; Edward Tuften *graphical integrity* -käsitettä erittelee muun muassa Kuusela (2000, 194). Tiedon visualisoinnin etiikasta kirjoittavien (esimerkiksi Pater 2016, 164) suosikki-esimerkkeinä tätä hyvettä vastaan rikkomisesta ovat puutteellisen visuaalisen vertailtavuuden tapaukset, joissa asteikkoja manipuloimalla on liioiteltu tai vähätelty muuttujien välisiä eroja.

Harhaanjohtavaa on myös käyttää erikoisasiantuntemusta vaativaa asteikkoa maallikoille suunnatuissa julkaisuissa (Koponen, Hildén & Vapaasalo 2016, 217). Uskaltaisin esimerkiksi itse väittää, että osa niinkutsutun 'Suuren Kiihtymisen' (*The Great Acceleration*)³ eksponentiaalinen eli jatkuvasti tuplaantuva, epäintuitiivisen nopeasti kiihtyvä luonne tulee edelleen monille yllätyksenä, sillä kasvua kuvataan usein logaritmisella asteikolla.

3 Eksponentiaalinen eli kiihtyvyydeltään kiihtyvä talous- ja väestönkasvu sekä niiden aiheuttamien ympäristökuormituksen lisääntyminen



KUVAT 8 & 9. Kaksi erilaista tapaa kuvata eksponentiaalinen kasvu: harmonisen tasainen sekä ahdistavan räjähtävä.

Valinta on hyvien tilastografiikkatapojen mukaan perusteltu: diagonalisointiperiaatteen mukaan aineiston ominaispiirteet – tässä tapauksessa kasvuvauhdin tasainen kiihtyminen – tulevat todennäköisesti parhaiten esiin, kun asteikot pyritään säätämään siten, että muutoskäyrät ovat noin 45 asteen kulmassa (Koponen, Hildén & Vapaasalo 2016, 211). Esitys ei kuitenkaan anna yliopistomatiikkaan perehtymättömälle oikeaa kuvaa meidän kaikkien elinympäristöjä ravistelevan kasvun huimaavasta äkillisyydestä.

3.7

EPÄMÄÄRÄISYYDEN ESITTÄMINEN

Datafeministit vaativat, että datan kontekstin esiintuomiseksi on kehitettävä parempia keinoja epätarkkuuden ja -määräisyyden sekä puuttuvan datan, aukkojen ja rajausten ilmentämiseen. Kehittämällä nokkelia keinoja näiden visuaaliseen esittämiseen onnistutaan purkamaan visualisaatioiden luomaa neutraaliuden ja objektiivisuuden illuusiota. Epämääräisyyden harkitulla ilmentämisellä saadaan visualisointi ikään kuin asettamaan itsensä so-

pivissa määrin kyseenalaiseksi ja kehoitetaan lukijaa valpastumaan ajattelemaan kriittisesti.

Myös esimerkiksi Koposen, Hildénin ja Vapaasalon (2016, 201–202) mielestä on tärkeää ilmentää, milloin kuviossa on kyse varsinaisista mitaustuloksista ja milloin arvioista tai ennusteista. Epävarmuutta voidaan periaatteessa kuvata samoilla visuaalisilla muuttujalla siinä missä muitakin aineiston ominaisuuksia, mutta jotkut keinot soveltuvat tähän intuitiivisemmin kuin toiset. Tiedeviestinnässä onkin kehitetty erityisesti todennäköisyyksien, vaihteluvälien ja muiden virhemarginaalien kommunikointiin lukuisia keinoja, kuten jana- tai viulukuvio. Todennäköisyyksien ja epävarmuuden hahmottaminen on kuitenkin usein jopa asiantuntijoille (Salmela 2003), saati sitten suuremmalle yleisölle haastavaa.

Myös sumeutta ja liukuvärejä, jotka ovat kehnon vertailtavuutensa vuoksi eksaktiin esittämiseen keltottomia visuaalisia muuttujia, on ajoittain kokeiltu epävarmuuden ja -tarkkuuden ilmaisutavoiksi (Liu, Mirzangar, Kirby, Whitaker & House 2015). Verkkojulkaisemisen ansiosta keinovalikoimaan on mahdollisesti vakiintumassa myös animaatio.

D'Ignazio ja Klein (2018, "Chapter Two: On Rational, Scientific, Objective Viewpoints from Mythical, Imaginary, Impossible Standpoints") esittelevät liikkeen hyödyntämisen esimerkkinä *The New York Timesin* grafiikkatoimituksessa huolella kehitetty Yhdysvaltain vuoden 2016 presidentinvaalien tulosta ennustavan todennäköisyysmittarin. Visualisaatio esitti hyvin kouriintuntuvasti sen, millainen epävarmuus vaalien tulosta koskeviin ennusteisiin liittyy: pelkän näkemisen sijaan katsojat saatiin myös tuntemaan epävarmuus.

Epäkonventionaalinen esitystapa keräsi paitsi kiitosta, myös pöyristynyttä kritiikkiä – väittivät-pä jotkut visualisoinnin jopa vaikuttaneen ihmisten äänestyskäyttäytymiseen. (D'Ignazio & Klein 2018 "Chapter Two – – ") Itse pidän *New York Timesin* valintaa perin onnistuneena – näen visualisoinnin nostattaman kritiikin lähinnä osoittavan sen, miten konservatiivisia odotuksia tiedon visualisointeihin liittyy. Visualisointien odotetaan lisäävän ymmärryksen ja varmuuden tunnetta harkittulla selkeydellään ja tyynellä kokonaisvaltaisella esityksellään, joten näiden odotusten rikkominen tai niiden kyseenalaistaminen on raikasta ja tervetullutta itsekritiikkiä.

Tyhjyyden visualisointiin on niin ikään laaja valikoima jo olemassaolevia keinoja. Datavisualisti Andy Kirk (2014a) erittelee sille esitelmässään neljä erilaista tyyppiä: puuttuvan tiedon (*"null"*), nollan (*"zero"*), tyhjän (*"blank"*) ja näkymättömän (*"invisible"*). Datafeministin kriittisen silmän on kiinnitettävä huomiota juuri puuttuvaan tietoon: aukko aineistossa voi ilmentää jonkin näkökulman laiminlyöntiä datan keruuprosessissa tai olla johtolankana suoranaisestä salailusta. Näin ollen aukkojen tulkitseminen ja niiden ilmentäminen visuaalisesti voi toimia kriittiseen ajatteluun herättelevänä disruptiona, häiriönä, ja kytkeä sitä tiiviimmin materiaaliseen tai historialliseen kontekstiin.

Tyhjä tila puolestaan paitisi antaa designille tilaa hengittää, on myös viisaasti käytettynä dramaattinen keino kiinnittää huomio jonkin poissaoloon tai käsiteltävän asian mittakaavaan. (Kirk, 2014b) Sääteliäästi hyödynnettynä se voi nähdäkseni olla myös erinomainen keino antaa käyttömälleen datasetille viimehetken visuaalista kritiikkiä jonkin näkökulman huomiotta jättämisestä.

Olen myös pohtinut tyhjän tilan hyödyntämistä asioiden mittakaavaan ravistelemiseen: miltä esimerkiksi äänestystulos näyttää, kun se rinnastetaan koko väestöön? Entä jos huomioon otetaan – Donna Harawayn (esimerkiksi 2016) koko elonkehän huomioivan ajattelun inspiroimana – kaikki valtion alueella elävät eläimet ja kasvit?

3.8 TUOTANTO JA VIIMEISTELY

Yksinkertaisimmillaan tuotantoon ja viimeistelyyn kuuluu lähinnä visualisaation siistiminen ja valmistelu esimerkiksi painamista tai verkkojulkaisua varten (Koponen, Hildén & Vapaasalo 2016, 315). Toisaalta tässä työvaiheessa saatetaan vielä tehdä monia visualisaation tunnelmaan ja äänensävyyn vaikuttavia esteettisiä valintoja, kuten visualisaatioon liittyvien tekstien muotoilua. Keskitynkin tämän vaiheen osalta erityisesti juuri tekstuaalisten selitteiden merkitykseen. Lisäksi käsittelen hieman interaktiivisuutta, sillä sekin toteutetaan tavallisesti tässä suunnitteluvaiheessa.

Kaikkien alan hyveiden mukaan minimalistiseksi ja geneeriseksi laadittu visualisointi on irti todellisuudesta, sillä se esittää vain abstraktien muuttujien välisiä suhteita. Tästä syystä huolellisella kielellisten selitteiden käytöllä on keskeinen rooli tiedon visualisoinnissa: ne kiinnittävät kuvan maailmaan selittämällä, mitä todellisuuden osia visuaalisen esityksen eri elementit esittävät eli representoivat. Sanavalinnoilla ja vaikkapa otsikoinnilla ohjataan lukijan huomio visualisoinnin tarjoamiin olennaisimpiin oivalluksiin ja annetaan myös laajempi tulkintakonteksti. Erityisesti kuvien viraalin levittämisen aikakaudella kaikki selitteet on syytä sisällyttää itse kuvatiedostoon, jottei asiayhteydestään irrotettua kuvaa voi ymmärtää väärin.

D'Ignazio ja Klein (2018, "Chapter Five: Numbers Don't Speak for Themselves") esittelevät esimerkkinä kaksi identtistä kuvaajaa, joista toinen on otsikoitu "Mental Health in Jail" ja toinen "Racism in Jail". Ensimmäinen, "neutraalimmin" otsikoitu kuvaaja jättää liikaa liikkumavaraa lukijan spekulatioille ja mahdollisesti stereotyyppittelevälle ajattelulle. Kuvioista ei lainkaan käy ilmi, että rodullistettua mielenterveysoireilijaa rangaistaan vankiloissa valkoista vankia useammin eristyssellillä sen sijaan, että hänet päästettäisiin terapiaan.

Digitaalisen julkaisemisen myötä vuorovaikutus visualisaation kanssa on keskeinen keino muun muassa visualisaation sisältämän tietomäärän kasvattamiseen selkeydestä tinkimättä. Erilaisten mahdollisten rinnastusten ja vertailujen määrä moninkertaistuu, kun tiedon tasoja tulee lisää. Tässä yhteydessä on toki kannettava graafiseen rehellisyyteen kuuluva vastuu siitä, että kaikki vertailut, jotka vuorovaikutteisessa visualisaatiossa ovat tehtävissä, ovat jollakin tavoin perusteltuja.

Vuorovaikutteisuus myös kutsuu entistäkin voimakkaammin vastaanottajaa omiin tutkimuksiin ja päätelmien tekemiseen: visualisaatiota ”pelatesaan” käyttäjä saattaa immersoitua eli imeytyä visualisoinnin maailmaan, jolloin sen ennakkovalinnat ja rajaukset alkavat tuntumaan hänestä itsestäänselviltä: esimerkiksi oletusvertailut, ennustava tekstinsyöttö, valikot ja muut valmiiksi tarjotut vaihtoehdot ohjaavat voimakkaasti tulkintaa (Hulman & Diakopoulos 2011, 2236–2237). Visualisaation on siis tässäkin yhteydessä vastuullista tehdä omat rajauksensa selkeästi tiettäväksi.

Itse ideoin esimerkiksi erilaisia leikkisiä glitch-estetiikkaan perustuvia ratkaisuja: visualisaation käyttöliittymä voisi osoittaa omat rajansa erilaisin häiriöin, esimerkiksi siten, että jotkin valinnat

menevät ilkkurisesti toistuvasti päälle tai vain kiel-
täytyvät toimimasta.

D'Ignazion (2015) peräänkuuluttama visualisoin-
nin ja aineiston kanssa käytävän dialogin vaatimus
ei myöskään täyty vain muutaman ennalta annetun
suotimen tai pudotusvalikon myötä. Hänen mieles-
tään tarvitaan nokkelampia tapoja osallistaa käyt-
täjää ja kenties antaa heille aidosti valtaa visualisaa-
tiossa. NykYTEknologialla visualisaatioihin lienee jo
verrattaen helppoa rakentaa erilaisia palauteratkai-
suja. Verkkosivun evästealgoritmit keräävät taustal-
la käyttäjästä joka tapauksessa dataa; kenties olisi
aika antaa käyttäjien ottaa kantaa aktiivisemmin ja
suostumukseen perustuen.

Myös Koponen, Hildén ja Vapaasalo (2016, 315)
korostavat, että tiedon visualisoiJien vastuu ei pää-
ty esityksen julkaisuun. On olennaista kerätä työs-
tä palautetta sekä muuten seurata visualisaation
vastaanottoa ja yleisöomääriä. Sähköisessä julkaise-
misessä on myös mahdollisuus grafiikan päivittä-
miseen ja laajentamiseen siitä saadun palautteen ja
muun karttuneen tiedon valossa.

3.9 SELKEYS JA USKOTTAVUUS

Käytyäni läpi datavisualisaation prosessin vaihe vaiheelta keskityn vielä erityisesti yhteen datavisualisaation laadinnassa keskeiseen ohjenuoraan: totuudenmukaisuuden ohella kaikista tärkeimpänä visualisaation laadinnassa korostetaan perinteisesti selkeyttä (Koponen, Hildén & Vapaasalo 2016, 32; Kuusela 2000, 194; Tufte 2006, 79). Kaikki muut ratkaisut ja valinnat ovat alisteisia sille, että saadaan aikaiseksi paitsi totuudenmukainen, myös rakenteeltaan selkeä esitys.

Visualisoinnin on siis syytä esittää muutoksia aineistossa, ei esitystavassa (Kuusela 2000, 194). Viusaalista hälyä on oltava siinä määrin vähän ja muutujien koodauksen niin yksiselitteistä ja hyvin järjestettyä, että aineiston ominaisuudet ovat kuvan pääosassa ja ilmenevät siitä mahdollisimman vähäeleisesti ja vaivattomasti. Visualisoinnin mielekäs järjestely jo itsessään saattaa antaa lisää tietoa aineistoista: Tufteen (esimerkiksi 1990, 46–47; myös Kuusela 2000, 23) lanseeraama data-muste-suhde -ajattelu (*data to ink ratio*) ihannoitilanteita, jossa sama mustepisara koodaa useampaa aineiston ominaisuutta. Myös esimerkiksi edellä mainittuja dra-

mattisia tyhjän tilan tapauksia voitaneen pitää tämän periaatteen ilmentymänä: kun kuvapinta-ala on älykkäässä käytössä, myös musteen poissaolo saa valtavan merkityksen.

Edward Tufte (muun muassa 1990, 34) minimalistisen eetoksen mukaisessa tiedon visualisoinnissa paheksutaan esityksen koristelua tiedonvälityksellisesti perustelemattomalla roskagrafiikalla (*chart junk*). Kolmiulotteisuus, räikeät värit tai kuvitukselliset elementit nähdään painomusteen tuhlaamisena ja harhaanjohtavana yrityksenä peitellä kuvaajan köyhää sisältöä. (Tufte *ibid.*) Pahimmillaan koristelu tosiaan aiheuttaa varsinaisen viestin tulkintaa: esimerkiksi kolmiulotteisten kuvioiden koon yksiselitteinen vertailu on yksinkertaisesti mahdotonta (Koponen, Hildén & Vapaasalo 2015, 80–81).

Visuaalisesta melun puuttumisesta on tullut myös jonkin sortin uskottavuuskriteeri ja statussymboli visualisoinneille. Tämä on visualisointien genressä niin tärkeää, että Hullman ja Diakopoulos (2011, 2234–2235) kertovat havainneensa tarkoitushakuisesti sotkuiseksi muotoiltua grafiikkaa käytettävän esimerkiksi poliittisten vastustajien kritisointiin ja jopa suoranaiseen naurunalaiseksi asettamiseen.

”Kuka uskoisi tilastoa, joka näyttää videopeliltä?” Tufte (1990, 34, kirjoittajan oma käännös) kysyy, mutta sosiaalisen median kyllästämässä huomiotaloudessa kyynistynyt datafeministi saattaa kysyä vastakysymyksen: kukapa toimisi sellaisen tilaston pohjalta, joka ei tunnu missään? Koponen, Hildén ja Vapaasalokin (2016, 78) muistuttavat, että tutkimuksissa on tosiaan havaittu, että datan selkeää esitystä vaarantamaton, teemaan sopiva koristeilu saattaa jopa auttaa visualisaation viestinnällisten päämäärien saavuttamisessa ja tehdä siitä entistä mieleenpainuvamman .

Designer-tutkija Kelly Dobson (D’Ignazio & Klein 2018, ”Chapter Two: On Rational, Scientific, Objective Viewpoints from Mythical, Imaginary, Impossible Standpoints”) haastaa puhumaan tiedon visualisoinnin sijaan kehollistamisesta (*”visceralization instead of visualisation”*). Datan kokemista ei pidä rajoittaa vain näköön, vaan sitä voi jäsentää myös muiden aistien ja koko kehon kautta – etenkin, jos tieto pitää saada tuntumaan luissa ja ytimissä.

Länsimaiseen ajatteluun juurtunut käsitys järjen ja tunteen, mielen ja kehon, jyrkästä erottelusta periytyy antiikista, tieteellisen vallankumouksen mekanistisista selitysmalleista ja yhteiskuntahierarkoita maallistuneesti pönkittävästä valistusfilosofiasta.

Erottelu on kuitenkin 1900-luvulta alkaen kumottu useasta näkökulmasta ja moneen otteeseen⁴: tehokas ja tarkoituksenmukainen ajattelu ja toiminta tapahtuu järjen ja tunteen erottamattomana yhteistyönä.

Esimerkkinä näistä virtauksista mainittakoon Domestic Data Streamersin (2016) yhteistyössä UNICEF:in kanssa YK:n huippukokoukseen toteutama lapsiin kohdistuvia riskitekijöitä elämyksellistävä installaatio. Häpeilemättömän sentimentaalinen aikakone herätteli moniaistisesti maailman johtajien omia lapsuusmuistoja ja esitteli samalla tietoa siitä, miten kokouksessa tehtävät päätökset vaikuttavat haavoittuvimmissa asemassa oleviin lapsiin, joita on nykyisellään vaikea kartoittaa saati auttaa.

Toisena esimerkkinä mainittakoon yhteiskunnalliseen vaikuttamiseen ja erityisesti haavoittuvassa asemassa olevien kolmannen maailman naisten auttamiseen tähtäävän M4ID-toimiston (2018) projektit. Heillä on käynnissä esimerkiksi "Images as Data" -hanke, jossa laadukkaita valokuvien dokumentoita terveydenhuollon fasiliteetteja ja asiakkaan palvelupolkua niiden läpi Intiassa, Indonesiassa,

4 Lisää aliheesta voi lukea esimerkiksi filosofien Simone de Beauvoir tai Jean-Luc Nancy sekä äivotutkija-neurologi Antonio Damasion teoksista.

Keniassa ja Ugandassa. Kuvat tarjoavat kouriintuntuva kontekstin ja huomattavasti rikasta informaatiota, jonka pohjalta voidaan jatkaa terveydenhuollon resurssien ja proserssien optimoimista.

3.10 UHTEENVETO

Tiedon visualisointi on erittäin kiperässä matematiikan, tietojenkäsittelyn, havaintopsykologian, journalismin sekä kulttuuristen piiloviestien leikkaus pisteessä, joten ei ole nähdäkseni kovinkaan yllättävää, että sen estetiikka on vakiintunut kovin pelkistetyksi ja melko konservatiiviseksi. Visualisointiratkaisuihin on paljon hyväksi havaituista perinteistä ammentavia nyrkkisääntöjä ja kikkoja: joka kerta ei ole syytä yrittää keksiä pyörää uudelleen (Koponen, Hildén & Vapaasalo, 2016). Aivan niin kuin typografiassa tai filosofiassa, on arrogantin aloittelijan virhe luulla ponnistavansa kylmiltään oman luovuutensa varassa pidemmälle kuin satojen vuosien visualisointitradition jättiläisen olkapäiltä.

Tuttuihin kuvaajatyyppeihin tukeutuminen myös lisää kuvion saavutettavuutta tekemällä vastaanottajan lukukokemuksesta sujuvamman: vieraan kuvaamistavan lukeminen vaatii aina opettelua

– aikaa ja energiaa – joten kuvaajan on syytä tarjota riittävän huomattava oivallus palkkioksi nähdystä vaivasta (Koponen, Hildén & Vapaasalo 2016). Liika tukeutuminen aiemmin nähtyihin ratkaisuihin saattaa kuitenkin rajoittaa visuaalista mielikuvitusta myös negatiivisessa mielessä: vakiintuneet konventiothan saattavat pönkittää jotakin vinoumaa tai muutoin vääristää ihmisten mielikuvia aiheesta, kuten väitän tapahtuneen esimerkiksi Suuren Kiihtymisen tapauksessa. Ehkä totunnaisuuksia tosiaan olisi syytä ravistella?

Lopulta niin klassikot kuin kriitikotkin ovat yhtä mieltä siitä, että jokainen visualisaation tekoprojektissa sisältämä valinta on tehtävä harkiten, kunkin tapauksen ominaislaatuun tarkoin perehtyneenä. Ei ole olemassa mitään patenttiratkaisua, joten sopiva esitystapa selviää yleensä aineistolla kriittisesti leikkitelemällä, luonnostelemalla sekä potentiaalisen kohderyhmän edustajien kanssa testaamalla. Feministisen näkökulman myötä visualisoija tulee entistä tietoisemmaksi työhönsä kuuluvasta vastuusta ja suhtautuu entistä kriittisemmin paitsi omiin ratkaisuihinsa, myös traditioon.



Pyrkimyksenäni tässä tutkielmassa oli selvittää, miten tehdä feminsitisempiä datavisualisaatioita. Paradoksaalisesti visualisaation käytännön tekoprosessin läpikäyminen osoittautui lopulta yllättävän abstraktiksi lähestymistavaksi. Tiedon visualisoinnin prosessi on niin laaja ja polveileva, että yksittäisiin työvaiheisiin ei voinut pureutua kovin syvällisesti. Olikin suuri haaste yrittää rajata aiheen käsittelyä siten, ettei siitä olisi tullut vain pinnallista alan niksien luetteloa. Käsittely on edelleen melko pirstaleinen, mutta katson yleiskatsausmaisuuuden sopivan kandidaatintutkielman tasoiseen työhön. Pyrin myös erityisesti analysoimillani esimerkeillä havainnollistamaan olennaisimpia oivalluksia, ja uskon niiden olevan kyllin konkreettisia, jotta myös lukijan saa uusia työkaluja visualisaatioiden kriitti-

seen tarkasteluun ja kenties jopa uusien ratkaisujen ideointiin.

Datan ja tiedon suhdetta ja näihin liittyviä valtarakenteita selvittäessä päädyttiin filosofisestian äärelle: huomasin toistuvasti palaavani epistemologian eli tietoteorian sekä tieteenfilosofian peruskysymyksiin totuuden, tiedon, objektiivisuuden ja tieteellisen metodin perusolemuksesta. On mielenkiintoista, miten poikkeuksellisessa roolissa tiedon visualisointi toimii rajapintana ihmsten ja tieteen välillä. Visualisointi kokoaa, jäsentää ja popularisoi tietoa esimerkiksi eksaktia tekstiä huomattavasti nopeammin ja intuitiivisemmin – erityisesti tässä suhteessa on valitettavaa, että niiden kriittinen lukutaito on edelleen harvojen etuoikeus.

Erityisen olennaista on huomioida datan ja tiedon ero – nämä tuntuvat ajoittain menevän sekaisin *big data* -huumassa¹. Ihmiset kokevat kartuttavansa visualisaatioilla ymmärrystään ja tietoaan, mutta

1 D'Ignazio ja Klein (2018, "Numbers Don't Speak For Themselves") antavat esimerkin muun muassa Wired-toimittaja Chris Andersonista, joka kiteyttää toiveet big datan suhteen seuraavasti: kun aineistoa on tarpeeksi, korrelaatio eli säännönmukaisuuksien löytäminen riittää, selityksiä ei tarvita. Ainakin tässä pelkistetyssä muodossaan linjaus viestii kylmäävän kritiikitöntä luottamusta datankeruumenetelmien poikkeukselliseen erehtymättömyyteen.

jos visualisaatio on vain nokkela datan jäsenitys ilman sen merkitysten pohdintaa ja kontekstualisointia, on kyseessä vain selkeä esitys verrattaen mielivaltaisesti kerätyistä luvuista, joihin vastaanottaja heijastaa omat ennakkotietonsa ja -luulonsa. On informaatiomuotoilijan vastuulla pitää huoli, että visualisaatio todellakin sisältää ja välittää tietoa tarjoamalla riittävän rikas tulkintakonteksti.

Tässä informaatiomuotoilun klassikot ja intersektionaalinen datafeminismi lyövät nähdäkseni lämpimästi kättä. Lopultakin molemmat tahot pyrkivät siihen, että visualisointi esittää ”tosia väitteitä todellisesta maailmasta” (Koponen, Hildén & Vapaasalo 2016, 32), tiivistää todenmukaisen esityksen kannalta juuri oikean määrän monimutkaisesta tietoa selkeäksi esitykseksi (Tufte 2006, 78–79). Feministinen näkökulma vain tarjoaa lisää perspektiiviä ja työkaluja arvioida, mikä tämä kunkin esityksen kannalta riittävä määrä tietoa ja näkökulmia kulloinkin on. Se vaatii myös arvioimaan, onko visualisoijien kuvaama käsitys todellisuudesta sellainen, jonka visualisaation lukijat tai ihmiset, joista data on kerätty, mielekkäästi jakavat. Feminismiä datan visualisoinnissa on siis ennen kaikkea pyrkimys ajatella tietoa aiempaa laajemmin ja kontekstisidonnaisemmin sekä tiedostaa oma vas-

tuullinen asemansa tietoa välittävänä, selittävänä ja jopa luovana vallankäyttäjänä.

Tutkimusta olisi voinut vielä syventää haastatteleamalla alan ammattilaisia, kuten suomalaisen tiedon visualisoinnin pioneeriteoreetikkoja Juuso Koposta ja Jonatan Hildéniä tai vaikkapa maan suurimpana sanomalehtenä huomattavaa valtaa käyttävän *Helsingin Sanomien* grafiikkatoimitusta. Myös M4ID-toimistoa olisi mielenkiintoista haastatella heidän kekseliäästä suhtautumisestaan dataan.

Tulevissa tutkimuksissa voisi myös tarkentaa käsittelyä vain muutamaaan yksityiskohtaan. Voisi esimerkiksi tutkia feminististä valta-analyysiä, semiootiikkaa, estetiikkaa ja diskurssianalyysiä yhdistellen juuri visuaalisen selkeyden vakuttavaa voimaa. Ei ole itsestään selvää, miksi niin tiedon visualisointi kuin web design ja korporaatiomaailmakin ovat takertuneet selkeisiin geometrisiin muotoihin ja sveitsiläiseen typografiseen tyyliin. Tällainen tutkimus olisi paitsi luonteva siirto henkilökohtaisella ristiretkelläni modernismia vastaan, myös niveltäisi ajankohtaisesti kapitalismia ja kolonialismia purkavaan akateemiseen keskusteluun.

Kriittisestä näkökulmastani huolimatta olen tutkimukseni myötä ymmärtänyt entistä paremmin käytössä hioutuineiden vakiintuneiden keinojen

eleganttiuden ja tarkoituksenmukaisuuden. Monet pelkistetyt ratkaisut on usein tehty monelta kantilta perustelluista syistä: on vaikuttavaa, miten paljon asiantuntevaa ajattelua yksinkertaiseen visuaaliseen esitykseen on kiteytynyt. Olenkin ollut vähällä nöyrytyä myöntämään, että monet aiemmat, kovin kokeelliset visualistointiyritykseni ovat olleet vain liikuttavan arroganttia erikoisuudentavoittelua.

Datafeminismiin perehtyminen on kuitenkin pelastanut suoranaiselta ramppikuumeelta ja lamaanukselta: on äärimmäisen inspiroivaa, että uusien, läpinäkyvämpien ja elämyksellisempien visualisoinnin keinojen kehittämiseen kannustetaan jopa railakkaan anarkistisesti. Feministinen näkökulma on onnistunut siis voimaannuttamaan ja laskemaan yrittämisen kynnystä tällä varsin konservatiivisella tiedon visualisoinnin alalla ainakin omalla kohdallani. Toivonkin datan visualisoinnin työkalujen ja taitojen olevan entistä useampien ulottuvilla tulevaisuudessa, jotta voimme yhdessä ottaa meistä louhitun datan entistä paremmin omaehtoiseen käyttöömme.

Tärkeä henkilökohtainen havainto tässä tutkimuksessa on ollut, että haluan palavasti jatkaa työskentelyä tiedon, erityisesti datan visualisoinnin parissa. Ala on kiinnostukseni kohteiden kannalta

täydellisessä yhteiskunnallisen vaikuttamisen, tutkivan journalistisen otteen, tiukkojen sääntöjen puitteissa operoimisen ja tietoteknisen näpertämisen leikkauspisteessä. Haluan päästä mukaan kehittämään uuden sukupolven kaikki mukaan ottavia, rehellisiä ja kaikin puolin hyviä visualisaatioita, sillä niitä tämä maailma tarvitsee.



KUVALÄHTEET

- 1 tuntematon taiteilija, tieteellinen kuvitus vesinokkaeläimestä (*Ornithorhynchus anatinus*), *The British Library*, ei tekijänoikeusrajoituksia, ladattu 22.3.2019 osoitteesta [flickr.com/photos/britishlibrary](https://www.flickr.com/photos/britishlibrary/).
- 2 Event Horizon Telescope Collaboration 2019, "First Image of a Black Hole", sivustolla [Eventhorizontelescope.org](https://eventhorizontelescope.org), kaikki oikeudet pidätetään, ladattu 12.4.2019 osoitteesta static.projects.iq.harvard.edu/files/styles/os_files_xlarge/public/eht/files/20190410-78m-800x466.png?m=1554877319&itok=mogPFSet.
- 3 Koponen, Juuso 2017, "Avoimet työpaikat lisääntyvät, mutta työttömyys ei vähene", artikkelissa "Valtiovarainministeriön työllisyysgrafiikka ei johda tarkoituksella harhaan" sivustolla [Informaatiomuotoilu.fi](https://informaatiomuotoilu.fi), Creative Commons BY4.0-käyttöoikeus, ladattu 2.4.2019 osoitteesta informaatiomuotoilu.fi/2017/03/valtiovarainministerion-tyollisyysgrafiikka-ei-johda-tarkoituksella-harhaan.
- 4 Saarinen, Ville Ojala, Juho 2018, "Flow Towards Europe", kuvakaappaus samannimisestä artikkelista sivustolta [Lucify.com](https://lucify.com), grafiikka perustuu avoimeen lähdekoodiin ja lisäksi kenellä tahansa on oikeus upottaa grafiikka omalle sivustolleen, otettu 4.4.2019 osoitteesta www.lucify.com/the-flow-towards-europe
- 5 Academic Torrents 2019, kuvakaappaus sivustolta academictorrents.com, otettu 10.4. osoitteesta academictorrents.com/browse.php?page=1.
- 6 tuntematon taiteilija, tieteellinen kuvitus erikoisesta sienestä, *The British Library*, ei tekijänoikeusrajoituksia, ladattu 22.3.2019 osoitteesta [flickr.com/photos/britishlibrary](https://www.flickr.com/photos/britishlibrary/).

- 7 Bertin, Jacques 1983, "Visual Variables", teoksessa *Semiology of Graphics* s. 43.
- 8 Waldir, "World population growth (lin-log scale)", sivustolla *Wikipedia.org*, Creative Commons Attribution-Share Alike 3.0 Unported -käyttöoikeus, ladattu 21.3.2019 osoitteesta [commons.wikimedia.org/wiki/File:World_population_growth_\(lin-log_scale\).png](https://commons.wikimedia.org/wiki/File:World_population_growth_(lin-log_scale).png)
- 9 Roser, Max & Ortiz-Ospina, Esteban 2013/2017, "World Population over the last 12,000 years and UN projection until 2100", kuvakaappaus artikkelista "World Population Growth" sivustolla *ourworldindata.org*, otettu 21.3.2019 osoitteesta ourworldindata.org/world-population-growth.
- 10 Koponen, Juuso 2017, "Yllätyksiä luvassa", artikkelissa "Näiden vaalien tuloksen ennustaminen on tavallistakin vaikeampaa" sivustolla *Informaatiomuotoilu.fi*, Creative Commons BY4.0-käyttöoikeus, ladattu 13.4.2019 osoitteesta informaatiomuotoilu.fi/2019/04/naiden-vaalien-tuloksen-ennustaminen-on-tavallistakin-vaikeampaa/.
- 11 Liu L., Mirzangar M., Kirby R.M., Whitaker R. & House D. H. 2015, kuva artikkelissa "Visualizing Time-Specific Hurricane Predictions, with Uncertainty, from Storm Path Ensembles", julkaisussa *Eurographics Conference on Visualization (EuroVis) 2015*, vol. 34, no. 3, luettu 27.3.2019 osoitteesta nations/www.cs.utah.edu/~kirby/Publications/EuroVis-2015.pdf.
- 12 Aisch, Gregor, Cohn, Nate, Cox, Amanda, Katz, Josh, Pearce, Adam & Quealy, Kevin 2016, vaalivisualisaatio The new York Timesiin, kuvakaappaus teoksessa Catherine D'Ignazio & Lauren Klein, *Data Feminism*, ladattu 10.4.2019 osoitteesta bookbook.pubpub.org/pub/8tjbs2x5.
- 13 *The Independent*, etusivu vuodelta 2006, käyttöoikeudet eivät tiedossa, ladattu 10.4.2019 osoitteesta hbr.org/2014/05/visualizing-zero-how-to-show-something-with-nothing.
- 14 D'Ignazio, Catherine 2015, "Two Potrayals of the Same Data", teoksessa Catherine D'Ignazio & Lauren Klein, *Data Feminism*, ladattu 10.4.2019 osoitteesta bookbook.pubpub.org/pub/6ui5n4vo.
- 15 Domestic Data Streamers 2016, kuva artikkelista "Making Visible The Invisible" sivustolla *domesticdatastreamers.com*, ladattu 10.4. osoitteesta domesticstreamers.com/makingVisible.

PAINETUT LÖHTEET:

- 1 Bertin, Jacques 1983/2011 *Semiology of graphics. Diagrams, networks, maps*, Esri Press, California.
- 2 Cairo, Alberto 2012, *The Functional art. An Introduction to information graphics and visualisation*, New Riders, Berkeley.
- 3 Koponen, Juuso, Hildén, Jonatan & Vapaasalo, Tapio 2016, *Tieto näkyväksi*, Aalto Arts Books, Helsinki.
- 4 Kuusela, Vesa 2000, *Tilastografiikan perusteet*, Edita, Helsinki.
- 5 McLuhan, Marshall 1969/1964, *Ihmisen uudet ulottuvuudet*, suomentanut Antero Tiusanen, WSOY, Porvoo.
- 6 Rolin, Kristina 2005, "Tutkimuksen näkökulmasidonnaisuuden paradoksi", teoksessa Liisa Husu & Kristina Rolin (toim.), *Tiede tieto ja sukupuoli*, Gaudeamus, Helsinki, s. 97–111.
- 7 Tufte, Edward, 1990, *Envisioning Information*, Graphics Press LLC, Cheshire (Connecticut).
- 8 Tufte, Edward 2006, *Beautiful Evidence*, Graphics Press LLC, Cheshire (Connecticut).
- 9 Veikkola, Eeva-Sisko 2005, "Sukupuoli ja tilastot" teoksessa Liisa Husu & Kristina Rolin (toim.), *Tiede tieto ja sukupuoli*, Gaudeamus, Helsinki, sivut 84–96.

SÄHKÖISET LÄHTEET

- Adriaans, Pieter 2012/2018. "Information", Stanford Encyclopedia of Philosophy, luettu 27.2.2019 osoitteessa plato.stanford.edu/entries/information.
- D'Ignazio, Catherine 2015, "What would feminist data visualization look like?" sivustolla *MIT Civic Media*, luettu 26.2.–12.4.2019 osoitteessa civic.mit.edu/2015/12/01/feminist-data-visualization.
- D'Ignazio, Catherine & Klein, Lauren (toim.) 2018 (luonnos), *Data Feminism*, MIT Press Open, luettu 30.1.–17.4.2019 osoitteessa bookbook.pubpub.org/data-feminism.
- Domestic Data Streamers 2016 "Making visible the invisible – A data experience experiment for the General Assembly of the United Nations" sivustolla *domesticstreamers.com*, luettu 1.4.2019 osoitteessa domesticstreamers.com/case-study/general-assembly-united.
- Gandomi, Amir & Haider, Murtaza 2015, "Beyond the hype: Big data concepts, methods, and analytics" Lehdessä *Journal of Information Management* vol. 35, no. 2, huhtikuu 2015, s. 137–144, luettu 21.3.2019 osoitteessa www.sciencedirect.com/science/article/pii/S0268401214001066.
- Haraway, Donna 1988, "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective" lehdessä *Feminist Studies* vol. 14, no. 3, syksy 1988, s. 575–599, luettu 20.3.–13.4.2019 osoitteessa philpapers.org/archive/HARSKT.pdf.
- Haraway, Donna 2016, "Tentacular Thinking: Anthropocene, Capitalocene, Chthulucene" verkkolehdeessä *e-flux*, Journal #75, syyskuu 2016, luettu 14.4.2019 osoitteessa www.e-flux.com/journal/75/67125/tentacular-thinking-anthropocene-capitalocene-chthulucene/.

- Hullman, Jessica & Diakopoulos, Nicholas 2011, "Visualization Rhetoric: Framing Effects in Narrative Visualization" lehdessä *IEEE Transactions on visualization and computer graphics* vol. 17, no. 12, joulukuu 2011, luettu 20.3.–31.3.2019 osoitteessa mucollective.northwestern.edu/files/2011-VisRhetoric-InfoVis.pdf
- Kirk, Andy 2014a, "The Design of Nothing: Null, Zero, Blank", Youtube-kanavalla *BocoupLLC*, katsottu 11.4.2019 osoitteessa www.youtube.com/watch?v=JqzAuqNPYVM.
- Kirk, Andy 2014b, "How to Show Something With Nothing", sivustolla *Harvard Business Review*, luettu 11.4.2019 osoitteessa hbr.org/2014/05/visualizing-zero-how-to-show-something-with-nothing.
- Koponen, Juuso 2019, "Näiden vaalien tuloksen ennustaminen on tavallistakin vaikeampaa", sivustolla *informaatiomuotoilu.fi*, luettu 13.4.2019 osoitteessa informaatiomuotoilu.fi/2019/04/naiden-vaalien-tuloksen-ennustaminen-on-tavallistakin-vaikeampaa.
- Liu L., Mirzangar M., Kirby R.M., Whitaker R. & House D. H. 2015 "Visualizing Time-Specific Hurricane Predictions, with Uncertainty, from Storm Path Ensembles", julkaisussa *Eurographics Conference on Visualization (EuroVis) 2015*, vol. 34, no. 3, luettu 27.3.2019 osoitteessa www.cs.utah.edu/~kirby/Publications/EuroVis-2015.pdf.
- M4ID 2018, "Images as data", sivustolla *m4id.fi*, luettu 10.4. osoitteessa m4id.fi/project/images-as-data.
- Saarinén, Ville & Ojala, Juho 2018, "Flow Towards Europe", sivustolla *Lucify.com*, luettu 4.4.2019 osoitteessa www.lucify.com/the-flow-towards-europe.
- Salmelin, Raili 2003, " Mistä on hyvät tilastokuvat tehty?", lehdessä *Duodecim*, no. 119, luettu 22.3.2019 osoitteessa www.terveyskirjasto.fi/xmedia/duo/duo93793.pdf
- Wang, Jen 2018, "Now You See It: Helvetica, Modernism and the Status Quo of Design", sivustolla *Lokidesign.net*, luettu 10.4.2019 osoitteessa lokidesign.net/journal/2018/5/14/now-you-see-it-helvetica-modernism-and-the-status-quo-of-design.

SUULLISET LÄHTEET

- Black, Alison 2017, "Representing Technical information in everyday communication", esitelmä tilaisuudessa *Tieto näkyväksi 2017*, 29.5.2017, Aalto-yliopisto, Helsinki.
- Joutsela, Markus & Hovinen, Safa 2018, *Graafinen suunnittelu III -kurssi*, luentosarja, tammi-maaliskuu 2018, Aalto-yliopisto, Espoo.