


# **VGI – KÄYTTÖ JA TULEVAISUUS MAANTIETEESSÄ**

Pia-Maria Johansson

Pro Gradu – tutkielma  
Maantieteen laitos  
Luonnontieteellinen tiedekunta  
Huhtikuu 2016

<b>Tekijä (Sukunimi ja etunimet)</b> Johansson Pia-Maria Erika	<b>Tutkielman Sivumäärä</b> 62
<b>Työn nimi</b> VGI - Käyttö ja tulevaisuus maantieteessä	
<b>Asiasanat:</b> Volunteered geographic information, VGI, GIS, Web 2.0, UGC	
<p>Tämän työn tarkoitus on saada kokonaisvaltainen kuva uudesta maantieteen ja geoinformatiikan ilmiöstä nimeltä volunteered geographic information, VGI. Vastaan kysymyksiin mitä on VGI, mistä ja miksi se on syntynyt, miten se sijoittuu maantieteen termistöön, miten sitä voidaan käyttää apuna tutkimuksenteossa ja vaihtoehtoisena datankeräysmenetelmänä ja mitä haasteita siihen liittyy. Pohdin ja tarkastelen myös sen mahdollista tulevaisuutta ja keinoja miten siitä saataisiin tulevaisuudessa suurempi hyöty. Tiivistettynä VGI on toimintaa, jossa ei-ammattilaiset osallistuvat maantieteellisen informaation tuottamiseen, mutta se on ilmiönä enemmän kuin vain uusi työkalu tai datankeruumuoto.</p> <p>Tutkimus on toteutettu tekemällä kattava kirjallisuuskatsaus monipuolisiin aihetta käsitteleviin artikkeleihin. VGI - termin otti ensimmäisenä käyttöön Michael Goodchild vuonna 2007 artikkelissa "<i>Citizens as sensors: the world of volunteered geography</i>" eli puhutaan melko uudesta ilmiöstä. Goodchildin artikkelin jälkeen VGI on levinnyt muidenkin tutkijoiden kirjoituksiin ja siitä on syntynyt kiinnostava uusi konsepti maantieteeseen. VGI on tällä hetkellä vaiheessa, missä se on alkanut herättää suurta kiinnostusta, mutta se ei ole vielä saanut kunnollista jalansijaa geoinformatiikan termistössä ja on vielä melko tuntematon esimerkiksi suomalaisessa maantieteessä.</p> <p>VGI tarjoaa kuitenkin täysin uudenlaista potentiaalia tulevaisuuden tiedonlähteisiin ja olisi mielestäni todella suuri menetys jos tämä ilmiö jäisi tutkimatta. Yhä kasvavassa määrin ihmiset omistavat tehokkaita tietokoneita, internetyhteyksiä ja mobiililaitteita GPS - vastaanottimella. Ihmisten internetkäyttäytyminen on myös muuttunut staattisesta tiedonkulutuksesta interaktiiviseen osallistumiseen, jossa käyttäjät osallistuvat sisällön tekemiseen. Nämä muutokset ovat tulleet jäädäkseen ja siihen on myös maantieteen ja geoinformatiikan reagoitava ja sitä se on tehnyt muun muassa VGI - ilmiön myötä.</p> <p>Kirjallisuuskatsaus VGI - ilmiöstä osoittaa, että lisätutkimukselle on ehdottomasti tarvetta. Aika näyttää onko VGI tulevaisuudessa ratkaisu uudenlaisten tiedonlähteiden saavutettavuudelle. VGI-lähteisiin ja menetelmiin liittyy vielä paljon ongelmia sen laadun varmistamiseksi, ja se onkin yksi tärkeimmistä haasteista joihin tulisi kohdistaa huomiota. Toinen tärkeä kohta, johon tutkimuksen tulisi jatkossa keskittyä, on ihmisten kiinnostuksen ja motivaation lisääminen, VGI:stä on tehtävä houkuttelevaa ja se onkin jossain määrin jo ottanut mallia sovelluksiinsa esimerkiksi pelimaailmasta.</p> <p>Kuten edellä totesin, korostan lisätutkimuksen tarvetta, vaikka on vielä epävarmaa mikä on VGI:n tulevaisuus, jääkö se liian epäluotettavana ainoastaan perinteisen informaation täydentäjäksi, mikä ei tietenkään ole vähäpätöinen tehtävä sekään, vai nouseeko siitä täysin uusia informaation muotoja. Siinä on kuitenkin niin merkittävä potentiaali, että olisi sääli jos menetämme sen jo lähtöruudussa.</p>	
<b>Muita tietoja</b>	
<b>Päiväys:</b> 15 / 04 2016	<b>Laatijan allekirjoitus:</b> 

## SISÄLLYSLUETTELO

1. JOHDANTO .....	4
2. VGI-TERMIN MÄÄRITYS JA ALKUPERÄ .....	5
3. VGI - ILMIÖN SYNTYYN JOHTANEET MUUTOKSET .....	8
4. TERMIN KRITIIKKI .....	11
5. VAPAAEHTOISET, KEITÄ HE OVAT, MITÄ HE TEKEVÄT? .....	14
6. KÄYTTÖ TUTKIMUKSESSA.....	19
6.1 Suunnittelu .....	19
6.2 Kriisin ja katastrofin hallinta .....	21
6.3 Kartoitus.....	24
6.4 Seuranta ja havainnointi.....	29
7. RISKIT JA LAADUNVARMISTUS .....	33
7. 1 Metadatan tärkeys .....	42
7. 2 Eettisyys.....	45
8. MILLAINEN ON HYVÄ VGI - SOVELLUS? .....	47
8.1 Gamification .....	52
9. YHTEENVETO .....	58
10. LÄHTEET.....	60

## JOHDANTO

Wilsonin (2014) mukaan on tapahtumassa uusi sijaintitietoisien tulevaisuuden nousu, jossa digitaalinen media on riippuvainen geospaatialisista teknologioista ja jossa GIS-tieteillä on vastuu tutkia yhä kasvavaa ja syvenevää suhdetta location based - palveluiden ja ihmisten jokapäiväisen elämän välillä. Ihmisen jokapäiväisen elämän tilat ovat yhä enemmän täytettyinä digitaalisella maantieteellisellä informaatioteknologialla. Geoinformatiikassa onkin tapahtumassa muutoksia vastaamaan tätä kehitystä.

Yksi näistä uusista ilmiöistä on nimeltään volunteered geographic information (VGI). VGI on ei-ammattilaisten vapaaehtoisten tuottamaa sisältöä ja toimintaa, joka on tullut mahdolliseksi vasta viimeaikaisen teknologiakehityksien myötä. Tässä työssä otetaan perustavanlaatuisesti selvää mitä tämä uusi termi tarkoittaa, mistä se on syntynyt ja mitä sen kanssa voi tehdä. Lisäksi pohditaan keitä nämä vapaaehtoiset ovat ja miten heidät saadaan tätä uudenlaista informaatiota tuottamaan. Työ pyrkii tuomaan lukijalle mahdollisimman selkeän kuvan ja ymmärryksen tästä uudesta ilmiöstä ja sen mahdollisuuksista mutta myös herättämään uteliaisuuden ajatuksesta millainen tulevaisuus sillä mahdollisesti voisi olla.

VGI kantaa mukanaan paljon potentiaalia, mutta myös haasteista. Ilmiön ongelmakohtia, kuin myös vahvuuksia tullaan käymään läpi monipuolisesti. Mukana on myös termin saama kritiikki ja katsaus mihin se sijoittuu muiden uusien maantieteellisten termien kanssa. Keskityn lisäksi tarkastelemaan millainen hyvän VGI - ohjelman tulisi olla, jotta tulevaisuudessa voisimme saada yhä laadukkaampaa informaatiota helposti ja vaivattomasti.

## VGI - TERMIN MÄÄRITYS JA ALKUPERÄ

Volunteered geographic information (VGI) on maantieteen yksi tuoreimmista ilmiöistä eikä sen syntymiseen johtaneet tekijätäkään jää kauas taakse historiaan. Ilmiö on itseasiassa niin uusi, että sen historiasta puhuminen vaikuttaa liioittelulta. Ennen kuin lähdetään tarkastelemaan tekijöitä, jotka johtivat VGI-ilmiön nousuun, määritetään mitä termi tarkoittaa. Termin määrittäminen ei tässä tapauksessa ole kovin helppoa, sillä samoja ilmiöitä, jotka kuuluvat VGI:n piiriin löytyy myös muista maantieteen termeistä ja raja eri termien välillä on häilyvä. Termi on vielä suhteellisen uusi, että väittelyä siitä, onko se yleisesti hyväksytty ja tunnistettu esiintyy eri tutkijoiden artikkeleissa. Koen että eri määrittelyjen ja rinnakkaistermien läpikäyminen uuden asian kanssa on tarpeen.

Mitä on VGI? Se on "laajalle levinnyt ja suurimääräinen yksityishenkilöiden osallistuminen maantieteellisen tiedon tuottamiseen ilman virallista pätevyyttä, toimintaa joka ennen oli varattu vain virallisille tahoille. Toiminta on lähes aina vapaaehtoista ja tulokset ovat tai eivät ole tarkkoja." Näin kirjoitti ja määritteli Michael Goodchild vuonna 2007 artikkelissaan *"Citizens as sensors: the world of volunteered geography"* ja antoi ilmiölle nimen VGI, volunteered geographic information. Goodchildin jälkeen muutkin tutkijat ovat antaneet ilmiöön johtaneille kehityksille määrittelyjä ja kuvauksia. Zhu (2015) kuvailee VGI:tä ilmiöksi jossa ihmiset voivat helposti jakaa georeferoituja havaintoja maailmasta interaktiivisten geovisuaalisten käyttöliittymien (esim. Google Maps, Google Earth) ja sosiaalisen median (esim. Twitter, Flickr, Instagram) kautta. Stevens & Pfeiffer (2015) kuvailee ilmiötä lauseella "wikification of GIS by the masses."

Mainittakoon heti alkuun, että vaikka monet pitävät Goodchildia VGI-ilmiön isänä ja nimenantajana on samankaltaisista muutoksista, jotka johtivat ilmiön syntyyn puhuttu muillakin termeillä aikaisemmin. (Esim. Ariffin ym. 2014, Blatt 2015, Cinnamon 2015, Mohammadi & Malek 2014). Ihan kaikki eivät myöskään miellä VGI-termiä ilmiön ainoaksi nimeksi ja esimerkiksi Upton ja muut (2015) toteavat ettei ilmiöllä ole yhtä ja ainoaa tunnistettua nimeä, mutta hekin kuitenkin käyttävät VGI-termiä tutkimuksissaan.

Rinner & Fast (2014) puhuvat "geospaatialisesta verkosta" tai "geowebistä", heidän mukaansa se on tietoverkosto joka mahdollistaa kaksisuuntaisen geospaatialisen datan kommunikaation kansalaisten ja organisaatioiden välillä. He käyttävät myös termiä

"participatory geoweb" joka sisältää jatkuvasti lisääntyviä kartoitusohjelmia jotka keräävät käyttäjien luomaa sisältöä. Nämä georeferoidut käyttäjien kontribuutiot esiintyvät monissa eri muodoissa, esimerkiksi koordinaatteina, geometrioina, erilaisina havaintoina, attribuuttitietoina, numeerisina mittoina, sisällön arviointina sekä multimediatiedostoina kuten videoina ja kuvina.

Rinner ja Fast (2014) toteavat, ettei näitä käyttäjien luomia sisältöjä voi laittaa yhteisesti hyväksytyyn termiin alle, mutta monet puhuvat niistä termillä VGI. He myös huomauttavat, että tutkijat ovat laittaneet merkille VGI:n syntyyn johtavien prosessien erilaisuuden ja moninaisuuden. Tämä on johtanut lisätermien, kuten involuntary geographic information, ambient VGI ja facilitated VGI, syntymiseen. Myös Cinnamon (2015) ja Tulloch (2008) käsittelee VGI-termin saamaa kritiikkiä ja siitä tullaan puhumaan tässä osiossa myöhemmin lisää.

Aikaisempia termejä jotka ovat käsitelleet samankaltaisia muutoksia ja ilmiöitä kuin VGI ovat neogeografia, crowdsourcing, PGIS ja PPGIS. Nämä termit eivät suinkaan ole hävinneet ja ne limittyvät VGI-termin kanssa (Ariffin ym. 2014). Neogeografia termiä on käytetty kuvaamaan ilmiötä, missä ei-ammattilaiset, tavalliset kansalaiset osallistuvat kartoitusprojekteihin heidän omalla lokaalilla tiedollaan (mikä kuvaakin myös VGI:tä). Crowdsourcing ("wisdom of the crowd") puolestaan viittaa näiden ei-ammattilaisten yhdistetyn lokaalin tiedon tuottavan laadukasta informaatiota (Rinner & Fast 2014). Brovelli ja muut (2014) toteavat, että vaikka VGI- ja crowdsourcing - termeillä on lievästi eriävä tarkoitus, koetaan ne yleensä synonyymeinä. Brovelli ja muut (2014) mainitsevat myös että koska web 2.0 työkalut mielletään tärkeäksi tekijäksi yhteisön osallistumisessa geoinformatiikkaan ja tutkimuksen tekemiseen ja koska monet PPGIS - käytännöt nojaavat näihin teknologioihin on raja VGI- ja PPGIS termin välillä sumuinen. He tiivistävät että vaikka VGI ja PPGIS ovat molemmat keskittyneet käyttäjän osallistumiseen ja web-teknologioiden käyttämiseen, VGI eroaa siinä että se keskittyy vapaaehtoisein yksilöinä ja on käyttäjälähtöinen ja PPGIS puolestaan vastaa yhteisön tai organisaatioiden tarpeisiin ja datan keräys lähtee virallisten tahojen toimesta.

Tulloch (2008) puhuu lisää VGI:n ja PPGIS-käsitteen suhteesta. Jotta VGI voisi saavuttaa laajaa tunnustusta geoinformatiikan piirissä ja tieteessä, tulee sen elementtien kuten crowdsourcing - käsitteen olla paremmin tunnettuja ja ymmärrettyjä. Niiden tulisi myös saada oma ja parempi jalansija maantieteen kentällä. Se pitää olla liitettävissä ja selitettävissä suhteessa tieteenalan muihin käsitteisiin ja ilmiöihin.

VGI voidaan nähdä innovaationa, joka nousi täysin itsenäisesti ja geoinformatiikan ulkopuolelta mutta monet esimerkit ja ideat sopivat selvästi geoinformatiikan sisälle ja osa sopii erittäin hyvin PPGIS:in alueelle (Tulloch 2008). Siksi on kannattavaa tehdä lyhyt katsaus, miten PPGIS ymmärretään geoinformatiikassa, jotta sen linkittymistä VGI:hin voisi paremmin ymmärtää.

Tullochin (2008) mukaan PPGIS ja VGI lomittuvat keskenään etenkin sen suhteen että molemmat tarkastelevat paikkoja, jotka ovat yksilöille tärkeitä. PPGIS:in kontekstissa tavallinen esimerkki voi olla yksilön pääsy julkiseen tietokantaan joka on osa projektia, jossa yksilöt osallistuvat heille tärkeän paikan muutoksia koskevaan suunnitteluun. VGI:n kontekstissa puolestaan yksilö voi vapaaehtoisesti luoda dataa hänelle tärkeästä paikasta. PPGIS:istä on vaikea puhua ilman kysymyksiä politiikasta, vallan siirtymisestä ja päätöksenteosta. Vastakohtaisesti VGI:n satunnaisuus ja viihteellisyys voi aiheuttaa että se nähdään vähäpätöisenä ja ei vakavasti otettavana toimintana. VGI keskittyy myös enemmän ohjelmistoihin ja informaatioon ja PPGIS puolestaan prosessiin ja sen tuloksiin.

Tulloch (2008) kysyykin tulisiko VGI käsittää PPGIS:in alakategoriana? Se helpottaisi ainakin VGI:n sijoittamista kirjallisuuteen ja tutkimukseen mutta VGI:n sijoittaminen sujuvasti PPGIS:in kehykseen voi olla haasteellista. On myös epäselvää, lomittuvatko termit suurilta osin vai ainoastaan vähäisessä määrin. Selvä tarve on ainakin VGI-ilmioon liittyvästä sanaston tarkastelusta. Etenkin termi "volunteered" on osoittautunut hyödylliseksi sanaksi kuvaamaan monia VGI:n aktiviteetteja mutta se ei välttämättä riitä kuvaamaan kaikkia VGI:hin liittyviä kokemuksia. Sana on myös saanut kritiikkiä, josta puhutaan myöhemmin lisää ja yksityiskohtaisemmin. Tulloch (2008) mielestä meidän tulee kysyä, keitä ovat vapaaehtoiset, mitä he tietävät ja mitä me tiedetään heistä. Tärkeä kysymys on, täytyykö vapaaehtoisten olla aina tietoisia luomastaan datasta vai onko hyväksyttävää käyttää dataa, jonka käytöstä sen alkuperäiset tekijät eivät ole tietoisia. Vapaaehtoisista puhutaankin yksityiskohtaisemmin lisää myöhemmin.

Tuleeko VGI-termistä tulevaisuudessa maantieteeseen yksi virallisesti tunnistettu termi lisää, vai kehittykö se eteenpäin ja sulautuu toisenlaiseen muotoon ja nimeen jää ennusteiden varaan. Ilmiö on kuitenkin todella mielenkiintoinen ja potentiaalinen valtti maantieteen ja geoinformatiikan tulevaisuudelle. VGI:n eduista, potentiaalista ja myös ongelmista tullaan puhumaan myöhemmissä osioissa ja ennen niitä on hyvä tehdä katsaus niihin muutoksiin jotka

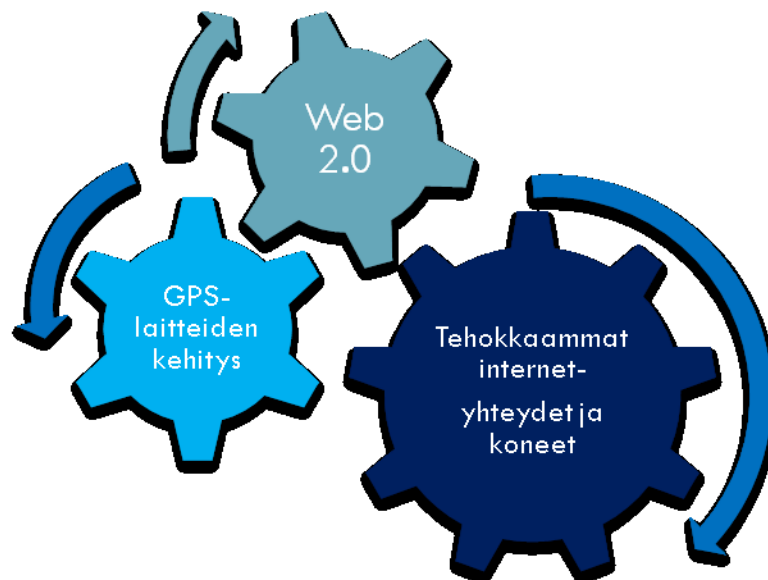
johtivat VGI:n syntyyn. Ilmiötä on vaikea ymmärtää, ellei katsota millaisista muutoksista ja kehityksistä se juontaa juurensa.

## VGI - ILMIÖN SYNTYYN JOHTANEET MUUTOKSET

Edellä todettiin että VGI on informaatiota, joka on luotu vapaaehtoisten, ei-ammattilaisten toimesta. Mutta mitkä tekijät ovat johtaneet tämän kaltaisen informaation syntymiseen? Mitkä tekijät ovat siirtäneet tehtävää mikä ennen oli mahdollista vain virallisten tahojen ja ammattilaisten toimesta tavallisten ihmisten toteutettavaksi?

Vuoteen 2004 mennessä oli tullut selkeäksi, että ihmiset voivat luoda omaa digitaalista maantieteellistä informaatiota korkealaatuisten online-karttojen muodossa ilman rahallisia kustannuksia. Tunnistettujen kohteiden koordinaatteja pystyi saamaan GPS:n avulla tai etsimään kohteita korkearesoluutioisista kuvista joita esimerkiksi Google Earth tarjoaa. Korkea taitotaso ei ollut enää vaatimuksena kartografiassa, sillä erilaiset open-source ohjelmat tarjosivat mahdollisuuden luoda tasokkaita karttoja helposti (Goodchild & Li 2012).

Tähän on pääasiassa johtanut kolme eri kehitystä (kuva 1) eri asioiden saralta: 1) web 2.0 - teknologioiden ilmestyminen, 2) GPS-laitteiden kehitys ja yleistyminen ja 3) tehokkaiden internetyhteyksien lisääntyminen (Esim. Zhu (2015), Ariffin ym. 2014, Haworth & Bruce 2015.)



Kuva 1. VGI:n syntymiseen johtaneet tärkeimmät kehitykset



Web 2.0 tarkoittaa ilmiötä, missä ihmiset eivät ole enää pelkästään internetsivujen ja palvelujen käyttäjiä ja kuluttajia vaan myös tekijöitä ja muokkaajia. Internetsivustoista on tullut interaktiivisempia, missä käyttäjät tulevat osalliseksi sisällön luomisesta (Cinnamon 2015). Web 2.0 - termin yhteydessä esiintyy usein toinen termi user-generated content (UGC), joka usein esimerkillistetään tarkoitamaan suosittuja sosiaalisen median sivustoja kuten Facebook ja Twitter (Haworth & Bruce 2015). Näiden sivujen idea perustuu käyttäjien luomiin teksti, kuva tai videopäivityksiin, joita yhteisö eli sivuston muut käyttäjät voivat kommentoida ja jakaa eteenpäin. UGC ei toki rajoitu sosiaalisen median sivustoihin. Feick & Robertson (2015) toteavat että informaatio tai media, jota ihmiset käyttävät, luovat ja jakavat internetissä on kehittynyt sellaisesta, jota ennen pidettiin lähinnä mielenkiintoisena kuriositeettina tai ohimenevänä teknologian villityksenä.

Nykyään käyttäjien luomaa sisältöä käytetään hyödyksi taloudessa, hallinnossa ja tieteessä. Informaation tuotanto ja jakaminen sekä tieto ovat lisääntyvässä määrin muuttumassa yhteisöllisemmäksi ja erot käyttäjien ja tuottajien välillä ovat kaventuneet. Pieni, mutta kasvava osio käyttäjien luomasta sisällöstä sisältää maantieteellistä informaatiota joko spatiaalisina koordinaatteina, jotka on saatu henkilökohtaisista paikannuslaitteista kuten matkapuhelimista tai vähemmän tarkkana tietoa kuten paikan, maamerkkien tai kaupunkien niminä (Feick & Robertson 2015).

Feick & Robertson (2015) jatkavat puhumalla miten jotkut spatiaalisen UGC:n muodot, kuten eläinhavainnot, tievauriohavainnot tai kansalaisten kommentit maankäytöstä sisältävät suoran yhteyden joihinkin datatuotteisiin ja yhteiskunnallisiin intresseihin. Yhteydet ovat yleensä vähemmän huomattavissa toisenlaisissa UGC:n muodoissa, jotka on usein luotu yksinkertaisesti ihmisten omista mielenkiinnon lähtökohdista tai netissä tapahtuvan sosiaalisen kommunikoinnin yhteydessä eivätkä sen tekijät useinkaan ajattele sen olevan samalla maantieteellistä informaatiota. Suuri osa sosiaalisista verkkosivuista on varustautunut sijaintitiedon mahdollistavilla teknologioilla ja netistä on muodostumassa sosiaalinen "geoweb." Maantieteilijät sekä muut ovat ilmaisseet suurta kiinnostusta tämän uuden datanlähteen mahdollisuuksia, mutta myös haasteita kohtaan.

Paikkatietojärjestelmissä saatavana olevan datan tyyppi voidaan jakaa spatiaaliseen ja ei-spatiaaliseen. Spatiaalinen informaatio jakautuu rastereihin ja vektoreihin. Edellä mainittua ei-spatiaalista dataa on attribuuttitieto, joka kuvaa spatiaalista informaatiota. Rasteridatan lähteenä on usein digitaalinen kaukokartoitus ja ilmakuvat. Vektoridatan lähteenä on tavallisesti GPS-mittaukset. (Ariffin ym. 2014)

Aiemmin spatiaalista dataa on luotu yleensä "top-down" lähestymistavalla, jossa viralliset tai kaupalliset tahot tekevät, testaavat ja muokkaavat tuotosta ennen kuin julkaisevat sen käyttäjille. Viime vuodet ovat tuoneet esimerkkejä tämän lähestymistavan kääntymisestä. Useat käyttäjät laajassa mittakaavassa osallistuvat spatiaalisen datan luomiseen (Upton ym. 2015).

Pitkään tavallisin datankeruumenetelmä oli kentällä tapahtuva GPS-mittaus tai pisteiden keräys jota suorittavat lähes yksinomaan pelkästään viranomaiset tai ammattilaiset. Menetelmä on kallis ja aikaa vievää ja siksi jotkut yritykset tai organisaatiot välttivät paikkatietojärjestelmien käyttämistä toiminnassaan vaikka se toisi merkittäviä etuja ja tehokkuutta päätöksen tekoon ja työvaiheisiin. Ennen datan kerääminen ja käyttäminen oli täysin riippuvaista ammattilaisista ja yhteisön käyttäminen maantieteellisen tiedon hankinnassa tuli mahdolliseksi vasta web 2.0 teknologioiden kehityksen myötä. Spatiaalisten teknologioiden kehittyminen ja internetin käytön lisääntyminen on tuonut uusia mahdollisuuksia organisaatioille kerätä dataa paikkatietojärjestelmien käyttöön. Ainakin Ariffin ja muut (2014) toteavat että VGI on nykyään globaalisti hyväksytty tiedonhankinta- ja päivitysmenetelmä.

GPS-laitteiden yleistyminen matkapuhelimissa on ollut merkittävässä osassa VGI-ilmiön kehittymisessä ja se on tuonut paikkatietojärjestelmän (GIS) lähemmäs tavallista kansalaista (Ariffin ym. 2014). Mobiililaitteista löytyvät GPS-vastaanottimet mahdollistavat lokaalin paikantamisen, lähes reaaliajassa tapahtuvan datan keräämisen ja levittämisen erilaisten mobiilisovellusten kautta (Haworth & Bruce 2015). Matkapuhelimet ovat suoraan yhteydessä internetiin ja niihin on saatavilla GPS-vastaanottimen lisäksi suuri määrä muita sensoreita, ne mahdollistavat helpon geospatiaalisen sisällön keräämisen ja jakamisen (Brovelli ym. 2014).

Käyttäjien tekemät kontribuutiot ovat olleet erittäin tärkeässä roolissa internetin kehityksessä. Internetin tuleminen osaksi ihmisten jokapäiväistä elämää vaikuttaa myös spatiaalisiin datan rakenteisiin (SDI, Spatial Data Infrastructure). SDI on yleisimmin aloitettu ja laitettu käyntiin sekä ohjattu valtion tai muiden auktoriteettien toimesta ja se tarjoaa virallista spatiaalista informaatiota käyttäjille. SDI:n datan valmistusprosessi sisältää yleensä laaduntarkistuksen ennen kuin se laitetaan saatavaksi julkisesti. VGI:n kehitys on kuitenkin haastamassa SDI:n. (Kalantari ym. 2014)

## TERMIN KRITIIKKI

Aikaisemmin jo mainittiin, että VGI - termi on saanut myös kritiikkiä. Kritiikki koskee erityisesti termin Volunteered geographic information "volunteered" osaa. Sillä vapaaehtoisesti kerätyn informaation käyttö tapahtuu usein ilman, että kyseiset vapaaehtoiset ovat siitä tietoisia. Hyvä esimerkki tästä on georeferoidun tiedon kerääminen sosiaalisesta mediasta kartoitushankkeisiin, jolloin miljoonat internetin käyttäjät, jotka ovat toimineet tiedon alkuperäisenä tuottajana siitä harvoin tulevat koskaan tietoisiksi, että heidän tuottama informaatio päättyy tutkijan käyttöön.

VGI:stä tuli kuitenkin Goodchildin vuoden 2007 artikkelin jälkeen tärkeä tutkimusaihe ja oli VGI sitten täysin "vapaaehtoisesti" tuotettua informaatio tai ei on se otettu käyttöön kuvaamaan kaikenlaista maantieteellistä sisältöä joka on yhteisön tuottamaa. Tämä sisältää esimerkiksi OpenStreetMappiin ladatun GPS-datan tai geotagattuja valokuvia tai tekstejä sosiaalisessa mediassa (esim. Twitter, Flickr) tai paikkatietoa mielenkiintoisista kohteista Wikimapiassa (Cinnamon 2015).

Cinnamonin (2015) kritiikki ulottuu myös VGI - ilmiöön liittyviin vastakkainasetteluihin. Hänen mielestä VGI:n saavuttaessa jatkuvasti kasvavassa määrin kiinnostusta, olisi aika siirtyä eteenpäin vastakkainasettelusta, jotta monimuotoisempia ja uudenlaisia datankeruumenetelmiä voisi kehittyä.

Cinnamonin (2015) tekemä kirjallisuuskatsaus VGI:stä paljasti, että se määritellään vastakohtana (binaries) perinteiselle tiedolle tai tiedonkeruumenetelmille. Sanaparit "top-down/bottom-up", "ammattilainen/amatööri", "virallinen/epävirallinen" toistuvat useasti. VGI sai terminä alkunsa Goodchildin artikkelista vuonna 2007 ja nämä toistuvat binäärimääritelmät sille ovat kulkeutuneet määritelmästä toiseen siitä lähtien. VGI:n käsitteellistäminen näiden vastakohtaisten binäärien avulla on ollut kuitenkin välttämätöntä, jotta ilmiötä on voitu kuvata.

Niistä voi kuitenkin koitua myös haittaa ilmiölle, jos niihin jymähdetään liian tiukasti. Jos jokin ilmiö tai menetelmä nähdään pelkästään "joko-tai" tyyppisenä, saatetaan ajatella että sen käyttö sopii vain tietynlaiseen toimintaan ja mahdolliset heterogeeniset ja monipuolisemmat käyttömahdollisuudet jäävät huomiotta. Myös uudet, monimenetelmiä hyödyntävät tiedontuotto- tai keruumenetelmät voivat jäädä kehittymättä jos aina ajatellaan vain esimerkiksi "ammattilaisten" tai "amatöörien" tuottamaa dataa eikä oteta huomioon niiden välissä olevia mahdollisuuksia.

Cinnamon (2015) vertaa kritiikissään OpenStreetMappia virallisiin kartoitusorganisaatioihin. Monet nykyaikaiset datantuottomenetelmät eivät usein myöskään mene sokeasti kahteen eri vastakkaiseen kategoriaan. OpenStreetMap on esimerkiksi tunnettu vastakohtana top-down tyyppisille yrityksille kuten Google Maps tai muut kansalliset yritykset jotka tuottavat karttoja. OpenStreetMap muistuttaa kuitenkin jossain määrin näitä "viranomaisia" kartoitusorganisaatioita. OpenStreetMapissa jokainen on vapaa tuottamaan ja muokkaamaan dataa, mutta siellä tuotettu informaatio on kuitenkin melko järjestelmällistä ja rakennettu joitakin sääntöjä noudattaen. OSM toimii osittain siten, että ammattitaitoiset jäsenet opastavat ja rekrytoivat uusia vapaaehtoisia keräämään, käsittelemään, lataamaan ja muokkaamaan dataa. OSM myös suunnittelee miten yhteisön tuottamaa dataa voisi paremmin käyttää kaupallisessa toiminnassa ja tämä vie sitä osittain "perinteisten" kartoitusorganisaatioiden suuntaan. On kuitenkin arveltu että juuri tämä bottom-up ja top-down strategioiden sulautuminen yhteen saattaa olla OpenStreetMapin menestyksen takana.

Ammattilainen-amatööri on myös sanapari joka usein esiintyy määriteltäessä VGI:tä. Lähempi tarkastelu kuitenkin paljastaa, että VGI:n tuottajat ovat taustoiltaan ja koulutuksiltaan hyvin kirjava joukko. GeoWiki on tästä hyvä esimerkki, miten VGI:n määrittelyminen "amatöörien" keräämäksi tiedoksi ei aina pidä paikkansa. GeoWiki-projektissa ihmisiä kannustetaan luokittelemaan maankäyttökarttoja omalta lähialueelta Google Earth Imagerya apuna käyttäen. Tämä projekti täyttää VGI:n kriteerit mutta kuitenkin valtaosa projektiin osallistuneista on tieteenharjoittajia joilla usealla on kaukokartoitus- tai paikkatietotaustaa, ei siis voida puhua täysin tavallisista kansalaisista joilla ei ole maantieteellisen tiedon tuottamiseen tai keräämiseen minkäänlaista koulutusta tai osaamista (Cinnamon 2015).

Moni tunnettu VGI-projekti nojaa ammattilaisiin tiedon tuotannossa, mukaan lukien projektit jotka luovat spatiaalista dataa luonnonkatastrofien jälkihallinnassa, kuten CrisisMappers ja Standby Task Force (SBTF). SBTF on kansainvälinen yhteistyöjärjestö joka koostuu yli 900 vapaaehtoisesta yli 70 eri maasta. SBTF:n vapaaehtoiset ovat kokeneita kriisikartoittajia jotka ovat lisäksi erikoistuneet erityiseen toimintaan järjestössä, kuten kokoamaan raportteja katastrofien aiheuttamista paikallisista vahingoista tai poliittisesta väkivallasta ja lisäämään ne internet-pohjaisille kartoille. Järjestäjät ja vapaaehtoiset viestivät sosiaalisen median avulla, joka mahdollistaa nopean osallistumisen ja karttojen tekemisen. Tyypiesimerkkinä voidaan pitää tapausta, jossa n. 250 vapaaehtoista keräsi ja kartoitti tietoa Libyan konfliktissa vuonna 2011.

Käytössä heillä oli Ushahidi crisis-mapping platform. Vaikka kyseessä on siis vapaaehtoisten ylläpitämä toiminta, spatiaalinen data ja kartat joita SBTF ja muut "ammattilaiset" vapaaehtoiset tuottavat tulisi ajatella virallisena/viranomaisena tiedonlähteenä "amatöörien" sijaan (Cinnamon 2015).

Nämä esimerkit VGI:stä osoittavat että tieto tai data, joka on tuotettu vapaaehtoisten toimesta, ei aina ole yksiselitteisesti amatöörien tuottamaa vaan liian yksinkertainen luokitus VGI-datasta. Myöskään virallinen tai viranomaisten tuottama spatiaalinen data ei aina ole täysin "ammattilaista." Valtion tai viranomaisten tuottamat paikkatieto tai karttatuotteet voivat olla vanhentuneita, yhteensopimattomia tai tehty liian pienellä rahoituksella. Joillakin yhtiöillä saattaa olla myös niin korkeatasoinen maine tuotteidensa laadukkuudesta ja ammattimaisuudesta, että niiden tuotteet otetaan vastaan "ammattilaisten" tuottamana automaattisesti. Esimerkin tästä mainitsee Cinnamon (2015) miten Apple teki Apple Maps sovelluksen iPhoneille ja iPadille, joka osoittautui hyvin huonolaatuiseksi. Vaikka Apple ei ole aiemmin tunnettu paikkatietotuotteista, otettiin sen tekemä karttasovellus vastaan laadukkaana.

Cinnamon (2015) ehdottaakin, että nykyaikainen spatiaalinen tiedontuotto tulisi uudelleen käsitteellistää jatkumona sen sijaan, että ajatellaan sitä vastakohtaisina sanapareina, kuten edellä esitettiin. Mikäli spatiaalinen datantuotto ajatellaan jatkumona, olisi esimerkiksi toisessa ääripäässä tavalliset yhteisön jäsenet, joilla ei ole koulutusta tai tietoutta spatiaalisen datan keräämisessä tai tuottamisessa, toisin sanoen amatöörit, jotka kuitenkin osallistuvat tiedontuottamiseen vapaaehtoisesti ns. "bottom-up" muotoisesti. Toisessa ääripäässä ovat koulutetut ammattilaiset jotka työskentelevät valtiollisissa kartoitusyhtiöissä ja tuottavat virallista paikkatietoa "top-down" muotoisesti. Näiden kahden ääripään välissä on kuitenkin tilaa tiedontuottamismuodoille, joissa on piirteitä kummastakin ääripäästä. VGI:n käsitteen ajattelemisen jatkumona voi auttaa väittelyssä siinä mikä on oikeasti vapaaehtoista ja mikä kerättyä tietoa.

Hybridi lähestymistapa spatiaaliseen tiedontuottoon luo toiveen suuremmalle määrälle laadukasta tietoa, jos kummastakin ääripäästä voidaan ottaa vahvimmat puolet: lokaali tieto, kustannustehokkuus, nopeus tiedon tuottamisessa ja leviämisessä sekä laaduntarkkailu ja järjestelmällisyys. Tämä tuo tietysti myös uusia haasteista. Spatiaalisen datan hallinta, jonka on tuottanut suuri määrä erilaisia ihmisiä, vaatii keskittymistä erityisiin avainkohtiin. Kuka omistaa datasetin, joka on tuotettu useamman tuottajan yhteisvoimin ja kuka sitä voi käyttää? Mihin

tarkoitukseen sitä voi käyttää ja millaisia hallinnollisia sääntöjä siihen liittyy? Kuka on vastuussa, jos datassa on virheitä ja niistä aiheutuu vahinkoa? Näihin kysymyksiin on vaikeampi vastata, jos tuote on tuotettu yhdistelemällä perinteistä ja uutta tiedontuotantomenetelmää kuin että se olisi tuotettu vain perinteisiä menetelmiä käyttäen (Cinnamon 2015).

Tämä kappale, missä esiteltiin VGI ilmiönä, kerrottiin siihen liittyvästä termistöstä ja sen saamasta kritiikistä, on hyvä lopettaa Goodchildin (2007) sanoihin siitä, miten hänen mielestään VGI on kohauttava innovaatio, jolla on väistämättä merkittäviä vaikutuksia paikkatietojärjestelmiin (GIS) sekä myös laajemmin maantieteen tieteenalaan ja etenkin sen suhteeseen julkisen yleisön kanssa.

## VAPAAEHTOISET, KEITÄ HE OVAT, MITÄ HE TEKEVÄT?

Seuraavaksi tehdään tarkempi katsaus vapaaehtoiisiin, johon koko VGI perustuu. Keitä vapaaehtoiset ovat ja mitä he tekevät? Tämä on tärkeää tietoa myös VGI-ohjelmistojen ja projektien suunnittelussa. Samalla tarkastellaan myös erilaisia projekteja joita voidaan luokitella muun muassa vapaaehtoisten mukaan. Bordogna ja muut (2014) ovat luoneet luokittelutaulukon perustuen erilaisiin VGI-projekteihin. Luokittelua on tehty tieteenalan, vapaaehtoisen tehtävän, VGI:n luonnin, VGI:n tarpeen ja vapaaehtoisten ominaisuuksien mukaan.

Citizens Science - projektit ja VGI ovat tulleet suosituiksi monella eri tieteenalalla, tähtitiede, avaruustieteet, tietokone- ja elektroniikka, kulttuuritieteet, luonnontieteet, lääketiede, biologia, sosiaalitieteet ja ympäristötieteet ovat hyviä esimerkkejä. Tieteenala johon projekti suuntautuu, on tärkeä huomionkohde, sillä se usein määrittää tietomuotojen (dataformaattien) piirteet. Laatuvaatimukset esimerkiksi sään ennustamiseen ovat hyvin erilaiset verrattuna kulttuurien alkuperätutkimukseen.

Erilaiset VGI projektit voidaan erottaa myös sen mukaan millainen tehtävä vapaaehtoisilla on projektissa. Bordogna ja muut (2014) ovat luoneet seitsemän eri ryhmää: massive computer time, specific human abilities, objects identification, observational measurements, transcriptions and corrections, user indications ja contextual and complementary information. Käyn tiivistetysti läpi jokaisen.

Massive computer time tarkoittaa tilannetta, jossa vapaaehtoisten omat tietokoneet osallistuvat laskentatoimeen jotta projektit, jotka vaativat suuria laskentatehoja suoriutuisivat nopeammin. Tällaisessa tapauksessa vapaaehtoiset eivät tuota maantieteellistä kontribuutiota ja tämä vaatii vapaaehtoiselta hyvin pientä osallistumismäärää. Esimerkkinä tästä on projekti, jonka tavoite on ennustaa maapallon ilmastoa vuoteen 2100 ja testata ilmastomallien tarkkuutta. Ihmiset ympäri maan voivat osallistua ja laskea malleja käyttämällä omien henkilökohtaisten tietokoneidensa laskentatehoa, tulokset lähetetään takaisin automaattisesti ja vapaaehtoiset voivat nähdä niistä tiivistelmän. Datan laatu ja tarkkuus riippuu ohjelmistosta ja laitteistosta.

Specific human abilities: jotkut projektit vaativat erityisesti ihmisen manuaalista osallistumista eikä tietokone yksinään riitä. Nämä projektit vaativat vapaaehtoiselta enemmän osallistumista verrattuna edelliseen. Tästä Bordogna ja muut (2014) antavat esimerkin Galaxy Zoo - projektista, jossa ihmiset tunnistavat galaksien muotoja. Tämän kaltaisessa tehtävässä ihmissilmä toimii edelleen paljon paremmin kuin kehittyneinkin tietokoneohjelma. Vapaaehtoisia saatetaan pyytää myös arvioimaan oma suorituksensa osallistumisen jälkeen.

Objects identification: näissä projekteissa ihmiset osallistuvat tarjoamalla informaatiota ja havaintoja heitä kiinnostavista objekteista, kuten esimerkiksi linnuista tai kasveista. He yleensä kertovat lajin sekä paikan missä havainto tapahtui. Ylimääräistä tietoa kuten kellon aika, havaintojen määrä, sukupuoli, kuvaus alueesta, valokuva ja niin edelleen voi myös olla mukana. Tällaiset projektit muodostavat suuren osan "Citizen Science" - projekteista luonnontieteen saralta. eBird - projekti on tästä hyvä esimerkki sekä Suomessa toimiva Tiira-tietokanta.

Observational measurements: näissä projekteissa julkiset tahot tai organisaatiot usein pyytävät vapaaehtoisia avuksi mittaamaan esimerkiksi liikennettä tai ilmansaastetta. Vapaaehtoisille usein tarjotaan spesifiset mittalaitteet ja nämä projektit vaativat yleensä laadultaan erinomaista dataa ja tutkijat ovat usein varovaisia käyttämään VGI-lähteitä. Esimerkiksi maanjäristyksestä aiheutuvia maanpinnan värähtelyjä joita vapaaehtoiset ovat ilmoittaneet havainneensa, voidaan käyttää avuksi tutkimuksessa. Toinen esimerkki on lumenpinnan syvyyden mittaaminen, jossa vapaaehtoisia pyydetään tavallista mittanauhaa käyttämällä mittaamaan lumen syvyyttä (Snowtweets project). AIR-projekti on julkinen, sosiaalinen kokeilu jossa ihmiset kuljettavat kannettavaa ilmanmittauslaitetta joka mittaa ilmansaasteita ja fossiilisia polttoaineen palamispaikkoja (hotspots). Myös matkapuhelimien monipuolistuminen ja kehittyminen on tuonut uusia mahdollisuuksia kansalaisten osallistumiselle tutkimuksissa.

Transcriptions and corrections: näissä projekteissa vapaaehtoisille annetaan jo olemassa olevaa dataa ja heitä pyydetään joko muuttamaan tai korjaamaan se toiseen muotoon tutkimusta varten.

User indications: nämä projektit vaativat joko suurta tai matalaa osallistumista vapaaehtoisilta. Pieni osallistuminen voi tarkoittaa esimerkiksi arvioinnin antamista sisällölle. Suurta osallistumista vaativa tehtävä voi pyytää vapaaehtoisia määrittämään tarve tutkimukselle tai sen tavoitteille.

Contextual and complementary information: nämä projektit pyytävät vapaaehtoisia lisäämään täydentävää informaatiota arkistoituihin sisältöihin.

Seuraava kategorisointi koskee sitä miten vapaaehtoiset luovat informaatiota. Luomistavalla on suora vaikutus informaation laatuun. Informaation luomistapa voi olla implisiittinen, eksplisiittinen, automaattinen tai manuaalinen. Niistäkin meille kertovat Bordogna ja muut (2014).

Automaattiset ja implisiittiset projektit, jotka vaativat suurta laskentatehoa ja aikaa tuottavat informaatiota automaattisesti ohjelmistojen avulla jotka vapaaehtoiset ovat asentaneet henkilökohtaisiin tietokoneihinsa, vapaaehtoiset eivät osallistu suoraan tiedon luomiseen. Tällä menetelmällä tuotettu informaatio on laadultaan riippuvainen ohjelmasta ja laitteistosta.

Manuaalisessa ja implisiittisessä tavassa vapaaehtoiset suorittavat tehtäviä tai pelaavat peliä manuaalisesti, vaikka heillä ei ole tieteellistä tietoa asiasta. Old Weather - projekti on tästä esimerkki, siinä vapaaehtoiset joilla ei välttämättä ole lainkaan kokemusta ilmaston mallintamisesta pelaavat peliä joka tuottaa tutkijoille hyödyllistä tietoa. Tästä gamification nimellä tunnetusta ilmiöstä tullaan puhumaan lisää myöhemmin. Toinen esimerkki on tapaus, jossa vapaaehtoiset eivät ole tietoisia olevansa hyödyksi tutkimukselle, esimerkiksi Google - hakukoneen tuloksia voidaan käyttää flunssan leviämisen ja tartuntamäärien ennustamiseen.

Manuaalisessa ja eksplisiittisessä mallissa vapaaehtoiset arvioivat itse kykynsä osallistua projektiin tiedontuottajana. Ennen kuin he saavat osallistua, heiltä voidaan kysyä joitakin asioita siitä, kuinka hyvin he arvioivat suoriutuvansa tehtävästä.

Erilaiset yhdistelmästrategiat ovat myös mahdollisia. Esimerkiksi Sickweather - projekti jossa algoritmi mahdollistaa ihmisten henkilökohtaisten terveystietojen seurannan ja kartoittamisen, jotka ovat julkisesti saatavilla sosiaalisessa mediassa. Vaihtoehtoisesti ihmiset voivat myös itse julkaista tietoja suoraan sivustolle nimettömänä.



Vapaaehtoisten ominaisuuksista Bordogna ja muut (2014) puhuvat puolestaan seuraavasti. He voivat olla tavallisia ihmisiä, jotka ovat motivoituneita osallistumaan projektiin josta on hyötyä yhteisölle, esimerkiksi veden laadunvalvonta. Nuoret ihmiset joita motivoivat pelimäiset käyttöliittymät, joiden kautta voi osallistua. Amatöörit ja kiinnostuneet jotka haluavat osallistua tieteen tekemiseen, esimerkiksi lintujen tarkkailijat tai luontoyhdistysten jäsenet. Vapaaehtoiset voivat tuottaa informaatiota myös olematta siitä itse tietoisia, kuten aikaisemmin jo tuli ilmi kun tarkasteltiin VGI - termin kritiikkiä.

Bordogna ja muut (2014) esittävät lisäksi 6 erilaista tyyppiä joilla vapaaehtoisia voi kuvata.

1. Neophyte: ei koulutusta tai tieteellistä taustaa tutkimusaiheesta. Hänellä on kiinnostusta, aikaa ja halua osallistua projektiin. Hän voi osallistua yksinkertaisiin tehtäviin jotka vaativat joko automaattista tai manuaalista osallistumista. Hän toimii yleensä yksin ja spontaanisti ja motivaation syynä voivat olla altruismi, uteliaisuus, kilpailuhenki tai tyytyväisyys ja ylpeys oman tuotoksensa näkemisestä julkisella alustalla tai sivustolla jossa myös muut voivat sen nähdä. Nämä vapaaehtoiset auttavat kattamaan etenkin lokaalia dataa. Jos vapaaehtoinen säilyttää kiinnostuksensa ja motivaationsa hän voi päättää parantaa taitojaan, erikoistua joihinkin alueisiin ja aloittaa yhteistyön toisten käyttäjien kanssa liittymällä ryhmään tai jakamalla kokemuksiaan yhteisössä.

2. Interested amateur: ihmiset jotka ovat kiinnostuneet projektista tai tieteenalasta, alkavat kerätä ja testata dataa hankkien kokemusta aiheesta. Joskus projekti tarjoaa heille koulutusta. Näiden vapaaehtoisten kontribuutiot ovat usein yksityiskohtaisempia ja tarkempia kuin kohdan 1. vapaaehtoisten mutta silti niissä saattaa esiintyä epätarkkuutta. Interested amateur vapaaehtoinen usein osallistuu vaativampiin tehtäviin ja voi muodostaa yhteistyöprojekteja toisten vapaaehtoisten kanssa. Näillä vapaaehtoisilla on korkeampi luotettavuus kuin edellä esitetyillä ja heidän arvokkain ominaisuus on jatkuva osallistuminen projektiin koko sen aikana.

3 ja 4. Expert amateur ja expert professional: expert amateur ja expert professional ovat molemmat tietotaitoisia jo ennen osallistumista. Nämä vapaaehtoiset voivat tarjota arvokasta dataa tai kommentoida, arvioida ja korjata muiden tuottamaa informaatiota. Heidän motivaationa voi toimia halu jakaa tietoa, ylpeys oman osaamisensa näyttämisestä ja täten oman maineensa nostattaminen, halu olla osana yhteisöä ja hyödyksi yhteisölle. On valitettavasti huomattu myös tapauksia, jossa osaavat vapaaehtoiset levittävät tahallaan väärää tietoa projektin sabotoimiseksi tai ekonomisten hyötyjen saamiseksi.

5. Expert authorities: joitain vapaaehtoisia voidaan kutsua osaaviksi viranomaisiksi sillä heillä on koulutusta ja kokemusta pitkältä ajalta projektissa. Heidän osaamistausta tarjoaa maksimaalisen luotettavuuden ja heidän osallistumisensa ja mielipiteensä vaikuttaa voimakkaasti projektissa.

6. Unaware volunteer: nämä vapaaehtoiset eroavat huomattavasti edellisistä sillä he eivät olet tietoisia, tai eivät välitä että heidän tuottamaa sisältöä käytetään tieteelliseen tarkoitukseen. Näitä vapaaehtoisia sanotaankin usein vapaaehtoisiksi väärin perustein. Tällaisen datan käyttämisessä suodatus olennaisen tiedon saamiseksi on välttämätöntä.

Vapaaehtoisten ominaisuudet ovat erittäin merkityksellisessä asemassa kun määritetään informaation laatua. Tavalliset motivoituneet kansalaiset ja amatööriä jotka haluavat tehdä oman osuutensa tieteessä voivat tarjota korkealaatuista informaatiota, ammattilaiset ja osaavat viranomaiset korkealla luotettavuudella voivat tarjota palautetta ja sisältöä projektien vetäjille. Yan ja muut (2016) kuitenkin tekevät huomion että yleensä ihmiset luottavat expert professional ja expert authorities vapaaehtoistyyppeihin. Vapaaehtoinen joka kuuluu "expert" kategoriaan voi tuntea projektin ominaisuudet hyvin mutta häneltä voi puuttua tietämys paikallistasolla olevista attribuuteista. Vapaaehtoinen joka mielletään neophyte - tai interested amateur -tyypiksi ei välttämättä tiedä paljoa VGI projektin teknisistä ja professionaalista puolesta mutta omistaa puolestaan loistavan tiedon tutkimusalueesta jolta dataa kerätään. Toisin sanoen joskus raja ammattilaisen ja amatöörin välillä on vaikea määritellä.

Tärkeimmät seikat jotka vaikuttavat VGI:n laatuun, ovat vapaaehtoisten taitojen ja osaamisen vaihtelevuus sekä heidän motivaatio ja sitoutuneisuus, median formaatit eri sosiaalisen median alustoilla, jotka johtavat vaihteleviin datarakenteisiin, vapaaehtoisten kontribuutioiden kontrollin puuttuminen, ja epäolennaisen tiedon suuri määrä olennaisen seassa. Bordognan ja muiden (2014) mielestä on tärkeää suunnitella yleisiä strategioita laadun parantamiseksi auttamalla ja rohkaisemalla itsearviointia, lisäämällä tiedon suodatusta ja estämällä tahallisia virheitä.

## KÄYTTÖ TUTKIMUKSESSA

Edellisessä luvussa esiteltiin VGI ilmiönä, tehtiin katsaus muutoksista, jotka johtivat sen syntyyn ja puhuttiin keitä ovat vapaaehtoiset. Tässä luvussa keskitytään tarkastelemaan tutkimuksien kautta, miten VGI käytännössä tarjoaa vaihtoehtoisen tiedonlähteen ja lähestymistavan ja miten sitä voidaan käyttää apuna tutkimuksen teossa. Käydessäni läpi tutkimuksia päädyin tekemään karkean jaottelun, joihin ne voidaan jakaa. Kategoriat löytyy tiivistettynä luvun lopussa olevasta kuvasta (kuva 2).

### **Suunnittelu**

Ensimmäinen kategoria (1) on nimetty suunniteluksi ja siitä on esimerkkinä kaksi eri tutkimusta. Ensimmäinen käsittelee VGI:n hyödyntämistä pyörätuolireittien luotettavuuden arvioinnissa. Toisessa käsitellään maisemareittien mallintamista, mukana on myös maininta VGI:n käyttämisestä kaupunkisuunnittelussa.

Neis'in (2015) artikkeli *Measuring the Reliability of Wheelchair User Route Planning based on Volunteered Geographic Information*, ilmaisee miten reitinsuunnittelu liikuntaesteisille on herättänyt yhä kasvavassa määrin tutkijoiden ja suunnittelijoiden kiinnostuksen. Ohjelmat tai sovellukset jotka suunnittelevat reittejä eivät useinkaan ota huomioon erikoistapauksia tai ryhmiä, koska saatavilla ei ole tarpeeksi tietoa niiden pohjaksi. Esimerkiksi pyörätuolilla liikkuvan vaatimukset tielle tai kulkualustalle ovat hyvin erilaiset autoilijaan tai kävelijään verrattuna. Pyörätuolilla liikkuvat eivät myöskään ole homogeeninen joukko, vaan heillä saattaa hyvinkin olla erilaisia tarpeita ja vaatimuksia reitille riippuen pyörätuolista tai siitä kulkeeko hän yksin vai avustajan kanssa.

Ongelmaa on lähestytty mm. Wheelmap projektin (2013) avulla, joka kerää tietoa paikkojen saavutettavuudesta pyörätuolilla liikkuville. Projektin tieto pohjautuu OpenStreetMap - dataan joka osoittaa että yhteiskunta on halukas ja kykenevä tuottamaan tarvittavaa tietoa. Etuna OpenStreetMapissa on että tiedon tuottaja voi suoraan lisätä objekteja jotka palvelevat hänen tarpeitaan. Pyörätuolireittejä suunniteltaessa pitää ottaa huomioon erilaisia parametreja tavalliseen kävelyreittiin verrattuna. Tien tai alustan pintamateriaali, leveys, kaltevuus, kynnykset tai portaat ovat tärkeitä tekijöitä pyörätuoliystävällistä reittiä suunniteltaessa.

Useimmat tutkimukset osoittavat että OSM:n tarjoama representaatio reaali maailmasta voi olla hyvinkin tarkka, jopa tarkempi ja täydellisempi joissain kaupungeissa kuin viranomaisten tarjoama vastaava dataset. Tässä täytyy kuitenkin huomioida, että kaupunkialueet ovat kuvattuna paljon paremmin ja tarkemmin kuin ruraalit alueet ja myös kaupunkien ja urbaanien alueiden välillä löytyy paljon tarkkuuseroja. Toisin sanoen OSM voi toimia korvaavana datan lähteenä vain joillekin alueille. Yksi suurin syy yhteisön kokoaman katuverkoston käytön vastustamiselle on sen attribuuttitietojen tarkkuus tai puuttuminen. Puuttuvat kadunnimet, kääntymisrajoitukset, osoitetiedot, tien pintamateriaali tai leveystiedot voivat hankaloittaa suunnittelua tai reitinhakuohjelmien tekemistä jos tarvittavia tietoja puuttuu.

Voikin todeta, että paras tapa arvioida sopiiko OSM - data tutkimuksen tai suunnittelun tiedonlähteeksi on tarkistaa datan laatu juuri kyseisen alueen kohdalla ja miettiä kuinka tarkkaa tietoa ja mitä sen tiedon pitää sisältää lopullista tuotetta ajatellen. Alueilla joilla datan tiheys on korkea ja tietoa on saatavilla tarkasti, voi yhteisön keräämää tietoa käyttää hyödyksi suunniteltaessa pyörätuoliystävällistä reitinhakuohjelmaa.

VGI:stä voi olla apua myös matkailureittien suunnittelussa. Siitä kertoo Alivandin ym. (2014) artikkeli *Analyzing how travelers choose scenic routes using route choice models*. Matkailutarkoituksessa tapahtuva liikkuminen paikasta toiseen ei aina tapahdu vain nopeinta tai helpointa reittiä. Matkailussa erilaiset maisemareitit saavat usein painoarvoa reittiä suunniteltaessa. Yhä enemmän hyödynnetään internetissä olevia reitinsuunnitteluovelluksia tai sivustoja, ne antavat käyttäjän valita reitin erillisten kriteerien mukaan, kuten matkustusaika, reitin pituus tai suorimman reitin valinta. Nykyään on siirrytty mallintamaan myös erilaisia maisemareittejä matkailijoille ja näiden reittien suunnittelu ei aina noudata vain tietynlaista kriteeriä, kuten aikaa, vaan kriteereitä on samalla kerralla useampia. Alivand ja muut (2015) käyttivät erilaisia internet-pohjaisia tiedonlähteitä jotka perustuivat VGI:hin. VGI data voi tässä yhteydessä paljastaa ihmisten matkustuskäyttäytymistä joiden avulla voidaan suunnitella parempia maisemareittejä.

Web 2.0 teknologia ja erityisesti VGI tarjoavat sopivan tiedonlähteen maisemareiteille. Valokuvanjakamissivustot kuten Panoramio ja Flickr ovat web 2.0 teknologian hyviä esimerkkejä. Ne sisältävät käyttäjien sinne lataamia geotagattuja valokuvia ja usein saatavilla on myös muutakin informaatiota ja tekstimuotoisia tageja. Suurin osa tästä informaatiosta saadaan merkittyjen API:n (Application programming interface) kautta.

Nettisivut joilta voidaan rakentaa maisemareittejä, voidaan Alivandin ja muiden (2015) mukaan jakaa kahteen kategoriaan: valokuvanjakamissivustoihin, joissa on geotagattuja kuvia ja matkapäiväkirjasivustoihin, jotka tarjoavat kokonaisia reittikuvauksia.

Alivandi ja muut (2015) totesivat että yksi heidän tutkimuksessa käytettävistä malleista vahvisti, että geotagatut valokuvat ovat hyödyllisiä ennusteita maisemareitin valinnalle. He myös uskovat että tulevaisuudessa VGI datan lisääntyminen, sisältäen geotagatut valokuvat, tulee lisäämään matkailureittimahdollisuuksia.

Myös kaavoituksessa ja urbaanissa suunnittelussa voidaan käyttää apuna VGI-lähteitä kuten valokuvanjakamissivustoja jotta saataisiin parempi käsitys mistä ihmiset pitävät rakennuksissa, maisemissa ja mitä paikkoja he urbaanissa tilassa pitävät tärkeinä. Alueet joilta on paljon geotagattuja kuvia olemassa, on mitä luultavammin suosittu kohde ja valokuvien laadullinen analyysi voisi auttaa paremmassa kaupunkisuunnittelussa. Huomattava on kuitenkin että ihmiset ovat yhä eriarvoisessa asemassa digitaalisten laitteiden käyttämisen suhteessa. (Cope 2015)

## **Kriisin ja katastrofin hallinta**

Toinen kategoria (2) käsittää kriisin ja katastrofin hallintaan liittyviä teemoja. Haworth ja Bruce (2015) tekevät kattavan katsauksen VGI:n käytöstä katastrofin hallinnassa artikkelissa *A Review of Volunteered Geographic Information for Disaster Management*. Luonnonkatastrofit muistuttavat meitä spatiaalisen datan tärkeydestä. Erityisesti sellaisen datan, jota pystytään tuottamaan lähes reaaliajassa ja jakamaan nopeasti. VGI tarjoaa uusia mahdollisuuksia katastrofeihin liittyvän maantieteellisen datan tuotannossa ja levittämisessä. Kuten GIS, VGI sisältää spatiaalisen datan kartoittamisen ja jakamisen. Menetelmänä se on kuitenkin erilainen sillä se tapahtuu yhteisön tuottamana vapaaehtoisesti.

Haworth ja Bruce (2015) toteavat että VGI tarjoaa sekä mahdollisuuksia mutta myös uhkia perinteisille datantuotantomenetelmille, joista osa liittyy läheisesti kriisinhallintaan. Mahdollisuutena on esimerkiksi lokaalin tiedon tuotanto siellä olevilta ihmisiltä ja uhkana että VGI saattaa vähentää virallisten tahojen tärkeyttä kartoituksessa.

Erilaiset kriisitilanteet tai luonnonkatastrofit luovat tarpeen reaaliajassa tuotettavalle tiedolle, johon perinteisesti kartoitus ei ole pyrkinyt eikä kyennyt. VGI siis täyttää

tarvittavan aukon spatiaalisen datan tuotannossa ja levittämisessä katastrofin hallinnassa. Kriisin ja katastrofin hallintaa tarvitaan sen jokaisessa vaiheessa, jotka ovat ehkäisy, valmistautuminen, toiminta sen aikana ja siitä elpyminen. Katastrofien negatiivisten vaikutuksien vähentäminen vaatii parempia lähestymistapoja sen jokaiseen edellä mainittuun vaiheeseen. VGI voi olla apuna jokaisessa vaiheessa, se mahdollistaa maantieteellisen informaation nopean kulkeutumisen viranomaisten ja ihmisten välillä kriisitilanteessa ja lisää yhteisöllisyyttä ja osallistumista katastrofien valmistautumiseen. Tämä on luonut kokonaisen uuden aihealueen maantieteellisen datan tuotannossa ja käytössä katastrofien hallinnassa, joka vaatii lisätutkimusta jotta sen käytön hyödyistä ja haasteista saataisiin yksityiskohtaisempi kuva (Haworth & Bruce 2015).

Vapaaehtoisten on nykyteknologian myötä helppoa luoda ja julkaista maantieteellistä informaatiota ja internetin kehittymisen, erityisesti sosiaalisen median ilmiön kehitys mahdollistaa myös tiedon nopean leviämisen. Tämä on luonut kriisin- ja katastrofinhallintaan täysin uudenlaisen pelikentän. Sosiaalinen media on muuttanut ihmisten tapaa luoda ja lukea informaatiota kriisitilanteissa.

VGI:n kasvu on muuttamassa teknologian käytön käytäntöjä kriisinhallinnassa ja lisäksi myös asenteita käyttäjien luomaa tietoa kohtaan. Esimerkiksi Queenslandin tulvissa Australiassa v. 2011, Australian Broadcasting Corporation käytti vapaaehtoisten kansalaisten luomaa dataa karttojen tekemiseen tulvista. Tämän sanottiin olevan kokeilu, jossa yhteisön luomaa tietoa käytettiin hyväksi. Ihmisiä on myös kehoitettu pysymään kuulolla sosiaalisessa mediassa kriisitilanteissa, kuten esimerkiksi vuoden 2013 pensaspaloissa Walesissa. Tämä kehitys heijastelee että ihmiset haluavat osallistua online-teknologioiden avulla saamaan ja vastaanottamaan tietoa katastrofitilanteissa sekä muiden ihmisten että viranomaisten välillä.

Viranomaiset voivat nopeasti kommunikoida ja kertoa lähes reaaliajassa maantieteellistä katastrofiin liittyvää tietoa suoraan yhteisölle paljon pienemmillä resursseilla ja kustannuksilla kuin ennen. VGI tarjoaa myös uniikkeja vaihtoehtoja kansalaisille tuottaa ja kartoittaa tärkeitä tietoja ja kommunikoida suoraan viranomaisille ja toisilleen uudenlaisilla tavoilla, vaikka he eivät sijaitsisikaan katastrofialueella.

Internet on rakenteeltaan yksilöiden välistä yhteistyötä ja kommunikointia tukeva, lisäten tiedon käyttöä ja saantia ja vähentäen siihen liittyviä kustannuksia. Suuri määrä saatavilla olevaa ajankohtaista tietoa lokaalilla tasolla on osoittautunut olevan kriittisessä asemassa kriisitilanteiden toiminnassa. VGI lähteiden on osoitettu tarjoavan vaihtoehtoisen tiedonlähteen

esimerkiksi Santa Barbaran tulipaloissa vuosina 2007 - 2009. Siirtymisellä pois perinteisistä kartografian käytännöistä ja protokollista on ollut tärkeitä vaikutuksia tiedontuotantoon katastrofien aikana ja maantieteellisen tiedon menettelytavoissa. Kriisiin liittyvän tiedon tuotanto ei ole enää vain ammattilaisten yksinoikeus vaan VGI on mahdollistanut sen että kuka tahansa pystyy osallistumaan, mikäli hänellä on pääsy tarvittavaan teknologiaan, joka nykyään tarkoittaa tavallista tietokonetta internetyhteydellä. Tietävistä käyttäjistä on tullut myös tietäviä tekijöitä.

Queenslandin tulvatapahtumia pidettiin ennen kokemattomana sosiaalisen median käyttönä katastrofitilanteessa kun niiden avulla raportoitiin informaatiota. Alustava tulvakartoitus tehtiin VGI:n avulla käyttäen geotagattuja kuvia ja sosiaalisen median sisältöä. ihmisillä joilla on kamera- ja GPS-puhelimia on yleensä eduksi olla ns. in situ eli sijaita täsmälleen siellä missä tapahtuu ja he voivat tallentaa tietoa lähes reaaliajassa. He eivät myöskään ole riippuvaisia muusta teknologiasta, kuten satelliittikuvista joita esimerkiksi pilvisuus voi haitata (Haworth & Bruce 2015).

Ihmisten voiman ja mahdollisuuksien on raportoitu vähenevän katastrofitilanteissa ja tutkimukset osittavat että VGI-teknologialla on mahdollisuus palauttaa ihmisille tämä menetetty valta. Vallalla tarkoitetaan tässä yhteydessä englanninkielen sanaa empowerment, joka tarkoittaa ihmisten kykyä hallita hänen henkilökohtaisia asioistaan ja kohdata kriisitilanne saaden välttämätöntä kriisinhallintatukea. Keski- ja etelä-Amerikan kansalainen on jo valmiiksi varustettu havainnointikyvyllä ja VGI antaa lisämahdollisuuden rekisteröidä havainnot ja ladata ne internetiin ja myöhemmin muodostaa niistä ymmärrettäviä raportteja ja karttoja. (Haworth & Bruce 2015).

Katastrofinhallinnassa VGI:n hyödyntämisen tutkimus keskittyy suurimmaksi osaksi toimenpiteisiin kriisin aikana ja jälkihallinnassa, mutta VGI:n käyttö kriisin estämiseksi ja ennakkosuunnittelussa on vähäisempää. Useat tutkijat ovat sitä mieltä, että VGI:ssä on potentiaalia olla apuna myös ennalta ehkäisyssä ja suunnittelussa katastrofien välttämiseksi. Tehokas kriisitilanteen kommunikointijärjestelmä toimii kaksisuuntaisesti ja on paikallinen, jotta riskinalaisuudessa olevat saavat keskitettyä informaatiota ja neuvoa valmisteluissa. Spesifinen lokaali tieto VGI - alustoilla jaettuna voi avustaa ihmisiä ja yhteisöä ymmärtämään oman lähiympäristönsä haavoittuvuutta ja riskejä ja kehittää tehokkaan suunnittelujärjestelmän mahdollisten hasardien varalta.

Kun suunnataan enemmän huomiota katastrofin ennaltaehkäisyn vaiheisiin voi VGI:ssä olla potentiaalia yhteisön mukaan saamisessa katastrofien ennaltaehkäisyssä. VGI:n kautta suuri määrä lokaalia tietoa voidaan kerätä ja jakaa kriisinhallinnan avuksi ja se voidaan toteuttaa paljon pienemmällä määrällä resursseja ja varoja kuin perinteisillä keinoilla, joilla dataa kerätään ja muodostetaan kartoiksi. Samaan aikaan voidaan potentiaalisesti lisätä yhteisön osallistumista katastrofien estämiseen, niihin valmistautumiseen, toimintaan niiden aikana ja jälkeen. VGI:n mahdollisuuksien lisäksi edessä on myös haasteita, erityisesti datan laatu, vain pienen osan soveltuminen käyttöön (vinouma), datan hallinta ja järjestely, yksilöiden yksityisyys ja turvallisuus vaativat lisähuomiota ja tutkimusta. VGI:n heikkojen kohtien ja rajoitteiden huomioiminen lisää kuitenkin sen luotettavuutta, kun asioita, myös niitä negatiivisia puolia tarkastellaan kehittämismielessä (Haworth & Bruce 2015).

Myös Blatt (2015) mainitsee VGI:n käytöstä kriisitilanteissa. Esimerkiksi Haitin 2010 maanjäristyksessä, se on ollut apuna myös metsäpaloissa, tulvissa, hurrikaaneissa ja muissa luonnonkatastrofeissa. Vapaaehtoiset keräävät ja tallentavat informaatiota omilla mobiililaitteillaan, yleensä GPS-varusteisilla matkapuhelimilla, ja lataavat datan internetiin sosiaalisen median sivustoille tai organisaatioiden nettisivuille jotka analysoivat, läpikäyvät ja julkaisevat tiedon karttamuotoisena. Jotkut viralliset tahot kyseenalaistavat VGI:n luotettavuuden mutta se on silti osoittautunut arvokkaaksi tiedonlähteeksi, sen lähes reaaliajassa tapahtuvan informaation keräämisen johdosta.

## **Kartoitus**

Kolmas (3) kategoria käsittää kartoitukseen liittyvät tutkimukset. Tästä toimii esimerkkinä Irlannissa tehty tutkimus virkistyskäyttöön sopivien metsien saavutettavuudesta ja kartoituksesta sekä maaperäkartojen validoinnista VGI:n avulla. VGI dataa voidaan käyttää apuna metsien, joilla on virkistyskäyttöarvoa, kartoittamiseen ja mallinamiseen. Upton ym. (2015) tarkastelivat VGI:n käyttöä Irlannissa virkistyskäyttöön soveltuvien metsien kartoittamisessa artikkelissaan *Combining conventional and volunteered geographic information to identify and model forest recreational resources*.

He kertovat että vaikka korkealaatuista dataa on saatavilla Irlannin metsäalasta, se ei kuitenkaan kerro mitään metsien saavutettavuudesta virkistyskäyttöä ajatellen ja tiedot metsien



omistajista ovat myös puutteelliset. Toisin sanoen kattavaa karttaa metsistä virkistysresurssina ei ole olemassa Irlannissa. Toinen esimerkki käsittelee VGI-datan käyttöä apuna maaperäkarttojen korjaamisessa ja tarkastuksessa.

Yksi parhaiten tunnetuista lähteistä VGI-datalle on OpenStreetMap, josta on jo aiemmin ollut puhetta. OpenStreetMapin data on peräisin vapaaehtoisesti muodostuneelta yhteisöltä, joka luo dataa ilmakuvien tai GPS:n avulla. OSM data Irlannista sisältää suuren määrän elementtejä, kuten suurimman osan tieverkostosta ja jonkun verran maankäyttödataa. VGI:n hyötyjä tutkijoille on sen hyvä saatavuus, halvat kustannukset ja sen jatkuva päivittyminen. On kuitenkin todennäköistä että sen laatu ja kattavuus vaihtelee melko voimakkaasti paikasta riippuen, mutta silti se tarjoaa tärkeän työkalun tutkimuksenteon avuksi etenkin silloin jos perinteisen spatiaalisen datan saatavuus on rajattua tai se puuttuu kokonaan.

Yksi suurimmista eduista VGI-datassa on usein sen tekijän paikallinen lokaali tieto, joka voi johtaa hyvinkin tarkkaan ja laadukkaaseen tietoon. Tämä voi olla erityisen hyödyllistä identifioidessa virkistysmahdollisuuksia. Vertailut VGI-datan ja perinteisen datan välillä ovat osoittaneet että VGI saattaa olla tarkempi joidenkin elementtien osalta, kuten esimerkiksi teiden. VGI:n tai UCG:n potentiaalinen arvo suunnittelu ja tutkimuskäytössä on hyvin tunnistettu ja otettu mukaan monella eri tieteenalalla (Upton ym. 2015).

Uptonin ja muiden (2015) tutkimus selvittää käyttäjien luoman open-source datan avulla kävelyreittejä, joiden avulla voidaan selvittää mitkä metsät ovat saavutettavissa ihmisille ja tätä tutkitaan miten kyseistä dataa voidaan käyttää luomaan virkistyskäytön vaatimusta kuvaava malli. Ensisijainen tavoite heillä on luoda kansallinen kartta Irlannin metsistä perustuen virkistysresursseihin.

Useita ilmaisia spatiaalisen datan lähteitä on saatavilla Irlannin metsistä, mutta esimerkiksi EU CORINE 2006 vektoridatan resoluutio on 25ha ja se on merkittävä rajoite tämän kaltaisessa tutkimuksessa jossa kohteet (virkistyskäyttöön sopiva metsä) voivat vaihdella kooltaan suuresti. Maankäyttödata on myös saatavilla mutta vaikka sitä päivitetään säännöllisesti, se on resoluutioltaan 6.25 ha ja voi kärsiä pilvipieitteen aiheuttamista epätarkkuuksista tai puutteista. Aineisto ei myöskään kerro mitään metsien saavutettavuudesta, eli siitä, kuinka helppoa sinne on kulkea tai löytää reittejä tai kuinka hyvin metsä soveltuu virkistyskäyttöön. Tämä lisää suunnitteluun selviä haasteita (Upton ym. 2015).

Yksinkertaisella lähestymistavalla, jossa yhdistetään VGI- ja perinteistä dataa on kyky tuottaa merkityksellistä ja hyödyllistä tietoa joka voi auttaa tunnistamaan metsiä, joilla on virkistyskäyttöpotentiaalia ja tarjota käyttökelpoisia muuttujia malleihin. Uptonin ja muiden (2015) tutkimuksessa tunnistettiin reilusti yli tuhat metsää, joita voidaan mahdollisesti käyttää virkistystarkoituksessa. Tämä voi tarjota tärkeän työkalun metsien käytön ja vaatimuksen/tarpeen ymmärryksessä. Se voi myös auttaa metsien hoidossa, sillä se tunnistaa missä resursseja puuttuu ja minne voisi suunnitella metsien levitystä tai jo olemassa olevien metsien avaamista käyttöön virkistysmielessä.

Mahdolliset VGI:n huonot puolet ovat tarkkuus, yksityiskohtaisuus ja spatiaalinen jakautuminen. Vaikka menetelmä todettiin tässä tutkimuksessa tehokkaaksi kun verrattiin metsiä referenssidataan, se ei takaa kuitenkaan että kaikki metsät identifioitiin. Tutkimus kuitenkin ehdottaa että virkistysmahdollisuuksien mallintajien tutkijat ja luonnonvarasuunnittelijat voivat hyötyä lisäämällä VGI:n tutkimukseen sekä tilanteissa missä muuta dataa ei ole saatavilla tai paikkamaan virallista dataa. Sen lisäksi että VGI data on laajalti saatavissa, se on ilmaista, on sen spatiaalinen kattavuus Irlannissa hyvä ja tulee luultavasti yhä paranemaan ajan kanssa.

Toisen esimerkin VGI:n käytöstä kartoituksessa tarjoaa maaperäkartat ja miten VGI voi toimia apuna niiden parantamisessa ja validoinnissa. Tästä esimerkistä kertovat Fonte ja muut (2015). Maaperäkartat (land cover maps) ovat keskeisiä aineistoja moneen eri tarkoitukseen ja sovellukseen kuten suunnittelulle, luonnon- ja biodiversiteetin suojelulle, ympäristön seurannalle, luonnonvararesurssien hallinnalle, ilmastonmuutokselle ja vesistöjen mallintamiselle. Maaperäkartat on usein laadittu kaukokartoituskuvista käyttäen automaattisia tai puoliautomaattisia lähestymistapoja. Maaperäkarttojen tarkkuus voi vaihdella suuresti, sillä niiden luonnissa käytetään monia eri menetelmiä. Karttojen tarkkuus on kuitenkin avainasemassa kun määritetään niiden sopivuutta spesifiseen käyttötarkoitukseen. Tarkkuuden määrittäminen koetaan nykyään oleellisena osana jokaisessa maaperäkarttaohjelmassa.

Korkealaatuinen ammattilaisen tekemä referenssidataset on aikaa vievää ja kallis prosessi tehdä. Prosessi on entistä vaikeampi, kun alue on suuri, suuriresoluutioista ilmakuvadataa ei ole saatavilla tai kun kartantekijät sijaitsevat kaukana itse kartoitettavasta alueesta. Näissä tapauksissa kenttätarkastukset ovat työläisiä, hankalia ja joskus mahdottomia toteuttaa ja siten on vain vähän lokaalia tietoa saatavilla alueesta (Fonte ym. 2015).

Vapaaehtoisten tarjoama, jatkuvasti lisääntyvä tietomäärä on saanut tutkijat kiinnostumaan VGI:n tarjoamista potentiaaleista halpana, tiedoiltaan ajankohtaisena ja määrältään runsaana referenssidatan lähteenä. Vapaaehtoisten tarjoama tieto voidaan kerätä monella eri tapaa ja eri formaateissa lähtien tekstikuvauksista tai valokuvista kokonaisuun karttoihin. VGI:llä on suuri potentiaali korvata tai täydentää virallista dataa joka on kallista ja rajattua tai täydentää puuttuvia kohtia etenkin globaalissa maaperän kartoituksen monitoroinnissa. Useampia kysymyksiä kuitenkin nousee esille pohdiskeltaessa VGI-datan käyttöä. Tärkeimpänä niistä lienee kysymys laadun varmistamisesta. Ongelmat laadun takaamisen ja tarkistamisen vaikeudesta voivat rajoittaa VGI:n käyttöä maaperäkartojen vahvistukseen ja tukemiseen etenkin koska perinteisessä karttojen tarkistuksessa ja korjauksessa referenssidatan oletetaan olevan erittäin laadukasta. VGI tarjoaa omat erikoisuutensa ja vahvuutensa, mutta tästä syystä se ei Fonten ja muiden (2015) mielestä kuitenkaan koskaan tule luultavasti 100 % korvaamaan virallista, ammattilaisten tekemää dataa tai saavuttamaan samaa tarkkuutta. VGI datan tarkkuudesta ja laadun varmistamisesta tullaan puhumaan myöhemmin lisää.

Useita VGI lähteitä voidaan käyttää apuna maaperäkartojen tekemisessä ja niiden laadun arvioinnissa. Pääasialliset lähteet tähän tarkoitukseen ovat valokuvat ja kuvaukset jotka Degree Confluence - projekti on kerännyt. Niistä ovat esimerkkinä vapaaehtoisten laittamat valokuvat sivustoille kuten Panoramio, Flickr, Geograph, OpenStreetMap, Geo-Wiki ja WIEV-IT projekti.

Kolme ensimmäistä lähdetä ei kerää dataa tarkoituksella maaperäkartoja ajatellen, mutta voi silti osoittautua hyödylliseksi, mutta Geo-Wiki ja WIEV-IT ovat projekteja jotka keskittyvät informaation keräämiseen joka sopii maaperäkartojen vahvistamiseen. Seuraavassa muutamia esimerkkejä erilaisista projekteista, jotka Fonte ja muut (2015) esittelevät.

Degree confluence kehitettiin v. 1996. Sen tarkoitus on kerätä vapaaehtoisten avulla valokuvia jokaisesta pituus- ja leveysasteen kohtaamiskohdasta siten että valokuvat otetaan jokaisesta ilmansuunnasta. Kuvan lisäksi kerätään kuvaus maisemasta. Joulukuussa vuonna 2014 sivuston tilastot raportoivat että 6328 yhdistymäkohtaa oli vierailtu. Joka kattaa 39 % prosenttia kaikista risteymäkohdista. Kuvia oli siihen mennessä 109,099 ja osallistujamaita 189.

Geograph on projekti jossa ihmisiä rohkaistaan lähettämään valokuvia jokaiselta neliökilometritä Isossa-Britanniassa. Projekti sittemmin levisi myös Irlantiin, Saksaan ja Kanaalinsaarille. Tähän mennessä projektiin on kertynyt 12,201 osallistujaa, jotka ovat

lähettäneet yli 4 miljoonaa valokuvaa kattaen 82,3 % Britannian ja Irlannin alasta. Saksassa projekti on paljon uudempi ja käyttäjiä on kertynyt 160. Jokainen vierailija sivulla voi nähdä informaatiota valokuvista kuten tagit jotka niihin on yhdistetty. Käyttäjät voivat myös katsella kuvien sijainteja Google Earth palvelussa.

Panoramio sivusto luotiin 2005 ja se kerää valokuvia maailmasta, johon on liitetty paikka ja aika. Sivuston tarkoitus on dokumentoida maailma valokuvien avulla, joten suurin osa valokuvista kuvaa erityisiä paikkoja. Kuvien lähettäjät voivat lisätä kuvaan lyhyen kuvauksen, tagin eli avainsanan ja lisätä kuvat eri kategorioihin. Kuvia voi katsoa myös Google Earth palvelussa. Kuvissa olevien paikkojen sijaintitiedot voi syöttää automaattisesti jos kuvissa käyttää EFIX tiedostoa ja kamerassa on sisäänrakennettu GPS. Kuvat voi myös manuaalisesti sijoittaa satelliittikuvien avulla.

Flickr aloitti v. 2004 ja se siirtyi Yahoo:n omistukseen 2005. Flickr mahdollistaa kuvien ja videoiden jakamisen käyttäjien kesken. Sivusto ei ole erikoistunut mihinkään erityiseen kuvatai videokategoriaan vaan käyttäjät voivat lisätä kuvia kaikista aiheista. Kuvien mukana tulee automaattisesti jotain metadattaa, kuten päiväys, kameran malli ja sijainti jos kamerassa on GPS. Käyttäjä voi lisätä kuviin avainsanoja (tageja) ja kuvauksia sekä lisätä ne teemakategorioihin. Vuoteen 2014 mennessä yli 150 miljoona geo-tagattua valokuvaa oli lisätty sivustolle.

Geo-Wiki projektin tarkoitus on helpottaa lokaalia/paikan päällä tapahtuvaa maaperädatan keräämistä ja auttaa maaperäkarttojen validointia käyttäen apuna korkearesoluutioista satelliittikuvaa joka on saatavilla Google Earth ohjelmassa. Geo-Wiki tarjoaa työkaluja joiden tarkoitus on auttaa hallitsemaan vapaaehtoisten datan laatua. Sivuston käyttöohjeiden lisäksi se tarjoaa videoita joissa opastetaan maaperän luokitusta.

Fonte ja muut (2015) kuvailevat kaksi eri tapaa jolla VGI:tä on käytetty maaperäkarttojen validointiin:

- 1) käyttäen valokuvia ja kuvauksia sivustoilta kuten CDP, Flickr ja Panoramio, joita sitten käytetään spesifiseen tarkoitukseen joko vapaaehtoisten tai ammattilaisten toimesta
- 2) hyödyntämällä käyttäjien, joille on annettu pääsy satelliittikuviin, valokuviin tai NDVI-arvoihin tekemiä luokituksia.

Alustavia tutkimuksia Flickr ja Panoramio valokuvien käytöstä on tehty. Yksi ongelma liittyy valokuvien epätasaiseen spatiaaliseen jakautumiseen, joko maantieteellisesti tai ajallisesti

tai maaperäluokallisesti. Myös niiden sijainnillinen tarkkuus vaihtelee. Joskus maaperän tyyppi on myös vaikea erottaa valokuvasta koska valokuvaa ei alun perin ole otettu siihen tarkoitukseen.

Fonte ja muut (2015) tiivistävät että se mikä tekee VGI:stä kuitenkin potentiaalisen avun maanperäkarttojen validointiin, on sen volyyymi, dataa on saatavilla suuret määrät ja se ei ole kallista. Kollektiivisesta crowdsourcing - menetelmästä saatiin sopivasti esimerkki vuonna 2014 Afrikan ebolakriisin yhteydessä. Tuolloin Guinean alueelta ei ollut saatavilla kuin muutama peruskartta mutta ei karttoja, jotka olisivat sopineet pohjaksi tarkkailu- ja hallintatoimille. Apua pyydettiin Humanitarian OpenStreetMap - Tiimiltä (HOT) kartoittamaan Guéckédoun kaupunki. Kahdenkymmenen tunnin kuluessa vapaaehtoiset olivat kartoittaneet kolme Guinean kaupunkia satelliittikuvien avulla ja yli 100,000 rakennusta. Tämä informaatio oli tärkeää ovelta-ovelle tiedusteluissa taudin leviämistä.

## **Seuranta & havainnointi**

Neljäs (4) kategoria sisältää seurantaan ja havainnointiin liittyvät tutkimukset ja projektit. Tästä esimerkkinä on tutkimus jossa ihmisten ruokailutottumuksia seurattiin sosiaalisen median avulla ja eBird joka perustuu vapaaehtoisiin lintujen tarkkailijoihin. Lopuksi tarkastellaan vielä VGI:n käyttöä sairausmaantieteen teemoissa.

VGI:n avulla voidaan tutkia ihmisten käyttäytymistä ja elintapoja. Seuraavassa esimerkki Chenin ja Yangin (2014) tutkimuksesta *Does food environment influence food choices? A geographical analysis through "tweets"* jossa yritettiin selvittää päivittäistavarakauppojen ja pikaruokaravintoloiden läheistä sijaintia ihmisten ruokailutottumuksiin. Tämä tutkimus ottaa innovatiivisen lähestymistavan tavoitella henkilökohtaista dataa ruokaan liittyvistä aktiviteeteista vapaaehtoisesti tuotetun maantieteellisen tiedon (VGI) kautta. Tässä tutkimuksessa tietoa on kerätty Twitter-sivustolta, joka on sosiaalisen median sivusto, jonne käyttäjät voivat kirjoittaa "twiittejä" (tweets) eli lyhyitä tekstejä. Kirjoitukset eli tweetit eivät sisällä ainoastaan päivityksiä ihmisten henkilökohtaisista tapahtumista vaan sisältävät myös ajan ja paikan josta ne on kirjoitettu. Puhutaan niin kutsutuista tageista eli tagaamisesta (eng. tags, tagging.) Ne tarjoavat idealin ratkaisun mittaamaan ihmisten altistumista erilaisille ruokaympäristöille juuri silloin kun se tapahtuu.

Kun ryhmä Twitter-käyttäjiä, jotka tekivät ostoksensa päivittäistavarakaupassa, verrattiin niihin, jotka ruokailivat pikaruokaravintoloissa, paljastui että päivittäistavarakaupan sijainnin läheisyys, joka tarjoaa tuoreita raaka-aineita voi merkittävästi vaikuttaa ihmisten tekemiin valintoihin ruokavaliossa. Pikaruokaravintoloiden määrä tai läheisyys ei kuitenkaan näyttänyt vähentävän terveellisten ruokavalintojen tekoa. Tämä tuki aikaisempia tutkimuksia siitä, miten tuoreita raaka-aineita myyvän kaupan läheisyys parantaa ihmisen ruokavaliota, mutta pikaruokaravintoloiden läheisyys ei kuitenkaan välttämättä lisää epäterveellisiä valintoja. Tällä tutkimuksella on suuri potentiaali antamaan lisätietoa terveystieteiden tutkijoille, miten sosiaalista mediaa voi käyttää hyödyksi yksilöiden käytöksen tutkimuksessa ja informoimaan asuinalueen vaikuttajille, miten tuoreita ja monipuolisia raaka-aineita tarjoava ruokakauppa vaikuttaa asuinalueen ihmisten terveyteen. Sosiaalisesta mediasta on muodostumassa uusi väline ihmisten terveyttä koskevassa tietämyksessä ja sairauksien ehkäisyssä.

Twitter-viestit ovat hyvä väline selvittämään yksilöiden ruokaan liittyvää saavutettavuutta koska ne ovat geotagattuja paikkaan ja aikaan, sisältävät tietoa josta yksilöiden ruokavalintoja voidaan heijastaa, ne ovat julkisia, vapaaehtoisesti tuotettuja ja kaikkien käytössä. Reaaliajassa tapahtuva tiedon kerääminen on mahdollista sosiaalisessa mediassa. VGI sosiaalisessa mediassa sallii analyyseja suuressa spatiotemporaalisessa resoluutiassa, tiedon saatavuus paikassa ja reaaliajassa sallii tarkan tutkimuksen ihmisten käyttäytymisaktiiviteeteista ennennäkemättömällä tarkkuudella. Menetelmä ei ole kuitenkaan rajoitteeton ja tulevaisuutta ajatellen ne on hyvä ottaa huomioon ja käsittelyyn. Ensinnäkin Twitter käyttäjät ovat vain osa koko väestöstä eikä se ole välttämättä realistinen representaatio. Nuoret aikuiset ja vähemmistöryhmät ovat yliesitettyjä. Huomioon on otettava myös tilanteet joista tai joissa ihmiset yleisimmin kirjoittavat sosiaaliseen mediaan ja ihmisen oleskelu paikassa, kuten ravintolassa, ei takaa että henkilö söisi siellä pikaruokaa. Myös Twitter-viestien valtavasta määrästä johtuen oleellisen sisällön etsintä ja filteröinti on aikaa vievää.

eBird on esimerkki kansalaisprojektista, joka havainnoi lintulajeja, niiden esiintymisiä ja määriä monien vapaaehtoisten lintuharrastajien toimesta. Lokaali tieto lajihavainnoista alueen ihmisiltä saattaa joskus olla ainoa keino saada informaatiota syrjäisiltä ja vaikeakulkuisilta alueilta esimerkiksi erilaisiin suojeluprojekteihin. (Zhu 2015).

Seurannasta ja havainnoista on hyvä esimerkki myös sairausmaantieteen teemat ja niissäkin VGI:llä saattaa olla suuri potentiaali. Suuria määriä reaaliaikaista tietoa tarttuvista

taudeista voi löytää internetin datavirroista. Näitä datavirtoja kutsutaan nimellä big data, nimi tulee saatavilla olevan datan suuresta määrästä. Se sallii meidän oppia asioita, joita ei voisi saada selville pienempiä datasettejä käyttämällä. Big dataa kuvataan myös sen kyvyllä muuttaa asioita dataksi, joita ei ole voitu laskea tai analysoida (Stevens & Pfeiffer 2015). Big datan nousu on luonut myös niin kutsutun cyberGIS - ilmiön. CyberGIS on tutkimuksen ja kehityksen osa-alue, joka keskittyy Big datan tuomiin haasteisiin säilöä, prosessoida ja visualisoida sitä. Tämä vaatii innovatiivista kehitystä laskentateholta moninaisista lähteistä (Wilson 2014).

Big datalla on kuitenkin rajansa joista etenkin tutkijoiden, jotka ovat tottuneet työskentelemään pienempien datasettien kanssa, tulisi olla tietoisia. Se ei ensinnäkään ole koskematonta tai käsittelemätöntä, toiseksi big data vaatii syy-seuraussuhteesta korrelaatioon siirtymiseen. Big data ei välttämättä kerro meille miksi jokin tapahtui vaan auttaa muodostamaan malleja ja kuvioita joiden avulla voidaan ennustaa tulevaisuuden esiintymiä. Big data usein pystyy havaitsemaan ensimmäiset merkit sairauden puhkeamisesta. Systemi perustuu olettamukseen että muutokset informaatioissa ja viestinnässä internetissä voi toimia aikaisena varoituksena populaation terveyden muutoksissa ja muodostaa automaattisen turvajärjestelmän joka jatkuvasti tiedustelee, erottelee, yhdistää ja visualisoi tartuntatauti-informaatiota suuresta määrästä erilaisia lähteitä.

Kaksi sivustoa on saanut erityistä huomiota, Google ja Twitter. Twitterin välitön nopeus tarjoaa viranomaisille suuren edun sekä valvonnassa ja tutkimuksessa. Esimerkiksi Bostonin hätäkeskukset saivat tietää vuoden 2013 maraton pommituksista Twitterin kautta ennen ilmoitusta median tai hätäkeskusten kanavien kautta.

Sen lisäksi että ihmiset laittavat informaatiota terveydentilanteestaan sosiaalisen median sivustoille, informaatiot hakukoneista ovat osoittautuneet myös hyviksi ennustajiksi. Esimerkiksi Google ja Yahoo kyselyitä on käytetty ennustamaan influenssan ja denguekuumeen leviämistä ja puhkeamista sekä lymen taudin esiintymistä. Internetpohjaisten seurantajärjestelmien välittömyyden lisäksi ne tarjoavat nopean kohdistamisen tartuntojen keskittymiin. Näin tehtiin mm. 2009 sikainfluenssan tapauksessa, jossa Google yhtiö oli suurena apuna. Google's Flu Trends (GFT) on parhaiten tunnettu internetpohjainen seurantasysteemi (Stevens & Pfeiffer 2015).

Crowdsourcer - seurantajärjestelmät, kuten Twitter viestien tiedonlouhinta (eng. data mining) käyttää algoritmeja viestien erotteluun avainsanojen avulla, määrittää niiden

oleellisuuden ja tarkkuuden, geotagaa ne ja vertaa niiden informaatiota muuhun valvonta- ja tarkkailudataan. Esimerkiksi NowTrending käyttää Twitteriä sairauksien seuraamiseen sekä kansallisella että globaalilla tasolla. Viimeaikaiset tutkimukset lisäksi osoittavat sosiaalisen median arvon kun se yhdistetään perinteisen sairausdatan kanssa tunnistamaan ja huomaamaan sairauksien kuten influenssa ja kolera, puhkeamista. Yksi pääasiallisista crowdsourcing seurantamenetelmistä on sen reaaliaikaisuus ja georeferoitu data mutta yksi suurin rajoite on sen sisältämä epäolennaisen tiedon suuri määrä. Algoritmeja epäolennaisen tiedon erotteluun on kuitenkin kehitetty (Stevens & Pfeiffer 2015).

Kuten Chen ja Yang (2014) jo totesivat, yksi Twitterin heikkouksista on se ettei sen käyttäjät edusta satunnaista osaa populaatiosta, suurin osa sivuston käyttäjistä on 18–50 vuotiaita. Internetpohjaiset seurantajärjestelmät voivat kuitenkin toimia loistavina aikaisina varoitusjärjestelminä ja näin vähentää sairauksien puhkeamisesta aiheutuva seurauksia (Stevens & Pfeiffer 2015).

Passiivisten internetpohjaisten seurantasysteemien lisäksi on olemassa seuransysteemejä joihin vapaaehtoiset voivat aktiivisesti osallistua lähettämällä informaatiota sairauden oireista. Nämä systeemit ovat osoittautuneet hyödyksi tunnistamaan riskiryhmiä ja rokotusmäärien tarvetta ja tehokkuutta. Ne ovat myös halvempia menetelmiä kuin perinteiset systeemit mutta niissä voi kuitenkin esiintyä vääristymää sillä ihmiset jotka valitsevat osallistua eivät välttämättä edusta satunnaista osaa populaatiosta (Stevens & Pfeiffer 2015).

Globalisaatio ja ilmaston lämpeneminen tuovat haasteita tutkijoille ja terveysviranomaisille, sillä ne levittävät tauteja alueille joilla niitä ei ole ennen tavattu ja tämä luo tarpeen uudelle informaatiolle, joka pystyy informoimaan riskeistä ajoissa. Spatiaalisen tautidatan suuri saatavuus voi auttaa meitä näissä uusissa haasteissa, olettaen että keskitymme sen potentiaaliin ja mahdollisuuksiin enemmän kuin sen ongelmiin (Stevens & Pfeiffer 2015).



<b>1. Suunnittelu</b>	Maisemareitit, liikuntaesteisille sopivat reitit, kaupunkisuunnittelu, ympäristösuunnittelu, logistiikka
<b>2. Kriisin ja katastrofin hallinta</b>	Luonnonkatastrofit, konfliktit Reaaliaikaista dataa, nopea levitys
<b>3. Kartoitus</b>	Virkistysalueet, saavutettavuus, maaperäkarttojen täydennys
<b>4. Seuranta &amp; havainnointi</b>	Lajihavainnot, vauriohavainnot, käyttäytyminen, sairaudet Reaaliaikainen tieto

Kuva 2. VGI:n hyödyntäminen eri tutkimusongelmissa

## RISKIT JA LAADUNVARMISTUS

Tähän mennessä on puhuttu VGI:n synnystä, määrytyksistä ja esitetty sen käyttämistä tutkimuksissa. Tämä luku keskittyy VGI:n hyviin ja huonoihin puoliin sekä erityisesti tarkastelemaan sen laadunvarmistusta, riskejä sekä mahdollisia ratkaisuja niihin. Luvun loppuun on koottu tiivistetty vertailutaulukko (kuva 3) VGI:n positiivista ja negatiivisista ominaisuuksista, jotta lukijalle tulisi mahdollisimman realistinen kuva VGI:n suuresta potentiaalista, mutta myös sen riskeistä.

Tiedon jakaminen internetissä mahdollistaa nopean ja laajan informaation liikkumisen joka onkin VGI:n sekä vahvuus että potentiaalinen uhka. VGI:n kerääminen ja jakaminen sosiaalisen median kautta antaa sille synnynnäisen kyvyn levittää viestejä ja informaatiota tehokkaasti. On kuitenkin syytä huomata nopean ja helpon tiedonlevityksen huonot puolet, jos hyödyllinen informaatio leviää nopeasti, leviää myös väärä tai puutteellinen tieto. Haworth & Bruce (2015) muistuttaakin että tämä pitää ottaa huomioon tutkittaessa uusia maantieteellisen tiedon tuotantomenetelmiä ja sen jakamista.

Turvallisuus ja datan laatu ovat suuria ongelmakohtia VGI:ssä. Yksilön fyysinen ja virtuaalinen turvallisuus saattaa olla uhattuna huonolaatuista VGI dataa käytettäessä. Data joka

on tuotettu ei-ammattilaisten toimesta saattaa olla puutteellista laadultaan ja erityisesti laaduntarkkailultaan. Tutkimukset ovat osoittaneet datan tarkistamisen tärkeyden etenkin kriisinhallinnassa käytettynä, mutta tietysti muuallakin. Näistä ongelmista huolimatta on huomattu että jossain yhteydessä, kuten vuoden 2010 Haitin maanjäristyksissä, jossa viranomaisten tuottamaa geospaatialista dataa ei ollut tarpeeksi saatavilla, oli yhteisön tuottama data kaikista kattavin ja ajankohtaisin (päivitetyn) ja ylipäätään saatavissa. Joskus huonolaatuisempikin data on parempi kuin ei dataa ollenkaan (Haworth & Bruce 2015).

VGI:n yleiseen luonteeseen tai käytäntöön kuuluu että se on yleensä julkista ja avointa kaikille. Tämä saattaa olla etenkin katastrofitilanteessa myös kyseenalaista, sillä geotagatuista valokuvista näkee yksityisten ihmisten omaisuuden heidän ollessa kriisin jäljiltä haavoittuvassa tilassa. Kuvia voidaan myös väärinkäyttää. Tutkimuksessa jossa selvitettiin Twitteriin ladattua tietoa eri kriisien osalta, kävi ilmi että vain 17 % kaikista viesteistä jotka koskivat kriisitapahtumia, sisälsivät luotettavaa tai uskottavaa tietoa. 13,5 % viesteistä oli puhdasta roskapostia. Loput olivat joko liian epäluotettavia, epäolennaisia tai henkilökohtaisia mielipiteitä (Haworth and Bruce 2015).

Internetissä olevan datan epävarmuus on yhteydessä myös online-turvallisuuteen. Epävarmuus ja pelko omien tietojensa jakamista kohtaan internetissä kriisien aikana saattaa rajoittaa ihmisten osallistumista joka puolestaan heikentää VGI-datan hyödyllisyyttä virallisille tahoille. VGI:n uskottavuus ja luotettavuus tiedonlähteenä on siis yksi haaste ja ongelmakohta sen käytössä katastrofinhallinnassa. Yksi ongelma digitaalisessa informaatioissa on sen lähteiden jäljittäminen, sillä metadata usein puuttuu. Mutta kuitenkin datan luomisen mahdollistaminen yhä monipuolisimmille ihmisryhmille on lisännyt informaation määrää runsaasti.

Kysymys kuitenkin kuuluu, miten voimme luottaa suureen määrään dataan jonka on luonut hyvin heterogeeninen joukko? Osalla saattaa olla ammattitaitoa mutta paljon on mukana myös niitä, joilla sitä ei ole. Twitter datasta tehdyt tutkimukset kriisitilanteissa ovat selvittäneet että jotkut ominaisuudet saattavat olla avuksi, jos yritetään automaattisesti tunnistaa ja luokitella hyödyllinen informaatio sosiaalisesta mediasta muiden viestien joukosta. Ne ovat yleensä pidempiä, niissä on URL ja niitä on kyseenalaistettu vähän muiden käyttäjien toimesta (Haworth and Bruce 2015).

Data voi sisältää virheellistä positiivista tai negatiivista informaatiota. Esimerkiksi öljyvuodosta raportointi voi olla positiivinen väärä informaatio, mikäli öljyvuotoa ei

ollutkaan tai negatiivinen jos informaatiota tapahtumasta ei ole. Epätosi positiivinen informaatio voi johtaa alueen evakuoimiseen turhan takia ja informaation puuttuminen, sen uskottavuuden epäily tai viivästyminen voivat puolestaan johtaa vaarallisiin tilanteisiin jos varotoimenpiteisiin ei ryhdytä ajoissa vaan odotetaan tietoa viranomaisilta. Tämä ei kuitenkaan tarkoita sitä että positiivinen väärä tieto olisi aina parempi kuin negatiivinen. Väärästä positiivisesta tiedosta saattaa aiheutua suuriakin kustannuksia ja evakuointiin liittyviä vaaroja, lisäksi jokainen virrehälytys vähentää systeemin uskottavuutta. VGI:n laadulla on ehdottomasti tärkeä rooli, huonolaatuinen data saattaa aiheuttaa hengenvaarallisia tilanteita ihmisille (Haworth and Bruce 2015).

Zhun (2015) mielestä spatiaalinen vinouma yhteisön luomassa datassa on suuri haaste muiden laatuun liittyvien ongelmien ohella. Yhteisön luoman datan spatiaalinen kattavuus on hänen mukaan vinoutunutta, sillä ihmisten tekemät havainnot ovat opportunistisia ja usein ne tapahtuvat vahingossa tai jonkun täysin muun toimenpiteen yhteydessä, eikä sen seurauksena että havainnoija olisi aktiivisesti ja tarkoituksen mukaisesti tehnyt havaintoja. Tämän seurauksena ne eivät ole systemaattisia mutta eivät myöskään täysin satunnaisia. Tämä voi hankaloittaa datan käyttämistä esimerkiksi ennakoivassa kartoituksessa.

Zhu (2015) mainitsee myös miten esimerkiksi OpenStreetMap data rakennuksista, kaduista, järivistä ja joista on usein tarkkoja sijaintitarkkuudeltaan, koska ne ovat paikallaan pysyviä kohteita ja digitoitu tarkkoista georeferoiduista ilmakuvista. Kuitenkin yhteisön luomassa datassa paikannus saattaa olla epätarkkaa riippuen siitä millainen objekti tai ilmiö on kyseessä sekä paikannukseen käytettävästä teknologiasta. Tilanteissa missä teknologiasta on puutetta ja kohde liikkuva voi tarkan paikannuksen varmistaminen olla haasteellista. Siinä missä jotkut biologit voivat nojata GPS-pantoihin eläinhavainnoissa, toteutuu usein vapaaehtoisten havainnointi paljon väljempään kuvaukseen siitä, missä esimerkiksi jokin eläinlaji oli havaittu.

Vinoumaa voi aiheuttaa myös se, että paikat joihin useilla paikallisissa on näkyvyys, tai jossa vierailaan usein, tulevat yliesitettyiksi. Esimerkiksi useat havainnot eläinlajeista yhdessä paikassa ei välttämättä tarkoita että paikka olisi kyseisen lajin erityisesti suosima. Tämä voi johtua pelkästään siitä, että paikkaa havainnoivat tai näkevät useat paikalliset ihmiset jonka johdosta aina kun sinne saapuu eläin, se havainnoidaan suurella todennäköisyydellä. Vastaavasti taas syrjässä olevat alueet jäävät havainnoimatta siitä yksinkertaisesta syystä ettei paikallisilla ole sinne näköyhteyttä tai siellä vierailaan vain harvoin.

Näitä vinoumia voi yrittää pienentää erilaisin metodein. Zhu ja muut (2015) esittelevät lähestymistavan minimoimaan spatiaalisen vinouman vaikutuksia ja epätarkkuuksia jotka liittyvät yhteisön luomaan dataan, jotta sitä voitaisiin paremmin käyttää ennakoivassa kartoituksessa. Esimerkkinä heillä on erään apinalajin (*Rhinopithecus bieti*) habitaattikartoitus. Heidän tutkimuksessa käytetään geospaatialista analyysitekniikkaa jota he kutsuvat nimellä "frequency-sampling strategy" vähentämään epätarkkuutta ja kompensoimaan vinoumaa. Tämä menetelmä vähentää spatiaalista vinoumaa yhteisön datassa ottamalla huomioon kuinka usein tietty alue on nähty paikallisen ihmisen toimesta ja näin kompensoi vinoumaa niitä alueita kohti, joihin näkyvyys on huono. Menetelmää testattiin vertaamalla paikallisten ihmisten havaintoja apinoista biologisten tekemiin kenttähavaintoihin. Heidän saamat tulokset vahvistivat, että menetelmä auttoi vinouman pienentämistä ja tarkkuuden paranemista.

Haworth and Bruce (2015) puolestaan mainitsevat Linuksen lain eli crowdsourcing-laaduntarkastuksessa. Joidenkin mielestä VGI lähestyy laatuvaatimuksiltaan viranomaista dataa. Sivustot kuten Wikipedia tarjoaa esimerkkejä siitä miten crowdsourcing on tehokas menetelmä poistamaan virheellistä tietoa, kun sitä on lukemassa ja muokkaamassa valtavat määrät ihmisiä. Mutta kuinka valtavia määriä yksilöitä tarvitaan osallistumaan, että tuloksia voidaan pitää tarkkoina? Linuksen laki (Linus law), joka tarkoittaa että mitä enemmän on havaitsijoita sitä vähemmän vihreitä jää huomaamatta ja datan laatu on parempi, näyttää pätevän ainakin OpenStreetMapissa. Lain sopiminen laajemmin VGI dataan on kuitenkin vaikea arvioida, etenkin katastrofitilanteissa, jotka ovat usein hyvin lyhytaikaisia luonteeltaan. Katastrofitilanteissa oikean tiedon filteröinti täytyy tapahtua myös hyvin nopeasti ja crowdsourcing voi tässä tapauksessa olla liian hidasta.

Haworth and Bruce (2015) mainitsevat myös Toblerin lain (Tobler's first law), joka tarkoittaa että paikat ovat yleensä enemmän samanlaisia lähellä olevien paikkojen kanssa kuin kauempana sijaitsevat. Maantieteellisesti runsaasti ympäristöään eroavat paikat ovat yleensä virheitä datassa. Toblerin lain mukaan paikkojen tulisi olla yhteensopivia sen tiedon kanssa, jonka jo ennestään tiedämme sitä ympäröiviltä alueilta. Raportoinnit esimerkiksi pensaspaloista on todennäköisesti oikea, jos havaintoja on ollut hiljattain myös lähialueilta.

VGI:n nopeus lisää myös sen tarkkuutta, se on useammin ajankohtaista (up to date) verrattuna perinteiseen (esimerkiksi kaukokartoitus) dataan ja uusi informaatio päivittyy lähes siinä ajassa kun se tapahtuu. Tämä ei toki tarkoita samaa kuin datan "oikeellisuus", mutta VGI:n

potentiaali tuoreelle datalle on erinomainen (Haworth and Bruce 2015). VGI:n laatua parantaa usein myös se, että monesti vapaaehtoiset ovat innokkaita tekemään havaintoja omasta lähialueestaan, joka usein tunnetaan hyvin. Näin saatu tieto on paikallista ja tarkkaa. VGI:n keräys on lisäksi paljon halvempaa kuin perinteisen tieteellisen datan keräys. Perinteinen datankeruu on usein myös paljon aikaa vievää ja sitä suorittavat vain alan ammattilaiset (Zhu 2015).

Datan laatua voi parantaa myös paikkatietosovellusten tarkkuuden parantuminen ja yhä useamman ihmisen pääsy käsittelemään georeferoitua dataa, mutta tämä ei tietenkään poista pelkästään käyttäjästä johtuvaa ihmisvirhettä. Paraskaan teknologia ei takaa sitä, että sitä käytetään joka tilanteessa oikein ja kaikki eivät aina huomaa, jos sovelluksessa ilmenee virheitä. Innovatiivisia menetelmiä tarvitaan tulevaisuudessa arvioimaan VGI datan laatua ja tarkkuutta (Haworth and Bruce 2015).

Yhteisön keräämä data luo haasteita tiedonkäsittelylle, jotka ovat erityisesti aiheellisia kriisinhallinnassa. Tärkeimmät haasteet ovat datan filteröinti ja tarkistaminen valtavasta määrästä ja lähteistä. Yksi este VGI - informaation käytölle on sen valtava tietomäärä, jonka läpikäymiseen ei ole vielä tarpeeksi tehokkaita ja nopeita keinoja katastrofitilanteissa. Tämä osoittaa tarpeen löytää keinoja joilla etsiä, filteröidä, verifioida ja yhteen vedota dataa ja datanlähteitä varmistaakseen tiedon laadun ja relevanttiuden (Haworth and Bruce 2015).

Esimerkiksi vain 36 % väestöstä omisti internetin kun Yolanda hirmumyrsky iski Philipien saarille. Tästä johtuen informaatio joka ilmestyi internettiin katastrofista paikallaolleiden ihmisten toimesta, oli osittain vinoutunut. Onkin hyvä muistaa että käyttäjien luoma informaatio (UGC) on aina valikoitunut esitys todellisuudesta ja osa ihmisistä jää täysin siitä ulkopuolelle eikä heidän tietonsa tule esille. Haitin maanjäristyksessä v. 2010 kävi puolestaan siten, että informaatiota sai lähettää tekstiviestein apupalveluun, mutta kaikki saapuneet viestit käännettiin englanniksi ja tulokset joita viestien perusteella saatiin, julkaistiin myös englanniksi. Informaatiota tai apua muulla kielellä lähettäneet tai pyytäneet eivät päässeet hyötymään omasta datastaan. Informaatio on usein vähiten saatavilla siellä missä sitä eniten tarvitaan. Katastrofien aikana ne yhteiskunnan osat, jotka ovat jo ennestään marginalisoituneita, ovat usein myös kaikkien haavoittuneimpia. VGI datan kehitys ei pidä tapahtua suuntaan, jossa jo ennestään haavoittuvaisia ihmisryhmiä jää hyödyn ulkopuolelle tai heidän informaatiota ei saada tietoon. Kriisinhallinnassa VGI pystyy olemaan hyödyllinen työkalu perinteisempien

menetelmien rinnalla ja spatiaalisten datalähteiden monipuolinen hyödyntäminen pitäisi olla päämäärä jokaisessa projektissa jossa käytetään käyttäjien luomaa dataa (Haworth and Bruce 2015).

Lisätutkimusta tarvitaan jotta erilaisia VGI ohjelmia ja alustoja ja teknologiaa voitaisiin hyödyntää ja käyttää tehokkaammin. Geotagien sisällyttäminen raportteihin esimerkiksi matkapuhelimista voi auttaa tietojen erottelussa paikkojen mukaan ja näin mahdollistaa kohdistetumpaa toimintaa ja parempaa spatiaalista suunnittelua. On kuitenkin havaittu, että vain 5 % käyttäjistä tarjoaa sijaintitietonsa näkyville yksityisyysuojan menettämisen pelossa tai yksinkertaisesti käyttäjä ei tiedä sijaintitiedon ilmoittamisen hyödyllisyyttä. Vain 1,5 % Twitter viesteistä ovat geotagattuja joka rajoittaa suuresti niiden käytön maantieteellisenä lähteenä ja ne ovat vain pieni osuus kaikista viesteistä ja niiden perusteella saatu informaatio saattaa olla vinoutunutta. Toinen esimerkki kehityksen tarpeesta on niin kutsuttujen hashtagien (hashtags) käyttö hallitsemaan suuria määriä dataa sosiaalisessa mediassa. Uusia menetelmiä tarvitaan rohkaisemaan VGI teknologian tehokkaampaa käyttöä ja kehittämään datan laatua ja tarkkuutta (Haworth and Bruce 2015).

Goodchild and Li (2012) puolestaan tarjoavat kolme vaihtoehtoa VGI:n laadun varmistamiseksi, jotka he nimeävät termeillä ”crowd-sourcing”, ”social” ja ”geographic”. Jokainen vaihtoehto tarjoaa mekanismeja VGI:n lajitteluun ja määrittämään sen paikkaansa pitävyyttä. Se kuinka automatisoituja prosessit ovat, vaihtelee, jossain tapauksessa se voidaan suorittaa täysin automaattisesti mutta jossain tapauksessa ihmisen tekemä tarkka työ on välttämätöntä.

Perinteisten kartoitusorganisaatioiden tarjoama laadunvarmistus sisältää kaksi vaihetta: prosessit jotka on suunniteltu kontrolloimaan laatua datan hankintavaiheessa ja prosessit jotka tarkkailevat laatua ottamalla näytteitä datasta ja vertaamalla niitä referenssilähteisiin. Jälkimmäinen vaihe yleensä dokumentoidaan ja laitetaan mukaan dataan joka on edellä opittua metadatan eli lisätietoa tiedosta.

Goodchildin ja Li'n (2012) Crowd-sourcing approach, joka konseptina on tullut jo aikaisemminkin esille, viittaa VGI:n yhteydessä yhteisön kykyyn huomata virheet jotka yksilöltä jäisi huomaamatta ja siihen että yhteisön luoma data havainnoista on luultavammin tosi jos vastaavanlaisia havaintoja on saatu useampia lähialueelta. Twitter-datan tutkimuksessa on käytetty mm. tätä lähtökohtaa apuna. On tietysti huomioitava, että tämä toimii hyvin vain

alueilla, jotka kiinnostavat suurta yhteisöä ja kaukaiset, harvaanasutut alueet saattavat sisältää vihreitä, sillä niille ei riitä tarpeeksi tarkkailijoita.

VGI yhteisön rakenne saattaa myös vaikuttaa systeemin toimivuuteen. Jos yhteisön jäsen, joka omaa korkealaatuisen "statuksen" ja kokeneen editoijan ja tiedontuottajan profiilin, ei hänen tekemiä kontribuutioita välttämättä epäillä lainkaan eikä oleteta niiden voivan olla virheellisiä. Myös ohjelman tai sivuston tapa palkita käyttäjien tekemää sisältöä voi vaikuttaa, jos tunnustusta saa vain uuden tiedon lisäämisestä, mutta ei editoimisesta, saattaa käyttäjien motivaatio jo tehtyjen muokkausten tarkistamiseen olla pieni. Sivuston tarjoama mahdollisuus poistaa tai muokata tietoa saattaa olla tehty myös sen verran vaikeaksi ja monimutkaisesti, että houkutus muokkaamiseen tai sen opetteluun on matala.

Goodchild and Li nimeää toisen lähestymistavan sosiaaliseksi (social approach) sillä se nojaa yhteisön hierarkiaan jossa osa jäsenistä muodostaa luotettavan joukon jotka toimivat moderaattoreina ja ylemmän tason käyttäjinä. Nämä profiilit ja tunnustukset voi ansaita toimimalla sivuston aktiivisena käyttäjänä ja tekemällä useita sisällöltään laadukkaita kontribuutioita. Käyttäjien tekemien muokkausten jäljittäminen ja tietynlainen palkintosysteemi myös lisää käyttäjien luotettavuutta. Käyttäjä jolla on voimassa niin sanottu "good standing" yhteisössä, voidaan ajatella tuottavan laadukasta sisältöä varmemmin kuin uusi, vasta hiljattain rekisteröitynyt käyttäjä. Näille kokeneille käyttäjille voidaan myös avata uusia toimintoja, joita uudet käyttäjät eivät pääse heti tekemään. Esimerkiksi OpenStreetMapissa on kaksi tasoa, tavalliset käyttäjät ja Data Working Group (DWG).

Maantieteellinen lähestymistapa (geographic approach) nojaa jo ennestään tunnettuihin maantieteellisiin faktoihin alueista ja sen perusteella voidaan arvioida pitääkö data paikkaansa. Esimerkiksi jo ennestään mainittu Toblerin sääntö "kaikki asiat liittyvät toisiinsa, mutta lähemmäs sijaitsevat asiat liittyvät toisiinsa enemmän kuin kaukaiset asiat" sisältää kaksi eri merkitystä. Toinen osa kuvailee spatiaalista riippuvuutta ja toinen sitä että esitetyt faktat paikasta tai alueesta pitäisi olla yhtenäisiä jo ennestään tiedossa olevien faktojen kanssa lähialueesta. Toisin sanoen esitetyn faktan pitäisi olla yhtenäinen sen maantieteellisen kontekstin kanssa.

Goodchild ja Li (2012) mainitsevat myös ominaisuuksia, joita on tullut esille jo aikaisemmin. VGI:llä on paljon hyötyä ja etuja, se on ilmaista, sitä on tarjolla paljon ja sitä voidaan luoda nopeasti, se on usein ajankohtaista ja päivitettyä ja se voi tarjota dataa, jollaista ei

perinteisin keinoin olisi useinkaan edes ajateltu tehdä. Kääntöpuolena sen laatu on hyvin vaihtelevaa ja ei dokumentoitua, se ei usein pysty noudattamaan tieteellisiä sääntöjä eikä se ole kaiken kattava. VGI ei välttämättä ole tarpeeksi riittävä vaihtoehtoiseksi datan lähteeksi tieteellisessä tutkimuksessa, mutta se voi olla hyödyksi tutkimuksen alkuvaiheessa (eksploratiivisessä hypoteesivaiheessa).

VGI:n laatua voi tarkastaa myös vertaamalla sitä virallisiin laadukkaaksi todettuun vertailudataan. Näin ovatkin monet tehneet ja Yan ja muut (2016) listaavat niistä osan. Esimerkiksi Saksan OpenStreetMap -dataa on verrattu TeleAtlas MultiNet - dataan ja Lontoon OSM - dataa on verrattu Iso-Britannian kansallisen kartoitusyhtiön Ordnance Surveyn dataan.

Yan ja muut (2016) lisäksi esittelevät systeemin joka varmistaa lajien valvomiseen kerätyn VGI - datan laadukkuutta. Systeemi käyttää hyödykseen "fuzzy set" teoria (sumea joukko) - nimistä toimintaa hoitamaan datassa olevaa epävarmuutta ja epäjärjestelmällisyyttä. Fuzzy set kehitettiin mallintamaan jatkuvaa ilmiötä ja maantieteessä se on otettu käyttöön mallintamaan spatiaalisissa dataseiteissä esiintyvää epävarmuutta. Esittääkseen fuzzy-systeemin toimivuutta VGI:n laadunmäärityksessä Yan ja muut (2016) käyttivät sitä tapaustutkimuksessa jossa tarkasteltiin satokasvien tuholaisvalvontaa.

Goodchildin ja Li'n (2012) kolme lähestymistapaa VGI:n laadun tarkastuksessa mainittiin edellä, ne olivat crowd-sourcing, sosiaalinen ja maantieteellinen. Näistä esimerkiksi käyttäjän arviointi (crow-sourcing), joka perustui siihen ideaan että kun on tarpeeksi monta tarkkailijaa, huomataan datassa olevat virheet, ei ole kovin hyvin soveltuva lajien tarkkailussa, sillä kohde on liikkuva tai se esiintyy vain rajoitettuna aikana. Me emme voi mennä takaisin tarkistamaan havaintoja paikkaan jossa se tehtiin. Maantieteellinen, sosiaalinen ja vapaaehtoisen lokaali tieto ovat tähän paremmin sopivia laadunmittareita.

Fuzzy systeemi auttaa VGI:n laadunvarmistusta käyttäjien luomissa lajivalvontaraporteissa ja yrittää paikata edellä mainittuja lähestymistapoja ja niiden ongelmia. Systeemi käyttää luottamusta laadun takeena, ottaen huomioon sekä käyttäjän historian (kuinka laadukkaita kontribuutioita hän on tehnyt aikaisemmin) sekä datan sopimisen maantieteelliseen kontekstiin. (Yan ym. 2016)

Heidän johtopäätös oli että käyttämällä fuzzy set teoriaa tuholaisvalvontadatassa voi johtaa haluttuihin tuloksiin VGI:n laadunvarmistustuloksissa. Fuzzy systeemi kehitettiin perustuen ideaan että VGI:n laatu voidaan varmistaa maantieteellisen kontekstin ja



käyttäjähistorian eli käyttäjän taidokkuuden avulla käyttäen luottamusta laadun takeena (eng. proxy of quality). He toteavat lopuksi että robustisten laskennallisten lähestymistapojen kehittäminen on välttämätöntä tämän kaltaisille projekteille missä määritetään VGI:n laatua valvontaohjelmissa johon vapaaehtoiset osallistuvat.

Vaikka vapaaehtoisten tai yhteisön luomaa dataa on käytetty hyväksi tutkimuksissa jo jonkun aikaa, on sen laatua edelleen kyseenalaistettava. Päälimmäinen huolenaihe on pystyvätkö usein ei-ammattilaiset vapaaehtoiset tuottamaan tarpeeksi laadukasta ja tarkkaa dataa tieteelliseen tutkimukseen käytettäväksi. Tämä kysymys ja huoli on johtanut moniin yrityksiin vertailla VGI-dataa perinteisempiin tiedonlähteisiin. On saatu selville esimerkiksi tuloksia, joissa ei-ammattilaiset olivat yhtä taidokkaita tunnistamaan ihmisten muokkaamia maankäyttömuotoja (land cover) Google Earth kuvista, mutta vähemmän tarkkoja tunnistamaan muita maan pintoja. Kuitenkin reittien ja muiden virkistysmahdollisuuksien kartoittaminen ei vaadi erityistaitoja ja vain hieman teknologiaosaamista (Upton ym. 2015), joten VGI datan käyttö saattaa kuitenkin olla hyödyllisempää laaturajoituksineen kuin että sitä ei käytettäisi lainkaan.

## **Metadatan tärkeys**

Seuraavaksi siirrytään metadatan rooliin informaatiossa josta Kalantari ja muut (2014) kertovat esimerkin metadatan puuttumisesta VGI-informaatiossa. He etsivät Royal Women's nimistä sairaalassa Melbournessa Wikimapiasta. Wikimapia antaa heille neljä eri hakutulosta, joista yksikään ei ole oikein. Kolme niistä on väärää ja neljännessä on vanhentunut osoite. Jos hakutuloksessa, jossa osoite oli vanhentunut, olisi toimiva metadata, helpottaisi se käyttäjää ymmärtämään hakutulosta paremmin.

Ongelma edellä esitetyissä on se että VGI-data on keskeneräinen ja tehty ei viranomaisten toimesta eikä se täytä riittäviä laatuvaatimuksia. Metadata voisi kuitenkin auttaa käyttäjää arvioimaan tiedon laatua ja käytettävyyttä. Metadatan luominen voi jossakin määrin parantaa ymmärrystä ja arvostusta informaation laadusta, joskin tarkkoja arvioita VGI-datan laadusta on silti vaikea selvittää. Kalantari ja muut (2014) ehdottavat että metadatan luominen VGI-informaatiolle on hyödyllistä yhteisöille jotka kasvavassa määrin nojaavat VGI:hin. He ehdottavat menetelmää, jossa VGI:n käyttäjät luovat itse metadatan VGI-tiedolle, tätä he kutsuvat nimellä Geospatial Metadata 2.0.

VGI:n laatu riippuu sen tekijästä ja kuinka tarkkaan tekijä on informaation luonut. Myös sillä on vaikutusta, montako tekijää on, mitä enemmän vapaaehtoisia, sitä paremmaksi sijainnillinen tarkkuus muuttuu. Tarkkuuden kasvu ei kuitenkaan ole lineaarista ja tarkkuutta pitäisi tarkastella lokaalilla, ei globaalilla tasolla. Ongelmana ovat vapaaehtoisten erilaiset taustat ja erilaiset metodit datan luonnissa, virheitä tapahtuu koska ei ole yhtenäistä menetelmää siitä, miten dataa luodaan eikä yhtenäisiä ohjeita. Myös relevantin tiedon löytäminen on haastavaa, ellei datan tekijä ole laittanut tarkkaa nimeä tai kuvausta datalle. Metadata voikin olla kriittisessä asemassa siinä, miten informaatiota löydetään VGI sovelluksista. Kalantarin ja muiden (2014) tutkimuksessa vertaillaan OpenStreetMappia ja Google Map Makeria tässä yhteydessä.

Google Map Makerissa, jos käyttäjä hakee haullla "Melbourne Central" nimistä tunnettua maamerkkiä saa tulokseksi väärän osoitteen ja ei minkäänlaista metadataa maamerkistä. Mutta jos hakuun lisää sanan "Dome" saa oikean osoitteen ja relevantin kuvauksen. OpenStreetMapissa Royal Women's - sairaalaa on mahdoton löytää, ellei käyttäjä kirjoita hakuun täsmällisesti "Royal Women's Hospital." Tämä johtuu OpenStreetMapin hakukoneen rajoitteista. OpenStreetMapin tietokannassa on kaksi featurea Royal Women's sairaalalle: sairaalan rajat ja rakennus. Tietokanta sisältää sairaalan nimen ja tiedon että se on kyseinen sairaala. Tämä tarkoittaa että mahdollisuus löytää sairaala muilla hakusanoilla on rajoittunut.

Google Maps - hakupalvelu on kehittyneempi ja se pystyy ehdottamaan vaihtoehtoisia hakuvaihtoehtoja, joten käyttäjän on todennäköisempää ja helpompaa löytää etsimä informaatio tai feature (ominaisuus). Google Map Makerin tietokanta on sisältörikkaampi ja sisältää kategorioita ja kuvaustietoa joka tekee hakuprosessista älykkäämmän. Tämä sisältää esimerkiksi erilaisia kuvauksia sairaalasta. Tätä kutsutaan metadataksi, se ei ainoastaan kerro lisätietoa datasta vaan auttaa myös sen löytämisessä.

Metadataan sanakirjaselitys on yksinkertaisesti data joka kuvailee tai antaa lisäinformaatiota toisesta datasta. Perinteiset geospaatialisten datojen (SDI) metadata noudattaa erilaisia standardeja, kuten ISO 11915. Perinteiset metadatatiedot on luotu lähinnä ammattikäyttöä ajatellen ja ovat käytöltään usein erikoisosaamista vaativia.

Ammattilaiset usein luovat metadatan ja useimmat geospaatialiset kirjastot ja luettelot ovat ammattilaisten tekemiä. Metadata luodaan tavallisesti noudattamalla erilaisia sääntöjä ja standardeja. Ammattilaisten tekemä metadata on usein laadultaan tarkkaa mutta sen

tekeminen on kallista ja aikaa vievää ja siksi sen tekeminen VGI-datalle ei olisi tehokasta, sillä VGI-informaatiota on valtavat määrät. Tällä hetkellä vapaaehtoiset ovat muuttumassa yhä kokeneimmiksi älypuhelimien, satelliittisysteemien- ja kuvien sekä karttaohjelmien käyttäjiksi ja siten tekevät enemmän geospaatialista dataa. Tästä syystä erilaisia joustavia metadata sisältöjä ja systeemejä olisi syytä alkaa suunnitella (Kalantari ym. 2014).

Kalantari ja muut (2014) ehdottavat, että vapaaehtoiset itse alkavat luoda metadataa ammattilaisten sijasta. OpenStreetMapissa käyttäjä voi lisätä jo nyt lisäinformaatiota, ylimääräisiä attribuuttitietoja datalleen tai muokkaukselleen. Metadatan luominen VGI:lle joko ammattilaisten tai vapaaehtoisten toimesta sisältää kuitenkin yhden puutteen: käyttäjä, eli datan kuluttaja ei ole mukana prosessissa. Metadataa tehdessä ei siis välttämättä palvella käyttäjän tarpeita tai ne eivät täyty. Vaihtoehtoinen tapa on luoda metadataa crowdsourcing menetelmällä, näin VGI-käyttäjät voivat käyttää omaa tulkintaansa jostain maantieteellisestä ilmiöstä ja kuvata sitä. Lisäksi käyttäjä voi ilmaista oman mielipiteensä datan kuvauksesta, laadusta ja sen sopivuudesta johonkin käyttötarkoitukseen.

Tämä metodi lisää käyttäjien osallistumisen prosessiin ja parantaa VGI:n laatua. Jos käyttäjät voivat luoda metadataa VGI:lle ja jakaa sen muiden käyttäjien kanssa, auttaa se VGI:n löytämistä ja lisää arvoa sen sisällölle.

Perinteiset luokittelu ja kuvausmenetelmät geospaatialiselle datalle perustuvat aina standardeihin ja sääntöihin. Standardit ja säännöt voivat luoda vain tietynlaisia luokitteluja kerralla VGI:lle. Esimerkiksi yksi sääntö voi luokitella tie-ominaisuuden yhteen rautatien kanssa ja tehdä niistä yhteisen tason "liikenne". Toinen sääntö voi tehdä näistä kaksi erillistä luokitusta, esim. tie ja julkinen liikenne. Tästä syystä perinteiset luokitusmenetelmät eivät täytä VGI:n tarpeita käyttäjän näkökulmasta, koska ei ole olemassa yhtä yhtenäistä systeemiä. Tästä syystä Kalantari ja muut (2014) väittävät, että crowdsourcing menetelmä voi olla tehokas luokittelija, koska siinä käyttäjien ei tarvitse tehdä päätöksiä erityisten sääntöjen mukaan ja näin rajoittaa luokittelua sellaisten sääntöjen mukaan, jotka eivät ole tuttuja heille. Erityisten sääntöjen sijaan, käyttäjät tuovat heidän oman kielensä luokitukseen, sellaisen joka palvelee juuri heitä ja on heille merkityksellinen. VGI käyttäjät voivat valita erilaisia tageja kuvaamaan samaa elementtiä tai asiaa. Mittakaavaan liittyviä elementtejä voidaan luokitella tageilla "500", "1:500" tai "1/500". Käyttäjä voi käyttää kuvausta, joka on hänelle itsellensä relevantti.

Kalantari ja muut (2014) uskovat että käyttäjien määrittelemä kokoelma tageja ja kuvauksia voi edistää metadatan luomista VGI:lle. VGI systeemissä, missä käyttäjät ovat vapaita tagaamaan maantieteellistä informaatiota, tämä tagien kokoelma voi muodostaa maantieteellisen folksonomian. Kalantari ja muut (2014) ehdottavat kaksi eri mallia metadatan luomiselle. Ensimmäisessä mallissa he luovat tietokannan metadatalle ainoastaan monitoroimalla käyttäjien käyttäytymistä heidän ollessa tästä tietämättömiä. Toisessa mallissa he tieteen tahtoen sallivat käyttäjien luoda metadataa.

Ensimmäisessä mallissa he monitoroivat käytettyjä hakusanoja, analysoivat ne ja sitten käyttivät niitä kuvauksina luomaan sisältöä VGI:n metadatalle. Tämä ns. implisiittinen malli jakautuu kolmeen vaiheeseen: hakusanojen monitorointi, hakusanojen tallennus ja niiden asettaminen. VGI systeemissä on yleensä mahdollisuus hakea tietoa paikan, featuren tai sijainnin mukaan ja systeemi sitten etsii ja palauttaa vastaavan tallenteen joka täsmää hakusanoihin. Käyttäjän on mahdollista tarkastella haun tuloksia ja päättää mikä niistä on relevantti. Esimerkiksi käyttäjä etsii "Black burn High School in Victoria, Australia" Wikimapiasta. Tällä yrityksellä käyttäjä ei saa yhtään hakutulosta ja ensivaikutelma on että Wikimapiassa ei ole Black burn nimistä koulua. Kuitenkin kun hakusanaa muuttaa, saa toisenlaisia tuloksia. Jos hausta poistaa sanan "Victoria" löytää haulla oikean koulun Wikimapian kartalta. Victoria on kuitenkin oletettu hakusana kyseiselle koululle ja pitäisi olla liitettyinä (tagattuna) elementtiin. Ongelman voisi korjata siten, että ylläpitäjä lisää sanan "Victoria" Black burn-koulun tietoihin, mutta tämän tekeminen manuaalisesti isoihin globaalikarttoihin olisi todella työlästä ja epäkäytännöllistä.

Kuitenkin monitoroimalla käyttäjien tekemiä hakuja, saadaan käytännössä tietää, millä sanoilla ja miten he mitäkin kohdetta etsivät. Jos sama hakusana esiintyy useasti, kertoo sen sanan hyödyllisyydestä ja siitä, että se olisi hyvä lisätä metadataan haun helpottamiseksi. Tallennetut, usein käytetyt sanat asetetaan kaikkien käyttäjien käytettäväksi. Tällä metodilla ajan kanssa voidaan luoda avainsanoista, käyttäjien eniten käyttämistä hakusanoista koostuva metadata.

Toisessa, ns. eksplisiittisessä mallissa luodaan metadataa suoraan käyttäjien tekemien kommenttien kautta eli päinvastoin kuin aikaisemmassa, jossa vain epäsuorasti tarkasteltiin mitä hakusanoja käyttäjät käyttivät eniten. Eksplisiittisessä mallissa käyttäjät luovat tageja featureihin perustuen heidän omaan tietoonsa ja ymmärrykseensä. Käyttäjä nimeää

maantieteellisen informaation sen mukaan missä yhteydessä ja kontekstissa sitä aikovat käyttää (Kalantari ym. 2014).

Käyttäjien lisääntyvä määrä, ohjelmien ja digitaalisen sisällön luomisen kasvu vaikeuttaa kuvauksien luomisen kaikelle VGI-informaatiolle. Metadatan luominen tagien avulla tekee käyttäjistä sen tekijöitä ja VGI on niin suurisisältöinen että sen pitäisi olla käyttäjien tukema myös luokittelultaan ja kuvauksiltaan. Toisin kuin SDI, joka on suunniteltu spesifiseen käyttöön ja tarpeisiin, on mahdotonta ennustaa miten käyttäjät etsivät VGI-informaatiota. Toisin sanoen, on mahdotonta ennustaa miten käyttäjät, joiden määrä lisääntyy jatkuvasti, ohjelmat ja sovellukset tunnistavat, etsivät ja käyttävät VGI-datasettejä

Yhteistyöhön perustuva tagien luonti muodostaa lopulta kokoelman kuvauksia, joista muodostuu metadata, joka on lisäksi helposti käyttäjien ymmärrettävissä. Käyttäjien tekemä metadata voi olla laadultaan kohtuullista vaikka siinä olisi virheitä, jos se on tehty tagien avulla. Tagien kautta luotu metadata ei ole vain objektien kuvaamista vaan myös tiedon lisäämistä paikkaa kohden (Kalantari ym. 2014).

## **Eettisyys**

Kuten yleensä uusien ideoiden noustessa esille, on VGI:ssäkin kysymyksiä jotka koskevat mm. sen laillisia ja eettisiä puolia, etenkin kun tätä informaatiota jaetaan ja tuotetaan julkisesti kaikkien nähtäville. Yksi iso huomionkohde on VGI:n tuottajien vapaaehtoisuus ja se, onko data oikeasti tuotettu vapaaehtoisesti tuottajan toimesta ja hänen ollessa siitä tietoinen. VGI voidaan erottaa tuotetusta maantieteellisestä tiedosta (contributed geographic information, CGI) määrittämällä onko kerätty data saatu käyttäjältä joka tietoisesti ja vapaaehtoisesti on sen luonut ja ladannut internettiin. Esimerkiksi käyttäjä on tietoisesti ladannut geotagatun valokuvan nettiin tai luonut havaintotietoa OpenStreetMappiin. Tämä on ns. vapaaehtoista (volunteered) tietoa. Tuotettua (contributed) tietoa on puolestaan tieto, joka kerätään automaattisesti käyttäjän itse ollessa siitä sivussa, ei-aktiivisena. Tämä voi tapahtua esimerkiksi jonkun sovelluksen kautta, joka kerää käyttäjän sijaintitietoja automaattisesti (Blatt 2015).

Digitaaliset kartat ja datakokoelmat jotka käyttävät VGI:tä ja CGI:tä lähteenään täytyy olla tietoisia tästä ja mainita asiasta metadatatiedoissaan. Vastuunalaisuuteen liittyviä kysymyksiä voi syntyä datan tarkkuudesta jos ilmenee potentiaalisia riskejä tai hengenvaaraa.

Tämä voi olla suuri huolenaihe digitaalisten karttojen tekijöille ja kokoelmille, etenkin silloin kun VGI:tä yhdistetään perinteisen datan kanssa, jotka on varastoitu ja käsitelty kartta- ja maantieteellisissä kokoelmissa. Koska VGI-datan tekijöillä ei usein tai aina ole ammatillista koulusta kartoituksesta tai maantieteestä voi VGI-dataan perustuvissa sovelluksissa olla virheitä, tämä voi johtaa haitallisiin päätöksiin jos niitä tehdään VGI-sovelluksien tai tiedon perusteella. Tämän vuoksi ohjelman tai nettisivun käyttöehdoissa on aina mainittava mahdollisista riskeistä eikä varmaa takuuta informaatiosta voida käyttäjälle luvata (Blatt 2015).

Toinen tärkeä asia VGI:n käytössä koskee LBS-laitteita, jotka tallentavat datan tekijän henkilötietoja, sijaintitietoja ja liikettä. Tämä johtuu siitä että crowdsourced menetelmällä saatu maantieteellinen informaatio on yleistynyt sosiaalisessa mediassa, yhä yksityiskohtaisempaa dataa kerätään älypuhelimista, etenkin jos käyttäjä ei ole sulkenut puhelintaan tai laittanut sijaintitietopalveluja pois käytöstä. On olemassa useampia esimerkkejä siitä miten crowdsourced paikkatietoa kerätään, usein tämä tapahtuu ilman käyttäjän tietoa datan keräyksestä tai sen jatkokäytöstä. Crowdsourced menetelmällä kerätty maantieteellinen tieto voi helposti lisätä ihmisten liikkeiden ja toimien ennustamista (Blatt 2015).

Hyvät ominaisuudet	Riskit ja haasteet
- lokaali tieto	-Laatu ja sen varmistus
- saatavuus, julkisuus, avoin data	- spatiaalinen jakautuminen
- halpaa	- oleellisen tiedon löytäminen
- päivittymistahti	- ei standardeja , infrastruktuuria
- kuka tahansa voi osallistua	- attribuuttitietojen puuttuminen
- tiedon määrä valtava	- metadatan puuttuminen
- virheet huomataan nopeasti	- yksityisyys
- nopea leviäminen	- Eettisyys (commercial gain)
- käyttäjien tekemä, käyttäjien kaltainen	- Osa ihmisistä jää ulkopuolelle

Kuva 3. VGI:n hyvät ominaisuudet, riskit ja haasteet.

## MILLAINEN ON HYVÄ VGI - SOVELLUS?

Jotta VGI voisi tulevaisuudessa olla vartenotettava vaihtoehto perinteiselle maantieteelliselle tiedonhankinnalle, tarvitaan lisätutkimusta sen saralta, millainen on hyvä ja motivoiva VGI-sovellus tai sivusto. Edelliset luvut ovat osoittaneet, että VGI osoittaa suurta potentiaalia tuomaan uusia mahdollisuuksia maantieteen ja tiedonkeräyksen kehitykselle, mutta se voi olla sitä vain jos suuria määriä motivoituneita vapaaehtoisia saadaan ilmiöön mukaan tai jos tiedonkeruu ja suodatusmenetelmät kehittyvät vastaamaan VGI:n vaatimuksia. Tämä luvun loppuun on koottu kuvatiivistelmä hyvän VGI - sovelluksen ominaisuuksista (kuva 4).

GIS on jo kauan tunnettu sen vaikeasta käytettävyydestään ja vapaaehtoiset jotka osallistuvat sen käyttämiseen eivät usein omaa tarpeeksi teknillistä tietoa tiedon tuottamiseen. Siitä syystä onkin tärkeää VGI - sovellusta luodessa, jossa ideana on vapaaehtoisten osallistuminen, tehdä ohjelmasta helppokäyttöinen. Käyttäjien pitäisi tuntea sovellus helppokäyttöiseksi sisällöntuoton suhteen ilman teknistä, logistista, laillista tai älyllistä estettä. Paikkatietojärjestelmissä piste on aina helpompi geometria digitoida kuin viiva tai polygoni (Ariffin ym. 2014).

Monet tutkimukset ovat pyrkineet löytämään mikä motivoi vapaaehtoisia tuottamaan tietoa ja mikä saa käyttäjät pysymään aktiivisina. Vaikka vapaaehtoiset eivät saa toiminnastaan rahallista palkintoa, saavat he positiivista mainetta ja arvostuksen tunnetta nettiyhteisössä. Tämä edellyttää oman profiilin luomista sovelluksessa, jotta voi tulla positiivisella tavalla tunnetuksi yhteisössä. Jonkunlainen palkinto, joskaan ei välttämättä rahallinen, olisi myös hyvä olla palvelussa, joka motivoi vapaaehtoisia. Jotta käyttäjä voidaan palkita, pitää olla mahdollista tietää kuka on tuottanut tai muokannut mitäkin ja tämä on syytä ottaa huomioon jo käyttöliittymän suunnittelussa (Ariffin ym. 2014). Myös pelimäiset elementit ja viihteellisyys usein lisäävät motivaatiota, tästä puhutaan lisää gamification - osiossa.

Mitä monivaiheisempi tiedon syöttäminen ohjelmaan on, sitä enemmän se saattaa karsia innostuneita käyttäjiä. Monimutkaiset attribuutit, jotka vaativat edistynyttä teknistä osaamista saattavat turhauttaa käyttäjiä. Mitä enemmän täytettäviä kohtia on, sitä kauemmin käyttäjällä kestää päivittää uutta tietoa. Nopea käyttöliittymien kehitys crowdsourcing-ohjelmissa tekee niistä tulevaisuudessa mahdollisesti hienostuneempia, nopeampia ja helppokäyttöisempiä. Käyttäjät yleensä pitävät helppoudesta ja yksinkertaisuudesta ja siksi on

hyvä selvittää, kuinka yksinkertainen paikkatietojärjestelmäsovelluksen pitäisi olla, jos tarkoitus on luottaa yhteisön tuottamaan sisältöön.

Jotkut käyttöliittymät ovat tehneet jopa rasteridatan käytön mahdolliseksi, vaikka yleensä sen katsotaan olevan liian edistynyttä, taitoa vaativa datamuoto ei ammattilaisten käyttöön. Monet VGI - sovellukset ovat erikoistuneet kaupunkien ja paikkojen kartoittamiseen ja isommassa mittakaavassa maailman kartoittamiseen. Kuitenkin vapaaehtoisilta tuleva informaatio saattaa olla harrastuksen tai muun mielenkiinnon innoittama. Vapaaehtoiset usein tuottavat sisältöä yhteisöihin joihin he tuntevat yhteenkuuluvuutta. Yhteisöillä ja ryhmissä on omat tarpeensa tiedon suhteen ja omat keinot ylläpitää käyttäjien motivaatiota. On syytä tehdä katsaus erilaisiin ryhmiin selvittääkseen millaiset ryhmät hallitsevat vapaaehtoisten osallistuvuutta (Ariffin ym. 2014).

OpenStreetMap tukee monenlaisia eri geometrioita kartoittamaan reaali maailman objekteja (real world objects) kuten supermarketteja, kioskeja, sairaaloita, postitoimistoja, hotelleja ja monia muita. Geometriatyyppejä ovat piste, viiva ja polygoni. Erilaiset valmiit kategoriat ovat helpottamassa käyttäjiä valitsemaan objekteja. Waze-sovelluksessa käyttäjät voivat päivittää tietoa liikenteestä lisäämällä pisteitä kertomaan onnettomuuksista tai poliiseista jotka näkyvät kartalla. Samoin kuin OSM:ssa, Wazessa on monia kategorioita objekteille jotka helpottavat niiden löytämistä. Toisin kuin Waze, Wikimapia rohkaisee käyttäjiään puolestaan kuvailemaan ympäristöään ja tieto jota vapaaehtoiset voivat lisätä linkittyy polygoneihin. Kokeneemmat käyttäjät voivat kuitenkin muokata myös viivageometrioita kuten jokia ja teitä Kokemus määritellään arvioinnilla tai luokituksella jonka käyttäjä on ansainnut profiiliinsa. Healthmap toimii puolestaan täysin OpenStreetMapista, Wazesta ja Wikimapiasta poiketen. Healthmap keskittyy saamaan vapaaehtoisilta tietoa tautien leviämisestä ja käyttäjä merkitsee kartalle paikan nimen josta tautia on havaittu. Koska tiedon muoto on jo valmiiksi hyvin spesifinen (taudin puhkeaminen) ei erilaisia kategorioita tarvita (Ariffin ym. 2014).

OpenStreetMap vaatii käyttäjää rekisteröitymään ennen kuin tämä voi aloittaa kartoittamisen. Rekisteröitymiseen ja kartoituksen aloittamiseen ei kuitenkaan vaadita kuin hyvin pieni määrä tietoja käyttäjästä. Wazessa käyttäjä voi saada tietoa sovelluksesta ilman rekisteröitymistä, mutta jos käyttäjä haluaa osallistua muokkaamiseen, on rekisteröityminen silloin pakollista. Tällä menetelmällä voidaan tehdä laaduntarkistusta, kun pystytään jäljittämään kuka käyttäjä on muokannut mitään. Wikimapia toimii samalla tavalla kuin Waze, käyttäjä voi



selata tietoa ilman rekisteröintiä, mutta jos tämä haluaa lisätä tietoa, tulee rekisteröinti pakolliseksi. Lisäksi Wikimapiassa pääsee muokkaamaan joitain kohtia vain jos on ansainnut tarpeeksi korkean arvostelun käyttäjäprofiiliinsa. Healthmap vaatii myös rekisteröitymisen jos haluaa osallistua tautien kartoittamiseen. Healthmapin kaltaisessa sovelluksessa rekisteröitymisen vaatiminen on välttämätöntä, sillä valheelliset tautimerkinnot voisivat luoda ihmisissä paniikkia (Ariffin ym. 2014).

OSM:ssa attribuuttien määrä, joka käyttäjän pitää täyttää tietoa lisätessään, ei ole aina sama vaan se vaihtelee objektien välillä. Wazessa attribuuttien määrä on puolestaan aina sama. Tämä parantaa sovelluksen käytettävyyttä ja yhtenäisyys (consistency) onkin avainasemassa hyvän käyttöliittymän suunnittelussa. Myös datan tyypillä on väliä, vektorin katsotaan olevan helpompi kuin rasteri. Datatyyppi jota vapaaehtoiset käyttäjät voivat tuottaa OpenStreetMapissa, Wazessa ja Wikimapiassa on vektoripohjainen. OSM kuitenkin tarjoaa toiminnon jolla käyttäjät voivat saada ladattua dataa rasterimuodossa. Prosessi on tehty yksinkertaiseksi, jotta käyttäjien ei tarvitse hallita kaukokartoituksen kuvankäsittelyprosesseja tai fotometriaa. Healthmapissa käyttäjän ei tarvitse piirtää mitään geometriaa sovellukseen, joten se ei käytä rasteria eikä vektoria (Ariffin ym. 2014).

OpenStreetMappia käytetään hyväksi keräämään maantieteellistä tietoa kartoitukseen. Tieto jota voidaan saada OSM:sta, käytetään virallisissa kartoitustehtävissä avuksi paikkamaan informaatioaukkoja. Waze on kuitenkin erilainen, vaikka se tarjoaakin vapaaehtoisille mahdollisuuden muokata karttoja, on sovelluksen pääasiallinen toiminto kuitenkin päivittää autoilijoille tietoa lähellä olevista ruuhkista. Informaatio käytetään reitinsuunnittelutarkoitukseen. Wikimapia taas keskittyy tarjoamaan lisätietoa kartalla oleville objekteille (Ariffin ym. 2014).

Rekisteröityminen parantaa datan laatua. Tutkimalla mitä geometriatyyppejä käyttäjät suosivat, voi auttaa luomaan parempia VGI-sovelluksia tulevaisuudessa. Geometriatyypin pitäisi edustaa oikeaa objektia, kuten maapalsta polygonia, tie polylinea eli viivaa mutta vapaaehtoiset eivät aina ole välttämättä tietoisia tästä. Google Mapsissa monet turistipaikat ja kaupungit ovat merkitty karttaan pisteinä polygonien sijaan joka kertoo vapaaehtoisten tietämättömyydestä mitä geometriatyyppejä kuuluisi kulloinkin käyttää. Hyvä idea onkin luoda valmiit pohjat joista käyttäjä voi valita aina oikean geometriatyypin eli osa

informaatiosta on jo valmiiksi syötetty ohjelmaan, tämä lisää helppoutta ja sitä kautta mahdollisesti osallistumista (Ariffin ym. 2014).

Vielä nykyään informaatio jota vapaaehtoiset tuottavat on melko yksinkertaista ja helposti ymmärrettävää myös ei-ammattilaiselle. Tulevaisuudessa saattaa kartoitus kuitenkin muuttua 3D-muotoon ja vapaaehtoisten tuottama tieto vaatia spatiaalista analysointia organisaatioiden tarpeisiin. Tämä voi johtaa siihen että vapaaehtoisilta vaaditaan myös laatua, joka saattaa tarvita teknistä osaamista paljon enemmän. Jos eri tahot alkavat luottaa yhä enemmän vapaaehtoisten ja yhteisön keräämään tietoon, voi myös vaatimukset tiedolle ja sen tyyppille nousta. Herää kysymyksiä kuinka paljon voidaan luottaa yhteisön keräämään tietoon? Kuinka luotettavaa se on ja miten taata että ihmiset pysyvät motivoituneina tuottamaan tietoa? Vaikka VGI - datan laatu on yhä huolenaihe sen käyttämisessä, on huonolaatuinen data kuitenkin joskus parempi kuin ei dataa lainkaan. Suurin osa ajattelee VGI:n olevan yhä vain täydentävä tiedonhankintamenetelmä eikä se pysty korvaamaan perinteisiä menetelmiä (Ariffin ym. 2014).

Wang ja muut (2014) puhuvat käyttäjäprofiilin ja erityisesti location field - kohdan tärkeydestä ja tämä on asia joka myös VGI - sovelluksissa tulisi huomioida. Käyttäjäprofiili, jossa on jonkinasteinen täydennettävä kohta/kenttä sijainnille, on hyvin tyypillistä online-yhteisöissä kuten esimerkiksi Facebookissa, Twitterissä ja Pinterestissä. Sellainen löytyy myös esimerkiksi eBaysta ja Githubista. Tutkijat ja operaattorit ovat huomanneet tämän ominaisuuden olevan hyödyllinen. Monet käyttäjät erilaisissa internetyhteisöissä ovat haluttomia tai eivät ole kiinnostuneet ilmoittamaan sijaintitietoa, eli lisäämään paikkatagin heidän luomaansa sisältöön. Esimerkiksi vain hyvin pieni osa Twitter viesteistä sisältää sijaintitiedon (geotagin) (Wang ym. 2014).

Sijaintitietodatan käyttämisessä voi olla myös riskejä ja ongelmia. Joillakin käyttäjillä on tapana ilmaista sijaintitietonsa ei-maantieteellisellä henkilökohtaisessa viestillä. Esimerkiksi joku voi kirjoittaa sen hetkiseksi sijaintitiedoksi lähettämänsä viestin yhteydessä että "olisinpä missä tahansa muualla kuin täällä" joka kuvailee humoristiseen sävyyn henkilön toivomusta olla muualla mutta se ei tutkijalle kerro maantieteellistä sijaintia. Sijaintitiedot saattavat olla myös hyvin suurimalkaisia, esimerkiksi kaupungin nimi.

Wang ja muut (2014) esittelevätkin artikkelissaan ideoita, miten sijaintitietokenttää (location field) voisi parantaa, jotta edellä mainittuja ongelmia olisi tulevaisuudessa vähemmän ja tutkijat saisivat käyttöönsä laadukkaampaa lähdemateriaalia. He toteavat erityisesti, että

tutkimusta sen suhteen millainen vaikutus sijaintikentän grafiikan suunnittelussa on siihen, kuinka laadukasta informaatiota käyttäjät siihen kirjoittavat, tulisi tehdä. Kuten edellä jo mainittiin, kolme eri tekijää voi vaikuttaa sijaintiedon laatuun: 1) käyttäjien aktiivisuus täyttää sijaintitietokenttä, 2) sijaintitiedon tarkkuus tai suuripiirteisyys 3) sijaintitiedon antaminen ei-maantieteellisesti, sanallisena kuvauksena joka ei välttämättä edes sisällä paikkatietoa.

Wang ja muut (2014) tarkastelivat 18 eri internetyhteisön sijaintitietokenttää ja tekivät havaintoja mm. sen otsikon, kentän pituuden, onko sijaintitieto julkinen kaikille vai voiko sen asettaa halutessaan yksityiseksi ja sen suhteen saako käyttäjä itse kirjoittaa kenttään sijainnin vai antaako ohjelma automaattiset vaihtoehdot josta valita oma sijainti. He tarkastelivat myös onko sijaintitietokentässä vain yksi täytettävä kohta vai esimerkiksi tämänhetkinen kotikaupunki ja synnyinkaupunki erikseen.

He huomasivat eroja eri sivustojen välillä kaikkien tarkasteltavien asioiden kohdalla. Jotkut sivut otsikoivat tietokentän "sijainti" sanalla ja toiset täsmällisemmin kaupunki, osoite tai postinumero sanoilla. Mahdollisen merkkien määrä vaihteli myös 29 - 49 välillä. Joissakin sivuissa sai paikan valita valmiiksi annetuista ja joissakin se oli jaettu moneen eri tasoon, kuten kaupunki ja kotikaupunki.

Jotta saataisiin ihmiset aktiivisemmin ilmoittamaan sijaintitietonsa ja laittamaan sen tutkimuskäytön kannalta laadukkaaseen maantieteelliseen muotoon ehdottavat Wang ja muut (2014) tutkimuksensa perusteella seuraavaa. He huomasivat, että jos sijaintitietokenttä otsikoidaan sanoilla "sijainti" tai "kaupunki" on aktiivisuusprosentti kentän täyttämässä suurempi kuin jos pyydetään osoitetta tai aluekoodia. Sivustojen suunnittelijoiden täytyy tehdä siis tietynlainen vaihtokauppa informaation tarkkuuden ja aktiivisuustason välillä. Myös kentän pituus vaikuttaa asiaan, kuinka tarkan osoitteen ihmiset ajattelevat siihen mahtuvan.

Sivustoilla jotka käyttävät sijaintitiedon verifiointisysteemiä, eli käyttäjä ei voi kirjoittaa sijaintiaan itse vaan järjestelmä hakee paikan omista tietokannoistaan ja käyttäjä voi ilmoittaa sijaintinsa vain jos vastaava löytyy tietokannasta. Tämä systeemi poistaa kokonaan ei-maantieteelliset sijaintitiedot. Wang ja muut (2014) eivät omassa tutkimuksessa havainneet suurta eroa kun he vertasivat aktiivisuustasoa sijaintitiedon ilmoittamisessa verifioidussa systeemissä ja ei verifioidun välillä.

Wang ja muut (2014) väittävät, että heidän löydöillään saattaa olla merkittäviä vaikutuksia sijaintikentän ilmeen suunnittelussa ja saattaa olla mahdollista saavuttaa parempia

sijaintitietoja kuitenkin vaikuttamatta aktiivisuustasoon millä ihmiset täyttävät sijaintietokenttiä. Verifiointi systeemi voisi mahdollisesti poistaa myös epätietoisuutta tilanteissa missä monella eri paikalla saattaa olla sama nimi eikä silloin voida olla varma mitä paikkaa käyttäjä on tarkoittanut.

Viimeisenä Wang ja muut (2014) tutkivat sijaintitiedon näkyvyyttä muille käyttäjille. Vaihtoehtoina he vertasivat miten halukkaita ihmiset ovat ilmoittamaan sijaintitietonsa jos se näkyy kaikille julkisesti heidän profiilissa, se on yksityinen tai jos ei ole ilmoitettu tuleeko tieto julkisesti näkyviin vai ei. Heidän tuloksensa olivat samankaltaisia verifiointikokeilun kanssa, eli näkyvyydellä ei ollut suurta merkitystä käyttäjien innokkuuteen ilmoittaa sijainti. Eroja löytyi kuitenkin sen suhteen, kuinka suurimalkaisesti ihmiset kertoivat sijaintitietonsa riippuen siitä tuliko se julkisesti näkyvästi kaikille vai ei. Mikäli sijaintitieto oli julkinen, laitettiin siihen mieluummin karkea sijainti, esimerkiksi vain kaupungin tai osavaltion nimi.

Loppupäätelmänä he mainitsevat että esimerkiksi Twitterin tapauksessa jos sivusto muuttaisi "sijainti" sanan tietokentässä tarkemmaksi, esimerkiksi osoite, aluekoodi ja kaupunki voisivat tutkijat jotka käyttävät sijaintitietokentän dataa saada paljon tarkempaa ja laadukkaampaa informaatiota. Tämä saattaisi vaihtoehtoisesti kuitenkin vähentää ihmisten aktiivisuutta ilmoittaa sijaintitieto ylipäättään. Jos Twitter puolestaan muuttaisi sijaintitietokentän verifiointia vaativaksi (kuten esimerkiksi Facebook tekee) niin sillä ei välttämättä olisi suurta vaikutusta ihmisten aktiivisuuteen ilmoittaa sijainti, mutta se poistaisi kaikki ei-maantieteelliset sijaintitiedot sekä sekaannukset tilanteissa missä monilla paikoilla voi olla sama nimi. Verifiointisysteemi helpottaisi myös huomattavasti tietojen geokoodausta.

## **Gamification**

Tavalliset internetin käyttäjät voidaan motivoida pelien avulla tekemään tehtäviä jotka ovat hyödyksi erilaisissa sovelluksissa, kuten hakukoneissa, ilman mitään muuta kannustusta kuin itse tekemisen viihteellisyys (Matyas ym. 2011).

Pelielementtien käyttäminen pelien ulkopuolella on vakiintunut konsepti. Gamification on systeemi, jossa parannetaan käyttäjäsysteemin houkuttelevuutta tekemällä siitä pelin kaltainen, jotta käyttäjät olisivat motivoituneita käyttämään enemmän aikaa ja resursseja tehtävään. Jotta peli tai pelimäiset ominaisuudet olisivat tarpeeksi viihdyttäviä motivoimaan

käyttäjää, on niissä yleensä erilaisia palkintosysteemejä. Palkinnot voivat olla esineitä (pelissä käytettäviä), pelimaailman rahaa, pääsy uusiin alueisiin tai informaatioon, valta tai koreampi status (Odobasić ym. 2013).

Peli on yleensä järjestelmällistä ja viihteellistä pelaamista ja joskus opetuksessa ja oppimisessa käytettävä työkalu. Gamification on puolestaan pelien elementtien käyttämistä pelin ulkopuolisessa kontekstissa. Gamification tarkoittaa myös prosessia jossa pelien ajatusmaailmaa ja mekanismeja käytetään käyttäjien motivoinnissa ja ongelmanratkaisussa. Pelit voivat olla myös ns. location based - pelejä eli ne jollakin tapaa toimivat ja etenevät perustuen pelaajan sijaintiin. Location based - pelit lähes aina tukevat jotain paikannusteknologiaa kuten GPS, WiFi tai bluetooth (Odobasić ym. 2013).

Citizen science on vain yksi esimerkki UGC:n ja gamificationin yhteensulautumisesta. Palvelut kuten Zooniverse ylläpitää lukuisia projekteja jotka sallivat käyttäjien osallistua tieteelliseen tutkimukseen joka vaihtelee galaksien luokittelusta ilmastodatan järjestelyyn. Projektit käyttävät pelimäisiä elementtejä lisätäkseen käyttäjien osallistumista. Peleillä on vahva kyky välittää pelaajille tuntemus että heidän päätöksillä on oikeasti merkitystä ja vaikutusta ympäröivään maailmaan ja että he voivat olla mukana muuttamassa asioita. Päätösvallan ja merkityksellisyyden tunne viittaa subjektiiviseen tietoisuuteen jossa henkilö suorittaa ja kontrolloi omia valinnaisia toimintoja maailmassa (Odobasić ym. 2013).

Palkintosysteemi on kaikista tavallisin elementti pelimalleissa. Odobasić ja muut (2013) jakavat ne neljään erilliseen kategoriaan: status, pääsy (access), valta ja materia. Status tarkoittaa sosiaalista asemaa peliyhteisössä, tämä on tärkeä etenkin sosiaalisissa verkkopeleissä. Pääsy (access) tarkoittaa että pelaajalla on mahdollisuus kokeilla uusia ominaisuuksia ennen muita. Valta tarkoittaa tässä yhteydessä esimerkiksi muokkausmahdollisuutta, pelaaja tai käyttäjä pääsee esimerkiksi muokkaamaan toisten käyttäjien luomia sisältöjä. Materia tarkoittaa mitä tahansa ilmaista objektia tai esinettä.

Pelaajan motivaation ajatteleminen ja huomioonottaminen on tärkeä osa pelin suunnittelua, koska ilman motivaatiota pelaaja ei ole kiinnostunut pelaamaan ja etenemään pelissä. Pelisuunnittelijat pyrkivät tekemään pelejä, jossa pelaaja on pysyvästi jatkuvassa flow-tilassa.

Pelimekanismit ovat sarja työkaluja, joiden avulla peli lupaa merkityksellisen kokemuksen pelaajalle. Pelimekanismeista voi erotella ainakin seuraavan laisia elementtejä: pisteet, tasot, tuloslistat, arvomerkit ja erilaiset haastetehtävät (Odobasić ym. 2013).

Foursquare on esimerkki location based - pelistä joka on lisäksi sosiaalisesti verkostoitunut. Foursquare kuvailee itseään seuraavan laisesti: ilmainen ohjelma joka auttaa sinua ja ystäviäsi saamaan kaiken irti paikasta, jossa sillä hetkellä sijaitset. Foursquareen voi tallentaa paikat joissa on vierailut ja jakaa ne muiden kanssa. Ohjelmasta voi myös etsiä ideoita ja inspiraatiota mitä kussakin paikassa voi tehdä ja ohjelma antaa personalisoituja suosituksia sijainnin perusteella. Foursquare on erinomainen esimerkki pelimäisestä kokemuksesta ja sen suosio on valtava. 30 miljoonaa käyttäjää ja yli 4 miljoonaa käyttöä päivittäin. Käyttäjät saavat statuspalkinnon yhdistelemällä ansioituja merkkejä, tuloksia listalla, saavuttamalla tasoja ja haasteita (Odobasić ym. 2013).

Aikaisemmissa kappaleissa esille tullut Waze sovellus noudattaa myös gamification - elementtejä. Se on ilmainen sosiaalinen mobiilisovellus joka käyttää reaaliaikaista tiedonsyöttöä kuljettajilta, joka auttaa liikenneolosuhteiden päivittämistä toisten ajajien käyttöön. Waze eroaa tavallisesta GPS navigointiohjelmasta koska se on yhteisölähtöinen joka oppii ja kehittyy käyttäjien ajonopeuksista. Waze palkitsee käyttäjiä statuksilla, vallalla ja tavaroilla. Ajaessa Waze kerää automaattisesti GPS-dataa joka käytetään reaaliaikaisessa navigoinnissa. Käyttäjät voivat kerätä pisteitä jokaisesta ajetusta metrissä, onnettomuusraporteista, liikennerruuhkista, nopeuskameroista, poliiseista tai muokkaamalla dataa.

Pelaajien ansaitsemat pisteet sijoittavat käyttäjän tuloslistalle jonne pääsemällä puolestaan voi ansaita uusia statuksia. Käyttäjät voivat myös ansaita valtuuksia jotka mahdollistavat yhteisön foorumin moderoinnin tai datan muokkausmahdollisuuksia. Waze tarjoaa myös lisäpisteitä, joita kerätä yksinkertaisesti ajamalla niiden kohdalle tai yli. Lisäpisteillä voi lunastaa esimerkiksi lahjakortteja tai saada alennusta polttoaineesta. Lisäpisteillä rohkaistaan käyttäjiä osallistumaan ja kilpailemaan joka puolestaan takaa laadukkaampaa informaatiota liikenteestä käyttäjille (Odobasić ym. 2013).

Ingress on tosielämään perustuva monin peli, joka on saatavilla Android laitteille. Pelin idea on että pelaaja kävelee paikasta toiseen pitäen pelin GPS:ssää päällä suorittaen samalla pelissä olevia tehtäviä. Peli muodostaa vaihtoehdoisen todellisuuden, jossa kaksi ryhmää kilpailee

niin kutsutuista control point - alueista. Eli ryhmän täytyy pitää kyseistä sijaintia hallussa. Voittaja on ryhmä jolla on hallussa eniten sijainteja.

Kort on OpenStreetMap peli ja web-palvelu joka on optimoitu mobiililaitteille ja joka yhdistää gamification konseptin VGI-ilmiöön ja on tarkoitettu ei-ammattilaisille. Pelissä pelaajat keräävät pisteitä tekemällä tehtäviä, esimerkiksi lisäämällä nimiä tunnettuihin paikkoihin, määrittämällä muiden käyttäjien tekemien sisältöjen kielen tai teiden nopeusrajoituksia. Pelielementtejä käyttämällä voidaan siis parantaa käyttäjien osallistumista tekemällä systeemistä viihdyttävän (Odobasić ym. 2013).

Gamification voidaan kuitenkin myös ajatella psykologiseksi tekniikaksi, joka mahdollistaa systeemien omistajien manipuloida käyttäjiä käyttäytymällä halutulla tavalla. Pelit sisältävät neljä ei ominaisuutta jotka motivoivat pelaajia pelaamaan. Ne ovat pelin oppiminen ja lopulta siinä taitavaksi tuleminen, stressin lievittäminen, viihteellisyys ja sosiaalisuus. Pelit luovat myös tunteita ja ne ovat tärkeitä flow:n ylläpitämisessä. Motivaatio on myös yksi tekijä joka pitää pelaajia jatkuvassa flow-tilassa. Tämä usein vaatii jatkuvien palkintojen tai ansioiden saamisen, käyttäjä menettää mielenkiintonsa jos on saavuttanut jo kaiken saavutettavan pelissä tai ohjelmassa.

Tulloch (2008) kertoo esimerkin Gamificationista Second Life nimisessä pelissä. Second Life on laajasti tunnettu ja suosittu videopeli, se muodostaa maailman jossa pelaajat käyttäjät virtuaalista rahaa, luovat virtuaalista taidetta, ja tekevät omia ilmaisia paikoista ja tiloista ja siitä miten ne koetaan ja joskus tämän voi rinnastaa VGI:hin. Yksi esimerkki Second Life pelistä koskee New Yorkin LaGuardia lentokentän laskeutumisvaloja ja puistoa sekä osoittaa miten peleissä voi olla todellista potentiaalia interaktiiviselle osallistumiselle ja ilmaisulle paikkoja ja tiloja koskevissa projekteissa. Tämä voi myös toimia esimerkkinä ruohonjuuritason datankeräämiselle. (Tulloch 2008).

Projektissa haluttiin tuoda mukaan yhteisön luovuus suunniteltaessa puistoa koneiden laskeutumisradan loppuun LaGuardia lentokentälle. Erityinen virtuaaliareena luotiin Second life peliin johon tehtiin malli puistosta ja sitä ympäröivistä rakennuksista sekä laskeutumisvalotorneista. Pelaajat pystyivät hahmoillaan menemään puistoon ja käyttämään erityisiä työkaluja ja suunnittelemaan oman persoonallisen vision puiston valoista. Virtuaaliareena sisälsi karttoja ja informaatiota vierailijoille ja mahdollisuuden keskustella niistä sekä vaihtoehtoisista suunnitelmista.

Vaikka virtuaalinen suunnittelu ei ole samanlaista VGI:tä kuin georeferoitujen kuvien lisääminen Google Earth palveluun, mutta rinnakkaisuus VGI:hin on siinä että nämä puistoa suunnittelevat vapaaehtoiset työskentelevät kartalla ja sijoittavat objekteja mihin he kokevat niiden sopivan. Kyky luoda virtuaalisia esityksiä ja keskusteluja lähetetyistä materiaaleista luo lisämahdollisuuden vapaaehtoisille erotella ja määrittellä luotua sisältöä. Osallistuminen vaatii kuitenkin navigoimaan Second Life pelin lävitse joka on sekä hyöty että huolenaihe. 3D kaupunkimallien määrä on lisääntynyt Second Lifessä huomattavasti. Tämä on merkki siitä että 3D mallit ovat saaneet laajempaa suosiota vaikka yleisö on sen suhteen edelleen rajoittunut teknologian ja taitojen suhteen. Second Lifen pelaajien keski-ikä on nuorekasta ja tämä voi vaikuttaa tuloksiin ja vanhemman väestön vähäinen määrä pelissä voi aiheuttaa vinoumaa (Tulloch 2008).

Mataya ja muut (2011) esittelevät Location Based - peleille laadunvarmistusstrategian. On tärkeää huomata, että sijainnillinen tarkkuus ei ole pääasiassa kiinni pelaajan taidoista käyttäen GPS-laitteita tai hänen tietämyksestä karttaprojektiosta tai koordinaattijärjestelmästä. Lähes kaikki ongelmat jotka vaativat teknillistä taustatietoa on peliohjelmiston vastuulla. Suuri ongelma maantieteellistä sijaintia tallennettaessa liittyy erilaisiin kohtiin, jotka vaativat monitulkintaa. Esimerkiksi rakennuksessa voi olla useampia sisäänkäyntejä mutta mitä niistä pitäisi käyttää merkitsemään rakennuksen sijainti?

Vaihtoehtoja on vielä enemmän laajemmille tai pidemmille objekteille kuten kaduille ja alueille. Esimerkiksi Google Maps palauttaa Otto-Friedrich Yliopistosta tehdyn haun koordinaattipareina 49.90763,10.90466. Mutta sama haku Bing Mapsilla johtaa koordinaatteihin 49,89378,10,88595. Jos ammatti-datankerääjällä on ongelmia monitulkinnassa, millainen mahdollisuus sen ratkaisuun voi sitten olla ei koulutetulla vapaaehtoisella? Keskittäen monta eri tallennetta ei myöskään toimi, sillä lopputulos on lähes aina sijainti joka ei vastaa mitään tosielämän paikkaa. Tämän ongelman pohjalta Mataya ja muut (2011) esittelevät seuraavanlaisen pelinsuunnittelumallin.

Suunnitelmassa he käyttävät jo työssä ennestään esille tullutta ja tuttua wisdom of the crowd konseptia. Pelaajat ottavat huomioon toisten pelaajien luomaa dataa. Suunnitelma on räätälöity tarjoamaan tietynasteinen implisiittinen laadunhallintamekanismi location based - peleihin jotka luovat geospaatialista informaatiota. Mekanismi on kuitenkin hieman erilainen kuin mitä tyyppillisesti liitetään wisdom of the crowd - ideaan. Se matkii arvostelustrategiaa



yrittämällä motivoida pelaajia luomaan dataa josta useimmat pelaajat ovat samaa mieltä. Tätä ei suoriteta ulkoisella palkinnolla tai rangaistuksella vaan spesifisesti suunnitellulla pelisäännöllä.

Matayas ja muut (2011) esittelevät pelisuunnitelmansa yksinkertaisella location based -pelillä GeoSnake, joka on uusi versio aikaisemmin tunnetusta ja suositusta pelistä Snake (käärmepele). Erotuksena alkuperäiseen peliin, tässä uudessa versiossa pelaajien täytyy vierailla eri paikoissa ja syöttää paikkaan kuuluvat koordinaatit. Pelaajat saavat pisteitä jokaisesta uudesta paikasta jossa he käyvät. Pisteitä vähennetään jos pelaaja kulkee uudestaan kohdasta mistä on jo aikaisemmin mennyt tai jos kävelee samaa reittiä kahdesti. Eli jokainen katu voidaan ymmärtää yksisuuntaisena ja pelaajan kävelemä reitti on ikään kuin käärmeen häntä joka kasvaa sitä mukaan kun pelaaja käyttää uusia reittejä.

He muotoilivat kaksi sääntöä sille, miten pelaajien tulisi valita koordinaatit vierailuille paikoille. Ensimmäinen sääntö kehottaa pelaajia valitsemaan koordinaatit joka tarkoittaa kyseistä sijaintia varmuudella, eli valitse koordinaatit niin läheltä paikkaa kuin suinkin mahdollista. Toinen sääntö sisältää wisdom of the crowd ajattelun, pelaajan tulee valita koordinaatit jotka he olettavat toisten pelaajien valitsevan. Pisteitä sai sen mukaan kuinka hyvin koordinaatit täsmäsivät paikan kanssa sekä siitä kuinka paljon muita pelaajia oli antanut samat koordinaatit. Heidän hypoteesi oli että pelaajat tuottavat tarkempaa geospaatialista dataa toisen säännön alaisena kuin ensimmäisen.

Matayas ja muut (2011) kertovat tuloksissaan että sääntö kaksi tuotti paljon suurempia koordinaattiklustereita kuin sääntö 1 jossa ne olivat enemmän hajallaan. He toteavat että näin pelaajia saadaan "hujattua" tuottamaan dataa josta suurin osa pelaajista on samaa mieltä. Tämä laittaa pelaajat myös ajattelemaan ei vain peliä, vaan myös heidän informaationsa laadukkuutta ja etenkin monitulkintakysymyksiä vaikeutta.

Isot yhtiöt kuten Google ja Foursquare ovat jo käyttäneet gamification tekniikkaa maantieteellisen datan keräämisessä ja hyödyntäneet sitä yritystoiminnassa. Avain on siitä että pelaaja ei keskity datankeräämiseen vaan on keskittynyt peliin ja datankeruu tapahtuu siinä sivussa. Maantieteellisen datankeruun pitäisi olla sivutuote pelaamisen ja pelikokemuksen ohella (Odobasić ym. 2013) mutta kuitenkin olisi hyvä saada pelaajat ajattelemaan tuotostensa laatua, kuten Matayas ym. (2011) toivat esille. Pelien käyttämät tekniikat käyttäjien motivoimisessa ja heidän kiinnostuksen ylläpitämisessä voivat tarjota tärkeän näkökulman ja ponnistuslaudan tulevaisuuden VGI-sovelluksien kehittämisessä.

Hyvän VGI – sovelluksen ominaisuudet	
1. Helppokäyttöisyys	Kynnys ohjelman käytön oppimiseen oltava matala, jotta yhä useampi innostuu lähtemään mukaan.
2. Pakollinen käyttäjätunnus	Oma profiiliin luonti. Lisäoikeuksia maineikkaille käyttäjille. Motivointi ja arvostus yhteisössä.
3. Valmiit kategoriat	Käyttäjät saavat valita valmiista pohjista ja kategorioista, helpottaa käyttöä ja vähentää virheitä.
4. Yhtenäisyys	Tiedonlisäyksessä täytettävät kohdat yhtenäisiä ja samoja joka kerralla.
5. Houkutteleva ja toimiva ulkoasu	Selkeät, tyylikkääät ja nykyaikaiset grafiikat. Ulkoasulla väliä datan laadussa! (Esim. Location field)
6. Yhteisö	Kaverilistat, kommentointi, mahdollisuus tykätä muiden muokkauksista. Kiinnostava ja sosiaalinen alusta yhteisölle.
7. Pelimäisyys (Gamification)	Peli-elementtien käyttäminen hyväksi käyttäjien motivoimisessa

Kuva 4. Hyvän VGI - sovelluksen ominaisuudet.

## YHTEENVETO

Olen tässä työssä käynyt läpi VGI - termin määrittämisen, sen syntyyn johtaneet tekijät, esimerkkejä VGI:n hyödyntämisestä tutkimuskäytössä, tarkastellut ilmiöön liittyviä ongelmia ja mahdollisia ratkaisuja sekä lopuksi pohtinut hyvän VGI - sovelluksen ominaisuuksia. Viimeisenä menttiin maantieteen alueelta pelimaailmaan joka minusta onkin samalla hyvä esimerkki siitä, miten VGI:llä, joka on osa geoinformatiikkaa, on loistava mahdollisuus linkittää maantiede mitä moninaisempiin osa-alueisiin, sekä tieteellisiin että viihdemaailmaan. Mielestäni VGI on jotain hyvin uniikkia, mistä geoinformatiikan ja maantieteen tulisi ottaa paremmin kiinni.

Pinnalta katsottuna VGI on moniosainen sarja työkaluja, digitaalinen teknologia joka mahdollistaa kenen tahansa helposti tuottaa ja käyttää spatiaalista informaatiota. Suositut sivustot kuten Google Maps, Twitter ja Facebook, joita monet käyttävät lähes päivittäin ovat muodostaneet kansikuvan tälle uudelle teknologialle. Mutta VGI on paljon enemmän kuin pelkästään teknologia (Poorthuis & Hook 2014).

VGI, kuten esimerkiksi Flickr kuvapalveluun ladattu valokuva jossa on sijaintitieto, digitaaliset liput ja matkakortit, Twitterin viestit ja niiden sijaintitiedot ja matkapuhelimella tehdyt varaukset ovat enemmän kuin pelkästään uusi datan muoto. Se muodostaa uuden paradigman sosio-spatiaaliselle tutkimukselle, jatkuvalla uusien maisemien ja käyttäytymismuotojen seurannalle, mielipiteille, asenteille ja sosiaalisille vuorovaikutuksille tässä nopeasti muuttuvassa urbaanissa yhteisöissä. Geopsatiaaliset teknologiat ovat merkittävästi aktivoineet sosiaalista mediaa, nettiä ja mobiililaitteita ja seurauksena suuri määrä georeferoitua informaatiota asioista ja ihmisistä on käytettävissä ja saatavilla, joka auttaa meitä paremmin ymmärtämään ympäristöä ja urbaaneja systeemejä (Jing & Thill 2015).

Mielestäni idea siitä, että maantiede on päässyt näinkin merkittäväksi osaksi sosiaalista mediaa, joka meidän nyky-yhteiskunnassa on ollut todella suuressa suosiossa ja mielestäni tullut jäädäkseen, on erittäin jännittävä. Jos VGI pystyy valjastamaan ihmismassat käyttöönsä tulevaisuudessa yhä tehokkaammin ja saamaan nämä suuret joukot motivoituneina mukaan tutkimuksenteekoon, saattaa meillä tulevaisuudessa olla pääsy aineistoihin, seurantajärjestelmiin ja tietokantoihin, jollaisia me ei osata vielä edes kuvitella.

## LÄHTEET

- Alivand M, Hochmair H & Srinivasan S (2015) Analyzing how travelers choose scenic routes using route choice models. *Computers , Environment Urban Systems* 50: 41-52.
- Ariffin I, Solemon B, Anwar RM, Din MM & Azmi NN (2014) Exploring the potentials of volunteered geographic Information as a source for spatial data acquisition. *Institute of Physics Publishing Conference Series: Earth and Environmental Science*. 20.
- Blatt AJ (2015) The benefits and risks of volunteered geographic information. *Journal of Map and Geographic Libraries* 11(1): 99-104.
- Bordogna G, Carrara P, Criscuolo L, Pepe M & Rampini A (2016) On predicting and improving the quality of Volunteer Geographic Information projects. *International Journal of Digital Earth* 9(2): 134-155.
- Brovelli MA, Minghini M & Zamboni G (2015) Public participation in GIS via mobile applications. *ISPRS Journal Photogrammetry and Remote Sensing* 114: 306-315.
- Chen X & Yang X (2014) Does food environment influence food choices? A geographical analysis through "tweets". *Applied Geography* 51: 82-89.
- Cinnamon J (2015) Deconstructing the binaries of spatial data production: Towards hybridity. *Canadian Geographer* 59(1): 35-51.
- Cope M (2015) Commentary: Geographies of digital lives: Trajectories in the production of knowledge with user-generated content. *Landscape and Urban Planning* 142: 212-214.
- Fast V & Rinner C (2014) A systems perspective on volunteered geographic information. *ISPRS International Journal of Geo-Information* 3(4): 1278-1292.
- Feick R & Robertson C (2014) A multi-scale approach to exploring urban places in geotagged photographs. *Computers , Environment and Urban Systems* 53: 96-109.

Fonte CC, Bastin L, See L, Foody G & Lupia F (2015) Usability of VGI for validation of land cover maps. *International Journal of Geographical Information Science* 29(7): 1269-1291.

Goodchild MF (2007) Citizens as sensors: The world of volunteered geography. *GeoJournal* 69(4): 211-221.

Goodchild MF & Li L (2012) Assuring the quality of volunteered geographic information. *Spatial Statistics* 1: 110-120.

Haworth B & Bruce E (2015) A Review of Volunteered Geographic Information for Disaster Management. *Geography Compass* 9(5): 237-250.

Jiang B & Thill J- (2015) Volunteered Geographic Information: Towards the establishment of a new paradigm. *Computers , Environment and Urban Systems* 53: 1-3.

Kalantari M, Rajabifard A, Olfat H & Williamson I (2014) Geospatial Metadata 2.0 - An approach for Volunteered Geographic Information. *Computers , Environment and Urban Systems* 48: 35-48.

Mohammadi N & Malek M (2015) Artificial intelligence-based solution to estimate the spatial accuracy of volunteered geographic data. *Journal of Spatial Science* 60(1): 119-135.

Neis P (2015) Measuring the Reliability of Wheelchair User Route Planning based on Volunteered Geographic Information. *Transactions in GIS* 19(2): 188-201.

Odobašić D, Medak D & Miler M (2013) Gamification of Geographich Data Collection. Conference Paper DOI: 10.1553/giscience2013s328 Conference: GI\_Forum 2013

Poorthuis A & Zook M (2014) Spaces of volunteered geographic information. *The Ashgate Research Companion to Media Geography* : 311-328.

Stevens KB & Pfeiffer DU (2015) Sources of spatial animal and human health data: Casting the net wide to deal more effectively with increasingly complex disease problems. *Spatial and Spatio-temporal Epidemiology* 13: 15-29.

- Tulloch DL (2008) Is VGI participation? From vernal pools to video games. *GeoJournal* 72(3-4): 161-171.
- Upton V, Ryan M, O'Donoghue C & Dhubhain AN (2015) Combining conventional and volunteered geographic information to identify and model forest recreational resources. *Applied Geography* 60: 69-76.
- Wang T-, Harper FM & Hecht B (2014) Designing better location fields in user profiles. *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work.* : 73-80.
- Wilson MW (2014) Geospatial technologies in the location-aware future. *Journal of Transport Geography* 34: 297-299.
- Yan Y, Feng C- & Wang Y- (2016) Utilizing fuzzy set theory to assure the quality of volunteered geographic information. *GeoJournal* : 1-16.
- Zhu A-, Zhang G, Wang W, Xiao W, Huang Z-, Dunzhu G-, Ren G, Qin C-, Yang L, Pei T & Yang S (2015) A citizen data-based approach to predictive mapping of spatial variation of natural phenomena. *International Journal of Geographical Information Science* 29(10): 1864-1886.