

Joukkoistaminen käytettävyyden arvioinnissa

Veera Kuusikko

Tampereen yliopisto
Viestintätieteiden tiedekunta
MDP in Human-Technology Interaction
Pro gradu -tutkielma
Ohjaaja: Päivi Majaranta
Toukokuu 2017

Tampereen yliopisto

Viestintätieteiden tiedekunta

MDP in Human Technology Interaction

Veera Kuusikko: Joukkoistaminen käytettävyyden arvioinnissa

Pro gradu -tutkielma, 66 sivua, 2 liitesivua

Toukokuu 2017

Käytettävyyden arvioinnin toteuttamista pidetään usein kalliina ja vaikeana. Varsinkin käytettävyydestaustuksen tapahtuessa käytettävyydelaboratoriossa tulee kuluja niin tilasta, laitteistosta kuin osallistujien saamisesta paikalle. Lisäksi monesti osallistujien rekrytointi on aikaa vievää ja osallistujien löytäminen vaikeaa. Näistä syistä käytettävyydestausta tehdään usein pienimuotoisesti ja vähäisellä osallistujamäärällä, mikä vaikeuttaa käytettävyysongelmien löytämistä. Yksi mahdollinen ratkaisu tähän ongelmaan on joukkoistaminen, jossa käytettävyydestausta ulkoistetaan tavallisten käyttäjien tehtäväksi. Joukkoistamisessa joukkoistaja eli yritys, organisaatio tai yksilö laatii testitehtävät ja julkaisee ne joukkoistamisalustan kautta avoimena kutsuna, johon joukkoistamisalustan käyttäjät eli osallistajat vastaavat ilmoittamalla haluavansa suorittaa tehtävät. Tämän kirjallisuuskartoituksen tarkoitus on selvittää, tekeekö joukkoistaminen käytettävyydestaustuksesta halvempaa ja helpompaa sekä ovatko sillä saatavat tulokset laadukkaita.

Kirjallisuuskartoituksen perusteella on viitteitä siitä, että joukkoistamisen avulla voidaan saavuttaa useita hyötyjä. Joukkoistetun käytettävyydestin järjestämisessä on kuitenkin omat haasteensa, joiden huomioon ottaminen on tärkeää hyötyjen saavuttamiseksi. Käytettävyydestaustuksen kannalta joukkoistamisen tärkeimmät hyödyt ovat osallistujien helpompi rekrytointi ja sitä kautta suuremman tietomäärän kerääminen, ajan ja rahan säästö, testausten joustavuus sekä eri taustan omaavien ja muualla asuvien käyttäjien testausten helpottaminen. Hyötyjen saavuttamista haittaavat haasteet voidaan voittaa ottamalla riskit huomioon jo suunnitteluvaiheessa. Lisäksi niiden voittamisessa auttavat sopivien tavoitteiden ja mittarien valitseminen käytettävyydelle, tehtävien huolellinen suunnittelu, pilottitestin suorittaminen tehtäville, valintakokeen järjestäminen osallistujille, osallistujien motiivointi ja sitouttaminen, laadunvarmistuksen tekeminen, palautteen kerääminen mahdollisimman laajasti sekä osallistujien tukeminen. Joukkoistajan on tärkeää myös asennoitua joukkoistamiseen siten, että ei odota siltä ihmeitä. Joukkoistetussa käytettävyydestaustuksessa on vielä useita avoimia kysymyksiä, joiden selvittämiseksi tarvittaisiin lisätutkimuksia.

Avainsanat ja -sanonnat: käytettävyys, käytettävyyden arviointi, joukkoistaminen, crowdsourcing, joukkoistettu käytettävyydestausta

Sisällys

1	JOHDANTO	1
2	TUTKIMUSMENETELMÄ JA MATERIAALIT	3
2.1	TUTKIMUSMENETELMÄ.....	3
2.2	MATERIAALIT.....	3
	<i>Andor</i>	3
	<i>Google Scholar</i>	4
	<i>Löytyneiden joukkoistusmateriaalien luokittelu</i>	4
3	JOUKKOISTAMINEN	6
3.1	VAHVUUDET.....	7
3.1.1	<i>Resurssien säästö</i>	8
3.1.2	<i>Osallistujat ja asiakkaat</i>	9
3.1.3	<i>Tuote ja tuotekehitys</i>	10
3.2	HAASTEET	11
3.2.1	<i>Motivointi</i>	11
3.2.2	<i>Osallistujat</i>	12
3.2.3	<i>Tehtävät ja ratkaisut</i>	13
3.2.4	<i>Laatu</i>	14
3.2.5	<i>Joukkoistajan haasteet, lait ja tietoturva</i>	14
3.3	JOUKKOISTAMISEN PROSESSI.....	15
3.4	JOUKKOISTAMISTAPOJEN LUOKITTELU	16
3.5	JOUKKOISTAMISEN SOVELLUTUKSET.....	17
4	KÄYTETTÄVYYS	21
4.1	KÄYTETTÄVYYDEN ARVIOINTI.....	22
4.2	KÄYTETTÄVYYSTESTAUS	23
4.3	KÄYTETTÄVYYDEN ETÄARVIOINTI.....	24
	<i>Automaattinen lokianalyysi</i>	25
4.4	KÄYTETTÄVYYDEN MITTAAMINEN.....	26
4.4.1	<i>Vaikuttavuus</i>	26
4.4.2	<i>Tehokkuus</i>	27
4.4.3	<i>Tyytyväisyys</i>	28
4.4.4	<i>Muita mitattavia osa-alueita</i>	29
4.5	JOUKKOISTETUN KÄYTETTÄVYYSTESTAUKSEN EROT VERRATTUNA MUIHIN TESTAUSMENETELMIIN	29
5	JOUKKOISTETUN KÄYTETTÄVYYSTESTAUKSEN JÄRJESTÄMINEN	36
5.1	MOTIVOINTI.....	36

5.1.1	<i>Motiivit</i>	36
5.1.2	<i>Kannustimet</i>	37
5.1.3	<i>Mekanismit</i>	38
5.1.4	<i>Motivointikeinojen valinta</i>	39
5.2	TEHTÄVIEN LAATIMINEN	42
5.2.1	<i>Sopivuus joukkoistamiseen</i>	42
5.2.2	<i>Tehtävien rakenne</i>	43
5.2.3	<i>Tehtävien suunnittelu ja toteutus</i>	44
5.3	KÄYTETTÄVYYDEN MITTARIT JA PALAUTE.....	45
5.4	LAADUNVARMISTUS	49
5.4.1	<i>Laadunvarmistuksen keinot</i>	49
5.4.2	<i>Yhteenveto laadunvarmistuksesta</i>	52
5.5	OSALLISTUJAT JA REKRYTOINTI.....	52
5.6	JOUKKOISTAJAN ROOLI.....	53
5.7	JOUKKOISTAMISALUSTAN VALINTA	54
6	POHDINTA JA KOOTUT SUOSITUKSET	56
6.1	RAJOITTEET	56
6.2	LISÄTUTKIMUS	57
6.3	EHDOTETUT SUOSITUKSET JOUKKOISTETUN KÄYTETTÄVYYDESTIN JÄRJESTÄMISEEN	59
7	YHTEENVETO	61
	VIITELUETTELO	63
	LIITE 1: KÄYTTÄJIEN MOTIIVIT, KANNUSTIMET JA KANNUSTINMEKANISMIT [KATMADA ET AL., 2016]	67
	LIITE 2: JEFF HOWEN [2009] OHJEET JOUKKOISTAMISEEN	68

1 Johdanto

Käytettävyyttä voidaan testata monilla eri menetelmillä sekä käyttäjien kanssa, että ilman. Ilman käyttäjiä tehtävää käytettävyyssarviointia kutsutaan asiantuntija- tai heuristiseksi arvioinniksi ja niissä asiantuntija arvioi käyttöliittymän löytääkseen käytettävyyssongelmat. Käyttäjien kanssa arviointia voidaan tehdä esimerkiksi haastatteluiden ja kyselyiden avulla, sekä suorittamalla käytettävyystestejä käytettävyysslaboratoriossa tai etänä. Näillä kaikilla menetelmillä on omat kulunsa ja järjestelyitä vaativat ominaisuudet. Käytettävyyden arvioinnin toteuttamista pidetäänkin usein kalliina ja vaikeana. Tästä syystä käytettävyyss-testaus saatetaan jättää kokonaan tekemättä tai toteutetaan hyvin pienimuotoisesti. Kitturin ja muiden [2008] mukaan käyttäjien mukaan saamisesta aiheutuvat kulut vaativatkin usein tasapainottelua otoskoon, aikarajoitusten ja rahallisten kustannusten välillä. Bruun ja Stagen [2015] mukaan resurssirajoitukset ovat erittäin tärkeitä varsinkin pienille ja keskikokoisille ohjelmistokehitystä tekeville yrityksille, joilla ei tyypillisesti ole rahoitusta kokoaikaiseen konsultaatioon tai käytettävyyssasiantuntijan palkkaamiseen. Otoskoon jäädessä pieneksi resurssien vähyyden takia, ei välttämättä löydetä kaikkia käytettävyyssongelmia. Spool ja Schroeder [2001] löysivät tutkimuksessaan 5 ensimmäisellä osallistujalla vain 35 % käytettävyyssongelmista, ja perustelivat vähyyttä sillä, että 5 osallistujaa ei voi käydä läpi kaikkia mahdollisia reittejä päämäärään.

Yksi mahdollinen ratkaisu näihin ongelmiin on joukkoistaminen, jota suositellaan projekteihin, joissa ei ole tarpeeksi rahaa palkata tehtävään tarvittavia työntekijöitä tai toisia yrityksiä suorittamaan tehtävää [Thuan *et al.*, 2013]. Joukkoistamista on käytetty monenlaisiin tarkoituksiin, kuten esimerkiksi tutkimukseen ja kehitykseen, markkinointiin, suunnittelu- ja ideointialustoihin, freelance-toimintaan, avoimeen innovaatioon, ohjelmistojen kehittämiseen, tuoteideoihin, yhteistuotantoon ja julkisten palveluiden kehittämiseen [Hosain, 2012a]. Sen on myös havaittu olevan tehokas malli ongelmanratkaisuun [Zhao ja Zhu, 2012]. Voisiko käytettävyyden arviointi olla yksi tarkoitus lisättäväksi tähän listaan? Joukkoistamisessa yksilö, yritys tai organisaatio laatii tehtävät ja jakaa ne joukkoistamisalustan kautta tuntemattomalle joukolle ihmisiä, jotka suorittavat tehtävät pientä palkkiota vastaan. Eli siis työ, joka ennen on suoritettu joko yrityksen sisällä tai annettu asiantuntijalle / toiselle yritykselle tehtäväksi, annetaankin tehtäväksi tuntemattomille, yrityksen ulkopuolisille henkilöille, jotka suorittavat sen omalla ajallaan ja omalla laitteistollaan. Tällöin kulut pienenevät, kun ei tarvitse palkata asiantuntijoita eikä järjestää testitiloja ja aikoja. Mutta tekeekö joukkoistaminen käytettävyyden testauksesta oikeasti helpompaa ja halvempaa? Ja ovatko joukkoistamisella saatavat tulokset yhtä luotettavia ja käytettävyyden kannalta merkittäviä kuin perinteisellä, asiantuntijavetoisella käytettävyyss-testauksella saatavat.

Tämän pro gradun tarkoitus on selvittää kirjallisuuskatsaukseen perustuen, soveltuuko joukkoistaminen käytettävyyden arviointiin, tekeekö se käytettävyyden arvioinnista helpompaa ja halvempaa sekä mikä on sillä saatavien tulosten laatu. Tutkimuskysymykset ovat siis:

1. Soveltuuko joukkoistaminen käytettävyyden arviointiin?
2. Tekeekö joukkoistaminen käytettävyyden arvioinnista helpompaa ja halvempaa?
3. Ovatko joukkoistamisella saadut tulokset yhtä laadukkaita kuin perinteisillä arviointimenetelmillä saadut tulokset?

Luvussa 2 esittelen käyttämäni menetelmän eli kirjallisuuskartoituksen, sekä mistä ja millä perusteilla olen valinnut taustamateriaalit. Luvussa 3 kerron mitä joukkoistaminen on, mitkä ovat sen vahvuudet ja haasteet. Lisäksi kerron, millainen on joukkoistamisprosessi ja kuinka joukkoistamista voidaan luokitella sekä millaisia sovellutuksia on aiemmin toteutettu joukkoistamalla. Luvussa 4 kerron mitä tarkoitan käytettävyydellä tässä pro gradussa, miten ja millä tavalla käytettävyyttä arvioidaan sekä miten käytettävyyttä mitataan. Lisäksi vertailen joukkoistettua käytettävyydestä muihin testausmenetelmiin saadakseni selville sen erot muihin menetelmiin nähden. Luvussa 5 kerron joukkoistetun käytettävyydestin järjestämisestä eli asioista, joihin tulee kiinnittää huomiota testiä järjestettäessä. Tällaisia asioita ovat motivointi, tehtävien laatiminen, käytettävyyden mittarit ja palaute, laadun varmistus, osallistujat ja rekrytointi, joukkoistajan rooli sekä joukkoistamisalustan valinta. Luvussa 6 pohdin pro graduni rajoitteita ja lisätutkimusta tarvitsevia aiheeni osa-alueita sekä yhdistän aiempien lukujen informaation suosituksiksi joukkoistetun käytettävyydestin järjestämiseen. Lopuksi, luvussa 7 esitän yhteenvedon.

2 Tutkimusmenetelmä ja materiaalit

2.1 Tutkimusmenetelmä

Tutkimusmenetelmäni on kokoava kirjallisuuskatsaus eli käyn läpi aiemmin tehtyjä tutkimuksia joukkoistamisesta ja teen niiden pohjalta johtopäätöksiä joukkoistamisen sopivuudesta käytettävyyden arviointiin. Joukkoistaminen on monelle vielä tuntematon käsite. Se kuitenkin on ajankohtainen aihe, koska digitalisaation sanotaan poistavan aikaan, tilaan, tiedonsaantiin sekä osallistumiseen liittyviä rajoitteita [Koiranen et al., 2016]. Digitalisaatio voisi siis olla käännekohta, joka saa organisaatiot tutkimaan joukkoistamisen hyödyntämistä, ja tämän takia tuleekin selvittää laajasti, mitä joukkoistaminen tarkoittaa. Joukkoistamista itsessään on jo tutkittu jonkin verran, joten siitä löytyy suhteellisen hyvin materiaalia. Tietojenkäsittelyn, ihmisen ja teknologian vuorovaikutuksen tai käytettävyyden näkökulmista joukkoistamista ei ole vielä tutkittu paljoa, ja olisikin tärkeää tutkia sitä näistä näkökulmista, jotta tiedettäisiin mitä hyötyä siitä on saatavilla.

Kirjallisuuskatsauksen tekemisessä oli neljä vaihetta, joista ensimmäinen vaihe oli hakusanojen valinta ja aineiston haku. Toinen vaihe oli aineiston läpikäynti ja materiaalien valinta tarkempaan tarkasteluun. Kolmannessa, eli tarkasteluvaiheessa tarkastelen materiaaleja tarkemmin ja karsin pois gradun fokuksen ulkopuolelle jäävät materiaalit. Viimeisessä eli neljännessä vaiheessa integroin eri lähteiden antamat tiedot. Pro gradu -tutkielmani materiaalit haettiin pääsääntöisesti Andor ja Google Scholar -palveluista. Yksi lähde-materiaali löytyi Tampereen teknillisen yliopiston järjestämän User experience – design and evaluation -kurssin materiaaleista. Seuraavaksi esittelen käytetyt hakusanat, niillä saadut tulokset ja perusteet, joilla taustamateriaalit valittiin.

2.2 Materiaalit

Andor

Andor on Tampereen yliopiston kirjaston uusi hakuliittymä, jossa voi hakea yhtä aikaa sekä kirjaston kokoelmista, että suuresta monitieteisestä, satoja tietokantoja, kirjoja, miljoonia lehtiartikkeleita ja muita julkaisuja sisältävästä ProQuestin Summon -tietokannasta [TaY, 2016].

Tehdyissä hauissa yleisinä rajauksina olivat julkaisuvuosi välillä 2000–2016 ja saatavuutena ”Kokoteksti verkossa”. Ensimmäisessä haussa materiaalin jossain kentässä tuli olla joku seuraavista: ”crowdsourcing”, ”crowd sourcing” tai ”crowd wisdom” ja lisäksi, materiaalin jossain kentässä tuli olla termi ”usability”. Tämä haku tuotti 1915 tulosta, joita rajattiin edelleen valitsemalla materiaalit, joiden avainsanoissa on termit ”crowdsourcing” ja ”usability”. Tällöin jäljelle jäi 242 hakutulosta, joista luettiin otsikot ja tiivistelmät ja niiden perusteella tehtiin päätös, otetaanko materiaali tarkempaan tarkasteluun. Tarkasteluun valittiin materiaalit, joissa kerrottiin käytettävyydestä tai joukkoistamisen toteuttamisesta.

Pois karsittiin materiaalit, jotka käsittelivät esimerkiksi puhekäyttöliittymiä, ympäristöä tai geolokaatiota, mobiilijoukkoistamista tai pelillistämistä. Näillä rajauserusteilla jäljelle jäi kaksikymmentä materiaalia.

Toisessa haussa tarkoituksena oli kerätä lisää materiaalia joukkoistamisesta yleisesti sekä tietokoneella suoritettavasta joukkoistamisesta. Tämän takia hakuehtoja karsittiin jättämällä termi käytettävyys pois, eli materiaalin jossain kentässä tulee olla joku seuraavista: "crowdsourcing", "crowd sourcing" tai "crowd wisdom". Tämä haku tuotti 9353 hakutulosta, joita rajattiin edelleen valitsemalla materiaalit, joiden asiasanoihin on määritelty termi "crowdsourcing", "crowd-sourcing" tai "crowd sourcing". Tällöin jäljelle jäi 1040 hakutulosta, joita rajattiin edelleen valitsemalla tieteenalaksi "tietojenkäsittely" tai "informaatiotutkimus", jolloin jäljelle jäi 280 hakutulosta. Näistä karsittiin pois ne, jotka oli jo valittu mahdollisiin lähdeviitteisiin ensimmäisessä haussa. Nämä hakutulokset käytiin läpi lue-
malla materiaalin otsikko ja tiivistelmä, ja tekemällä niiden perusteella päätös siitä, ote-
taanko materiaali tarkempaan tarkasteluun. Päätökseen vaikutti materiaalin aihealue, eli
esimerkiksi opetukseen, pelikehitykseen, videoiden arviointiin ja historiaan liittyvät mate-
riaalit karsittiin pois, sekä käytettävät laitteet, eli mobiililaitteisiin ja niiden tiedonkeruuseen
(esim. geolokaatio) liittyvät materiaalit jätettiin pois. Näiden rajausten tarkoituksena on ra-
jata aineisto niin, että se tarjoaa yleistä tietoa joukkoistamisesta sekä tietoa tietokoneella tai
selaimella tehtävästä joukkoistamisesta. Näiden rajausten jälkeen jäljelle jäi 12 materiaalia.

Google Scholar

Haussa käytettiin hakulausekkeena "crowdsourcing" AND "usability", sekä aiempien hakujen tapaan aikarajauksena 2000–2016. Näillä hakuehdoilla saatiin 9000 hakutulosta. Hakutuloksia ei ollut mahdollista rajata pidemmälle esimerkiksi avainsanojen avulla, mutta ne järjestettiin relevanttiuden perusteella ja mukaan otettiin kolmekymmentä ensimmäistä hakutulossivua. Näiden sivujen hakutuloksista karsittiin pois ne, jotka olivat jo mahdollisten lähdeviitteiden joukossa aiempien hakujen tuloksena sekä ne materiaalit, joiden aihepiiri oli fokuksen ulkopuolella. Tällöin jäljelle jäi 17 materiaalia, joista vielä neljä karsiutui pois, koska ne eivät joko olleet saatavilla yliopiston kautta tai olivat liian yleistasoisia eli eivät tarjonneet tarvittavaa tietoa.

Löytyneiden joukkoistusmateriaalien luokittelu

Kaikki 45 materiaalia luettiin läpi ja niistä osa karsiutui pois, koska ne eivät tarjonneet aiheen kannalta relevanttia tietoa. Lisäksi joitakin materiaaleja tuli lisää, kun joihinkin luku-
ihin tarvittiin täydentävää tietoa. Pro graduni lopullisena taustamateriaalina on 41 mate-
riaalia. Taulukossa 1 on esitetty tärkeimmät lähteet luokiteltuina sen mukaan, käsittelevätkö
ne pelkästään joukkoistamista vai sekä joukkoistamista että käytettävyyttä.

Joukkoistamiseen liittyvät lähteet	Joukkoistamiseen ja käytettävyyteen liittyvät lähteet
Moktar Hossain: Users' motivation to participate in online crowdsourcing platforms	Mortem Hertzum <i>et al.</i> : What do think-aloud participants say? a comparison of moderated and unmoderated usability session
Allan Afuah: Crowdsourcing and open innovation	Bill Albert <i>et al.</i> : Beyond the Usability Lab: Conducting Large-scale Online User Experience Studies
Moktar Hossain ja Ilkka Kauranen: Crowdsourcing: a comprehensive literature review	Brian Fitzgerald ja Klaus-Jan Stol: The dos and don'ts of crowdsourcing software development
Yuziang Zhao ja Qinghua Zhu: Evaluation on crowdsourcing research: current status and future direction	Victor H. M. Gomide <i>et al.</i> : Affective crowdsourcing applied to usability testing
Aikaterini Katmada <i>et al.</i> : Incentive mechanisms for crowdsourcing platforms	Di Liu <i>et al.</i> : Crowdsourcing for usability testing
Gareth J.F. Jones: An introduction to crowdsourcing for language and multimedia technology research	
Nguyen Hoang Thuan <i>et al.</i> : Factors influencing the decision to crowdsource	
Jeff Howe: Crowdsourcing	

Taulukko 1: Tärkeimmät lähteet

3 Joukkoistaminen

Joukkoistamisen käsitteen esitteli ensimmäisen kerran Jeff Howe vuonna 2006 [Howe, 2009], vaikka ilmiö itsessään on tunnettu jo vuosisatojen ajan. Eräs ensimmäisiä esimerkkejä on Yhdistyneen kuningaskunnan hallituksen vuonna 1714 järjestämä kilpailu ”The Longitude Prize”, jonka voittaja sai 20 000 puntaa palkinnoksi [Hossain ja Kauranen, 2015]. Kilpailuun saivat osallistua kaikki ja sen tarkoituksena oli löytää ratkaisu pituuspiirien laskemiseksi luotettavasti [Hossain ja Kauranen, 2015]. Romanovsin ja muiden [2014] mukaan englanninkielinen termi ”crowdsourcing” eli *joukkoistaminen* johdetaan sanoista ”crowd” eli *joukko* ja ”outsourcing” eli *ulkoistaminen*. Howe [2009] määritteli joukkoistamisen toimena, jossa yritys tai instituutio ulkoistaa aiemmin työntekijöiden suorittaman toimen ennalta määrittämättömälle joukolle avoimen kutsun kautta.

Vaikka joukkoistaminen sisältääkin tehtävien ulkoistamista, se ei kuitenkaan ole sama asia kuin *ulkoistaminen* (outsourcing). Zhaon ja Zhun [2012] mukaan joukkoistamisen ja ulkoistamisen suurin ero on sopimuksen laatiminen. Ulkoistamisessa yritys etsii palveluntarjoajan ja määrittää tarpeensa, jonka jälkeen palveluntarjoaja tarjoaa sopimuksen mukaisesti palveluitaan tai tuotteitaan. Joukkoistamisessa yritys esittää avoimen kutsun ja joukon yksilöt tarjoavat työsuoritteitaan vapaaehtoisesti, eli osapuolien välille ei laadita sopimusta.

Ulkoistamisen lisäksi toinen läheinen käsite on *avoin lähdekoodi* (open source). Joukkoistamisella ja avoimen lähdekoodin kehittämisellä on yhteisiä piirteitä: molemmissa internet mahdollistaa toiminnan ja molemmissa yksilöt voivat vapaasti osallistua tietyn tehtävän tekemiseen [Župič, 2013]. Joukkoistamisen sanotaankin olevan keino soveltaa avoimen lähdekoodin periaatteita muihin toimialoihin. Tämä ei kuitenkaan ole täysin paikkansa pitävä väite, koska samankaltaisuuksien lisäksi niillä on tärkeitä eroavaisuuksia. Ensinnäkin joukkoistaminen ei ole niin avointa toimintaa kuin avoimen lähdekoodin toiminta [Zhao ja Zhu, 2012]. Tällä tarkoitetaan sitä, että joukkoistamisella kerättyjen vastausten ja ratkaisujen immateriaalioikeudet ovat joukkoistavalla organisaatiolla. Toiseksi avoimen lähdekoodin toiminnassa ongelman paikallistaminen ja siihen ratkaisun kehittäminen on itsessään riittävä palkkio, kun taas joukkoistamisessa osallistujat usein palkitaan rahallisesti [Zhao ja Zhu, 2012] tai rahattomilla palkkioilla (esimerkiksi pisteet tai kunnia). Joukkoistamisessa osallistujat tuottavat vastauksia tai ratkaisuja joko itsenäisesti tai yhteistyössä, kun avoimen lähdekoodin toiminnassa joukot toimivat usein yhdessä luodakseen jotain ja tuotosten välillä on tärkeä riippuvuus (tai siis ratkaisun osat ovat riippuvaisia toisistaan) [Zhao ja Zhu, 2012]. Lisäksi joukkoistamisessa kilpailu on usein normi, kun taas avoimen lähdekoodin kehitystyössä yhteistyö on pääasia [Župič, 2013].

Howen [2009] määritelmä on toiminut pohjana muillekin määrittelyille. Simulan [2012] mukaan joukkoistamisen neljä perusulottuvuutta ovat osallistujien esivalinta, vertaiskonttribuutioiden saatavuus, osallistujien yhdistäminen sekä palkkio kontribuutioista. Thuan ja

muut [2013] päätyivät hyvin samankaltaisiin tuloksiin ja määrittelivät joukkoistamisen kahdeksan pääpiirrettä seuraavasti: selkeästi määritelty joukko, selkeällä päämäärällä oleva tehtävä, selkeä korvaus joukolle, tunnistettu joukkoistaja, määritelty kompensatio joukkoistajalle, verkossa oleva prosessi, avoin kutsu ja Internetin käyttö. Nämä peruspiirteet toistuvat määritelmässä usein. Simulan [2012] mukaan joukkoistaminen on osallistavan, verkossa tehtävän aktiviteetin muoto, jossa yksilö, instituutio, organisaatio tai yritys ehdottaa avoimella kutsulla yksilöiden joukolle mahdollisuutta suorittaa tehtäviä vapaaehtoisesti ja kokea tyytyväisyyttä työpanoksestaan.

Myös Hammon ja Hippner [2012] esittivät määritelmänä joukkoistamisen olevan alun perin yrityksen sisällä tai ulkoiselle bisneskumppanille ulkoistettujen tehtävien ulkoistamista määrittämättömän laajalle, heterogeeniselle, potentiaalisista ratkaisijoista koostuvalle massalle avoimella kutsulla. Thuanin ja muiden [2013] mukaan joukkoistaminen on verkkostrategia, jossa organisaatio ehdottaa tehtäviä joukon jäsenille avoimella kutsulla ja jossa osallistujat panostavat työtä, osaamista, taitoja ja/tai kokemusta korvausta vastaan. Tämän jälkeen yritys vastaanottaa osallistujien panostukset ja käyttää ratkaisuja määriteltyihin päämääriin.

Edellä mainittujen lisäksi löytyy hieman erilaisia määrittelyjä: Jonesin [2012] mukaan joukkoistaminen on ihmisen tekemän työn (human computing) eräs muoto. Ihmisen tekemä työ määritellään työnä, joka normaalisti mielletään tietokoneen automaattisesti suorittamaksi, mutta sen suorittaakin ihminen.

Tässä pro gradu -tutkielmassa joukkoistamisella tarkoitetaan internetin välityksellä tapahtuvaa toimintaa, jossa yritys tai organisaatio ulkoistaa tehtäviä ulkopuoliselle joukolle. Joukkoistamista suorittavasta organisaatiosta, yrityksestä tai yksilöstä käytän nimitystä *joukkoistaja* (voidaan kutsua myös osallistajaksi) ja joukkoistamiseen käytettävästä palvelusta nimitystä *joukkoistamisalusta*. Joukkoistamisalustan kautta tehtäviä suorittavista yksilöistä käytän nimitystä *osallistuja*. Seuraavissa kohdissa kerron joukkoistamisen vahvuuksista ja haasteista. Kohdassa 3.3 kerron mistä vaiheista joukkoistamisprosessi koostuu ja mitä näissä vaiheissa tapahtuu. Kohdassa 3.4 kerron, miten joukkoistamista voidaan luokitella, eli mitä eri joukkoistamisen lajeja on olemassa, ja kohdassa 3.5 kerron esimerkkejä siitä mihin kaikkeen joukkoistamista on käytetty.

3.1 Vahvuudet

Taustamateriaalien perusteella joukkoistamisella on useita vahvuuksia. Vahvuudet on luokiteltu kolmeen alalukuun sen mukaan, mihin tai keneen vahvuus osoittaa. Alakohdassa 3.1.1 kerron resurssien säästöön liittyvistä vahvuuksista ja alakohdassa 3.1.2 kerron joukkoistamiseen osallistuvia käyttäjiä sekä kehitettävien tuotteiden loppukäyttäjiä eli asiakkaita koskevia vahvuuksia. Alakohdassa 3.1.3 kerron joukkoistamalla kehitetyn tuotteen ja joukkoistettuun tuotekehitykseen liittyvistä vahvuuksista.

3.1.1 Resurssien säästö

Thuanin ja muiden [2013] mukaan kulujen vähentäminen on yksi pääsyy joukkoistamisen valintaan. Tätä väitettä tukee myös se, että useassa taustamateriaalissa mainittiin vahvuutena erilaisissa kuluissa säästäminen, kulujen pieneminen tai kustannustehokkuus [Zhao ja Zhu, 2012; Hossain ja Kauranen, 2015; Župič, 2013; Afuah, 2011; Thuan *et al.*, 2013]. Joukkoistaminen muun muassa pienentää markkinointikuluja [Hossain, 2012a], tuotantokuluja [Hossain, 2012a], kehityskuluja [Fitzgerald ja Stol, 2015; Thuan *et al.*, 2013], koulutuskuluja [Gomide *et al.*, 2014] sekä palkkakuluja [Simula, 2012].

Lisäksi joukkoistaminen auttaa muuttamaan kiinteitä kuluja vaihtuviksi kuluiksi [Hossain, 2012a]. Tällä tarkoitetaan sitä, että aiemmin testaukselle on laskettu kiinteät kulut (tietyllä summalla voidaan ottaa tietty määrä osallistujia käytettävyysslaboratoriossa tehtävään testaukseen), mutta joukkoistamisen myötä kulut vaihtelevat tilanteen mukaan: joskus tarvitaan enemmän osallistujia ja joskus vähemmän. Joukkoistaminen tarjoaakin pääsyn laajaan *käyttäjävarantoon* (user pool), joka suorittaa tehtävät nopeammin ja alhaisemmillä kustannuksilla [Hammon ja Hippner, 2012; Jones, 2012]. Nämä alhaiset kustannukset mahdollistavat sen, että tarvittaessa osallistujien määrää pystytään lisäämään ilman suuria lisäkuluja [Gomide *et al.*, 2014].

Tuotanto-, kehitys- ja palkkakuluihin joukkoistaminen vaikuttaa jo toimintaperiaatteellaan. Joukkoistaminen mahdollistaa sen, että työ delegoidaan isolle määrälle ihmisiä [Simula, 2012; Hossain ja Kauranen, 2015] ja tästä joukosta voidaan kerätä osaamista, informaatiota, taitoja ja ratkaisuja [Thuan *et al.*, 2013]. Tällä tarkoitetaan sitä, että joukkoistamiseen osallistuvilla käyttäjillä saattaa olla joukkoistamista tekevää yritystä korkeampi *tietämys* (knowledge) aiheesta, eli osaaminen on syvempää ja laajempaa [Afuah, 2011]. Tämä on yleistä tilanteissa, joissa yritys alkaa kehittää tuotetta heille ennalta tuntemattomalle toimialalle. Joukkoistamisen kautta yritys voi siis vähentää koulutuskuluja, kun avuksi otetaan internetin käyttäjien tietämys [Hammon ja Hippner, 2012] eli oman työväen kouluttamisen sijasta etsitään asiantuntijuus joukkoistamisen kautta. Župič [2013] kutsui tätä myös *kollektiivisen viisauden* (collective intelligence) valjastamiseksi. Joukkoistaminen on siis väline helppoon ja kustannustehokkaaseen tiedonhakuun, koska se mahdollistaa satojen käyttäjien rekrytoinnin interaktiivisiin tehtäviin marginaalisilla kustannuksilla nopealla varoitusaajalla [Kittur *et al.*, 2008]. Lisäksi joukkoistaminen on ehtymätön resurssi ongelmien ratkaisemiseen [Romanovs *et al.*, 2014], ja potentiaalisten ratkaisijoiden määrä kasvaa koko ajan [Afuah, 2011].

Joukkoistaminen auttaa tehostamaan ja lisäämään tuotantoa [Župič, 2013], koska se mahdollistaa nopean tehtävien suorittamisen nopealla aikataululla [Jones, 2012]. Se myös tehostaa toimintaa säästämällä aikaa ja rahaa, jos ratkaisu onkin jo olemassa [Afuah, 2011]. Tällä tarkoitetaan sitä, että yrityksellä on jokin tuotetta koskeva ongelma, jota he lähtevät

itse ratkaisemaan ja käyttävät siihen paljon aikaa ja rahaa. Samaan aikaan joukossa olisi kuitenkin osallistuja tai osallistujat, joilla olisi suoraan tarjota ratkaisu kyseiseen ongelmaan, jolloin ratkaisua ei tarvitsisi keksiä uudelleen. Tämä vaikuttaa suoraan palkkakuluihin, koska tällöin yrityksen ei tarvitse hallita sisäistä ongelmanratkaisuprosessiaan joukon ratkaistessa ongelmia [Afuah, 2011]. Näiden lisäksi joukkoistaminen auttaa tavoittelemaan tehokkuutta tehtävissä, joissa ihmisen syöte on tärkeä [Kazai *et al.*, 2013].

Joukon potentiaalinen voitokas käyttö voi vahvistaa yrityksen kilpailuasemaa [Hammon ja Hippner, 2012]. Tällä tarkoitetaan sitä, että joukkoistamisen kautta yritys saa tuotteensa paremmin potentiaalisten asiakkaiden näkyville, eli kun yritys kehittää tuotteitaan joukkoistamalla, se samalla markkinoi tuotteitaan. Tämä lisää joustavuutta hinnoitteluun [Afuah, 2011], koska osallistujat tietävät enemmän tuotteesta ja tietävät millaista vastinetta rahoillensa saavat tuotteen ostaessaan. Tällöin tuotteen hinnoittelua ei tarvitse enää miettiä sen mukaan, mihin hintaan asiakkaat tuotteen ostaisivat, vaan mikä olisi tuotteen hyödyllisyyttä vastaava hinta.

Yhteenvetona voidaan todeta joukkoistamisen säästävän resursseja vähentämällä erilaisia kustannuksia. Lisäksi se säästää aikaa, kun yrityksen ei tarvitse hallita ongelmanratkaisuprosessia ja osaaminen haetaan sieltä, mistä se löytyy valmiiksi. Myös kehittämiseen tarvittava aika vähenee, koska ongelmanratkaisu ei vie työntekijöiden aikaa, ja osallistujilla voi olla ratkaisut jo valmiina.

3.1.2 Osallistujat ja asiakkaat

Joukkoistaminen tarjoaa osallistujille ennennäkemättömiä uusia mahdollisuuksia jakaa tietämystään ja havaintojaan [Gomide *et al.*, 2014]. Se tarjoaa osallistujille mahdollisuuden saada työkokemusta sekä *näkyvyyttä* (exposure) työmarkkinoilla [Zhao ja Zhu, 2012]. Lisäksi joukkoistaminen tarjoaa osallistujille mahdollisuuden hyödyntää taitojaan, tutkia sekä muuttaa harrastuksensa *hyödylliseksi* (beneficial) [Zhao ja Zhu, 2012]. Tällä tarkoitetaan sitä, että osallistuja voi joukkoistamisalustan kautta hyödyntää ja syventää oppimiansa asioita, kokeilla miltä työskentely tuntuu ja saada siitä vielä pienen korvauksenkin. Joukkoistaminen poistaa erot tutkinnon suorittaneiden ammattilaisten ja ilman muodollista koulutusta olevien itseoppineiden amatöörien väliltä [Rockwell, 2012]. Joukkoistamiseen pystyy osallistumaan kuka tahansa eli osaaminen on pääasia, ei se, miten osaamisen on hankkinut. Rockwellin [2012] mukaan joukkoistaminen tarjoaa mahdollisuuden osallistua isoja joukkoja eri kulttuureista ja käyttäjäryhmistä ilman suuria kuluja. Edellä mainittujen syiden vuoksi joukkoistamisen sanotaankin yhdistävän työn ja työntekijän [Simula, 2012] ja siksi se voi olla hyvä rekrytointityökalu yrityksille [Afuah, 2011]. Lisäksi joukkoistaminen mahdollistaa sen, että yhteistyötä tehdessä osallistujat voivat neuvotella vastauksensa niin, että

heidän äänensä ei katoa yhteenvedossa [Rockwell, 2012]. Tällä tarkoitetaan sitä, että yhteenvedossa on kaikkien ajatukset ja mielipiteet mukana sen sijasta, että joku muu koostaisi yhteenvedon ja poimisi vain osan ajatuksista ja mielipiteistä.

Joukkoistamisesta on hyötyä myös asiakkaille, koska se mahdollistaa ison asiakasmäärän helpon tavoittamisen sekä helpottaa tiedonjakoa [Hossain, 2012a; Župič, 2013]. Lisäksi se mahdollistaa toimintastrategian viestimisen sidosryhmille [Afuah, 2011]. Tällä tarkoitetaan sitä, että jos kilpailija ilmoittaa tuovansa jonkin tuotteen markkinoille, voi yritys vastata siihen tuomalla testattavaksi oman tuotteensa samankaltaisilla ominaisuuksilla. Tällöin yritys tuo julki sen, millä strategialla se vastaa kilpailuun. Muun ohella joukkoistaminen elvyttää käsitystä asiakkaan tärkeydestä suunnitteluprosessissa [Simula, 2012]. Jos joukkoistamisen kohteena on esimerkiksi kunnan tarjoamien palveluiden kehittäminen, helpottaa joukkoistaminen avointa dialogia kansalaisten ja päättäjien välillä [Hossain ja Kauranen, 2015]. Kansalaisilla on siis mahdollisuus saada äänensä kuuluville ja päättäjät kuulevat, mitä kansalaiset oikeasti haluavat. Tiedonjaon parantuessa asiakkaat saavat paremmin tietoa yrityksestä ja sen tuotteista. Tämä voi tuoda yritykselle enemmän uskollisia asiakkaita [Hossain, 2012a].

Yhteenvetona voidaan todeta joukkoistamisen tarjoavan osallistujille mahdollisuuksia jakaa tietämystään, saada työkokemusta, näkyvyyttä työmarkkinoilla ja jopa työpaikkoja. Joukkoistamisen ansiosta nämä mahdollisuudet ovat melkein kaikkien saavutettavissa. Se myös parantaa tiedonjakoa yritysten ja asiakkaiden välillä. Tämä tarkoittaa sitä, että asiakkaat tietävät paremmin mikä yritys on ja mitkä sen tuotteet ovat, ja yritys tietää mitä asiakkaat haluavat.

3.1.3 Tuote ja tuotekehitys

Joukkoistamisella on positiivinen vaikutus joukkoistamisprosessin lopputuloksena saataviin ratkaisuihin ja tuloksiin tai tuotteisiin. Joukkoistamisen avulla kehitetyt tuotekonseptit ovat innovatiivisempia [Hossain, 2012a; Fitzgerald ja Stol, 2015] ja laadukkaampia [Fitzgerald ja Stol, 2015; Afuah, 2011; Hammon ja Hippner, 2012], parempia sekä sisällöltään rikkaampia [Zhao ja Zhu, 2012]. Usein ratkaisut ovat myös *luovia* (creative) [Fitzgerald ja Stol, 2015], tehokkaita [Hammon ja Hippner, 2012], *uusia* (novelty) ja tarjoavat enemmän *hyötyä asiakkaille* (customer benefit) [Simula, 2012]. Tällöin tuotteet usein vastaavat paremmin käyttäjän tarpeisiin ja asiakkaat ovat halukkaampia ostamaan [Afuah, 2011]. Lisäksi joukkoistaminen auttaa yrityksiä löytämään tyhjiä markkinapaikkoja ja tuomaan tuotteita markkinoille paremmalla ajoituksella [Afuah, 2011]. Tämä tarkoittaa sitä, että yritys saa osallistujilta tietoa siitä, minkälaisia tuotteita he haluaisivat tai tarvitsisivat. Tällöin yrityksen on helppo löytää tuotteita, joita ei vielä ole markkinoilla, ja yritys pystyy auttamaan asiakkaitaan tuomalla markkinoille kyseiseen tarpeeseen vastaavaan tuotteen silloin, kun

sellaiselle on kysyntää. Kun asiakkaiden tarpeet ja mieltymykset ovat selvillä, ovat uusien tuotteiden tuomisen riskit myöskin pienemmät [Hossain, 2012a].

Tuotekehityksen näkökulmasta joukkoistaminen nopeuttaa toimintaa esimerkiksi lyhentämällä innovaatio sykliä [Hossain, 2012a] ja tarjoamalla nopeamman pääsyn markkinoille [Fitzgerald ja Stol, 2015]. Tällä tarkoitetaan sitä, että osallistujien keksiessä ideat tai ratkaisut ei yrityksen tarvitse käyttää aikaa innovaatio sykliin. Kehityksen ollessa nopeampaa myös pääsy markkinoille on nopeampaa. Markkinoille pääsyä nopeuttaa se, että niin sanottu kriittinen massa testaa ja kehittää tuotetta [Fitzgerald ja Stol, 2015], jolloin se on asiakkaiden näkökulmasta jo hyväksytty markkinoille tullessaan. Kriittinen massa tarkoittaa tuotteesta tai tuoteryhmästä kiinnostuneita asiakkaita, jotka osallistuvat joukkoistamiseen nähdäkseen minkälaisia tuotteita olisi tulossa ja haluaa vaikuttaa niihin. Kun tämä kriittinen massa hyväksyy tuotteen, on se todennäköisesti silloin käyttäjäryhmän halujen ja vaatimusten mukainen.

Markkinoiden kannalta yksi vahvuus on joustava *toimituksellinen toimintatapa* (editorial approach) [Alonso ja Baeza-Yates, 2011], mikä tekee joukkoistamisesta hyvän keinon uusien, nopeasti muuttuvien ideoiden testaamiseen. Tällä tarkoitetaan sitä, että osallistujilta saatua ideaa pystytään muuttamaan nopeasti ja testaamaan uudelleen, kun uudistettu idea julkaistaan joukolle arvioitavaksi. Joukkoistamisen kautta saadaan helposti ratkaisuja sellaisiin ongelmiin, joita vain joukko osaa ratkaista [Afuah, 2011]. Tuotekehityksen kannalta eräs joukkoistamisen vahvuus on isojen projektien kehittäminen suurella tarkkuudella [Hossain ja Kauranen, 2015]. Tällä tarkoitetaan sitä, että joukkoistamisen avulla on mahdollista saada laaja osallistujajoukko, jossa on edustettuna eri kulttuureja ja tämän ihmisjoukon vastausten tai ratkaisujen hallinta on helpompaa.

Yhteenvetona voidaan todeta, että joukkoistamisella saadaan luovia, uusia ja laadukkaampia ratkaisuja, jotka vastaavat asiakkaiden tarpeita paremmin. Lisäksi se auttaa yrityksiä löytämään uusia markkinapaikkoja. Joukkoistaminen auttaa tuotteiden markkinoille pääsyä tekemällä kehityksestä nopeampaa ja pienentämällä markkinoille tuloon liittyviä riskejä.

3.2 Haasteet

Joukkoistamiseen liittyy useita haasteita, joista osan voittaminen on erittäin tärkeää joukkoistamisen onnistumisen kannalta. Seuraavassa esittelen taustamateriaaleissa mainittuja haasteita, jotka on jaettu viiteen alalukuun aihepiirin mukaan.

3.2.1 Motivointi

Joukkoistamisprojekti epäonnistuu täysin, jos kukaan vapaaehtoisista ei osallistu [Rockwell, 2012]. Tämän takia osallistujien motivointi onkin kriittinen osa joukkoistamista, ja on

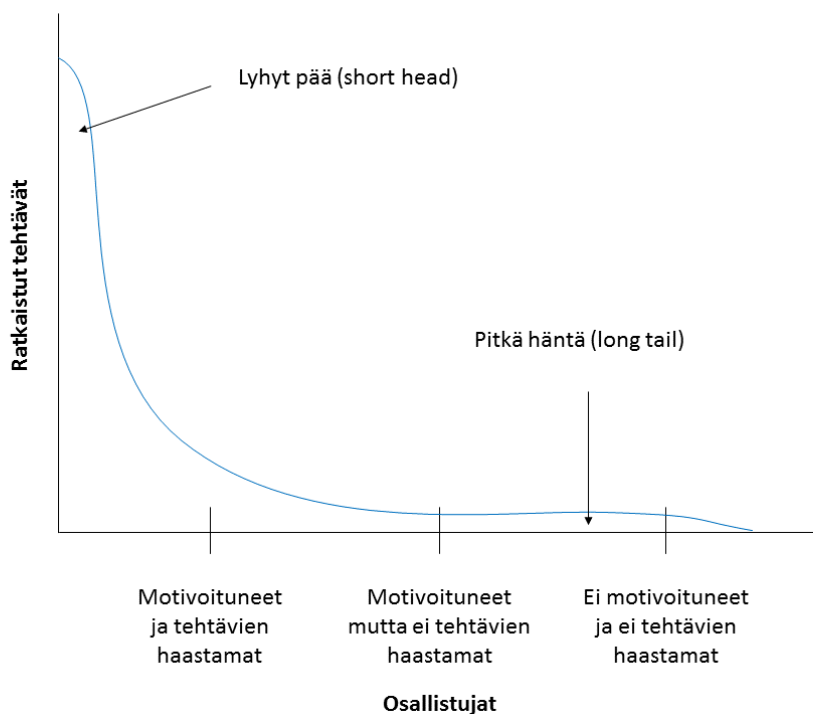
tärkeää tietää mikä motivoi osallistujia ja kuinka suunnitella toiminta motivoivaksi [Rockwell, 2012]. Motivointi täytyy aloittaa jo siinä vaiheessa, kun mietitään, miten mahdolliset osallistujat saavat tiedon mahdollisuudesta osallistua [Simula, 2012], ja kuinka osallistujat rekrytoidaan [Bruun ja Stage, 2015]. Tämän jälkeen motivoinnin täytyy jatkua, jotta saadaan varmistettua se, että osallistujat osallistuvat eli suorittavat tehtäviä ja myös pysyvät mukana toiminnassa [Simula, 2012; Bruun ja Stage, 2015]. Tällöin jokaiseen työvaiheeseen ei tarvitse rekrytoida osallistujia uudelleen.

Motivaatioon vaikuttaa lisäksi yrityksen kyky luoda luottamus osallistujien ja yrityksen välille [Zhao ja Zhu, 2012] eli jos yritys ei vaikuta luotettavalle, voi sillä olla suuri vaikutus motivaatioon. Motivointi ei kuitenkaan ole helppoa, sillä erilaisia osallistujia motivoi erilaiset asiat. Tämän takia *kannustinmekanismien* (incentives) valinta onkin haaste [Zhao ja Zhu, 2012] ja niiden valinta tulee tehdä harkiten.

3.2.2 Osallistujat

Yksi suuri haaste on joukon monimuotoisuus ja luonteen tuntemattomuus [Kittur *et al.*, 2008]. Tällä tarkoitetaan sitä, että vaikka joukossa voi olla osallistujia eri kulttuureista ja eri käyttäjäryhmistä, voi osallistujaryhmä silti muodostua vain yhden kulttuurin/käyttäjärühmän jäsenistä. Tällöin saadut vastaukset eivät edusta kaikkia käyttäjiä. Sekään ei ole varmaa, että saatavilla olisi juuri aihepiirin kannalta relevantti joukko [Hossain ja Kauranen, 2015]. Tällöin tuotetta/palvelua eivät testaa henkilöt, jotka sitä todennäköisesti tulisivat käyttämään, ja tämä voi ohjata kehitystä väärään suuntaan.

Osallistujiin liittyvä haaste on myös niin kutsuttu "long tail" -efekti, jolla tarkoitetaan sitä, että pieni joukko osallistujia tekee suurimman osan suorituksista [Rockwell, 2012; Afuah, 2011]. Kuvassa 1 on esitetty uudelleen piirrettynä ja suomennettuna Afuahia [2011] mukaillen ratkaistujen tehtävien jakautuminen "long tail" -efektissä. Kuvan 1 pystyakselilla esitetään ratkaistujen tehtävien määrä ja vaak akselilla esitetään erilaiset osallistujat. Kuvassa vasemmalla ovat ensimmäisenä ryhmänä yksilöt, jotka ovat motivoituneita ja tekevät haastavat heitä. Keskellä ovat yksilöt, jotka ovat motivoituneita, mutta tekevät eivät haasta heitä. Oikealla ovat yksilöt, jotka eivät ole motivoituneita eikä tekevät myöskään haasta heitä. Kuvassa 1 näkyy selkeästi se, että motivoituneet ja tehtävien haastamat ratkovat eniten tehtäviä (*lyhyt pää* eli "short head"-ilmiö). Niin kutsutun *pitkän hännän* (long tail) muodostavat vastaukset osallistujilta, jotka ovat motivoituneita, mutta ei tehtävien haastamia tai eivät ole motivoituneita eikä tehtävien haastamia. Tällöin vastaukset kertovat pääasiassa sen, mitä mieltä lyhyen pään vastaajat ovat arvioitavasta tuotteesta tai palvelusta, jolloin muiden osallistujien vastaukset "hukkuvat" joukkoon ja lopullinen ratkaisu ei välttämättä vastaa heidän tarpeitaan. Yksi onnistuneen joukkoistamisen piirteitä onkin se, että myös pitkän hännän osallistujat onnistutaan rekrytoimaan mukaan ja pyritään sitä kautta saamaan parempia ratkaisuja [Rockwell, 2012; Afuah, 2011].



Kuva 1: Ratkaistujen tehtävien jakautuminen long tail -efektin mukaisesti

3.2.3 Tehtävät ja ratkaisut

Yksi joukkoistamisen suurimmista haasteista on sen sopivuus suoritettaviin tehtäviin. Joukkoistaminen ei välttämättä sovellu tehtäviin, jotka vaativat niin sanottua *hiljaista tietämystä* (tacit knowledge) [Afuah, 2011]. Tällä tarkoitetaan tietämystä, joka karttuu tekemällä ja kokemuksen kautta eli niin sanottua käytännön kokemusta. Tämän takia joukkoistaminen ei välttämättä ole sopiva pitkiin ja monimutkaisiin tehtäviin, kuten esimerkiksi lentokoneen rakentamiseen [Afuah, 2011]. Myös Kittur ja muut [2008] ovat samoilla linjoilla ja toteavat joukkoistamisen toimivan parhaiten tehtäviin, joihin on *aito* (bona fide) vastaus, jolloin osallistuja ei pysty huijaamaan järjestelmää *järjettömällä* (non-sense) vastauksilla. Tällaisia tehtäviä ei kuitenkaan pystytä laatimaan kaikista kerättävistä tiedoista (esimerkiksi arvostelut tai mielipiteet), jolloin tahallisten, väärin vastausten poissulkeminen on hankalaa [Kittur *et al.*, 2008]. Päinvastaisesti, Bruun ja Stage [2015] sanovat joukkoistamisen sopivan mihin tahansa *monimutkaiseenkin* (non-trivial) tehtävään, kun tehtävien taso ei ole kognitiivisesti liian vaativia.

Vaikka tehtävät olisivatkin joukkoistamiseen sopivat, eivät haasteet kuitenkaan lopu vielä siihen. Omat haasteensa tuo vastausaika [Hossain ja Kauranen, 2015], osallistujilta tulevien ratkaisujen/ehdotusten hallinta [Zhao ja Zhu, 2012], tulosten yhdistäminen [Bruun ja Stage, 2015] ja ratkaisun liittäminen organisaatioon ja/tai tuotteisiin ja palveluihin [Afuah, 2011]. Viimeisellä kohdalla tarkoitetaan sitä, että jos yrityksellä ei ole tietämystä aiheesta, he eivät välttämättä osaa lisätä ratkaisua nykyisiin tuotteisiin siten, että se *lisäisi tuotteen arvoa*

(value creation) asiakkaille ja *haltuun ottaisi* (capturing activities) ratkaisun tehostamaan toimintaansa [Afuah, 2011]. Tällöin yritys ei osaa hyödyntää kunnolla saamaansa ratkaisua vaan se on vain yksi osa järjestelmää tai organisaatiota eikä siitä saada mahdollista lisähyötyä.

3.2.4 Laatu

Toinen suuri haaste on laatu eli kuinka varmistaa vastausten laadullisuus [Zhao ja Zhu, 2012; Hossain ja Kauranen, 2015; Simula, 2012]. Laatua voi heikentää osallistujien antama epätarkka tai virheellinen tieto [Hossain ja Kauranen, 2015], tai vastaukset eivät vain ole hyviä (esimerkiksi eivät auta ratkaisemaan ongelmaa) [Afuah, 2011].

Laatuun vaikuttaa myös osallistujien yritykset huijata järjestelmää. Osallistujat voivat yrittää huijata parantamalla tuloksiaan (esimerkiksi äänestämällä ideaa useamman kerran) tai vitsinä äänestämällä huono/epäedullinen ratkaisu voittajaksi [Simula, 2012]. Lisäksi huijausta voidaan tehdä järjettömillä vastauksilla: vastaukset ovat joko hölynpölyä tai niiden vastausarvo on olematon (esimerkiksi pelkkä ok vastauksena) [Kittur *et al.*, 2008]. Tällä tavoin osallistujia yrittää vain vastata nopeasti jotain saadakseen tehtävät suoritettua ja saadakseen niistä mahdollisen palkkion. Huijausta voi myöskin tehdä kilpailija, joka voi ohjata yrityksen ongelmiin osallistumalla joukossa ja tarjoamalla tahallisesti ideana toimimattoman ratkaisun [Afuah, 2011].

Ratkaisuna huijaamiseen käytetään usein osallistujien ja tulosten/ratkaisujen arviointia [Bruun ja Stage, 2015]. Tämä tuo kuitenkin mukanaan omat haasteensa, koska arviointikriteerien (esimerkiksi osallistujien rajaaminen tai automaattinen/manuaalinen suodatus) valinta ei ole helppoa.

3.2.5 Joukkoistajan haasteet, lait ja tietoturva

Joukkoistamiseen tuovat omat haasteensa immateriaalioikeudet, tietoturva ja lait, ja joukkoistamisen käyttöönotto lisää näihin liittyviä riskejä [Hossain ja Kauranen, 2015]. Joukkoistettaessa tulee ottaa huomioon työläinsäädäntö, keksijän oikeus, tietoturva, tekijänoikeudet sekä henkilötietojen hallinta [Simula, 2012]. Lisäksi, joukkoistajan täytyy muistaa suojella omia immateriaalioikeuksiaan olemalla tarjoamatta liikaa tietoa strategioistaan, jotta tämä tieto ei päädy kilpailijoille [Afuah, 2011].

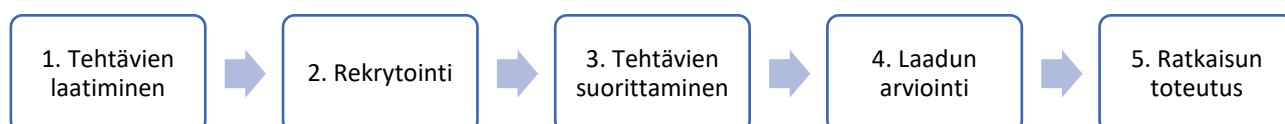
Joukkoistamista tekevä yritys voi tuntea menettäneensä tuotekehityksen hallinnan [Zhao ja Zhu, 2012]. Samalla kuitenkin osallistujien suuri määrä luo lisää tarvetta hallinnalle [Hossain ja Kauranen, 2015]. Tällöin yrityksestä voi vaikuttaa siltä, että sillä ei ole hallintaa siihen millainen lopputulos on, mutta samanaikaisesti yrityksen pitää hallita osallistujia ja ohjata heitä. Tämän ohjauksen kautta yritys pystyy kuitenkin vaikuttamaan lopputulokseen ja ohjaamaan joukkoa niin sanotusti oikeaan suuntaan, jotta lopputulos on toivottu.

Myös vastuu lopputuloksesta on haasteellinen kysymys. Joukkoistamisessa lopputuloksen alhaisesta laadusta tai epäonnistumisesta voidaan syyttää joukkoa, joka sen laati [Simula, 2012]. Tällä tarkoitetaan sitä, että periaatteessa joukkoistajalla ei ole mitään hävittävää, kun lopputulos voidaan aina jättää julkaisematta. Tämä ei kuitenkaan ole mielestäni hyvä tapa toimia, koska joukkoistajan tehtävä on hallita ja ohjeistaa joukkoa, jotta lopputulos olisi onnistunut. Lisäksi haastetta tuo se, että jotkut voivat nähdä joukkoistamisen työvoiman hyväksikäyttönä [Simula, 2012], jolloin he eivät pidä sitä hyväksyttävänä toimintatapana.

Myös tulosten ja ulkoisen tietämyksen käyttö ja lisääminen yrityksen toimintaan tuo omat haasteensa. Organisaatio voi olla haluton liittämään ulkoisia ideoita tai ratkaisuja toimintaansa, koska niitä ei ole keksitty yrityksen sisällä ("not invented here" -dilemma) [Simula, 2012] tai yritys ei yksinkertaisesti osaa arvostaa, tuoda ja käyttää ulkoista tietämystä [Afuah, 2011]. Tällöin tietämys jää hyödyntämättä ja ratkaisu voi olla laadultaan heikompi.

3.3 Joukkoistamisen prosessi

Seuraavassa esitetään joukkoistamisen prosessi perustuen Thuanin ja muiden [2013], Jonesin [2012], Zhaon ja Zhun [2012] ja Afuahin [2011] esittämiin jaotteluihin. Joukkoistamisprosessissa on yhteensä viisi vaihetta, joista ensimmäinen vaihe on tehtävien laatiminen. Tässä vaiheessa joukkoistaja tunnistaa testattavat toiminnot tai ratkaistavat ongelmat ja jakaa ne pieniin tehtäviin. Ensimmäisessä vaiheessa myöskin päätetään tehtävien suorittamisesta saatava palkkio. Toisessa vaiheessa nämä tehtävät julkaistaan käytettävän joukkoistamisalustan kautta potentiaalisille ratkaisijoille, jotka sitten päättävät haluavatko osallistua ja suorittaa tehtävät. Tässä vaiheessa siis rekrytoidaan osallistujat. Kolmannessa vaiheessa hyväksytyt tai kaikki mukaan ilmoittautuneet osallistujat suorittavat tehtävät ja palauttavat ne joukkoistamisalustan kautta. Neljännessä vaiheessa joukkoistaja arvioi palautettujen ratkaisujen laadun, ja laadun ollessa kunnossa, hyväksyy vastauksen ja antaa palkkion suorituksesta. Muutoin ratkaisu hylätään ja osallistuja ei saa palkkiota. Viidennessä eli viimeisessä vaiheessa joukkoistaja vastaanottaa kaikista ratkaisuista koostetun lopullisen ratkaisun ja toteuttaa sen. Kuvassa 2 on esitetty yllämainitut vaiheet kaaviona.



Kuva 2: Joukkoistamisprosessin vaiheet

3.4 Joukkoistamistapojen luokittelu

Thuan ja muut [2013] laativat taustamateriaaliensa perusteella joukkoistamistapojen luokittelun, joka on esitetty taulukossa 2 suomeksi alkuperäistä mukailleen. Taulukon 2 sarakkeet esittävät *osallistumisen muotoa* (participation mode) ja rivit tehtävien *monimutkaisuutta* (complexity) [Thuan *et al.*, 2013]. Osallistuminen voi tapahtua *yksilöllisesti* (individual) tai *kilpailullisesti* (competitive). Yksilönä osallistuminen voi olla *yhdistävää* (integrative) eli vaikka tehtävät tehdään yksin, koskevat kaikki vastaukset samaa ongelmaa ja vastauksista yhdistetään yksi ratkaisu. Tätä kutsutaan myös *yhteistyöperustaiseksi* (collaboration-based) joukkoistamiseksi [Afuah, 2011]. Kilpailullisessa osallistumisessa kaikista ratkaisuksista valitaan voittaja. Tätä tapaa kutsutaan myös *turnausperustaiseksi* (tournament-based) joukkoistamiseksi [Afuah, 2011].

Osallistumisen muoto	Yksilöllinen (yhdistävä)	Kilpailullinen (valikoiva)
Monimutkaisuus		
Yksinkertainen	Kaupalliset joukkoistamisalustat <ul style="list-style-type: none"> • AMT 	Yksinkertaiset kilpailut <ul style="list-style-type: none"> • Yahoo Answers • Amazonin Ask Ville
Taitoa vaativa	Kollektiivinen äly <ul style="list-style-type: none"> • Wikipedian artikkeleiden kirjoittaminen 	Ongelmien ratkaisukilpailut <ul style="list-style-type: none"> • NextStopDesign • Innocentive • Threadless • iStockPhoto

Taulukko 2: Joukkoistamisen jaottelu

Thuanin ja muiden [2013] mukaan joukkoistamisen tehtävät voidaan luokitella *helppoksi* (simple), *monimutkaisiksi* (complex) tai *luoviksi* (creative). Helpot tehtävät ovat sellaisia, että niiden ratkaisemiseen riittää perustaidot, kun taas monimutkaisten tehtävien ratkomiseen tarvitaan asiantuntijuutta ja ongelmanratkaisutaitoja. Monimutkaisten tehtävien ratkaisuun tarvitaan usein myös luovuutta, ja monesti luovien ja monimutkaisten tehtävien erot ovat hyvin pienet. Tämän takia ne on yhdistetty taulukossa 2 riviotsikolle *taitoa vaativa* (skilled). Helppoja, yksilönä suoritettavia tehtäviä ovat kaupallisilla joukkoistamisalustoilla (esimerkiksi AMT tai mTurk eli Amazon Mechanical Turk -palvelu) suoritettavat lyhyet tehtävät. Esimerkiksi AMT on markkinapaikka, joka auttaa yrityksiä etsimään ihmisiä suorittamaan tehtäviä, joihin tietokoneet eivät välttämättä pysty (esimerkiksi asioiden tunnistaminen kuvista, dokumenttien luokittelu, tuotekuvausten kirjoittaminen tai podcastien puhtaaksikirjoitus). Yksinkertaisia kilpailullisia tehtäviä ovat esimerkiksi erilaiset kyselyalustat, joihin

käyttäjät voivat lähettää kysymyksiään. Yleensä nopein tai parhaimman vastauksen antanut saa pisteitä, jotka nostattavat hänen *asemaansa* (reputation tai vastaava) sivustolla. Taitoa vaativia, yksilöllisiä tehtäviä ovat kollektiivista viisautta vaativat tehtävät, kuten esimerkiksi Wikipedian artikkeleiden kirjoittaminen. Taitoa vaativia, kilpailullisia tehtäviä ovat erilaiset *ongelmien ratkaisukilpailut* (problem solving contest). Esimerkiksi Innocentive on ongelmien ratkaisukilpailuja järjestävä palvelu, jonka kautta ihmisille tarjotaan ratkaistavaksi reaali maailman haasteita, ja niiden ratkaisusta saa rahallisen palkinnon. NextStopDesign oli puolestaan joukkoliikenteeseen liittyvä ratkaisukilpailu, jossa haettiin ratkaisuja joukkoliikenteen pysäkkien parantamiseen. Threadless on palvelu, johon osallistujat voivat lähettää vaatetukseen painatettavia luonnoksia, joista muut palvelun käyttäjät ja osallistujat voivat äänestää parhaimmat. Voittaneet ratkaisut päätyvät palvelun kauppaan myyntiin. iStockPhoto taas on kuvapankki, johon kuvaajat julkaisevat ottamiensa kuvia ja niistä parhaimpia myydään palvelun kautta.

Taulukossa 2 esitetyistä tyypeistä pro graduni keskittyy joukkoistamisen tyyppeihin, jotka ovat yksilöllisiä ja yksinkertaisia.

3.5 Joukkoistamisen sovellutukset

Joukkoistamista voidaan käyttää erilaisiin tarkoituksiin. Tässä kohdassa esittelen näitä tarkoituksia esittämällä Hossainin ja Kaurasen [2015] sekä Romanovsin ja muiden [2014] antamia esimerkkejä joukkoistamisen luokista taulukossa 3. Luokat on koottu taulukon ensimmäiseen sarakkeeseen. Lisäksi esitän jokaiselle luokalle Hossainin [2012a] laatiman joukkoistamisalustojen kategorisoinnin mukaisen kategorian, osa luokista voi kuulua useampaankin kategoriaan. Kategoriat ovat *luovuus* (creativity), *tietämys* (knowledge), *työnteko* (labour) ja *sisältö* (content). Katteoria kertoo, minkälaisiin tehtäviin joukkoistamisalustaa voidaan käyttää. Kategorian lisäksi annan jokaiselle luokalle esimerkkejä joukkoistamisalustoista, tai siitä minkälaisiin projekteihin sitä on hyödynnetty. Luokkia kuvataan tarkemmin taulukon 3 jälkeen. Esimerkit on kerätty Hossainin [2012a], Liun ja muiden [2012], Hossainin ja Kaurasen [2015] sekä Hammonin ja Hippnerin [2012] materiaaleista.

Luokka	Kategoria	Esimerkit
<i>Ideoiden luonti</i> (idea creation)	Luovuus, tietämys, sisältö	Dell Ideastorm, Threadless, Lego DesignByMe, Wilogo.com, iStockphoto, Spreadshirt, P & G connect and develop
<i>Mikrotaskaus</i> (microtasking)	Työnteko	Clickworker, Amazon Mechanical Turk (mTurk), CrowdFlower, uTest, Testbirds, Atizo

<i>Avoimen lähdekoodin ohjelmistot</i> (open source software, OSS)	Tietämys, sisältö	Linux, Mozilla Firefox, Thunderbird, OpenOffice, Ubuntu
<i>Julkinen osallistuminen</i> (public participation)	Tietämys	Google Earth, julkiset suunnittelu projektit (public planning projects)
<i>Kansalaistiede</i> (citizen science)	Tietämys, luovuus	Googlen reCAPTCHA -projekti, Innocentive
<i>Kansalaisjournalismi</i> (citizen journalism)	Tietämys, sisältö	Digg, Newsvine
<i>Wikiti</i> (wikis)	Tietämys, sisältö	Wikipedia, OpenStreetMap, Geo-Wiki.org, Wikimapia, WikiTerra, GLOBE-ohjelma

Taulukko 3: Joukkoistamisen sovellutukset luokkineen, kategorioineen ja esimerkkeineen

Seuraavassa esittelen eri luokat mukailen Hossainin ja Kaurasen [2015] keräämää tietoa heidän kirjallisuuskartoituksensa taustamateriaaleista. Ideoiden luonti voi esiintyä kahdessa eri muodossa: ideakilpailuna tai ideointina kollektiivisella tietämyksellä. Ideakilpailussa organisaatio tai yksilö kutsuu joukon lähettämään uusia ideoita, joista parhaimmat valitaan. Ideointi kollektiivisella tietämyksellä tarkoittaa sitä, että joukon jäsenet yhdistävät älynsä ja luovat yhdessä ideoita. Annetuista esimerkeistä Wilogo.com on lähinnä ideakilpailu, koska sitä kautta yritys voi ilmoittaa haluavansa uuden logon ja osallistujat laativat logoehdotukset, joista yritys valitsee parhaan. Esimerkeistä Spreadshirt ja Threadless ovat taas kollektiivisella älyllä ideointialustoja, koska osallistujien suunnitelmassa uusia paitamalleja äänestävät muut osallistujat ideoita ja parhaimmat mallit pääsevät myyntiin. Tällöin parhaat ideat valitaan niin sanotulla kollektiivisellä päätöksellä.

Mikrotaskauksella tarkoitetaan järjestelmää, jossa käyttäjät voivat valita ja suorittaa pieniä tehtäviä rahallista tai rahatonta korvausta vastaan [Kittur *et al.*, 2008]. Mikrotaskausalusta koostuu yleensä kolmesta toimijasta: työnantaja (eli joukkoistaja), ylläpitäjä sekä *mikrotyöntekijät* (microworker) eli osallistujat. Työnantajan vastuulla on tehtävien luominen ja rahoituksen hoitaminen. Alustan ylläpitäjä tarjoaa tarvittavan infrastruktuurin ja työskentelee välikätenä työnantajan ja mikrotyöntekijöiden välillä. Mikrotyöntekijät puolestaan suorittavat *mikrotehtäviä* (microtask) ja saavat siitä rahallisen korvauksen. Taulukossa 3 mikrotaskauksen luokasta esimerkkeinä mainitut alustat toimivat kaikki tällä periaatteella.

Avoimen lähdekoodin ohjelmistojen kehityksessä vapaaehtoiset koodaavat ja kehittävät ohjelmistoja kollektiivisesti yhteisen hyvän vuoksi. Joukkoistamisen myötä ohjelmiston laatu voi parantua, kun joukko arvioi ohjelmistoa jatkuvasti. Suurin ero miten avoimen lähdekoodin ohjelmistojen kehitys on muuttunut joukkoistamisen myötä, on se, että on siirrytty yhteisön yhteisen hyvän vuoksi jakamasta koodista siihen, että osallistujia on paljon,

mutta edunsaajia vain muutama. Tällä tarkoitetaan sitä, että ennen vapaaehtoiset kehittivät ohjelmia saadaksensa käyttöönsä parempia ohjelmistoja, mutta nykyään kehittäminen on enemmän ”kaupallista”. Osallistujat kehittävät ohjelmaa saadakseen siitä korvauksen ja yritys hyötyy saamalla parempilaatuisia ratkaisuja ohjelmistoonsa.

Joukkoistamista käytetään julkisissa suunnitteluprosesseissa varmistamassa julkista osallistumista. Tällä halutaan varmistaa avoin keskustelu kansalaisten ja päätöksentekijöiden välillä. Lisäksi se auttaa kehittämään isojakin projekteja suurella tarkkuudella. Esimerkki tästä on Google Earth, mihin käyttäjät pystyvät merkitsemään mielenkiintoisia paikkoja ja ilmoittamaan väärin sijainteihin merkittyjä paikkoja, jolloin tiedon tarkkuus on parempi.

Kansalaistiede on yhteistyöperustaista tutkimusta, jossa joukko valjastetaan ratkomaan reaali maailman ongelmia. Kansalaistieteen avulla voidaan esimerkiksi kerätä ja prosessoida tietoa tieteelliseen kyselyyn, ja vapaaehtoiset yleensä osallistuvat siihen mielenkiinnon, uteliaisuuden tai velvollisuudentunnon vuoksi. Esimerkki tästä on Google reCAPTCHA -projekti, jossa osallistujat testaavat reCAPTCHA:n toimivuutta auttaakseen kehittämään sitä.

Kansalaisjournalismi on vaihtoehtoismedia, jossa kansalaisjournalistit kilpailevat toisiansa vastaan ammattilaisten ja akateemikkojen kanssa. Kansalaisjournalismin joukkoistamisalustat kuten Digg.com ja Newsvine tarjoavat käyttäjille mahdollisuuden lähettää, arvostella, suositella ja kommentoida uutisia. Kansalaisjournalismi perustuu osallistujien monimuotoisuuteen, mikä voi myös parantaa uutisten luotettavuutta.

Taulukon viimeinen luokka eli wikit tarkoittavat verkkosivuja, joiden sisältöjä käyttäjät pystyvät lisäämään, poistamaan ja muokkaamaan. Käyttäjät pystyvät myös linkittämään muita sivuja ja resursseja sisältöön. Wikit toteutetaan yleensä tekniikoilla, jotka mahdollistavat yhteistyön, ja tämä seikka tekee wikien toiminnasta joukkoistamista. Tunnetuin esimerkki tällaisesta wikistä on Wikipedia, mutta samalla periaatteella voidaan kerätä myös esimerkiksi maantieteellistä dataa. Yksi tällainen esimerkki on OpenStreetMap, jonka karttatiedot perustuvat käyttäjien lähettämään dataan.

Taulukossa esitettyjen luokkien lisäksi on olemassa myös muita sovellutuksia. Tällaisia ovat esimerkiksi niin kutsuttu havainnointi ja joukkorahoitus. *Havainnoinnissa* (sensing) ihmiset osallistuvat keräämällä dataa esimerkiksi mobiililaitteella tai muilla sensoreilla ja lisäävät siihen omia havaintojaan [Romanovs *et al.*, 2014]. Havainnointia käytetään esimerkiksi säätietojen keräämisessä, jolloin erilaisista säätutkista ja sensoreista saadaan dataa ja lisäksi käyttäjät lähettävät omia säähavaintojaan. *Joukkorahoituksessa* (crowdfunding) yksilö tai yritys esittelee tuotteensa tai palvelunsa ja pyytää joukolta rahoitusta sen toteuttamiseen. Pitäessään ideasta joukon jäsenet osallistuvat yleensä pienellä summalla ja näistä pienistä summista koostuu idean/tuotteen rahoitus eli joukko rahoittaa toteutuksen. Tällainen joukkorahoitus alusta on esimerkiksi Kickstarter.com [Hammon ja Hippner, 2012], jossa yksilöt

hakevat joukkorahoitusta luoville projekteilleen (esimerkiksi musiikki, kuvataide, peli, elokuva tai tuotesuunnittelu).

Näistä luokista pro graduni keskittyy mikrotaskaus-luokkaan eli keskityn alustoihin, jotka mahdollistavat osallistujien työskentelyn käytettävyydestin osallistujina.

4 Käytettävyys

Käytettävyydelle löytyy useita määritelmiä ja monesti käytettävyyden kohdalla pitää paikkansa sanonta, että määritelmiä on yhtä monta kuin määrittelijöitä. ISO-9241-11 -standardi [ISO, 1998] määrittelee käytettävyyden seuraavasti:

“ Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.”

Seuraavassa esitän standardin osa-alueet suomennettuna alkuperäiseen standardiin [ISO, 1998] pohjautuen. *Tarkkuudella* (effectiveness) tarkoitetaan käyttäjän päämäärien saavuttamisen *tarkkuutta* (accuracy) ja *täydellisyyttä* (completeness). Tämä tarkoittaa sitä, että käyttäjä pystyy suorittamaan palvelussa sen, mitä hän tuli sinne tekemään. *Tehokkuus* (efficiency) puolestaan tarkoittaa resurssien suhdetta päämäärien saavuttamisen tarkkuuteen ja täydellisyyteen. Tämä tarkoittaa, että jos käyttäjän tavoite on löytää tietoa nopeasti, täytyykö tämä tavoite vai meneekö kyseisen tiedon löytämiseen kauan. Viimeisenä, *tyytyväisyys* (satisfaction) tarkoittaa vapautta *epämukavuudesta* (discomfort) ja positiivista asennetta tuotteen käyttöä kohtaan. Käytännössä tämä tarkoittaa mielestäni sitä, että tavoitteen saavuttamiseen tarvittavat laitteet ja järjestelmät/palvelut ja vuorovaikutus toimivat sujuvasti. Tällöin tuotteen käyttäminen on mukavaa ja käyttäjä saavuttaa tavoitteensa resurssien rajoissa ja on tyytyväinen vuorovaikutusprosessiin.

Toinen vakiintunut, paljon käytetty määritelmä on Nielsenin [1993] laatima viisi osaluetta sisältävä määritelmä. Nielsenin [1993] mukaan käytettävyys koostuu *opittavuudesta* (learnability), *tehokkuudesta*, *muistettavuudesta* (memorability), *virheiden määrästä* (errors) sekä *tyytyväisyydestä*. Näistä opittavuus tarkoittaa sitä, että kuinka nopeasti käyttäjä suoriutuu perustason tehtävästä saavuttuaan ensimmäistä kertaa palveluun, eli yksinkertaistettuna kuinka nopeasti käyttäjä oppii käyttämään palvelua. Tehokkuus puolestaan tarkoittaa kuinka nopeasti käyttäjä suoriutuu tehtävistä. Muistettavuus tarkoittaa sitä, että käyttäjän palatessa palveluun pitkän ajan päästä, kuinka nopeasti hän muistaa miten palvelu toimii. Virheiden määrä keskittyy siihen, kuinka monta virhettä käyttäjä tekee tehtäviä suorittaessa, kuinka vakavia nämä virheet ovat ja kuinka helposti hän toipuu niistä, eli pystyy jatkamaan palvelun käyttöä. Tyytyväisyys kertoo, miten miellyttävää palvelun käyttäminen oli.

ISO-standardin ja Nielsenin [1993] määritelmissä on paljon samankaltaisuutta. Taulukossa 4 on esitettyä ISO-standardin osa-alueet ja niitä vastaavat Nielsenin määrittämät osa-alueet.

ISO-standardi	Nielsenin osa-alueet	Huomion kohteet
Tarkkuus	Opittavuus, muistettavuus	Tarkkuus ja suorituksen täydellisyys
Tehokkuus	Tehokkuus, muistettavuus, virheiden määrä	Resurssien käyttö tavoitteen saavuttamiseksi
Tyytyväisyys	Tyytyväisyys	Palvelun toimivuus ja käytön mukavuus

Taulukko 4: ISO-standardin osa-alueet ja niitä vastaavat Nielsenin osa-alueet

Opittavuus ja muistettavuus vastaavat mielestäni tarkkuutta, koska jos palvelun käyttö on vaikeaa oppia sekä muistaa, vaikuttaa se suorituksen tarkkuuteen ja täydellisyteen. Tällöin käytössä voi tulla virheitä tai suorituksen vaiheita voi jäädä pois, jolloin suoritus ei ole täydellinen ja tavoitteeseen ei välttämättä päästä.

Molemmissa määritelmässä mainittiin tehokkuus. Mielestäni myös muistettavuus ja virheiden määrä kuuluvat samaan kategoriaan tehokkuuden kanssa, koska niilläkin on suuri vaikutus resurssien käyttöön. Jos palvelun käyttö on vaikeaa muistaa tai käytön aikana esiintyy runsaasti virheitä, kuluu silloin paljon resursseja kuten esimerkiksi aikaa. Tällöin resurssien käyttö suhteutettuna tavoitteen saavuttamiseen ei ole tasapainossa. Viimeinen osa-alue on tyytyväisyys, joka mainittiin molemmissa määritelmässä, ja molemmissa painotettiin käytön toimivuutta ja miellyttävyyttä.

Tässä pro gradu -tutkielmassa käytettävyydellä tarkoitetaan ISO-standardin määritelmää ja osa-alueita täydennettynä Nielsenin [1993] määrittämällä osa-alueilla taulukon 4 mukaisesti. Nielsenin [1993] osa-alueita käytetään selittämässä ja täydentämässä ISO-standardin määritelmää. Pro gradun tarkoituksena on tutkia, miten näitä osa-alueita voitaisiin arvioida joukkoistetulla käytettävyydestä.

4.1 Käytettävyyden arviointi

Ovaskan ja muiden [2005] mukaan käytettävyydestutkimukseen käytettävät menetelmät voidaan jakaa kolmeen pääluokkaan: suunnittelu, määrittely ja arviointi. Näiden luokkien välillä ei ole selviä rajoituksia vaan jotkin menetelmät voivat kuulua useampaan luokkaan. Menetelmä tarkoittaa työvaiheittain tai tavoitetuloksina kuvattua tehtäväjoukkoa, jonka taustalla voi olla jokin erityinen teoria, malli tai viitekehys. Menetelmiä voidaan käyttää esimerkiksi aineiston keruuseen ja/tai analysointiin sekä käyttöliittymän suunnitteluun liittyvien konkreettisten ehdotusten tuottamiseen. Käytettävyyden arviointiin on olemassa monia menetelmiä, jotka valitaan usein tutkimuksen aiheen ja tarvittavan tiedon perusteella. Koska pro graduni aihepiiri on käytettävyyden arviointi, keskityn Ovaskan ja muiden [2005] määrittämään pääluokkaan arviointi. Johdatus erilaisiin arviointimenetelmiin löytyy läh-

teistä Ovaska ja muut [2005] sekä Nielsen [1993]. Arviointi-pääloukka jaetaan yleensä kahden alaluokkaan sen mukaan osallistuvatko käyttäjät toimintaan. Alaluokat ovat *tarkistusmenetelmät* (inspection methods) ja *testausmenetelmät* (user testing) [Ovaska *et al.*, 2005]. Esimerkiksi asiantuntija-arviointi kuuluu tarkistusmenetelmiin ja käytettävyydestä testausmenetelmiin.

4.2 Käytettävyydestä

Käytettävyydestä on menetelmä, jossa pyritään aitoja käyttötilanteita muistuttavien tehtävien avulla selvittämään mahdollisimman objektiivisesti, kuinka kohderyhmään kuuluvat käyttäjät käyttävät tuotetta [Koskinen, 2005]. Käytännössä tämä tarkoittaa sitä, että osallistujiksi värvätään kohderyhmään kuuluvia henkilöitä, jotka suorittavat ennalta määritellyjä testitehtäviä. Tehtävät laaditaan siten, että ne vastaavat tehtäviä, joita kohderyhmän henkilöt tuotteella oikeastikin tekisivät. Esimerkiksi tuotteen ollessa verkkosivu tehtävät voivat koskea tiedonhakua, eli löytääkö käyttäjä tarvitsemansa tiedon. Käytettävyydestä sopii menetelmäksi tuotteen käytettävyyden arviointiin, kun halutaan saada paljon tietoa tuotteen käytettävyydestä sen todellisen käyttäjän näkökulmasta. Tämän takia käytettävyydestä tavoitteena on löytää tuotteen käytöstä ongelmakohtia, jotta jatkokehityksessä voidaan kiinnittää huomiota juuri näiden tiettyjen asioiden korjaamiseen ja parantamiseen. Toisaalta testi tarjoaa tietoa myös siitä, mitkä tuotteen ominaisuudet ovat jo käytettävyydeltään hyvällä tasolla.

Käytettävyydestä toteutetaan yleensä siten, että kerrallaan paikalla on yksi testikäyttäjä ja yhdestä kolmeen tarkkailijaa, joista osa voi olla käyttäjälle näkymättömissä teknisenä tarkkailijana tai taustahavainnoitsijana [Koskinen, 2005]. Kuitenkin aina yksi tarkkailijoista on testin valvoja eli moderaattori, joka valvoo ja johtaa testitilannetta. Joskus käytettävyydestä voidaan toteuttaa myös ns. pari- tai ryhmätestauksena, jolloin tuotetta käyttää samanaikaisesti useampi testikäyttäjä. Tietoa tuotteen käytettävyydestä saadaan analysoimalla testeistä kerättyä aineistoa, kuten muistiinpanoja ja videokuvaa käyttäjien toiminnasta sekä nauhoitteesta kirjoitettua lokia.

Käytettävyydestä voidaan käyttää lisäksi erilaisia tiedonkeruutapoja, kuten esimerkiksi äänenajattelua ja katseenseurantaa, joiden avulla kerättyä tietoa voidaan käyttää apuna testauksen tulosten tulkinnassa. Äänenajattelu on tiedonkeruutekniikka, jossa käyttäjää pyydetään ajattelemaan ääneen tehtäviä tehdessä [Ilves, 2005]. Sen avulla saadaan tietoa testattavan tuotteen ongelmakohtista sekä käyttäjän mentaalimalleista eli käsityksistä, joilla käyttäjät selittävät tuotteen rakennetta tai toimintatapaa itselleen. Ongelmakohtien lisäksi tekniikalla saadaan myös selville syitä ongelmakohtien taustalla. Käytännössä tämä tarkoittaa sitä, että saadaan tietoa esimerkiksi siitä, miten käyttäjä odottaa tuotteen toimivan ja miten käyttäjä reagoi, kun tuote ei toimikaan kyseisellä tavalla. Monet ihmiset kuitenkin

kokevat ääneenajattelun epämiellyttäväksi, mikä voi vaikuttaa myös tuloksiin. Lisäksi ääneenajattelu vaikeutuu, kun osallistujan kognitiivinen kuormitus kasvaa, joten tekniikka ei sovellu monimutkaisten ja vaikeakäyttöisten tuotteiden testaukseen.

Toinen, usein käytettävyydestä lisänä käytettävä, tiedonkeruutapa on katseen seuranta. Katseenseuranta on menetelmä, jossa käyttäjän silmänliikkeitä tarkastelemalla saadaan tietoa käyttäjän tarkkaavaisuuden suuntautumisesta ja kiinnittymisestä käyttöliittymässä [Lehtinen, 2005]. Katseenseurantaa voidaan käyttää sekä käyttöliittymien suunnittelussa, että tutkimuksessa. Sen soveltamista kuitenkin rajoittaa vaatimus erityisistä laitteista ja tyypillisesti katseenseurantatutkimus suoritetaan käytettävyyslaboratoriossa. Menetelmän vahvuuksiin kuuluu muun muassa käyttäjän tiedostamattomista kognitiivisista prosesseista saatava tieto; esimerkiksi tehtävän kognitiivinen kuormittavuus saattaa heijastua silmänliikkeisiin. Katseenseuranta voi vahvistaa muilla menetelmillä saatuja tuloksia, koska sillä saadaan määrällistä ja pätevää tietoa osallistujan käyttäytymisestä. Ongelmallista menetelmässä on ollut laitteiden kömpelyys ja kalleus, sekä datan analysoimisen hitaus ja metodologiset ongelmat. Lisäksi heikkoutena on osallistujan liikkeiden rajallisuus ja tutkimustilanteen vastaavuus todelliseen käyttötilanteeseen.

4.3 Käytettävyyden etäarviointi

Cartan ja muiden [2011] mukaan etäarvioinnin avulla pystytään analysoimaan käyttäjiä heidän päivittäisessä ympäristössään ja samalla pienentämään arvioinnin järjestämisen kuluja. Lisäksi sen avulla voidaan kerätä laajalti tietoa tuotteen käytöstä ja käyttötilanteista, koska käyttäjät suorittavat tehtävät erilaisilla laitteilla, käyttöjärjestelmillä ja selaimilla. Kuten käytettävyyden arviointiin, myös sen etäarviointiin on tarjolla monia menetelmiä, joilla kerätään tietoa käyttäjiltä. Cartan ja muiden [2011] mukaan menetelmät voidaan jaotella sen mukaan, tapahtuuko tiedon tallentaminen palvelimella vai käyttäjän tietokoneella. Tällä valinnalla on vaikutus siihen, minkälaista tietoa saadaan tallennettua.

Myös käytettävyydesti voidaan suorittaa etänä, jolloin testiin osallistuva käyttäjä ja testin moderaattori eivät ole samassa tilassa vaan käyttäjä suorittaa testitehtävät esimerkiksi kotonaan tai töissä omalla ajallaan [Bastien, 2010]. Etänä tehtävällä käytettävyydestestissä (tästä eteenpäin etätesti) tarkoitan tässä pro gradussa käytettävyydestestejä, jotka tehdään käyttäjän normaalissa käyttöympäristössä joko ilman moderaattoria tai moderaattorin valvomana. Etätesti voidaan suorittaa joko *synkronisena* (synchronous) tai *asynkronisena* (asynchronous) [Bastien, 2010]. Synkroninen etätesti tarkoittaa sitä, että siihen osallistuvat yhtäaikaaisesti sekä käyttäjä että moderaattori jonkin internetin välityksellä käytettävän *yhteiskäyttöisen* (collaborative) työkalun avulla. Bastien [2010] määrittelee synkronisen lähestymistavan siten, että testin *tekninen asiantuntija* (facilitator) ja *arvioijat* (evaluators) keräävät dataa ja hallitsevat arviointisessiota reaaliajassa etänä osallistuvan osallistujan kanssa. Se voidaan ajatella eräänlaisena moderoituna videokonferenssina, jossa resurssien salliessa

voidaan käyttää myös erilaisia virtuaalitekouhuoneita tai muita virtuaalituloja. Teknologian pääpainona on se, että arvioija näkee, mitä käyttäjän näytöllä tapahtuu [Bastien, 2010].

Asynkronisessa etätetstissä ei ole moderaattoria lainkaan, vaan käyttäjä tekee testitehtävät omaan tahtiinsa ja yleensä jokin elektroninen työkalu ohjaa testiä sekä kerää tarvittavan tiedon (logit, klikkaukset, tehtävään käytetyt ajat yms.). Tällöin *valvojilla* (observers) ei ole pääsyä tietoon reaaliaikaisesti, ja testin aikana ei ole välikättä, joka olisi vuorovaikutuksessa käyttäjän kanssa [Bastien, 2010]. Tällä tavoin toteutettu etätetstaus on joustavaa, resursseja säästävää sekä tarjoaa mahdollisuuden testata laajemmin erilaisilla loppukäyttäjillä (esim. eri maista olevat käyttäjät, erikoisryhmät jne.). Asynkronista testausta voidaan tehdä käyttämällä mittalaitteita sisältävää selainta, joka kerää käyttäjän klikkauspolut ja ruutukaappaukset ja lähettää nämä tiedot arvioijan sivustolle analysoitavaksi tai käyttämällä *välityspalvelinta* (proxy) [Bastien, 2010]. Tällöin käyttäjiä ohjeistetaan menemään tietyille verkkosivulle ja seuraamaan ohjeita, jolloin he siirtyvät arvioitavalle verkkosivulle. Tällä sivulla heidän käyttäytymistään tallennetaan, yhdistellään ja visualisoidaan siten, että saadaan selville mitä verkkosivuja käyttäjät tutkivat [Bastien, 2010].

Etäarvioinnissa voidaan käyttää apuna erilaisia automaattisia työkaluja, joilla tarkoitetaan käytettävyyden arvioinnissa suoritettavien toimenpiteiden (aineiston keräys – analysointi – ratkaisujen esittäminen) osittaista tai täydellistä automatisointia erilaisia tietokoneohjelmistoja käyttämällä [Heimonen, 2005]. Tällaisia työkaluja ovat muun muassa mallinnus- ja simulointimenetelmät sekä erilaiset automaattiset tarkistukset ja analyysit, kuten esimerkiksi automaattinen lokianalyysi.

Automaattinen lokianalyysi

Automaattisessa tarkistuksessa asiantuntija käyttää heuristiikkoja ja muita kriteerejä tunnistaakseen käyttöliittymässä olevia ongelmia. Tällaisilla työkaluilla voidaan käsittääkseen tarkistaa muun muassa verkkosivujen saavutettavuutta, linkkien toimivuutta, värien kontrastia tai koodin oikeellisuutta. Yksi esimerkki tällaisesta automaattisesta tarkistustyökalusta on W3C – Mark Validation Service (validator.w3.org), jota käytetään HTML-koodin oikeellisuuden tarkistukseen.

Yksi automaattisen tarkistuksen muoto on automaattinen lokianalyysi, jossa käyttäjän käyttötilanteesta kerätystä tiedosta (esim. tehtävään käytetty aika, klikkaukset jne.) etsitään vihjeitä mahdollisia käytettävyysongelmia. Nämä voivat ilmetä esimerkiksi poikkeuksellisen pitkänä suoritusaikana (lasketusta keskiarvosta selvästi poikkeava suoritusaika) tai käyttäjän käyttämä reitti tietyille sivulle poikkeava keskiarvosta (esim. keskiarvo 2 klikkausta, käyttäjä teki 6 klikkausta). Nämä tiedot voivat tarjota vinkkejä mahdollisista käytettävyysongelmista [Carta *et al.*, 2011].

Automaattisen lokianalyysin, kuten myös yleisesti automaattisen arvioinnin, suurimpana heikkoutena voidaan pitää laadullisen ja subjektiivisen informaation puutetta. Niillä

on vaikeaa tallentaa käyttäjien mieltymyksiin ja käsityksiin liittyvää informaatiota, jonka kerääminen on mahdollista perinteisiä käytettävyydestestauksen menetelmiä käyttämällä [Heimonen, 2005].

4.4 Käytettävyyden mittaaminen

Käytettävyyden määritelmien ja arviointimenetelmien tapaan myös käytettävyyden mittaamiseen on useita ratkaisuja. Perinteisiä käytettävyyden mittareita ovat erilaiset käytettävyyden osa-alueisiin eli vaikuttavuuteen, tehokkuuteen ja tyytyväisyyteen liittyvät arvot. Seuraavassa esittelen muutamia näiden osa-alueiden mittareista perustuen Hornbækin [2006] tekemään kirjallisuuskatsaukseen, joka koostaa yhteen käytettävyydestutkimuksissa käytettyjä mittareita ja niiden mitta-arvoja.

4.4.1 Vaikuttavuus

Hornbækin [2006] kirjallisuuskartoituksen mukaan vaikuttavuuden mittareita ovat *binäärinen tehtävien suorittaminen* (binary task completion), *tarkkuusmitat* (accuracy measures), *mieleen palauttaminen* (recall), *kokonaisuus* (completeness), *lopputuloksen laatu* (quality of outcome) sekä joitakin muita mittareita. Taulukossa 5 esitetään edellä mainitut mittarit, niiden selitykset sekä esimerkkejä mitta-arvoista mukailen Hornbækin [2006] määritelmiä.

Mittari	Selitys
Binäärinen tehtävien suorittaminen	<i>Saiko käyttäjä tehtävät suoritettua loppuun asti?</i> Mitta-arvoja ovat esimerkiksi oikein menneiden, keskeneräisten ja keskenjätettyjen tehtävien määrät
Tarkkuusmitat	<i>Tekikö käyttäjä virheitä testitehtäviä tehdessä tai onko testitehtävään loppuratkaisussa virheitä?</i> Mitta-arvoja ovat esimerkiksi virheiden määrä, osoitinlaitteiden testauksessa <i>etäisyys käyttäjän osoittaman ja oikean kohteen välillä</i> (spatial accuracy) sekä tiedonhaussa <i>oikeiden dokumenttien suhde kaikkiin haettuihin dokumentteihin</i> (precision accuracy).
Mieleen palauttaminen	<i>Kuinka paljon tietoa käyttäjä pystyy palauttamaan mieleen käytettyään käyttöliittymää?</i> Mitataan usein selvittämällä muistaako käyttäjä nähneensä jonkin verkkosivun elementin aiemmin. Voidaan myös mitata kysymällä käyttäjältä testin jälkeen montako ominaisuutta hän muistaa [Nielsen, 1993].
Kokonaisuus	<i>Mitkä tehtävät käyttäjä ratkaisi?</i> Mitta-arvoja ovat esimerkiksi suoritettujen toisarvoisten tehtävien määrä tai tiedonhaussa löydettyistä materiaaleista relevanttien materiaalien osuus.

Lopputuloksen laatu	<i>Mikä oli tehtävien lopputulos?</i> Mitta-arvoja ovat esimerkiksi tuotteen tai oppimisen laatu. Myöskin ymmärtämistä mitataan usein, esimerkiksi testaamalla mitä on opittu ohjaavasta käyttöliittymästä (esim. tietokonepohjaiset tutorit).
Muut	Esimerkiksi käyttäjän kyky ennakoita, kuinka käyttöliittymä toimii, sekä muutoksen käyttäjän elintoiminnoissa (esimerkiksi sydämen syketaajuus).

Taulukko 5: Vaikuttavuuden mittarit ja niiden selitykset

4.4.2 Tehokkuus

Hornbækin [2006] mukaan tehokkuuden mittareita ovat *aika* (time), *syöttönopeus* (input rate), *mentaalin vaivannäkö* (mental effort), *käyttökaavat* (usage patterns), *vaivannäkö viestintään* (communication effort), *oppiminen* (learning measures) sekä joitakin muita mittoja. Taulukossa 6 esitetään edellä mainitut mittarit, niiden selitykset sekä esimerkkejä mitta-arvoista mukaillen Hornbækin [2006] määritelmiä.

Mittari	Selitys
Aika	<i>Kauanko käyttäjällä menee tehtävään suorittamiseen?</i> Mitta-arvoja ovat esimerkiksi koko tehtävään tai tehtävän osaan tarvittu aika, jossain käyttöliittymän osassa vietetty aika tai aika ensimmäiseen näppäimen painallukseen. Lisäksi voidaan mitata aikaa, jolloin käyttäjä ei vuorovaikuta järjestelmän kanssa (ns. tyhjä aika), eli joko käyttäjä odottaa järjestelmää tai järjestelmä odottaa käyttäjää [Nielsen, 1993].
Syöttönopeus	<i>Kuinka nopeasti käyttäjä kirjoittaa?</i> Mitta-arvoja ovat esimerkiksi käyttäjän syöttönopeus (sanoja per minuutti, korjattuja sanoja per minuutti) tai <i>suoritusteho</i> (throughput).
Mentaalinen vaivannäkö	<i>Kuinka paljon mentaalista vaivannäkää käytetään vuorovaikutukseen?</i> Mitta-arvoja ovat esimerkiksi sykevaihdelu ja <i>subjektiivinen ajan arviointi</i> (subjective time estimation).
Käyttökaavat	<i>Kuinka käyttöliittymää käytetään?</i> Mitta-arvoja ovat esimerkiksi, kuinka monta kertaa jokin toiminto suoritetaan, kuinka paljon tietoa käyttäjä käyttää tehtävää ratkaistessaan sekä kuinka paljon käyttäjän ratkaisu eroaa optimaalisesta ratkaisusta. Lisäksi voidaan mitata, monestiko käyttäjä ajautui sivuraiteille tehtävästä eli ei keskittynyt enää tehtävään, kuinka montaa komentoa ja ominaisuutta käyttäjä käytti, sekä komennot ja ominaisuudet joita käyttäjä ei käyttänyt kertaakaan [Nielsen, 1993].

Vaivannäkö viestintään	<i>Kuinka paljon käyttäjä käyttää resurssejaan viestintään? Mitta-arvoja ovat esimerkiksi käytetyt vuorot keskustelussa sekä yhteiskäyttöisessä selausympäristössä käytettyjen tarkistuskysymysten (grounding questions) määrä.</i>
Oppiminen	<i>Muuttuuko käytön tehokkuus eli oppiiko käyttäjä? Mitta-arvoja ovat esimerkiksi tehtävien suorittamiseen tarvittava aika sekä muut mitat, joilla voidaan viitata tehokkuuden muutokseen.</i>
Muut	<i>Esimerkiksi minuutissa luettujen sanojen määrä, tehtävän suorittamisen helppous sekä rajallinen suoritus aika.</i>

Taulukko 6: Tehokkuuden mittarit ja niiden selitykset

4.4.3 Tyytyväisyys

Hornbækin [2006] mukaan tyytyväisyyden mittareita ovat *vakioidut kyselyt* (standard questionnaires), *mieltymykset* (preferences), *helppokäyttöisyys* (ease-of-use), *asenteet* (specific attitudes, attitudes towards other persons, attitudes towards the content of the interface), *käsitykset lopputuloksesta* (perception of outcomes), *käsitys vuorovaikutuksesta* (perception of interaction) sekä muut mittarit. Taulukossa 7 esitetään edellä mainitut mittarit, niiden selitykset sekä esimerkkejä mitta-arvoista mukaillen Hornbækin [2006] määritelmiä.

Mittari	Selitys
Vakioidut kyselyt	<i>Käytettiinkö tyytyväisyyden mittaamiseen vakioituja kyselyitä?</i>
Mieltymykset	<i>Mitä käyttöliittymää käyttäjät suosivat? Mitataan pakottamalla käyttäjät valitsemaan käyttöliittymä, jota he suosivat. Usean käyttöliittymän tapauksessa käyttäjät järjestävät ne mieltymysten mukaisesti arvojärjestykseen tai antavat niille arvosanat</i>
Helppokäyttöisyys	<i>Onko käyttäjä yleistasoisesti tyytyväinen käyttöliittymään? Mitataan antamalla käyttäjälle kysely, jossa kysytään termejä tyytyväisyys/tyytyväinen/käyttäjän tyytyväisyys sisältäviä kysymyksiä ja neuvotaan käyttäjää arvioimaan tyytyväisyyttään johonkin ominaisuuteen (esim. ominaisuuden käyttö) annetulla asteikolla.</i>
Asenteet	<i>Mitkä ovat käyttäjän asenteet käyttöliittymää/vuorovaikutusta ja yhteistyökumppaneita/käyttöliittymän sisältöä kohtaan? Asenteita käyttöliittymää kohtaan arvioidaan esittämällä käyttäjälle attribuutteja, jotka hän arvioi annetulla asteikolla (esimerkiksi hauskuus, pitäminen tai</i>

	ärsyttävyys). Vuorovaikutusta ja yhteistyökumppaneita voidaan arvioida mm. attribuuteilla läsnäolo, luottamus, <i>yhteinen tietopohja</i> (common grounds) ja kommunikoinnin helppous. Käyttöliittymän sisältöä arvioidessa kysymykset koskevat sisällön laatua kuten kohtia, joihin käyttäjillä oli mielenkiintoa sekä tiedon organisointia.
Käsitykset lopputuloksesta	<i>Kuinka käyttäjä arvioi käsityksensä vuorovaikutuksen lopputuloksesta?</i> Mitataan kysymällä käyttäjältä luottamuksesta omia ratkaisujaan kohtaan kysymyksillä, jotka koskevat hänen ymmärrystään aiheesta, oppimistaan ja tyytyväisyyttä omaan suoritukseensa.
Käsitykset vuorovaikutuksesta	<i>Kuinka käyttäjä arvioi käsityksensä vuorovaikutusprosessista?</i> Mitataan kysymällä käyttäjältä hänen käsitystään tehtävien monimutkaisuudesta ja suoritusajoista.
Muut	Esimerkiksi <i>kauneus</i> (beauty), kuinka <i>sotkuiselta</i> (cluttered) näkymä tuntui sekä käyttäjän <i>nolostuminen</i> (embarrassment).

Taulukko 7: Tyytyväisyyden mittarit ja niiden osa-alueet

4.4.4 Muita mitattavia osa-alueita

Yllä esiteltyjen ”perinteisten” osa-alueiden lisäksi käytettävyyden mittareina on käytetty muun muassa *hedonista laatua* (hedonic quality), *hauskuutta* (fun), *estetiikkaa* (aesthetics), *ilmeistä käytettävyyttä* (apparent usability), *seurallisuutta* (sociability) sekä *virtausta* (flow) [Hornbæk, 2006].

4.5 Joukkoistetun käytettävyydestauksen erot verrattuna muihin testausmenetelmiin

Tässä kohdassa esittelen neljän eri arviointimenetelmän (heuristinen arviointi, etätetaus, käytettävyydestaus laboratoriossa ja joukkoistettu käytettävyydestesti) hyvät ja huonot puolet sekä koostan niistä taulukon. Lisäksi vertaan näitä menetelmiä keskenään saadakseni selville joukkoistetun käytettävyydestauksen edut.

Asiantuntija-arvioinnit ovat nopeita, kustannustehokkaita ja helppoja oppia, lisäksi ne eivät vaadi suuria etukäteisjärjestelyitä, ja nopeimmillaan sen voi suorittaa jopa yhdessä päivässä. Mielestäni nämä ovat ne tärkeimmät syyt, jotka nostavat asiantuntija-arvioinnin käytetyimpien menetelmien joukkoon. Yksi asiantuntija-arviointien muoto on heuristinen arviointi. Sen suurin heikkous on todellisten käyttäjien osallistumattomuus, koska tällöin menetelmässä ei nouse esille todellisen käyttäjän ja tuotteen välistä vuorovaikutusta [Korvenranta, 2005]. Koska testikäyttäjät eivät ole mukana arvioinnissa, tuloksista puuttuvat ongelmat, jotka tulevat esiin ainoastaan todellisessa tuotteen ja käyttäjän vuorovaikutustilanteessa. Tästä syystä asiantuntija-arviointia suositellaankin tehtävän yhdessä käytettävyyss-

testauksen kanssa, jolloin testauksessa käydään läpi tuotetta todellisen käyttäjän vuorovai-
kutuksen näkökulmasta sekä asiantuntijan näkökulmasta. Asiantuntija-arvioinneista mu-
kaan valittiin heuristinen arviointi, koska sitä käytetään usein käytettävyyden arvioinnissa,
ja koska se voisi hyötyä joukkoistamisesta.

Käytettävyydestä vahvuus on se, että se antaa joihinkin muihin menetelmiin (esi-
merkiksi asiantuntija-arviointiin) verrattuna selvästi enemmän tietoa tuotteen käytettävyy-
destä sellaisena kuin tuotteen todellinen käyttäjä sen kokee [Koskinen, 2005]. Lisäksi käy-
tettävyydestä voidaan tehdä lähes koko tuotteen kehitysprosessin ajan, koska testejä
voidaan tehdä niin alkeelliselle prototyypille kuin valmiille tuotteellekin. Vahvuutena on
myös joustavuus, sillä käytettävyydestä voidaan käyttää monipuolisesti yhden tuot-
teen kertaluontoiseen testaamiseen, tuotteen käytettävyyden kehittämiseen testaamalla
useaan kertaan pitemmällä aikavälillä, tai useamman version tai jopa eri tuotteiden vertai-
luihin testauksen avulla. Käytettävyydestä voidaan löytää yleensä paljon käytettä-
vyysongelmia, ja varsinkin kriittisimmät ongelmat löydetään tehokkaasti. Testauksessa saa-
daan paljon tietoa tuotteen käytettävyydestä: numeerisen aineiston (esimerkiksi virheiden
määriä, tehtäväaikoja ym.) lisäksi saadaan tietoa myös esimerkiksi tuotteen hyödyllisyy-
destä ja miellyttävyydestä käyttäjän kannalta. Nämä voidaan mielestäni laskea myös vah-
vuuksiksi, koska käytettävyydestä tavoitteena on löytää mahdollisimman paljon käy-
tettävyysoongelmia, ja mieluiten vielä kriittisimpiä ongelmia. Lisäksi tuotekehityksen kan-
nalta on erittäin hyvä, kun on mahdollisimman paljon tietoa saatavilla. Käytettävyydes-
tauksen suurin heikkous on sen työläs toteutus. Tällä tarkoitetaan sitä, että alan ammattilai-
sia tarvitaan niin testauksen suunnittelussa, toteutuksessa ja raportoinnissa. Myöskin osal-
listujien värväämiseen menee usein runsaasti resursseja. Kuitenkin, käytettävyydestä
suurin heikkous on se, että testaus toteutetaan usein käytettävyydelaboratoriossa, jolloin ympä-
ristö on keinotekoinen, eli testistä saatava tieto ei välttämättä vastaa todellisia käyttöti-
lanteita. Käytettävyydestä valittiin mukaan vertailuun, jotta saataisiin selville, onko jouk-
koistamisella etuja siihen verrattuna.

Etätestauksen vahvuudet ovat hyvin samankaltaiset kuin käytettävyydestä vah-
vuudet, varsinkin jos testaus tehdään synkronisesti. Tällöin moderaattori ja osallistuja osal-
listuvat yhtäaikaaisesti: moderaattori kuulee käyttäjän verbalisoimat ajatukset ja kommentit.
Asynkronisessa testauksessa moderaattori ei yleensä saa näitä tietoja, vaan numeerista ai-
neistoa tehtävien suorituksesta (esimerkiksi tehtäväajat, klikatut kohteet ja käyttäjän kulke-
mat polut). Etätestauksen vahvuus laboratoriossa tehtävään käytettävyydestä verrattuna
on se, että käyttäjien ei tarvitse matkustaa laboratorioon. Käyttäjä suorittaa tehtävät joko
omalla tietokoneellaan ja oheislaitteillaan tai saatavilla olevassa konferenssitilassa. Tällöin
testin järjestäminen on helpompaa ja halvempaa, koska ei tarvitse varata laboratoriota ja
sopia osallistujien matkustamisesta sinne. Etätestauksen heikkoutena on se, että arvioinnin

järjestäjän täytyy itse rekrytoida osallistujat, mikä vie paljon aikaa, eikä osallistujien rohkaiseminen osallistumaan etätestiin ei ole helppoa [Chynal ja Sobecki, 2015]. Lisäksi toteutustapa tuo omat heikkoutensa etätestaukseen: synkronisessa testauksessa testien toteuttamiseen menee yhtä paljon aikaa kuin laborioritestauksessa, koska moderaattori tekee testin käyttäjän kanssa. Asynkronisessa testauksessa saatava tieto on puolestaan usein vain numerista, koska toteutustapa ei mahdollista tiedon keruuta havainnoimalla ja nauhoittamalla käyttäjän spontaaneja verbalisointeja [Bastien, 2010]. Etätestaus valittiin mukaan vertailuun, koska se on samankaltainen menetelmä kuin joukkoistaminen. Näin nähdään, onko näiden menetelmien välillä suuria eroja.

Joukkoistetussa käytettävyydestestauksessa käyttäjät suorittavat usein käytettävyyssarviointia ratkaisemalla joukon ennalta määritellyjä tehtäviä ja raportoivat niitä tehdessä kohtaamansa ongelmat [Bruun ja Stage, 2015]. Joukkoistetun käytettävyydestestauksen vahvuudet ovat suurempi osallistujamäärä, helpompi rekrytointi ja se, että se toteutetaan käyttäjän laitteistolla hänen normaalissa käyttöympäristössään. Joukkoistetun käytettävyydestestin toteutukseen ei siis tarvita erityislaitteistoa, käytettävyysslaboratoriota tai myöskään moderaattoreita. Myös osallistujien rekrytointi on helpompaa, jos testeissä käytetään jo valmiiksi olemassa olevaa joukkoistamisalustaa, jonka käyttäjistä eli joukosta valitaan sopivat osallistujat. Joukkoistamisen suurin heikkous onkin moderaattorien puute eli käyttäjiltä ei saada ajantasaista, suullista palautetta siitä, mitä he ajattelevat tehtäviä tehdessä tai mitä mieltä he ovat ominaisuuksista. Tosin, käyttäjille voidaan antaa mahdollisuus antaa kirjallista palautetta ja vastata kysymyksiin, joilla arvioidaan tyytyväisyyttä johonkin ominaisuuteen. Taulukkoon 8 on kerätty koosteena näiden neljän edellä mainitun menetelmän ominaisuuksia, joissa ne eroavat toisistaan. Taulukon riveillä esitetyt vertailtavat ominaisuudet on muodostettu keräämällä taustamateriaaleista menetelmien vahvuuksia ja heikkouksia ja sen jälkeen ryhmittelemällä ne yhteenkuuluviksi kokonaisuuksiksi.

	Heuristinen arviointi	Käytettävyydestesti laboratoriossa	Etätestaus	Joukkoistettu käytettävyydestesti
Osallistujien lukumäärä	Vain asiantuntija tai asiantuntijat	Vähintään 5	Vähintään 5	Kymmeniä
Osallistujien rekrytointi	Ei	Kyllä, arvioinnin järjestäjä rekrytoi	Kyllä, arvioinnin järjestäjä rekrytoi	Kyllä, joukkoistamisalustan kautta

Moderaattorit	Ei	1 moderaattori ja mahdolliset tarkkailijat	1	Ei
Tarvittava laitteisto	Ei erityisvaatimuksia	Useita tietokoneita, ruudun- ja äänentallennus, tarvittaessa katseen seurantalaite	Käyttäjän oma laitteisto, joskus myös valvojan laitteisto	Käyttäjän oma laitteisto
Tarvittava tila	Ei erityisvaatimuksia	Käytettävyyslaboratorio	Tarvittaessa virtuaalikonferenssitila	Ei erityisvaatimuksia
Testiintarvittava aika	Vain asiantuntijalta läpikäyntiin kuuluva aika	Ohjeistukseen, esitietolomakkeeseen, testitehtäviin sekä loppuhaastatteluun kuuluva aika.	Riippuu toteutustavasta (synkroninen vai asynkroninen)	Annettuun ohjeistukseen tutustumiseen ja tehtävien suorittamiseen kuuluva aika
Saatava palaute	Heuristiikkoja rikkovat kohdat, asiantuntijan kommentit	Käyttäjien ajatukset ja mielipiteet, numeerinen aineisto, myös katseenseurannalla saatu data, videonauhoite ja siitä laadittu loki	Tehtäviin kulunut aika, klikkaukset ja reitit. Myös ajatukset ja kommentit, jos toteutus synkronisesti	Tehtäviin kulunut aika, osallistujien antamat arvioinnit, osallistujien raportoidut käytettävyysongelmat. Myös mielipiteitä ja kommentteja.

Taulukko 8: Eri arviointimenetelmien vertailu

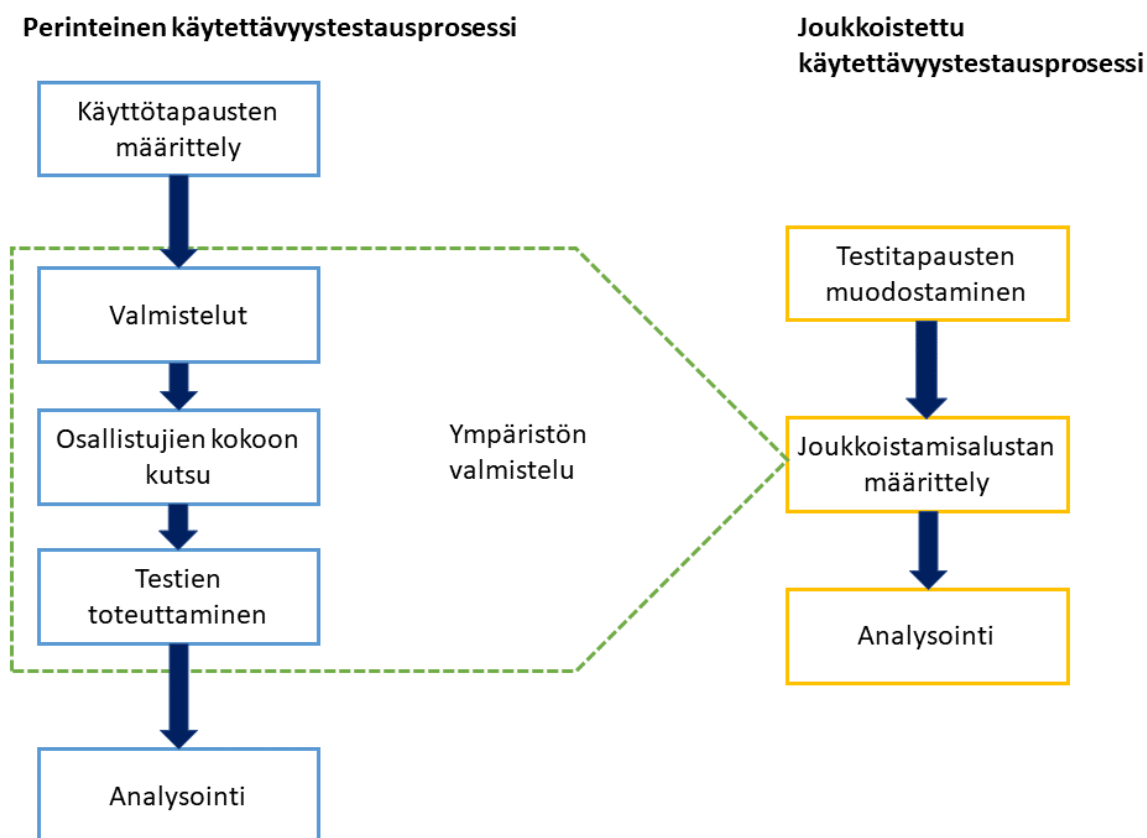
Edellä esitetyt menetelmien eroina ovat osallistujien määrä, osallistujien rekrytointi, moderaattorien/asiantuntijoiden määrä, tarvittava laitteisto, tarvittava tila, testiin kuuluva aika sekä saatavalla palaute. Taulukon perusteella voidaan todeta joukkoistamisella olevan muutamia etuja verrattuna muihin menetelmiin. Seuraavaksi perehdyn tarkemmin näihin etuihin.

Joukkoistettu käytettävyystesti toimii samankaltaisella periaatteella kuin asynkroninen etäarviointi, eli käyttäjä suorittaa tehtävät omalla laitteistolla silloin kun hänelle parhaiten sopii. Tällöin käyttäjiä toimintaa pystytään arvioimaan heidän päivittäisessä ympäristössään ja samalla pystytään pienentämään arviointikuluja, kun ei tarvita erityistä laboratorioympäristöä eikä käyttäjän tarvitse matkustaa [Carta *et al.*, 2011; Chynal ja Sobecki, 2015]. Laboratoriossa tehtävän käytettävyysarvioinnin järjestämisestä aiheutuu monia kuluja ja

paljon vaivaa, kun täytyy rekrytoida osallistujat ja suorittaa testit, värvätä tarkkailijat ja mahdollisesti hankkia tai vuokrata tarvittava laitteisto [Liu *et al.*, 2012; Kittur *et al.*, 2008]. Liu ja muut [2012] listaavatkin monia hyötyjä joukkoistetusta käytettävyydestestauksesta: se helpottaa rekrytointia ja täten mahdollistaa suuremman tietomäärän keräämisen, säästää aikaa (matkustaminen, tervehtiminen ja valmistelu) ja rahaa (teknisistä asiantuntijoista, laboratorion, laitteiston ja matkoista aiheutuvat kulut), tekee testauksesta joustavampaa eli testausta voidaan tehdä helposti koko kehityksen ajan, helpottaa eri taustan omaavien sekä muualla asuvien käyttäjien testausta. Käytettävyyden etätetaustakin toimii muualla asuvien osallistamiseen testaukseen, mutta siinä on lisäksi testin järjestämiseen ja suorittamiseen tarvittava aika [Liu *et al.*, 2012]. Albertin ja muiden [2010] mukaan laboratoriossa tehtävään testisessioon voi helposti kuluu tunnista puoleentoista, kun verkossa tehtävään käytettävyydestiin menee yleensä vain 15-45 minuuttia. Myös Gomide ja muut [2014] sanovat perinteisen käytettävyydestestauksen olevan kallista ja aikaa vievää. Perinteisellä käytettävyydestestauksella tarkoitan tässä yhteydessä käytettävyydelaboratoriossa suoritettavaa testausta, ja perinteisellä arviointiprosessilla tarkoitan tällaisessa testissä käytettävää prosessia.

Gomide ja muut [2014] esittivät perinteisessä käytettävyydestestauksessa olevan useita vaiheita, joita joukkoistettu käytettävyydestestaus pystyy helpottamaan. Kuvassa 3 esitetään uudelleen piirretty ja suomennettu Gomiden ja muiden [2014] perusteella laadittu vertailu perinteisen ja joukkoistetun käytettävyyden arviointiprosessista.

Kuten kuvassa 3 näkyy, perinteisessä testauksessa ensimmäinen vaihe on *käyttötapausten määrittely* (use case definitions), jossa kartoitetaan erilaiset toiminnot, joita käyttäjä voi käyttöliittymässä tehdä. Näitä käyttötapauksia käytetään pohjana testitehtäville. Seuraavaa kolmea vaihetta Gomide ja muut [2014] kutsuvat *ympäristön valmisteluksi* (environmental preparations). Nämä vaiheet ovat *valmistelut* (preparations), *osallistujien kokoon kutsuminen* (participant convocation) ja *testien toteuttaminen* (tests execution). Tässä vaiheessa prosessia siis valmistellaan fyysinen testitila (käytettävyydelaboratorio), valitaan ja rekrytoidaan osallistujat käyttötapausprofiilien pohjalta sekä valmistellaan tallennuslaitteet ja tarkistetaan, että osallistujat ovat kohderyhmää ja vastaavat profiileja. Viimeinen vaihe on *analyysi* (analysis), jossa analysoidaan testitulokset.



Kuva 3: Perinteinen testausprosessi verrattuna joukkoistettuun testausprosessiin

Joukkoistetun testauksen prosessi on kolmivaiheinen, joista ensimmäinen vaihe on *testitapausten muodostaminen* (test case elaboration). Tässä vaiheessa määritellään testattavat tapaukset, eli esimerkiksi käyttöliittymän toiminnot tai osat, jotka halutaan testata ja joita käytetään pohjana testitehtävien laatimiseen. Toinen vaihe on *joukkoistamisalustan määrittely* (crowdsourcing platform definition), joka vastaa perinteisen testausprosessin ympäristön valmistelu -vaihetta. Tässä vaiheessa prosessia valitaan joukkoistamiseen käytettävä alusta, määritellään ja laaditaan testitehtävät, julkaistaan testitehtävät ohjeistuksineen joukolle ja annetaan joukon suorittaa tehtävät. Nämä kaikki toiminnot tapahtuvat joukkoistamisalustan kautta eli se saattaa nopeuttaa testausprosessia. Joukkoistamisessa oletetaan usein, että alusta on jo käyttäjille tuttu (esimerkiksi osallistuja on osallistunut ennemminkin tai tutustunut alustaan ennen testiä). Tehtävissä itsessään on ohjeistukset kyseisen kohdan suorittamiseen eli prosessissa ei tarvitse varata aikaa käyttäjien opastukseen. Kuten perinteisessäkin testauksessa, myös joukkoistetussa testauksessa viimeinen vaihe on testauksen tulosten analysointi.

Jotkin ongelmat ja virheet ovat sellaisia, ettei niitä löydetä pienellä osallistujamäärällä [Kittur *et al.*, 2008]. Monesti käytettävyystestaus onkin tasapainottelua otoskoon, käytettävän ajan ja kustannusten välillä. Pienellä osallistujamäärällä löydetään yleensä kaikkein il-

meisimmät käytettävyysoongelmat [Albert *et al.*, 2010]. Käytettävyytestauksessa käytetäänkin usein niin sanottua pienen otoskoon sääntöä, jonka mukaan 80% käytettävyysongelmista löytyy viidellä testikäyttäjällä [Anttonen, 2005]. Tätä sääntöä käytetään yleensä perusteluna mietittäessä kustannuksissa säästämistä ja testauksen kustannustehokkuutta. Joukkoistamisesta aiheutuvat kulut osallistujaa kohden ovat paljon pienemmät, jolloin osallistujien määrää pystytään helposti lisäämään ilman kulujen suurta kasvamista.

Yhteenvetona joukkoistamisen voidaan todeta tekevän käytettävyyden arvioinnista helpompaa ja halvempaa. Joukkoistaminen pienentää käytettävyytestauksesta aiheutuvia kuluja ja tekee testin järjestämisestä vaivattomampaa niin järjestäjän kuin osallistujan näkökannalta. Lisäksi se säästää aikaa nopeuttamalla testausprosessia ja vähentämällä valmisteluun tarvittavaa aikaa.

5 Joukkoistetun käytettävyydestauksen järjestäminen

Kuten kohdassa 3.2 kerrottiin, on joukkoistamisen järjestämisessä useita haasteita. Kuitenkin monet näistä haasteista pystytään voittamaan oikeilla järjestelyillä [Afuah, 2011]. Tässä luvussa kerron kuinka kohdassa 3.2 mainitut haasteet voidaan voittaa. Luvun tarkoitus on tarjota tietoa siitä, mitä tulee ottaa huomioon joukkoistettua käytettävyydestausta järjestäessä, miksi nämä asiat tulee ottaa huomioon ja miten ne voitaisiin hoitaa.

5.1 Motivointi

Motivointi on tärkeä osa joukkoistamista, koska se vaikuttaa suuresti ratkaisujen määrään ja laatuun [Hossain, 2012b]. Jos osallistujat eivät ole motivoituneita osallistumaan, ei vastauksia välttämättä tule paljoa eivätkä saatavat vastaukset ole laadullisesti hyviä. Hossainin [2012a] mukaan motivaatio voidaan nähdä prosessina, jolla käynnistetään, hallitaan ja ylläpidetään fyysisiä ja henkisiä tehtäviä. Tällä tarkoitetaan sitä, että kun henkilöllä on motivaatio suorittaa jokin tehtävä, hän myös aloittaa asian tekemisen, huolehtii, että se tulee tehtyä ja varsinaisesti suorittaa tehtävän. Motivaatio ilmenee tiettyssä tilanteessa vuorovaiikutuksena eri motiivien ja kannustimien joukossa [Hossain, 2012a; Hossain 2012b]. Eli tiettyssä tilanteessa (esimerkiksi työtarjous) ihminen arvioi tilannetta motiiviansa (esimerkiksi rahan tarve tai halu auttaa) ja kannustimiensa (esimerkiksi palkka tai tieto toiminnan vaikutuksesta) perusteella ja osallistuu, jos motiivit ja kannustimet täsmäävät. Motivointi koostuu siis kolmesta osasta: motiiveista, kannustimista ja mekanismeista.

Motivaatio voidaan jakaa karkeasti kahteen luokkaan: *sisältöpäin tulevaan* (intrinsic) ja *ulkoapäin tulevaan* (extrinsic) [Hossain, 2012a]. Tästä eteenpäin kutsun näitä sisäiseksi ja ulkoiseksi motivaatioksi. Sisäinen motivaatio viittaa motiiveihin, jotka osallistuja kokee saavansa suoraan itse tehtävästä [Hossain, 2012a]. Näissä tilanteissa tehtävän suorittaminen riittää tekemään osallistujan tyytyväiseksi eikä mitään muuta palkkiota anneta tehtävästä [Hossain, 2012a]. Ulkoinen motivaatio puolestaan on aina ulkoisten kannustimien ajamaa [Hossain, 2012b]. Seuraavassa esittelen motivoinnin osat ja lopuksi esitän koontitaulukon, josta näkyy näiden osien linkittyminen toisiinsa. Sekä koontitaulukossa että osien esittelyissä on käytetty apuna esitettyä Katmadan ja muiden [2016] laatimaa kaaviota (ks. liite 1), jossa on esitelty motiivit, kannustimet ja mekanismit, niiden väliset yhteydet sekä esimerkkejä eri joukkoistamisalustojen sijoittumisesta näiden osien välille. En kuitenkaan esittele joukkoistamisalustoja tässä luvussa lainkaan, sillä esimerkkejä niistä on esitelty kohdassa 3.5.

5.1.1 Motiivit

Motiivit ovat sisäisiä tai ulkoisia syitä, *miksi käyttäjät osallistuvat joukkoistamiseen*. Ulkoiset motiivit voidaan luokitella kolmeen ryhmään: taloudelliseen, sosiaaliseen ja organisaatio-

riseen [Hossain, 2012b]. Taloudellisia motiiveja ovat esimerkiksi hyödyt, raha, tyytymättömyys, työmahdollisuudet, ongelma ja tuotto. Sosiaalisia motiiveja ovat esimerkiksi velvollisuus, tunnustus vertaisilta, maine, valta, taitojen kehittäminen, kokemus, tiedonkeruu, sosiaaliset suhteet ja vuorovaikutus, verkostoituminen, yhteistyö, ego, julkisuus, *oikeuksien* (privileges) saaminen, suuttumus ja avuliaisuus. Organisatorisia motiiveja ovat esimerkiksi vastuut, vaikutusvalta, ura ja rekrytointi.

Sisäisiä motiiveja ovat esimerkiksi hyväntekeväisyys, nauttiminen, mielihyvä, omahyväisyys, halu ratkaista, kompetenssi, hauskuus, autonomia, itsemääräämisoikeus, tunnistaminen, altruismi, kuuluminen yhteisöön, yhteisöllisyys, ylpeys, ajan kuluttaminen, oppiminen ja itsensä toteuttaminen [Hossain, 2012b]. Katmada ja muut [2016] muodostivat taustamateriaaliensa perusteella seitsemän motiivia: oppiminen/henkilökohtainen saavuttaminen, altruismi, nautinto/älyllinen uteliaisuus, sosiaaliset motiivit, itsensä markkinointi, epäsuora työnteko ja suora kompensatio. Oppiminen/henkilökohtainen saavuttaminen kertoo siitä, että henkilö osallistuu joukkoistamiseen oppiakseen uutta ja tunteakseen saavuttaneensa jotain. Altruismissa henkilöä motivoi yhteisön auttaminen ja nautinnossa/älyllisessä uteliaisuudessa henkilöä motivoi pelkkä nautinto tekemisestä tai mahdollisuus tavata uusia ihmisiä ja nähdä uusia paikkoja ja tilanteita. Sosiaaliset motiivit liittyvät nimensä mukaisesti sosiaalisuuteen eli niissä henkilöä motivoivat sosiaaliseen statukseen liittyvät asiat. Itsensä markkinoinnissa ihmistä motivoi mahdollisuus saada itsellensä ja osaamisellensa näkyvyyttä. Epäsuoraa työntekoa kutsutaan myös *passiiviseksi joukkoistamiseksi* (passive crowdsourcing), eli osallistuminen on joko *sivuvaikutus* (side effect) tai *tiedon luovuttamista* (contributing information) kolmannelle osapuolelle. Sivuvaikutus tarkoittaa sitä, että osallistuja antaa tietämättään vastauksen johonkin toiseen ongelmaan suorittaessaan annettua tehtävää. Tiedon luovuttamisessa osallistuja luovuttaa tietoa kolmannen osapuolen verkkosivulle (esimerkiksi sosiaaliseen mediaan) joskus jopa tiedostamattaan, ja tätä tietoa käytetään apuna testauksessa. Tällöin osallistuja siis osallistuu tarjoamalla tietoa, mutta ei kuitenkaan varsinaisesti osallistu testaamiseen. Suorassa kompensatiossa henkilöä motivoivat osallistumaan erilaiset palkkiot. Näistä neljä ensimmäistä voidaan luokitella sisäisiksi motiiveiksi. Vaikka Katmada ja muut [2016] ja Hossain [2012b] luokittelivat motiivit eri tavoilla, on luokitteluissa kuitenkin paljon yhtäläisyyksiä (esimerkiksi altruismi on mainittu molemmissa). Motiivit ovat esitetty yhdistetysti taulukon 9 ensimmäisessä sarakkeessa.

5.1.2 Kannustimet

Kannustimet ja motiivit liittyvät toisiinsa tiiviisti [Hossain, 2012b]. Motivoidakseen käyttäjät osallistumaan, täytyy kannustimien olla heidän motiivejansa vastaavia. Kannustimet ovat siis *”palkkioita”, jotka vastaavat osallistujien motiiveja* (esimerkiksi jos motiivi on raha, kannustin voi olla palkka tai rahallinen bonus).

Hossain [2012b] jakaa kannustimet samoihin kolmeen kategoriaan kuin motiivitkin. Taloudellisia kannustimia ovat esimerkiksi maksu, premium-tuotteet, tuotelisenssi, ilmaiset tuotteet tai palvelut, bonukset ja kupongit. Sosiaalisia kannustimia ovat esimerkiksi palkinnot, kunnia, ylpeys kehittämisestä ja luotettavuus. Organisatorisia kannustimia ovat esimerkiksi lisäoikeudet, uramahdollisuudet ja pääsy tietoon.

Katmada ja muut [2016] esittivät motiiveilensa seuraavat kannustimet. Oppimisen/henkilökohtaisen saavuttamisen kannustimet ovat pääsy tietoon ja palaute asiantuntijoilta tai vertaisilta. Altruismiin kannustimia ovat puolestaan mahdollisuus osallistua hyväntehtävään ja palaute oman työ vaikutuksesta. Nautinnon/älyllisen uteliaisuuden kannustimet ovat pelkkä nauttiminen ja tyytyväisyys toimintaan sekä uusien ihmisten tapaaminen ja uusien paikkojen/tilanteiden löytäminen. Sosiaalisten motiivien kannustimet ovat statuksen hankkiminen, kunnioitus, sosiaalisen kuvan esittäminen ja sosiaalisen tiedon esittäminen, itsensä markkinoinnin kannustin on uran valinta. Epäsuoran työnteon kannustin on niin sanottu passiivinen joukkoistaminen, eli toista tehtävää tehdessä voi tarjota tietoa tai ratkaisuja toisiinkin tehtäviin. Suoran kompensaation kannustimet ovat pieni tai isompi rahallinen palkkio tai ”*poletti*” (token). Myös kannustimissa on nähtävissä yhtäläisyyksiä, ja ne esitetään yhdistetysti taulukon 9 toisessa sarakkeessa.

5.1.3 Mekanismit

Mekanismit kertovat *minkälaisia keinoja motivointiin käytetään*. Katmadan ja muiden [2016] mukaan mekanismit voidaan jakaa neljään kategoriaan: *mainejärjestelmät* (reputation systems), *pelillistäminen* (gamification), *sosiaaliset kannustinjärjestelmät* (social incentive mechanisms) sekä *taloudelliset palkkiot ja uramahdollisuudet* (financial/career rewards). Mainejärjestelmissä keinoja ovat esimerkiksi käyttäjien keräämät pisteet (esimerkiksi vastauksista tai äänestysaktiivisuudesta), erilaiset *suorat* (explicit) palkkiot, hyväksyminen tehtäviä suorittamaan, *merkit* (badge) yhteisöön osallistumisesta, oikeudet hallita sivustoa tai äänestää, erilaiset *epäsuorat* (implicit) palkkiot (kunnioitus yhteisössä, uramahdollisuudet, tietämys omasta uskottavuudesta). Lisäksi keinoina ovat myös erilaiset rangaistukset (estot, julkaisukiellot, tilin jäädytys). Pelillistämisen keinot ovat taitojen kehittäminen, luovuus, itsensä tunteminen taitavaksi, mielekäs tekeminen, *tasot* (level), pisteet, tarrat tai merkit sekä saavutukset. Sosiaalisten kannustinjärjestelmien keinot ovat taitojen esitys ja sosiaalisen statuksen saaminen (toimintoja esimerkiksi kontaktit/ystävät, tapahtumat, viestien ja kehujen vaihto, tutustuminen samankaltaisiin henkilöihin). Taloudellisten palkkioiden ja uramahdollisuuksien keinot ovat halu voittaa, rahallinen palkkio, vapaa pääsy palveluihin ja tuotteisiin, bonukset hyvästä työstä sekä oman työn myynti. Mekanismit on esitetty taulukon 9 kolmannessa sarakkeessa siten, että se kertoo mitä kannustimia ne käyttävät.

Taulukossa 9 on tiivistettynä aiemmin esiteltyt motiivit, kannustimet ja mekanismit luokun ottamatta motiiveissa mainittua epäsuoraa työtä. Epäsuora työ jätettiin pois, koska se on käytännössä sivuvaikutus, eli osallistujat eivät varsinaisesti osallistu toimintaan.

Motiivi	Kannustimet	Mekanismit
Oppiminen / henkilökohtainen saavutus, tiedonkeruu	Pääsy tietoon, palaute asiantuntijoilta tai vertaisilta	Pelillistäminen
Altruismi	Mahdollisuus osallistua hyvään, palaute oman työn vaikutuksesta	Pelillistäminen, sosiaaliset kannustinjärjestelmät
Nauttiminen / älyllinen uteliaisuus	Nauttiminen, tyytyväisyys toimintaan, uusien ihmisten tapaaminen, uusien paikkojen ja tilanteiden löytäminen	Pelillistäminen
Sosiaaliset motiivit	Statuksen hankkiminen, kunnioitus, sosiaalisen kuvan ja tiedon esittäminen, sosiaaliset kannustimet	Sosiaaliset kannustinjärjestelmät, mainejärjestelmät
Itsensä markkinointi / organisatoriset motiivit	Uran valinta, organisatoriset kannustimet	Sosiaaliset kannustinjärjestelmät, mainejärjestelmät
Suora kompensatio / rahalliset motiivit	Rahallinen palkkio, poletti, taloudelliset kannustimet	Mainejärjestelmät, taloudelliset palkkiot ja uramahdollisuudet

Taulukko 9: Motiivit, kannustimet ja mekanismit

5.1.4 Motivointikeinojen valinta

Motivointikeinojen valintaa varten joukkoistajan tulee ymmärtää keitä osallistujat ovat ja mikä heitä motivoi [Župič, 2013]. Nämä seikat vaikuttavat kannustimien ja kannustinmekanismien valintaan ja osittain myös niiden suunnitteluun. Alakohdassa 5.1.3 esitettyjä kannustinmekanismeja suunnitellessa tulee ottaa huomioon erilaisia asioita. Seuraavassa eri keinojen huomioon otettavat asiat ja ratkaisuja niihin.

Mainejärjestelmää suunnitellessa on tärkeää huomioida epäreilusti tuotetut positiiviset tai negatiiviset *mainevääristymät* (reputation bias) [Katmada *et al.*, 2016]. Tällä tarkoitetaan tilanteita, joissa maine muodostuu tahallisen toiminnan perusteella, eli esimerkiksi jokin ratkaisu saa joukolta vain *negatiivisia ääniä* (down-vote) vaikka ratkaisu olisikin hyvä. Näitä tilanteita varten tulee järjestelmässä olla estomekanismeja jotka suodattavat nämä vääristy-

mät pois. Tällaisia mekanismeja ovat esimerkiksi tilastollisten ominaisuuksien tai arvostelijoiden maineen tarkistus sekä *toimialueen tietämykseen* (domain knowledge) perustuva suodatus. Yksi mahdollinen ratkaisu tähän on se, että käyttäjien henkilöllisyydet ovat vain mainejärjestelmän tiedossa ja tahallisten tekojen ja käyttäjien tunnistamiseen käytetään yhteisöllistä moderointia [Katmada *et al.*, 2016]. Lisäksi on tärkeää kannustaa tulokkaita palkitsemalla ensimmäisestä suorituksesta ja rakentaa maine siten, että se kuvaa suoritusten laatua määrän sijasta.

Katmadan ja muiden [2016] mukaan pelillistämistä suunniteltaessa on tärkeintä varmistaa *pistekeskeisyyden* (pointsification) estäminen, eli ettei osallistumisen päätavoitteeksi muodostu pelkkä pisteiden kerääminen. Tämä takia myös tulostaulukoita tulee käyttää harkiten, koska ne voivat lisätä tarpeetonta kilpailua. Ratkaisu tähän on käyttää lyhyen aikavälin tulostaulukoita (esimerkiksi viikoittain nollautuva) tai tulostaulukoita, joita pystyy järjestämään usealla perusteella (esimerkiksi tehdyt tehtävät tai laatu). Yhteistyön lisäämiseksi alustaan voidaan lisätä yhteisöllisiä päämääriä ja sosiaalisia saavutuksia yksittäisen käyttäjän päämäärien lisäksi, jolloin alusta kannustaa yhteistyöhön ja sosiaalisuuteen.

Sosiaalisia kannustinmekanismeja suunniteltaessa tulee Katmadan ja muiden [2016] mukaan päämääränä pitää ystävällisen sosiaalisen kanssakäymisen ja yhteistyön tukemista. Tämä voidaan tehdä lisäämällä erilaisia elementtejä, jotka vahvistavat osallistavaa käytöstä ja mahdollistavat käyttäjien verkostoitumisen (esimerkiksi keskusteluryhmät, sosiaaliset verkot ja toiminnot palautteen antamiseen). Vuorovaikutuksella ja palautteella on suuri vaikutus varsinkin tulokkaille: ensimmäiset vuorovaikutuskokemukset sekä vanhemmilta jäseniltä saatu palaute vaikuttavat tulokkaan sitoutumiseen tuleviin osallistumisiin [Katmada *et al.*, 2016]. Jos tulokkaalle jää niin sanotusti huono maku ensimmäisistä kokemuksista, ei hän ole välttämättä enää innokas osallistumaan jatkossa. Tämän takia alusta tulee toteuttaa niin, että tulokkailla on mahdollisuus sosiaaliseen oppimiseen, eli mahdollisuus oppia havainnoimalla mikä on sopivaa ja sosiaalisesti hyväksyttävää käytöstä. Myös osallistujien tietoturva tulee ottaa huomioon. Yksi keino tähän on tarjota käyttäjille mahdollisuus pitää heidän profiilinsa ja toimintansa yksityisenä ja/tai erillään heidän sosiaalisen median tileistään [Katmada *et al.*, 2016].

Taloudelliset palkkiot ovat erittäin yleisiä mikrotaskauksessa. Taloudellinen palkkio ei kuitenkaan ole aina paras vaihtoehto, koska suorat rahalliset palkkiot vähentävät sisäistä motivaatiota [Katmada *et al.*, 2016] eivätkä ne myöskään välttämättä takaa parempia tuloksia. Amazon Mechanical Turk -alustalla tehdyissä tutkimuksissa on Katmadan ja muiden [2016] sekä Jonesin [2012] mukaan huomattu, että korkeampi rahallinen palkkio houkuttelee enemmän osallistujia ja sitä kautta kasvattaa suoritettujen töiden määrää, mutta ei tulosten laatua ja tarkkuutta. Päinvastaisesti, Kazain ja muiden [2013] mukaan maksu vaikuttaa laatuun positiivisesti parantaen sitä. Korkeampi maksu kannustaa parempaan työskente-

lyyn ja parantaa tarkkuutta. Isompi palkkio saattaa kuitenkin houkuttaa osallistujia huijamaan kasvattaakseen palkkiosummaa [Katmada *et al.*, 2016; Kazai *et al.*, 2013]. Jonesin [2012] mukaan sopivan rahamäärän valitseminen ei ole helppoa. Jos korvaus on liian pieni verrattuna osallistujalta vaadittuun aikaan tai osaamiseen, ei se houkuttele osallistumaan ja jos korvaus on liian suuri, voi se houkuttaa pelkästään palkkion perässä olevia osallistujia [Jones, 2012]. Joukkoistajan täytyykin tarjota sopiva korvaus, joka houkuttelee sopivan tasoisia osallistujia ja motivoi heitä suorittamaan tehtävän hyvin, mutta ei kuitenkaan houkuttele osallistujia, jotka haluavat palkkion suorittamatta tehtävää kunnolla. Kazai ja muut [2013] neuvovatkin, että palkkiota miettiessä tulee ottaa huomioon tehtävään vaadittava vaiva (effort) ja valita palkkio sen mukaan. Toinen keino taloudellisten palkkioiden käyttöön on se, että pieniä rahallisia korvauksia käytetään alustavana motivointina ja sen lisäksi käytetään muita konkreettisia palkkioita (esimerkiksi palkintoja) yhdistettynä pelillistäviin saavutuksiin [Katmada *et al.*, 2016].

Seuraavassa yhteenvetona Katmadaa ja muita [2016] mukaillen suomennettuna ohjeita kannustinmekanismien suunnitteluun.

Mainejärjestelmät

- Toteuta estomekanismeja ei-toivotun ja tahallisen käytöksen varalle.
- Kannusta tulokkaita bonuspisteillä ensimmäisestä kontribuutiosta.
- Peilaa maineessa kontribuutioiden määrän lisäksi myös niiden laatua.
- Älä paljasta kaikkia yksityiskohtia mainealgoritmista ja/tai käyttäjien henkilöllisyyksiä.

Pelillistäminen

- Sisällytä yhteisöllisiä päämääriä ja sosiaalisia saavutuksia, vältä tulostaulukoita.
- Suosi lyhyen aikavälin ja monen järjestelyperusteen tulostaulukoita kannustaaksesi tulokkaita.
- Lisää pelillistäviä elementtejä varovasti ja mieti niiden hyödyllisyyttä välttääksesi pistekeskeisyyden.

Sosiaaliset kannustinmekanismit

- Varaudu osallistujien minäkuvaa koskeviin ongelmiin.
- Ota huomioon osallistujien yksityisyys ja tietoturva.
- Tee sosiaalinen oppiminen, vuorovaikutus ja palautteen saaminen helpoksi.
- Tue tulokkaita erilaisilla käytöksillä ja sitoutumisen tasoilla.
- Lisää sosiaalista vuorovaikutusta ja positiivista palautetta tukevia ominaisuuksia.

Taloudelliset palkkiot ja uramahdollisuudet

- Muista, että maksu voi vähentää sisäistä motivaatiota.
- Toteuta mainekaavio ja kartoita virtuaali-identiteetit yhdistäen ne oikeisiin identiteetteihin.

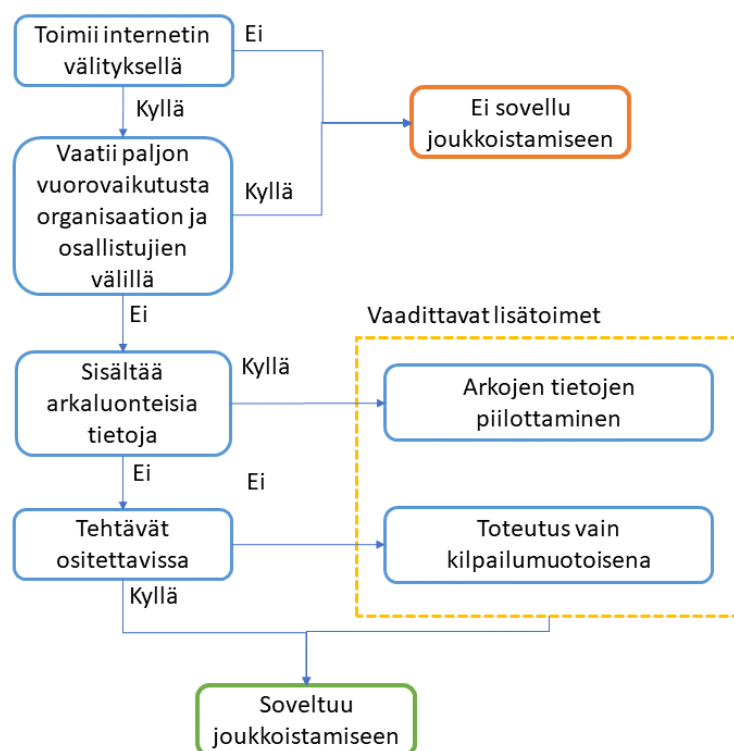
- Käytä pientä rahallista palkkiota alustavana motivaationa ja muita konkreettisia palkkioita ylläpitääksesi pitkän aikavälin sitoutumista.

Muun ohella tehtävien tyypillä on vaikutusta motivointikeinojen valintaan. Hossainin [2012b] mukaan sisäinen motivaatio lisää luovuutta, kun taas ulkoinen motivaatio vähentää sitä eli ulkoinen motivaatio ei ole välttämättä paras valinta luoviin tehtäviin. Käytettävyystestausta ajatellen taloudellinen korvaus voisi olla paras valinta, koska se houkuttelee osallistujia mukaan, mutta summan ollessa maltillinen se houkuttelee tekemään tehtävät hyvin. Lisäksi apuna voisi käyttää Katmadan ja muiden [2016] ohjeiden mukaisesti konkreettisia palkkioita ja saavutuksia. Tällöin testauksella voitaisiin saada runsaasti laadukasta testidataa.

5.2 Tehtävien laatiminen

5.2.1 Sopivuus joukkoistamiseen

Joukkoistamiseen sopii parhaiten tehtävät, jotka pystytään jakamaan helposti pienempiin autonomisiin tehtäviin tai vaiheisiin [Kazai *et al.*, 2013; Rockwell, 2012; Thuan *et al.*, 2013; Alonso ja Baeza-Yates, 2011; Howe, 2009]. Pienet tehtävät ovat usein helpompia ymmärtää ja pienentävät riskiä osallistujan *väsymiseen* (fatigue) [Alonso ja Baeza-Yates, 2011]. Joukkoistamisen kohdalla on tärkeää varmistaa, että annettu tehtävä ja siihen osallistujan tarjoamat vastaukset pystytään jakamaan ja keräämään Internetin välityksellä [Thuan *et al.*, 2013]. Liioin joukkoistaminen ei toimi tehtävissä, joissa vaaditaan paljon vuorovaikutusta organisaation ja osallistujien välillä [Thuan *et al.*, 2013], koska toiminta tapahtuu monesti välikäden eli joukkoistamisalustan kautta. Se ei myöskään toimi suoraan tehtävissä, jotka sisältävät arkaluonteisia tietoja (esimerkiksi yksityisyyteen, tietoturvaan ja älylliseen omaisuuteen liittyvä tieto) [Thuan *et al.*, 2013]. Lisätoimien jälkeen joukkoistamiseen soveltuvat myös tällaiset tehtävät. Kuvassa 4 esitetään suomennettuna Thuanin ja muiden [2013] taulukkoa mukaillen vuokaaviona päätökseen vaikuttavat tehtävien ominaisuudet ja niiden mahdollistamat toiminnat.



Kuva 4: Tehtävien ominaisuuksien vaikutus joukkoistamispäätökseen

5.2.2 Tehtävien rakenne

Tässä alakohdassa kerrotaan mikrotaskauksessa käytettyjen tehtävien rakenteesta. Alakohdassa annettu ohjeistus voi toimia myös muissakin joukkoistamisen lajeissa, mutta niissä voi olla myös muita osioita, joihin en ota tässä kantaa.

Kazain ja muiden [2013] mukaan yksittäinen tehtävä koostuu kolmesta osasta: ohjeet, tehtävä ja kysely. Ohjeet-osa sisältää tehtävän kuvauksen, aihepiirin/aihepiirit sekä mitä pitää tehdä. Tehtävä-osa sisältää suoritettavan tehtävän, tekstikenttä tehtävää koskeville kommentteille sekä mahdollisesti jokin varmistuskysymys tehtävän suorituksesta [Kazai *et al.*, 2013]. Kysely-osassa kerätään osallistujilta taustatietoa (esimerkiksi tausta ja kokemus) sekä palautetta tehtävästä [Kazai *et al.*, 2013]. Lisäksi käyttäjiltä voidaan kysyä kysymyksiä koskien motivaatiota, mielenkiinnon määrää, aiheen *tuttuutta* (familiarity), tehtävän vaikeutta sekä maksun suhdetta ratkaisuun tarvittuun vaivaan. Jones [2012] esitti samankaltaisen listan tehtävän rakenteesta: tehtävä sisältää lyhyen esittelyn, tiedon karsivista vaatimuksista (esimerkiksi vaatimus aiemmin suoritetuista testeistä), suoritukseen annetun ajan, palkkion, erässä/sarjassa olevien tehtävien lukumäärän sekä tiedon siitä, paljonko tähän mennessä on maksettu tuntipalkkaa osittain suoritetusta sarjasta. Myös Alonso ja Baeza-Yates [2011] ovat samoilla linjoilla ja kertovat tehtävän sisältävän selkeän ohjeistuksen, otsikon, kuvauksen tehtävästä sekä avainsanat. Liu ja muut [2012] painottavat, että varsinkin joukkoistetussa käytettävyydestä ohjeet ja tehtävät tulee olla tarkasti kuvattu ja yksiselitteisiä, koska testisession aikana ei välttämättä ole mahdollisuutta tarjota lisäselvennystä niihin. Kazain

ja muiden [2013] mukaan tehtävien laatimista voidaan helpottaa laatimalla pohja, joka sisältää tehtävissä toistuvat tiedot. Tällaisia tietoja ovat osallistujalle tarjottava palkkio, joukkoistajalle aiheutuvat kulut, tehtävää varten annettava suoritusaika sekä tehtävän ominaisuudet eli otsikko, kuvaus ja avainsanat [Kazai *et al.*, 2013]. Edellä mainittujen lähteiden yhteenvedona voidaan siis todeta yksittäisen tehtävän sisältävän seuraavat tiedot:

- otsikko
- tehtävän kuvaus
- aihepiiri tai -piirit (voidaan esittää myös avainsanoina)
- ohjeistus, eli mitä pitää tehdä
- tehtävien lukumäärä
- suoritukseen annettu aika
- maksettava palkkio
- varsinainen tehtävä (lisänä mahdollisuus kommentoida tehtävää).

Edellä mainittujen lisäksi voi olla muitakin tietoja tai osioita (kuten esimerkiksi palauteosio), mutta edellä mainitut ovat mielestäni tehtävän tärkeimmät tiedot, jotka tulee olla esitettynä osallistujalle. Lisäksi myös aihepiiri vaikuttaa tietoihin: esimerkiksi aihepiirin ollessa ohjelmistokehitys tai muu kehitystoiminta, tulee osallistujille tarjota selkeä dokumentaatio, jotta osallistuja tietää mitä ja miten kehitetään [Fitzgerald ja Stol, 2015]. Ohjeita ja muita tietoja laadittaessa tulee pitää mielessä tiedon määrä, eli tietoa ei saa olla tarjolla liikaa, mutta ei myöskään liian vähän [Fitzgerald ja Stol, 2015]. Liika tieto voi karkottaa mahdolliset osallistajat ja liian vähäiset tiedot voi johtaa virheellisiin vastauksiin [Fitzgerald ja Stol, 2015].

5.2.3 Tehtävien suunnittelu ja toteutus

Tehtäviä suunnitellessa tärkein kysymys on *kuinka kysyä oikeat kysymykset parhaalla mahdollisella tavalla* [Alonso ja Baeza-Yates, 2011]. Houkuttaakseen osallistujia tehtävien tulisi olla jollain tavalla houkuttelevia ja kuulostaa mielenkiintoiselle, mutta samalla olla myös rehellisiä [Jones, 2012]. Lisäksi Kitturin ja muiden [2008] mukaan on erittäin tärkeää, että osana tehtävää on varmistettavia kysymyksiä, joihin on selkeä, oikea vastaus. Tällainen kysymys voi olla esimerkiksi verkkosivua arvioitaessa kysymys, johon vastataan kirjoittamalla verkkosivun ensimmäinen tai viimeinen sana, jolloin tiedetään, että osallistuja on oikeasti käynyt sivulla. Kitturin ja muiden [2008] mukaan tällaiset kysymykset johtavat subjektiivisissa tehtävissä johdonmukaisempiin vastauksiin, jotka myös vastasivat paremmin asiantuntijoiden arviointeja tai vastauksia. Lisäksi tehtävät tulisi suunnitella niin, että niiden tekeminen oikein ja niin sanotusti hyväntahtoisesti (eli tarkoitus vastata oikein ja suoriutua hyvin) vaatii yhtä vähän tai vähemmän vaivaa kuin *satunnaisvastauksen* (random answer) antaminen tai vastaaminen tahallaan virheellisesti tai väärin [Kittur *et al.*, 2008]. Jos

testin osana käytetään kyselyitä, tulisi nekin suunnitella siten, että kyselyyn vastattaessa huijaaminen on vaikeaa [Liu *et al.*, 2012].

Kuten perinteisessäkin käytettävyytestauksessa tehdään ensiksi pilottitesti testataksien valitut testausmenetelmät, kysymykset ja muut testissä käytettävät välineet, myös joukkoistamisessa käytettäville tehtäville kannattaa suorittaa ”pilottitesti”. Jonesin [2012] mukaan kannattaa valita pieni erä kysymyksiä ja suorittaa niillä testiajo, ja jos kaikki on kunnossa, voi koko testin julkaista. Kuitenkin käytettävyydestin pilottitestissä testataan aina kaikki tehtävät ja sama käytäntö kannattaa ottaa käyttöön myös testikysymysten kanssa. Kaikki kysymykset tulee testata, jotta tiedetään, että osallistujat ymmärtävät ne oikein ja osaavat tehdä ne annetulla ohjeistuksella. Tämä on joukkoistamisessa erittäin tärkeää, koska osallistujilla ei välttämättä ole mahdollisuutta saada apua, jos he eivät ymmärrä tehtävänantoa. Pilottitestissä tulee Jonesin [2012] mukaan kiinnittää huomiota seuraaviin asioihin:

- Kauanko tehtävien suorittamiseen menee?
- Ymmärtävätkö osallistujat tehtävät?
- Tarvitaanko laadunhallintaa?
- Onko saatava lopputulos sopivaa ja oikeellista?
- Näkyykö järjestelmää hyväksikäyttäviä tai huijaavia osallistujia?
- Millaisia kommentteja ja palautetta saadaan osallistujilta?
- Ovatko osallistujat tyytyväisiä palkkion tasoon?

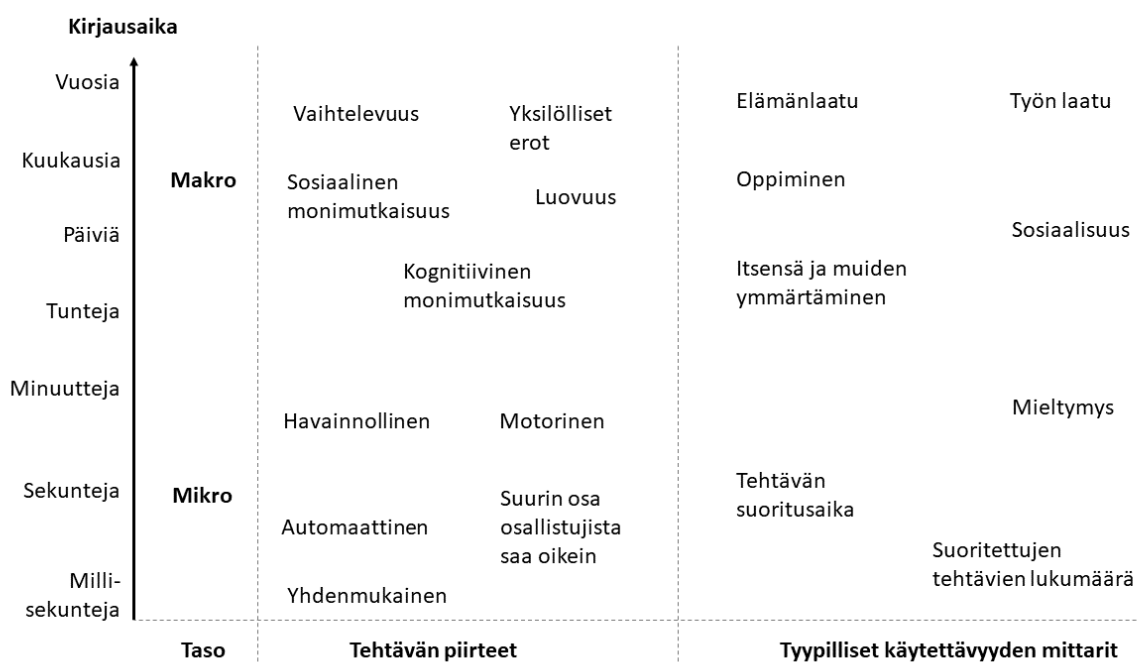
Näiden osa-alueiden perusteella on tarvittaessa helppo tehdä muutoksia tehtäviin ja sen jälkeen ajaa testit uudelleen. Tätä vaihetta kannattaa toistaa, kunnes kaikki on kunnossa ja sen jälkeen siirtyä varsinaisiin testeihin. Testauksessa kannattaa testata pienet tehtävät ennen isompia, koska silloin niiden palautteella voidaan tehdä korjauksia ennen isompia tehtäviä ja näin välttää suuremmat ongelmat [Alonso ja Baeza-Yates, 2011]. Koska joukkoistetua käytettävyydestausta on tutkittu hyvin vähäisesti, ei tutkittua tietoa kysymysten rakenteesta ja toimivuudesta löydy juurikaan. Kysymyksiä laatiessa kannattaakin käyttää apuna perinteiseen käytettävyydestaukseen liittyviä julkaisuja, joissa neuvotaan testitehtävien laatimista.

5.3 Käytettävyyden mittarit ja palaute

Joukkoistamisessa tulosten laatu ei ole yhtä hyvä kuin laboratoriossa tehdyissä käytettävyydestesteissä, mutta osa tärkeistä, kriittisistä käytettävyysongelmista voidaan tunnistaa joukkoistetulla käytettävyydestillä [Liu *et al.*, 2012]. Käytettävyydestin tulosten laatuun ja uskottavuuteen vaikuttavat suuresti käytettävyyden mittarit sekä testin aikana saatava palaute. Tällä tarkoitan sitä, että tulosten täytyy olla luotettavia ja laadukkaita, jotta käytettävyysongelmat löydetään ja pystytään korjaamaan. Kohdassa 4.4 kerrottiin erilaisista yleis-

sesti käytetyistä käytettävyyden mittareista. Monet näistä mittareista (esimerkiksi tehtävään käytetty aika tai virheiden määrä) voi olla vaikeita toteuttaa joukkoistetussa käytettävyydestestauksessa, koska osallistuja suorittaa tehtävät itsekseen eli raportoituja tietoja ei voida helposti varmistaa. Joukkoistetussa käytettävyydestestauksessa myös varmistavien kysymysten laatiminen [Liu *et al.*, 2012] sekä tahallisten väärin vastausten löytäminen voi olla vaikeaa [Kittur *et al.*, 2008], koska käytettävyydestestauksessa ei tehtäviin välttämättä ole yhtä oikeaa vastausta. Tällöin vastauksista on vaikea erottaa, onko väärä vastaus oikeasti väärä vastaus (tahallinen tai tahaton), vai onko kyseessä käytettävyysongelma. Tällöin myös huijarien tunnistaminen on hankalaa [Liu *et al.*, 2012].

Varmistavien kysymysten laatiminen on kuitenkin mahdollista, ja ne voitaisiin sijoitella testitehtävien joukkoon samaan tapaan kuin laboratoriossa tehtävässä käytettävyydestestauksessa. Käytettävyydesti aloitetaan yleensä helpolla tehtävällä, jotta osallistuja rentoutuu ja tutustuu testitehtävien tekemiseen. Testi lopetetaan helppoon kysymykseen, jotta osallistujalle jää onnistumisen tunne ja hyvä mieli testin jälkeen. Varmistuskysymykset voisivat olla samaan tapaan helppoja tehtäviä, kuten esimerkiksi erilaiset laskenta- tai tiedonhakutehtävät testattavassa käyttöliittymässä (esimerkiksi kuvien tai otsikoiden lukumäärä, jonkin osion ensimmäinen tai viimeinen sana tai täydennystehtävä, jossa täytyy täyttää tekstissä olevat aukot käyttöliittymästä löytyvän tiedon avulla). Näiden edellä mainittujen seikkojen takia on mielestäni tärkeää, että joukkoistetulle käytettävyydestestaukselle löydettäisiin sopivat käytettävyyden tavoitteet ja mittarit, jolloin tiedettäisiin, onko tulokset luotettavia, vaikka joukossa näyttäisikin olevan virheellisiä vastauksia. Kuvassa 5 esitetään suomennettuna ja alkuperäistä mukailien uudelleen piirrettynä Hornbækin [2006] esittämä luokittelu mikro- ja makrotason käytettävyydsmittareista. Näitä voidaan mielestäni käyttää apuna mittoja valittaessa, jotta saadaan aihepiiriin sopivat mittarit ja sitä kautta luotettavampia ja laadukkaampia tuloksia.



Kuva 5: Käytettävyyden mikro- ja makromittarit

Kuvassa 5 pystyakselilla on esitetty kirjausaika, eli kuinka pitkän ajan kuluttua tapahtumasta kirjaus tehdään. Vaaka-akselilla on esitetty taso, tehtävän piirteet ja tyypilliset käytettävyyden mittarit. Mikrotason tehtävät ovat yleensä lyhyitä (sekunneista minuutteihin), monimutkaisuus on hallittavissa eli suurin osa osallistujista saa ne oikein, ne keskittyvät havainnollisuuteen ja motorisiin näkökulmiin (esimerkiksi visuaalinen skannaus, syöte hiirellä) ja aika on usein kriittinen resurssi [Hornbæk, 2006]. Makrotason tehtävät ovat yleensä pidempiä (tunteja, päiviä, kuukausia), ovat kognitiivisesti tai sosiaalisesti monimutkaisia (vaativat ongelmanratkaisua, oppimista, kriittistä ajattelua tai yhteistyötä), sisältävät suuria yksilöllisiä eroja vuorovaikutusprosessissa ja lopputuloksissa sekä käyttävät tyytyväisyyttä ja tarkkuutta kriittisinä parametreina.

Hornbæk [2006] suosittelee, että käytettävyydsarvioinnissa käytettäisiin makrotasoa aina kun se on relevanttia ja käytännössä toteuttamiskelpoista. Mielestäni joukkoistaminen jää usein kuitenkin mikrotasolle, koska suoritettavat tehtävät ovat yleensä lyhyitä ja nopeita suorittaa ja tehtävät ovat hyvin samankaltaisia eli yhdenmukaisia. Lisäksi mielestäni mikrotason ehdotetut mitat (mieltymys, tehtävän suoritus aika ja suoritettujen tehtävien lukumäärä) toimivat joukkoistamisessa ja olisivat helppoja toteuttaa. Järjestelmä voisi itse mitata suoritus aikaa, jolloin siitä jäisi pois osallistujan epävarmuus, kun osallistujan ei tarvitse itse seurata aikaa ja raportoida sitä. Lisäksi järjestelmä pitäisi kirjata suoritettujen tehtävien määräästä ja testin lopuksi osallistujalle annettaisiin täytettäväksi kysely, jolla selvitetäisiin mieltymys testattuun käyttöliittymään.

Käytettävyydestin aikana osallistujalta saatu *palaute* (feedback) kuten esimerkiksi mielipiteet ja kommentit, vaikuttaa tulosten laatuun siten, että mielestäni laatu ei ole hyvä, jos saatua palautetta on vähäisesti tai palautteen laatu on heikkoa (esimerkiksi mielipiteenä vain "ok"). Kuten kohdassa 4.5 kerroin, perinteisellä käytettävyydestillä saadaan palautteena muun muassa käyttäjien ajatuksia ja mielipiteitä, erilaista numeerista aineistoa sekä videonauhoite ja siitä laadittu loki, ja lisäksi voidaan kerätä katseenseurannalla saatua dataa. Tärkeä osa laboratoriossa tehtävän käytettävyydestauksen palautetta on ääneenajattelulla saatava tieto, eli mitä käyttäjä ajattelee sillä hetkellä tehtävää tehdessään. Tämän avulla saadaan usein kriittistä tietoa siitä, miksi jokin on käytettävyysongelma (esimerkiksi käyttäjä voi ajatella ääneen etsivänsä vaadittua tietoa jostain tietystä valikosta, koska se hänen mielestään kuuluu sinne). Lisäksi laboratoriossa on helppo kysyä osallistujalta lisätietoa hänen kommentistaan ja mielipiteistään testin aikana. Nämä edellä mainitut seikat vaikuttavat suuresti joukkoistamisen laatuun, koska joukkoistamisessa ei yleensä käytetä ääneen ajattelua eikä osallistujilta ole mahdollista kysyä lisätietoja palautteesta [Liu *et al.*, 2012].

Yksi mahdollinen ratkaisu palauteongelmaan on Hertzumin ja muiden [2015] ehdottama moderoinnaton ääneenajattelu, jolloin osallistuja suorittaa tehtävät kotonaan ja hänen käyttäytymisensä (behaviour) ja verbalisointinsa videoidaan analysointia varten. Moderoinnattomalla ääneenajattelulla saatavan palautteen määrä on pienempi kuin laboratoriossa suoritettujen testien aikana saatu, mutta kuitenkin löydettyjen käytettävyysongelmien määrässä ei ole eroa. Lisäksi moderoinnattomien osallistujien antamasta palautteesta suurempi osuus oli *korkean relevanssin* (high-relevance) verbalisointeja (moderoinnattomat 21% ja moderoidut 11%) [Hertzum *et al.*, 2015]. Tällä ei kuitenkaan ollut vaikutusta löydettyjen käytettävyysongelmien määrään ja itse asiassa moderoinnattomilta saatiin vähemmän tietoa vakavista käytettävyysongelmista. Mielestäni palautteen määrään vaikuttaa kuitenkin se, että laboratoriossa moderaattori usein kannustaa ja muistuttaa osallistujaa ajattelemaan ääneen, kun taas moderoinnattomat ovat "oman onnensa nojassa". Silti, mielestäni moderoinnaton ääneenajattelu on yksi joukkoistamisen tulevaisuudensuunnista varsinkin käytettävyyden arvioinnin saralla. Lisäksi, kun testeistä on videotallenteet, auttaa se myös saamaan lisätietoa palautteesta, kun osallistujaan voi ottaa yhteyttä ja viitata videon kohtaan ja kysyä mitä tässä tapahtui.

Osallistujien lisätieto-ongelmaan puolestaan yksi ratkaisu on tarjota joukkoistamisalustan kautta testille valvoja, johon voi ottaa yhteyttä ja kysyä lisätietoja. Esimerkiksi käyttökokemukseni perusteella Testbirds-alustalla tämä on toteutettu siten, että testin aikana pystyy ottamaan yhteyttä Birdmasteriin eli testin valvojana toimivaan työntekijään sivuston sisäisellä chatilla ja kysyä tehtäviin liittyvistä ongelmista. Tällöin joukkoistamisella saatu vastaus ei olisi enää osallistujan paras arvaus, koska hänellä olisi mahdollisuutta kysyä lisätietoa tehtävästä [Liu *et al.*, 2012].

5.4 Laadunvarmistus

Joukkoistamista on kritisoitu sen lopputuotoksen sekalaisen laadun takia [Kazai *et al.*, 2013]. Syitä laatuongelmiin ovat erilaiset tavat huijata järjestelmää sekä väärinkäytökset. Erityyppinen vandalismi voi tuottaa todellisia ongelmia joukkoistaessa, ja siksi siihen kannattaa valmistautua ja varautua erilaisiin väärinkäytöksiin keinoilla, jotka eivät heikennä osallistumiselle avointa ilmapiiriä [Rockwell, 2012]. Tässä kohdassa kerron, minkälaisilla keinoilla vandalismin voidaan varautua ja täten varmistaa lopputuotoksen laatu. Alakohdassa 5.4.1 kerron erilaisista menetelmistä, joilla laatua voidaan varmistaa ja alakohdassa 5.4.2 esitän yhteenvedon laadunvarmistuksesta sisältäen myös kohdassa 5.3 esitellyt käytettävyyden mittaamisen ja palautteen, koska ne ovat tärkeitä laadun osa-alueita käytettävyydestä.

5.4.1 Laadunvarmistuksen keinot

Laadunvarmistukseen on ehdotettu erilaisia mekanismeja kuten esimerkiksi tulosten tarkastuttaminen asiantuntijoilla, joukon jäsenten käyttäminen arvioinnissa ja kolmannen osapuolen käyttö arvioinnissa [Thuan *et al.*, 2013; Zhao ja Zhu, 2012]. Laadunvarmistuksen keinot voidaan löyhästi jaotella sen mukaan, tehdäänkö varmistusta ennen testiä, testin aikana vai testin jälkeen. Usein onkin hyvä olla useita keinoja havaita epäilyttäviä vastauksia [Kittur *et al.*, 2008], jolloin varmistusta voidaan tehdä useamassa vaiheessa ja keinoja voidaan valita tai yhdistellä tavoitteiden, tehtävien ominaisuuksien ja kerättävän palautteen mukaan [Zhao ja Zhu, 2012]. Taulukossa 10 on esitelty eri keinoja sekä ajankohdat, jolloin kutakin keinoa voidaan käyttää.

Keino	Ajankohdat
Ansakysymykset (esimerkiksi HoneyPot, Golden Unit ja ELICE)	Testin aikana
Suodatus (esimerkiksi osallistujien tai vastausten)	Ennen testiä, testin aikana ja testin jälkeen
Valintakokeet / esivalinta	Ennen testiä
Tehtävien ominaisuudet (esimerkiksi palkkion määrä, tehtävien muodostaminen ja lisätiedot)	Ennen testiä, testin aikana ja testin jälkeen
Muut keinot (esimerkiksi vastauksen tärkeys ja bonukset)	Ennen testiä, testin aikana ja testin jälkeen

Taulukko 10: Laadunvarmistuksen keinot

Ansakysymykset (trap questions) tai *karsintakysymykset* (qualification questions) ovat kysymyksiä, joihin on olemassa varmistettava, oikea vastaus. Näitä kysymyksiä ripotellaan

satunnaisesti testitehtävien joukkoon ja jos osallistuja ei vastaa määritellyyn lukumäärään kysymyksiä oikein, voidaan hänen luokitella huijaavan ja poistettavaksi testistä [Nguyen *et al.*, 2013; Alonso ja Baeza-Yates, 2011]. Menetelmiä jotka käyttävät ansakysymyksiä ovat esimerkiksi Honeypot, Golden Unit sekä ELICE (Expert Label Injected Crowd Estimation) [Nguyen *et al.*, 2013; Jones, 2012]. ELICE eroaa kahdesta ensin mainitusta sillä, että siinä arvioidaan osallistujan asiantuntijuuden tasoa mittaamalla osallistujan vastausten suhdetta oikeisiin vastauksiin [Nguyen *et al.*, 2013].

Ansakysymykset voidaan valita suhteellisen vapaasti aihepiirin mukaan, ainoa vaatimus on oikea, varmistettava vastaus. Kitturin ja muiden [2008] mukaan esimerkiksi artikkeleiden luokittelutehtävissä ansakysymyksiä voivat olla esimerkiksi viittausten, kuvien tai osioiden lukumäärä sekä lisäksi voidaan vaatia 4-6 avainsanaa artikkelille. Liu ja muut [2012] puolestaan käyttivät verkkosivun arvioinnissa ansakysymyksinä täydennystehtäviä, jotka velvoittivat osallistujia menemään testattavalle verkkosivulle etsimään tietoa vastatakseen tehtävään ja päästäkseen jatkamaan testausta. *Huijarien* (spammers) tunnistaminen on tärkeää laadunvarmistuksessa [Jones, 2012] ja varmistavat kysymykset auttavat tähän, eli kuten jo alakohdassa 5.2.2 kerrottiin, varmistavia kysymyksiä on hyvä olla ja ne johtavat johdonmukaisempiin vastauksiin.

Suodatusta käytetään apuna, kun halutaan suodattaa mahdolliset huijaavat osallistajat sekä virheelliset vastaukset pois. Osallistujien suodatuksessa käytetään usein apuna joukkoistamisalustan mainejärjestelmää, jolloin voidaan helposti suodattaa ennen testiä pois osallistujat, jotka tekevät huonoa työtä (esimerkiksi keskenjääneitä testejä tai huonoja tuloksia aiemmista testeistä). Tämä mahdollistaa sen, että osallistujiksi saadaan henkilöitä, joiden tiedetään tekevän hyvää ja tarkkaa työtä [Kazai *et al.*, 2013]. Yksi tällainen mitta on AMT:ssa (Amazon Mechanical Turk) käytetty *hyväksymisprosentti* (approval rate), joka kertoo, kuinka suuri osuus osallistujan tekemistä töistä on hyväksytty [Alonso ja Baeza-Yates, 2011]. Tehtäviä voidaan puolestaan suodattaa esimerkiksi tehtäviin käytetyn ajan mukaisesti sekä vertaamalla suoritettujen varmistusten määrää vaadittujen varmistusten määrään [Kazai *et al.*, 2013]. Tehtävien suodatusta voidaan tehdä testien aikana, jolloin lopputuloksiin ei tule virheellisiä vastauksia, tai testin jälkeen vastauksia läpikäydessä. Kazain ja muiden [2013] mukaan tehtävien jälkeistä laadunvalvontaa voidaan suorittaa tunnistamalla *kohinaa* (noise) ja *harhoja* (bias), analysoimalla käytöstä, kehittämällä *työntekijämalleja* (worker models) sekä käyttämällä *luottamusalgoritmeja* (trust algorithms) ja *suodatusta* (spam filters). Kohinalla tarkoitetaan tilanteita, joissa satunnainen vaihtelu tai virhe vaikuttaa tuloksiin (tavanomaisesta poikkeavat vastaukset) ja harhoilla tarkoitetaan tilanteita, joissa systemaattiset virheet vaikuttavat tuloksiin (vastausten keskiarvo poikkeaa suuremman keskiarvosta, johtuu valitusta mittarista) [UIAH, 2007]. Näitä kaikkia edellä mainittuja keinoja voidaan käyttää apuna suodattamassa virheellisiä vastauksia pois, jolloin lopputuloksesta tulee laadukkaampi.

Valintakokeita tai esivalintaa käytetään osallistujien suodatukseen, kun mukaan halutaan osallistujia, joilla on jo tietty osaaminen (esimerkiksi osaa jo käyttää alustaa, on jo tietämystä aiheesta tai tehnyt samankaltaisia testejä aiemmin). AMT esimerkiksi tarjoaa kaksi tapaa tähän tarkoitukseen: esisuodatus tilastojen perusteella ja alustava harjoittelu tai valintakoe [Kazai *et al.*, 2013]. Nämä tilastot sisältävät tietoa esimerkiksi suoritettujen tehtävien määrästä ja hyväksymisprosentista, ja näiden tietojen perusteella joukkoistaja voi muodostaa suodatuskriteerit ja täten määritellä valittavat osallistujat [Kazai *et al.*, 2013]. Valintakoe sisältää usein testitehtävien kaltaisia kysymyksiä, joihin täytyy vastata oikein saadakseen hyväksynnän osallistua [Alonso ja Baeza-Yates, 2011]. Tämä parantaa arvostelun tasoa. Lisäksi valintakokeen tarkoitus on usein opettaa myös joukkoistamisalustan käyttöä esittelemällä ominaisuuksia, joita osallistuja käyttää valintakoetta tehdessä. Valintakoetta tai esivalintaa suunniteltaessa on tärkeää miettiä vaatimustasoa, koska jos taso on liian korkea, ei osallistujia välttämättä löydy tarpeeksi. Liian alhaisella tasolla laatu voi kärsiä, kun osallistujat eivät esimerkiksi ymmärrä aihepiiriä [Jones, 2012]. Valintakoe voidaan esittää osallistujille myös harjoitustehtävänä, jolloin sen tekeminen tuntuu vapaaehtoiselle ja osallistujat voivat olla motivoituneempia sen suorittamiseen.

Muun ohella tehtävien ominaisuuksilla on vaikutusta laatuun. Esimerkiksi tehtävien hinnoittelu tulisi toteuttaa siten, että tehtävästä maksetaan siihen vaaditun *vaiivan* (effort) mukaisesti [Kazai *et al.*, 2013]. Yksi laatua parantava keino on tarjota lisäpalkkiota hyödyllisestä palautteesta, jolloin oikeaa, kehityksen kannalta arvokasta palautetta antava osallistuja saa lisäpalkkion kiitoksena palautteestaan [Liu *et al.*, 2012]. Tämä voi kannustaa osallistujaa jatkossakin antamaan hyödyllistä palautetta ja voi saada muut osallistujat pyrkimään samaan saadakseen lisäpalkkion. Hinnoittelussa tulee myös ottaa huomioon se, että jos maksaa osallistujille lisää, saa työn tehtyä nopeammin, muttei välttämättä paremmin [Alonso ja Baeza-Yates, 2011]. Jossain tapauksissa maksu voi kuitenkin lisätä laatua [Kazai *et al.*, 2013]. Eli kuten jo kohdassa 5.1.4 kerrottiin, kannattaa palkkion taso pitää sopivana.

Hinnoittelun lisäksi ohjeet ja dokumentaatio ovat myös tärkeitä laadun kannalta. Jos haluaa laadukasta lopputulosta, ohjeiden ja dokumentaation tulee olla laadukasta ja selkeästi esitettyä [Alonso ja Baeza-Yates, 2011]. Kuten kohdassa 5.2.3 kerrottiin, laadun parantamiseksi tehtävät tulee suunnitella siten, että niiden tekeminen oikein vaatii vähemmän vaihua kuin tahallaan väärin tekeminen [Kittur *et al.*, 2008]. Yksi ratkaisu tähän on pyytää osallistujia kirjoittamaan lyhyt selitys tehtävästä. Alonso ja Baeza-Yates [2011] pyysivät tutkimuksessaan osallistujia kirjoittamaan vapaaehtoisesti lyhyen selityksen suorittamastaan tehtävästä. Lopputuloksena he saivat kerättyä enemmän dataa ja tunnistivat tehokkaammin huijareita, jolloin tulosten laatu parani.

Muita keinoja laadun varmistamiseen ovat esimerkiksi osallistujien informointi tulosten käytöstä ja tärkeydestä. Jos käyttäjille kerrotaan heti testin aluksi, mihin tuloksia käytetään

ja kuinka tärkeitä ne ovat, voi se motivoida osallistujia tekemään parempaa työtä ja sitä kautta laatu paranee [Liu *et al.*, 2012].

5.4.2 Yhteenveto laadunvarmistuksesta

Yhteenvetona voidaan todeta, että testitehtävien joukkoon kannattaa sijoittaa varmistavia kysymyksiä (esimerkiksi täydennystehtävä, johon etsitään tietoa testattavasta käyttöliittymästä tai joidenkin käyttöliittymän elementtien lukumäärä). Lisäksi tulee miettiä osallistujien taso eli tarvitaanko esivalintaa, suunnitella tehtävät ja ohjeistukset huolella, kertoa mihin tuloksia käytetään ja maksaa osallistujille tehtävään vaaditun vaivan mukaisesti. Kannattaa myöskin harkita lisäbonuksen tarjoamista hyvästä työstä. Käytettävyydestä laatuun vaikuttavat myös käytettävyyssarvioinnin tavoitteet ja saatu palaute (Katso kohta 5.3). Käytettävyyden mittaukseen voi käyttää perinteisiä mittareita (esimerkiksi tehtävään kulunut aika, tehtyjen tehtävien lukumäärä, virheiden määrä) ja testin lopuksi tehtävää tyytyväisyyslomaketta. Palautteen laadun kannalta kannattaa harkita moderoimattoman ääneenajattelun toteuttamista.

5.5 Osallistajat ja rekrytointi

Jonesin [2012] mukaan osallistujia rekrytoidessa kannattaa kiinnittää huomiota osallistujien sopivuuteen tehtäviin. Yksi keino tähän on varmistaa osallistujien taso valintakokeella. Kuten alakohdassa 5.3.1 jo kerrottiin, kannattaa valintaa tehdessä pitää vaatimustaso kohtuullisena saadakseen sopivia osallistujia mukaan. Joukkoistamisen ja myöskin rekrytoinnin onnistumisen kannalta molempien osapuolien maine on tärkeää [Jones, 2012]. Rekrytoinnissa onnistuvat todennäköisesti osallistajat ja joukkoistajat, joilla on vakiintunut, vahva maine. Jos joukkoistajalla on huono maine (esimerkiksi alhainen palkkio tai ei maksa suorituksista ollenkaan), eivät osallistajat halua tehdä hänen tehtäviään, ja jos osallistujalla on huono maine, ei joukkoistaja halua ottaa häntä mukaan.

Rockwell [2012] listasi tutkimuksessaan viisi vinkkiä, joita voidaan käyttää rekrytoinnin apuna. Ensimmäinen vinkki on, että pyydä apua eli pyydä osallistujilta auttamaan ja ole valmis kertomaan, millaista apua tarvitaan [Rockwell, 2012]. Tällä vedotaan osallistujien auttamishaluun ja se voi innostaa osallistujia paremmin kuin pelkkä pyyntö suorittaa tehtävät. Toinen vinkki koskee long tail-efektin tukemista. Epätasapaino on normaalia joukkoistamisessa, eli usein pieni osa joukosta tekee paljon tehtäviä ja iso osa vähän tehtäviä, ja kaikkia osallistujia pitää tukea [Rockwell, 2012]. Joukkoistamisalustoilla osallistajat voivat olla esimerkiksi työnantajia, käyttäjiä/työntekijöitä, johtajia ja suunnittelijoita, ja heidän roolinsa työelämässä vaikuttavat toimintaan ja käytökseen [Zhao ja Zhu, 2012]. Kolmas vinkki koskee erilaisuuden tukemista: jokainen osallistuja on erilainen ja alustan tulisi toimia siten, että heidän erovaisuutensa eivät huku yhteistyössä [Rockwell, 2012]. Tällä tarkoitetaan sitä, että jos osallistujilta vaaditaan liikaa yhteistyötä, voi mielipiteiden yhdistäminen yhdeksi

vastaukseksi poistaa osallistujien väliset mielipide-erot, jolloin vastaus ei edusta kaikkia osallistujia. Neljäs vinkki on tunnista yhteisö eli joukkoistaminen voi luoda yhteisön ja sitä tulisi tukea sopivilla kommunikointikeinoilla [Rockwell, 2012]. Yhteisön tunnistaminen ja tukeminen ovat tärkeitä myös senkin takia, että yhteisöllisyyden tunne on yksi motivaation lähde. Viimeinen vinkki koskee joukon kokoa. Rockwell [2012] muistuttaa, että myös pienet joukot toimivat eli menestys ei riipu isosta osallistujamäärästä. Tällä tarkoitetaan sitä, että pieni ahkera joukko voi toimia paremmin kuin iso laiska joukko, koska isosta joukosta ei ole hyötyä, jos kukaan ei tee töitä tai työn laatu on heikkoa.

5.6 Joukkoistajan rooli

Joukkoistajan on tärkeä tuntea joukkoistamisprosessi, jotta ohjelmistosuunnittelijat, kehittäjät ja projektipäälliköt voivat valmistautua ja keskustella sujuvan vuorovaikutuksen vaatimuksista [Fitzgerald ja Stol, 2015]. Rockwellin [2012] mukaan tämä tietämys täytyy ajatella prosessina tuotteen sijasta, eli sen kerääminen alkaa vähästä tiedosta, joka lisääntyy ja tarkentuu ajan myötä. Joukkoistamisen järjestämiseen ja suorittamiseen tarvitaan tiimi [Rockwell, 2012] ja joukkoistaminen onnistuu, jos organisaatiolla on sopivaa asiantuntijuutta ja kokemusta aiheesta [Thuan *et al.*, 2013]. Nämä seikat aiheuttavat sen, että esimerkiksi ohjelmistokehitykseen tulee muutama uusi rooli [Fitzgerald ja Stol, 2015] liittyen vuorovaikutuksen järjestämiseen ja alustan toimintaan. Bruun ja Stagen [2015] mukaan joukkoistettujen käytettävyyssarvioinnit perustuvat minimalistiseen loppukäyttäjien koulutukseen, jonka pohjalta he osaavat suorittaa arviointeja ja raportoida ongelmia etänä. Koulutuksen ollessa vain käyttäjille, ei joukkoistettu käytettävyydestä sovellu käytettäväksi konteksteissa, joissa ei ole HCI-osaamista, koska siihen tarvitaan loppukäyttäjien koulutuksen sekä tulosten analysoinnin osaava tiimi [Bruun ja Stage, 2015].

Toinen joukkoistajan rooliin liittyvä osa-alue on osallistujien tukeminen. Osallistujilla voi usein olla kysymyksiä joukkoistamiseen liittyen, mikä voi olla aikaa vievää [Fitzgerald & Stol, 2015]. Ratkaisu tähän on asettaa työntekijä vastaamaan joukon kysymyksiin, jolloin osallistuja saa vastauksen henkilöltä, joka on perehtynyt teknologiaan ja prosessiin ja tarjoaa tarkempia vastauksia kuin yleistasonen UKK-palsta [Fitzgerald & Stol, 2015]. Osallistujan tukemista on myös organisaation toimiminen omalla nimellä tehtävänänoissa. Joukkoistava organisaatio voi päättää piiloutua anonymiteetin taakse ja käyttää salanimeä kuvauksessa yrittäessään mielestään esimerkiksi suojella yrityssalaisuuksiaan [Fitzgerald & Stol, 2015]. Tällöin joukkoistamisprojekti ei kuitenkaan välttämättä kiinnitä osallistujien huomiota eikä ole houkutteleva, jolloin osallistujat eivät lähde mukaan [Fitzgerald & Stol, 2015]. Lisäksi joukkoa tulisi kohdella *kumppanina* (partner) ja joukon tarpeet, tavoitteet ja motivaatiot olisi otettava huomioon [Zhao ja Zhu, 2012].

Joukkoistajan tulee myös ottaa huomioon resurssien kuluminen sekä joukkoistamista kohtaan olevat odotukset. Organisaatioille päätös joukkoistaa on merkittävä, koska epäonnistuminen tuhlaa resursseja [Thuan *et al.*, 2013]. Monelle organisaatiolle joukkoistaminen on vielä tuntematon käsite, eli siitä ei ole aiempaa kokemusta. Tällöin joukkoistaminen voi tuntua niin sanotusti ”hypyltä tuntemattomaan” ja mietitään, onko se sen arvoista eli onko saatava hyöty suurempi kuin mahdollinen riski? Toinen joukkoistamispäätökseen vaikuttava tekijä on odotukset. Joukkoistamiselta odotetaan usein hyötynä korkealaatuisia ja innovatiivisia ratkaisuja nopeasti ja alhaisilla kustannuksilla [Fitzgerald & Stol, 2015]. Näin ei kuitenkaan aina ole, koska osallistujia voi olla vähäisesti tai vastausten taso voi olla heikko. Fitzgerald ja Stol [2015] ohjeistavatkin, että älä odota ihmeitä joukkoistaessa. Tällä tarkoitetaan sitä, että vaikka joukkoistaminen pienentääkin kuluja, ei se silti poista vaatimuksia asiantuntijuudesta ja tee yhtäkkiä testauksesta ongelmatonta ja melkein ilmaista. Fitzgerald ja Stol [2015] neuvovatkin olemaan aliarvioimatta kuluja. Testausbudjettia miettiessä voisikin olla parempi ajatella joukkoistamisen tarjoavan paremmin vastinetta rahalle sen sijasta, että ajattelee sen suoraan säästävän rahaa.

Viimeinen joukkoistajan roolin osa-alue on lopputulosten ja riskien hallinta. Kuten kaikissa projekteissa, myös joukkoistamisprojekteissa tulee riskien kartoituksen ja riskinhallinnan olla kunnossa, jotta järjestäminen onnistuu [Thuan *et al.*, 2013]. Esimerkiksi sitoutumisen puute organisaation ja osallistujien välillä luo riskejä [Thuan *et al.*, 2013]. Omat riskinsä luo lopputulosten hyväksyminen, koska päätös lopputuotoksen hyväksymisestä tai hylkäämisestä tehdään usein hyvin nopealla aikataululla eikä laatu välttämättä ole odotettua tasoa [Fitzgerald ja Stol, 2015]. Jos organisaatiolle muodostuu maine vastauksia helposti hylkäävänä, eivät osallistujat halua jatkossa osallistua, koska he pelkäävät käyttävänsä turhaan aikaansa tehtäviin saamatta palkkiota [Fitzgerald ja Stol, 2015], mikä lisää joukkoistamisen epäonnistumisen riskiä. Fitzgerald ja Stol [2015] kehottavat olemaan hylkäämättä vastauksia, jos ne ovat helposti korjattavissa eli hyväksymään käyttäjiltä tulevat vastaukset sellaisenaan ja korjaamaan pienet viat ennen ratkaisun käyttöönottoa. Tämä neuvo ei toimi kuitenkaan käytettävyydestänsä kanssa, koska kaikki käytettävyysongelma tulee aina korjata ennen ratkaisun käyttöönottoa.

5.7 Joukkoistamisalustan valinta

Joukkoistamisalustoista ei ole tehty paljoa tutkimusta varsinkaan joukkoistetun käytettävyydestänsä kannalta. Tämän takia on vaikea antaa suosituksia, ja siksi tässä kohdassa esitetään vaan yleisiä ohjeita koskien joukkoistamisalustoja. Erilaisia joukkoistamisalustoja on olemassa runsaasti ja usein ei kannata kuluttaa resursseja ja ruveta niin sanotusti ”keksimään pyörää uudelleen”, vaan käyttää olemassa olevia työkaluja ja/tai muokata niitä [Rockwell, 2012]. Lisäksi Thuanin ja muiden [2013] mukaan jo saatavilla oleva joukkoista-

misalusta voi pienentää kehityskustannuksia, mikä puoltaa joukkoistamisen valintaa. Olemassa olevan alustan saatavuus on tärkeä osa-alue joukkoistamispäätöksessä erityisesti pienille ja keskikokoisille yrityksille, joilla on vähemmän resursseja ja heikompi tekninen osaaminen. Näiden lisäksi sopiva, jo olemassa oleva alusta on arvokas jäsenten saatavuuden takia [Thuan *et al.*, 2013], koska tällöin alustalla on jo oma joukkonsa eikä rekrytointia tarvitse aloittaa niin sanotusti tyhjästä. Joukkoistamisalustalla valmiina oleva joukko ei kuitenkaan välttämättä ole juuri testin kohderyhmää, mikä on käytettävyydestä kannalta erittäin huono asia. Jos käytettävyydestä ei tehdä oikealla kohderyhmällä, eivät tulokset auta parantamaan käyttöliittymää vastaamaan kohderyhmän tarpeita.

Toinen alustan valinnassa huomioonotettava asia on sen käytettävyys. Osallistujat tuottavat virheellistä tietoa, jos tehtävät ja käyttöliittymä ovat huonolaatuisia, esimerkiksi jos tehtävät ovat epäselviä ja ohjeet monitulkintaisia sekä lomakkeet rajaavat syötteitä ja asteikot vääristävät vastauksia [Kazai *et al.*, 2013]. Nämä täytyy aina määrittää arviointiprojektikohtaisesti, koska eri projekteissa voi olla esimerkiksi käytössä erilaiset asteikot. Alonso ja Baeza-Yates [2011] kiteyttävät, että käyttöliittymän tulisi tehdä tehtävistä helpompia niiden vaikeuttamisen sijasta. Käyttöliittymän tulisi olla houkutteleva ja intuitiivinen sekä esittää tulokkaille eri tavat osallistua, ja osallistumisen hyödyt selkeästi ja perusteellisesti. Lisäksi annettavan palautteen tulisi olla asiaankuuluvaa (esimerkiksi eteneminen, kannustaminen ja osallistumisen ylläpito) [Katmada *et al.*, 2016]. Lisäksi käyttäjäystävällinen käyttöliittymä voi houkutella enemmän osallistujia ja täten kasvattaa korkealaatuisen lopputuloksen todennäköisyyttä [Gomide *et al.*, 2014].

6 Pohdinta ja kootut suositukset

6.1 Rajoitteet

Pro graduani lukiessa tulee ottaa huomioon, että monet taustamateriaaleistani olivat eri aihepiireistä ja tieteenaloista, kuten esimerkiksi kirjallisuudesta ja informaatiotieteistä, eli niiden sisältämän tiedon paikkansapitävyyttä ja toimivuutta joukkoistetussa käytettävyytestauksessa ei ole varmistettu. Joukkoistamisesta itsestään löytyy suhteellisen hyvin materiaalia, mutta ei kuitenkaan tietojenkäsittelyn, ihmisen ja teknologian vuorovaikutuksen tai käytettävyyden näkökulmista. Tämä vaikutti tietysti tuloksiin jättäen paljon avoimia kysymyksiä ja samalla myös vaikeutti pro gradun tekemistä. Oman haasteensa tekemiseen toi myöskin vähäiset ennakkotiedot joukkoistamisesta, eli aiheen ymmärtämiseen vaadittiin paljon opettelua. Materiaaleissa esitetyistä tutkimuksista osa oli esimerkiksi kirjallisen materiaalien luokittelua, eli testitehtävissä osallistujille annettiin katkelma jostain materiaalista ja osallistujien tuli kuvailla materiaali ja antaa sille avainsanoja. Näihin tehtäviin annettu ohjeistus ei auta laatimaan käytettävyytestaukseen sopivia testitehtäviä, koska vaatimukset ovat erilaiset. On ihan erilainen tehtävä kuvata jokin materiaali kertomalla sen aihepiiri ja sisältö tiivistetysti kuin testata esimerkiksi lomakkeen käytettävyys.

Myös materiaaleille asettamillani rajauksilla oli varmasti vaikutusta lopputuloksiin. Rajauksieni perusteella minulla oli mukana vain materiaaleja, joissa puhuttiin joko joukkoistamisesta tai joukkoistamisesta ja käytettävyydestä. Tällöin saattoi jäädä pois materiaaleja, joissa puhuttiin esimerkiksi etätestauksesta ja joukkoistamisesta. Näissä materiaaleissa olisi voinut olla pro graduni kannalta hyödyllistä tietoa, koska etätestaus on toimintatavoiltaan hyvin lähellä joukkoistamista. Rajasin pois myös materiaaleja, joiden tieteenala eroaa kovasti tietojenkäsittelystä (esimerkiksi lääketiede). Tällöin sain rajattua aiheitani, mutta koska muillakin aloilla tietotekniikan käyttö on lisääntymässä, olisi näissäkin materiaaleissa voinut olla aiheeni kannalta relevanttia tietoa. Materiaalien valintaan vaikutti lisäksi aiheen tuntemattomuus minulle, eli osa materiaaleista kertoo hyvin yleistasoisesti mitä joukkoistaminen on, koska minun täytyi perehdyttää itsenikin aiheeseen. Näiden lisäksi materiaalien valintaan vaikutti myös Andor-hakupalvelu. Materiaaleja hakiessa Andor oli juuri tullut käyttöön ja ei toiminut mielestäni täysin oikein, eli välillä tasan samoilla hakuparametreilla tuli erilaiset hakutulokset.

Motivaationi tähän aiheeseen oli se, että käytettävyys jätetään usein taka-alalle ja sen ajatellaan olevan vain isojen yritysten juttu. Halusin perehtyä aiheeseen nähdäkseni, olisiko joukkoistaminen ratkaisu tähän eli saataisiinko sen avulla käytettävyyssarvioinnin hintaa alaspäin ja suosiota ylöspäin. Olemassa olevan tiedon perusteella joukkoistetun käytettävyytestauksen hyödyllisyys on kuitenkin tulkinnanvarainen. Kuten aiemmin kerroin, joukkoistaminen tekee kyllä testauksesta helpompaa ja halvempaa, mutta se ei silti poista tarvetta asiantuntijuuteen. Joukkoistamisen kannalta on tärkeää, että joukkoistaja tuntee

joukkoistamisprosessin, ja tulosten tulkinnessa tarvitaan puolestaan osaamista käytettävyydestä. Jos asiaa ajatellaan pienten ja keskikokoisten yritysten kannalta, ei joukkoistaminen siis suuresti helpota tilannetta. Tähän asti käytettävyydestä on tehty pieniä muotoisesti tai jätetty tekemättä käytettävyyden osaamisen tai rahan puutteen takia ja nyt pitäisi vielä lisäksi olla osaamista joukkoistamisesta. Joukkoistaminen tuskin lisää suoraan intoa lähteä testaamaan käytettävyyttä. Ratkaisu tähän olisi mielestäni joukkoistamisalusta, joka opastaisi joukkoistajaa tehtävien muodostamiseen ja käytettävyyden mittaamiseen. Alusta voisi olla käytettävyyden asiantuntijoiden ylläpitämä ja heillä tulisi myös olla tietämystä ja osaamista joukkoistamisesta. Tällöin alusta olisi helppokäyttöinen ja ohjeistaisi joukkoistajaa tehtävien muodostamisessa ja tulosten tulkinnessa. Joukkoistamisalustan tulisi olla mahdollisimman automaattinen, eli joukkoistajan tarvitsisi periaatteessa tietää vain testattava tuote ja kohderyhmä, ja alusta hoitaisi loput. Tällaisen toteuttaminen vaatisi kuitenkin sen, että joukkoistettua käytettävyydestä tutkittaisiin kunnolla ja saataisiin sen toimivuus varmistettua sekä laadittua ohjeistus siihen. Tämän jälkeen olisi helpompi miettiä sen sovellettavuus pienille ja keskikokoisille yrityksille.

6.2 Lisätutkimus

Taustamateriaalini perusteella on selvää, että joukkoistamisessa on runsaasti avoimia lisätutkimusta vaativia kysymyksiä varsinkin käytettävyydestä kannalta. Esimerkiksi testitehtävistä tarvitsisi tutkia millaiset kysymykset toimisivat, miten ne tulisi muotoilla ja minkä tyyppisiä kysymyksiä voidaan esittää. Lisäksi tulisi tutkia varmistavia kysymyksiä eli missä väleissä niitä esitetään, kuinka paljon ja minkälaisia niiden tulisi olla sekä auttaisiko laadunvarmistukseen tehtävän lopuksi tehtävä kysely. Taustamateriaalini tutkimuksissa oli toteutettu tällaisia kyselyitä, joissa kysyttiin muun muassa tehtävän vaikeutta ja aiheen tuttuutta. Materiaalini perusteella on siis viitteitä siitä, että kyselyiden toteutus onnistuisi olemassa olevilla joukkoistamisalustoilla. Käytettävyydestä kyselyssä voitaisiin kysellä esimerkiksi tehtävän helppoutta ja aiheen kannalta hyödyllistä taustatietoa. Koska tietoa tehtävistä on niin vähän tarjolla, kannattaa joukkoistettua käytettävyydestä suunnittelevan ottaa avuksi perinteisen käytettävyydestä ja etätestauksen teokset, muodostaa testitehtävät niiden ohjeiden mukaisesti ja katsoa, minkälaisia vastauksia niillä saa pilottitestissä, ja muokata sen jälkeen tehtäviä tarpeen mukaan.

Yksi tutkittava asia on myös käytettävyyden mittarit. Taustamateriaaleissani ei ollut yhtään tutkimusta, joissa olisi kerrottu joukkoistetun käytettävyydestä käytettävyyden mittareista, eli sillä saralla riittää myös tutkittavaa. Kohdassa 5.3 kerroin käytettävyyden mikro- ja makrotasoisista ja siitä, miten joukkoistettu käytettävyyden arviointi jää usein mikrotasolle. Hornbækin [2006] mukaan esimerkiksi oppimiseen ja mieleen palautukseen tulisi kiinnittää enemmän huomiota ja muuttaa mittareita niin, että nämä näkökulmat tulisivat paremmin esille. Käytettävyydestä on vaarana se, että testauksessa etsitään vain selviä virheitä

sen sijasta, että otettaisiin huomioon oppiminen ja mieleen palautus. Tähän vaikuttaa osittain myös rekrytoinnin työläys, eli testiosallistujiksi ei välttämättä saada sekä noviisikäyttäjiä, joiden testin kautta nähtäisiin oppimisen helppous, että kokeneempia tai toista kertaa osallistuvia käyttäjiä, joiden suorituksiin voitaisiin noviiseja verrata. Kohdassa 5.3 neuvoin joukkoistajia käyttämään perinteisen käytettävyyden mittareita ja niillä on mielestäni hyvä aloittaa. Aiheesta tarvitaan kuitenkin tutkimusta, jotta nähtäisiin, toimivatko nämä mittarit oikeasti joukkoistettuun käytettävyydestäukseen vai tarvitsisiko se omat mittansa.

Käytettävyyden kannalta myös osallistujilta saatava palaute on tärkeää, kuten aiemmin jo kerroin. Tälläkin saralla on paljon tutkittavaa, kuten minkälaista palautetta joukkoistamisella voidaan kerätä, miten palautetta kannattaa pyytää sekä minkälainen palaute on hyödyllistä käytettävyyden kannalta. Yksi tutkimista vaativa osa-alue on moderoimaton ääneenajattelu. Taustamateriaalien perusteella se voisi olla yksi ratkaisu palautteen alhaiseen tasoon, mutta sen toteutus ja toimivuus tarvitsisi lisää tutkimusta. Tutkittavaa olisi muun muassa, että toimiiko se oikeasti, pystyykö sitä toteuttamaan olemassa olevilla alustoilla, suostuvatko osallistujat siihen sekä onko siitä saatava hyöty suurempi kuin sen vaatima lisävaiva.

Kuten aiemmissa kohdissa mainitsin, olisi tärkeää selvittää, onnistuuko edellä mainittujen asioiden toteutus olemassa olevilla alustoilla. Lisäksi tulisi kartoittaa, mitä joukkoistamisalustoja ylipäänsä on olemassa ja toimisivatko ne joukkoistettuun käytettävyydestäukseen. Toimivuuden kannalta tärkeää olisi kartoittaa minkälaiset tehtävät alustan kautta onnistuvat sekä erityisesti, millaisia osallistujia alustalla on ja kuinka ominaisuuksien perusteella voi suodattaa testiosallistujiksi sopivat. Käytettävyydestäuksessa testin kohderyhmä on tärkeä ottaa huomioon. Testattavalla käyttöliittymällä on oma käyttäjäryhmänsä ja testaus tulisi suorittaa tähän ryhmään kuuluvilla henkilöillä. Materiaaleissani kerrottiin kyllä osallistujien suodattamisesta joukkoistamisalustan mainejärjestelmän perusteella, mutta ei muiden osatekijöiden perusteella. Perinteisessä käytettävyydestäuksessa osallistujien rekrytointia varten laaditaan persoonia, jotka esittävät käyttöliittymän erilaisia käyttäjiä ja niitä käytetään apuna kohderyhmän edustuksen varmistamiseen. Samankaltaista toimintoa tarvittaisiin myös joukkoistamisalustoihin, jotta saataisiin varmistettua kohderyhmä myös joukkoistetussa käytettävyydestäuksessa. Käytettävyydestäuksen kannalta tämä on tärkeä kysymys ja olemassa olevissa alustoissa tulisikin tarkastella, että onnistuuko niissä tällainen suodatus. Kaikkien näiden edellä mainittujen asioiden selvittämiseksi tulisi laatia tutkimus, jossa suoritettaisiin käytettävyydestä samoilla tehtävillä sekä käytettävyydelaboratoriossa että joukkoistamisalustalla, jotta näitä kahta voitaisiin verrata kunnolla toisiinsa ja nähtäisiin, toimiiko joukkoistaminen käytettävyyden arviointiin.

6.3 Ehdotetut suositukset joukkoistetun käytettävyydestin järjestämiseen

Tässä kohdassa esittelen suosituksia joukkoistetun käytettävyydestauksen järjestämiseen mikrotaskauksella. Ne on muodostettu perustuen lukuun 5 sekä Howen [2009] esittämien ohjeiden perusteella. Howen [2009] laatimat ohjeet on esitetty suomennettuna lyhyiden esitelyiden kera liitteenä 2.

Suosituksukset ovat kronologisessa järjestyksessä sen mukaan, mitkä asiat pitää olla tehtynä ennen seuraavaan kohtaan etenemistä. Poikkeuksena ovat suositukset 6-9, jotka ovat rinnasteisia keinoja riskien hallintaan, eli niitä toteutetaan samanaikaisesti.

1. Valitse sopivat käytettävyydestavoitteet ja mittarit.

Yleisesti käytetyt mittarit sopivat myös joukkoistettuun käytettävyydestaukseen. Mittoja esimerkiksi tehtävään kulunut aika, löydettyjen käytettävyydestongelmien lukumäärä, suoritettujen tehtävien lukumäärä.

2. Suunnittele tehtävät huolellisesti.

Tee tehtävistä lyhyitä ja yksinkertaisia, sekä sellaisia, että niiden tekeminen oikein vaatii yhtä vähän tai vähemmän vaivaa kuin tahallaan väärin tekeminen. Lisäksi tee tehtävänannoista selkeitä ja ohjeista tehtävät huolellisesti ja yksiselitteisesti.

3. Tee pilottitesti.

Suorita pilottitesti osalla testikysymyksistä, tee tarvittavat korjaukset ja julkaise koko testi vasta sen jälkeen.

4. Järjestä valintakoe, joka on läpäistävä ennen testiä.

Sisällytä kysymyksiä, jotka esittelevät joukkoistamisalustan toimintoja ja raportointiohjeita sekä kysymyksiä, jotka koskevat käytettävyydestongelmia ja miltä niiden esiintyminen voi näyttää. Näillä keinoilla varmistetaan, että testin alkaessa alusta, raportointitavat sekä konteksti ovat osallistujalle tutut.

5. Ota riskit huomioon ja varaudu niihin.

Kuten kaikissa projekteissa, myös joukkoistamisprojekteissa riskienhallinta on tärkeää. Riskejä voidaan pienentää useilla toimilla, kuten esimerkiksi motivoinnilla, valintakokeilla ja ansakysymyksillä.

6. Motivoi pienellä rahallisilla palkkioilla ja sitouta konkreettisilla palkkioilla.

Rahallinen palkkio houkuttaa osallistujia, mutta pieni summa ei kuitenkaan houkuta pelkästään rahan perässä olevia osallistujia. Maksa kuitenkin osallistujille tehtävään vaaditun vaivan perusteella.

Sitouta osallistujat pysymään mukana konkreettisilla palkinnoilla (esimerkiksi tuotepalkinnot).

Joukkoistaminen toimii parhaiten, kun joukkoistaja antaa joukolle jotain mitä se tahtoo.

7. Varmista laatu testin aikana ansakysymyksillä.

Sisällytä testiin ansakysymyksiä, joilla varmistetaan, että osallistuja ei vaan arvaa vastauksia vaan on oikeasti tutustunut testattavaan käyttöliittymään. Ansakysymyksiä voi kysyä esimerkiksi joidenkin elementtien kuten kuvien lukumäärää, verkkosivun ensimmäistä tai viimeistä sanaa tai muita varmistettavia tietoja.

8. Kerää palautetta mahdollisimman laajasti.

Pyydä käyttäjiä kertomaan tehtävän suoritukseen liittyviä tietoja, kuten esimerkiksi kuvaus tehtävästä, kuljettu reitti ja tehtävän koettu helppous. Sijoita tehtäväsivulle kenttä, johon käyttäjä voi kirjoittaa vapaasti palautetta, kuten huomioita ja mielipiteitä tehtävää koskien.

Jos mahdollista, käytä myös moderoimatonta ääneenajattelua.

9. Tue osallistujia.

Aseta joukkoistamiseen ja tekniikkaan perehtynyt henkilö vastaamaan osallistujien kysymyksiin.

Ole läpinäkyvä eli esitä testin kuvauksessa joukkoistajan (henkilö, yritys tai organisaatio) oikea nimi. Tämä lisää luottamusta ja kannustaa osallistumaan.

10. Älä odota ihmeitä.

Joukkoistaminenkin tarvitsee resursseja ja tietämystä, se ei muuta testausta ilmaiseksi ja helpoksi.

Joukkoistamisen tuloksena odotetaan usein korkealaatuisia ja innovatiivisia ratkaisuja, mutta näin ei aina käy.

7 Yhteenveto

Joukkoistamisessa yksilö, yritys tai organisaatio laatii tehtävät ja jakaa ne joukkoistamisalustan kautta tuntemattomalle joukolle ihmisiä, jotka suorittavat tehtävät pientä palkkiota vastaan. Joukkoistamisella saavutetaan ehkä useita hyötyjä, kuten esimerkiksi rahan ja ajan säästö tai luovia, uusia ja laadukkaampia ratkaisuja, voidaan kartoittaa uusia markkinoita ja saadaan tuotteet nopeammin myyntiin. Joukkoistaminen myöskin tarjoaa osallistujille mahdollisuuksia jakaa tietämystä, saada työkokemusta ja näkyvyyttä sekä löytää työpaikkoja. Lisäksi se voi parantaa tiedonjakoa yrityksen ja asiakkaiden välillä. Joukkoistamisen toteuttamisessa on kuitenkin omat haasteensa, mutta ne voidaan voittaa kattavalla suunnittelulla. Seuraavaksi esittelen joukkoistetun käytettävyydestä hyötyjä ja haasteita antamalla vastaukset tutkimuskysymyksiini.

Pro graduni tutkimuskysymykset olivat:

1. Soveltuuko joukkoistaminen käytettävyyden arviointiin?
2. Tekeekö joukkoistaminen käytettävyyden arvioinnista helpompaa ja halvempaa?
3. Ovatko joukkoistamisella saadut tulokset yhtä laadukkaita kuin perinteisillä arviointimenetelmillä saadut tulokset?

Joukkoistaminen soveltuu käytettävyyden arviointiin, mutta siinä on monia huomioonotettavia asioita. Tällaisia ovat muun muassa joukkoistamisalustan valinta, osallistujien valinta, tehtävien rakenne sekä laadunvarmistus. Erilaisia joukkoistamisalustoja on olemassa runsaasti ja usein kannattaakin valita valmis alusta ja tarvittaessa muokata sitä sen sijasta, että käyttää resursseja oman alustan tekemiseen. Joukkoistamisalustoilla on yleensä valmiiksi käyttäjiä eli osallistujia, mikä helpottaa ja nopeuttaa rekrytointia. Joukkoistaessa osallistujat valitaan alustan käyttäjistä suodattamalla jollain perusteella, kuten esimerkiksi suoritettujen tehtävien lukumäärällä. Varmistaakseen, että osallistujat osaavat suorittaa testitehtävät, tehtävien tulee olla selkeitä, ja rakenteeseen kannattaakin hakea apua perinteisen käytettävyydestä kirjallisuudesta. Lisäksi ohjeistuksen tulee olla selkeää, yksityiskohtaista ja huolellisesti laadittua.

Joukkoistaminen tekee käytettävyyden arvioinnista helpompaa ja halvempaa, koska se pienentää erilaisia kuluja, kuten esimerkiksi tila- ja laitevuokrat, palkat sekä erilaiset kehittämisestä aiheutuvat kulut. Lisäksi se auttaa säästämään aikaa vähentämällä valmisteluun tarvittavaa aikaa yhtä osallistujaa kohden sekä vähentämällä turhaa ajankäyttöä tilanteissa, joissa ratkaisut löytyvät valmiina osallistujilta. Joukkoistaminen siis tekee testin järjestämisestä vaivattomampaa niin järjestäjän kuin osallistujankin näkökannalta. Muun lisäksi joukkoistaminen myös nostaa otoskokoja ja sitä kautta tilastollista luotettavuutta. Tiivistetysti joukkoistettu käytettävyydestä testaus helpottaa osallistujien rekrytointia, säästää aikaa ja rahaa sekä tekee testauksesta joustavampaa ja helpottaa etätestausta.

Joukkoistamisella saatavat tulokset eivät ole suoraan yhtä laadukkaita kuin perinteisillä menetelmillä saatavat, mutta aihetta ei ole tutkittu paljoa. Tulosten laadukkuuteen vaikuttavat joukkoistamalla saatavan palautteen vähäisyys sekä laadunvarmistuksen hankaluus. Laboratoriossa tehdyllä käytettävyydestillä saadaan runsaasti tietoa testattavasta tuotteesta. Erilaisien numeeristen tietojen (esimerkiksi tehtäviin käytetty aika, virheiden tai klikkausten määrä) lisäksi saadaan tietoa siitä, mitä osallistuja ajattelee (ääneenajattelu), mitä mieltä osallistuja on sekä testisession videonauhoite ja siitä laadittu loki. Lisäksi tietoa voidaan kerätä katseenseurannalla, jolloin saadaan selville, mihin osallistujan katse kohdistuu ja miten hän etsii tarvitsemaansa tietoa. Joukkoistetussa käytettävyydestissä osallistujien ajatukset ja mielipiteet jäävät usein saamatta, koska käyttäjä suorittaa testin omalla ajallaan ja omalla laitteistollaan ilman moderaattorin läsnäoloa. Joukkoistamisella kerättävä tieto jääkin usein vain numeeriselle tasolle, ja sen perusteella voi olla hankalaa tunnistaa käytettävyysoongelmia. Joukkoistamisen ongelmana on myöskin se, että osallistujalta ei voida kysyä tarkentavia kysymyksiä eikä osallistuja voi kysyä apua sitä tarvitessaan. Ratkaisuksi näihin ongelmiin on ehdotettu moderoimatonta ääneenajattelua, eli osallistujan testisuoritus nauhoitetaan ja osallistuja ajattelee ääneen tehtäviä suorittaessaan. Tällöin suorituksesta saadaan numeerisen tiedon lisäksi videonauhoituksen kautta käyttäjän ajatuksia ja mielipiteitä. Videonauhoituksen mukaanotto tuo mukanaan kysymyksiä käyttäjän yksityisyydestä. Perinteisessä käytettävyydestäuksessa osallistujalle kerrotaan, että testisessio videoidaan, mitä videolle tallentuu, ja pyydetään osallistujalta kirjallisesti lupa videointiin. Nämä vaiheet tulisi myös toteuttaa joukkoistamiseen, jos osallistujaa tai hänen suoritustaan videoidaan.

Laboratoriossa suoritettavassa käytettävyydestissä laadunvarmistus on helpompaa, koska osallistuja on koko ajan moderaattorin seurassa ja osallistujalta pystytään kysymään tarvittaessa selvennystä kommenttiin. Joukkoistamisessa laadunvarmistus täytyy tehdä varmistavien kysymysten ja kattavan ohjeistuksen avulla. Kysymyksillä varmistetaan, että osallistujat ovat tutustuneet testattavaan tuotteeseen ja suorittavat tehtävät hyvántahtoisesti, eli eivät anna vain satunnaista tai väärää vastausta palkkion saadakseen. Varmistavat kysymykset voivat olla esimerkiksi käyttöliittymän elementtien laskemista (kuvien tai vastavien lukumäärä) tai täydennystehtäviä, joihin tieto haetaan testaavasta käyttöliittymästä. Nämä voidaan sijoitella testitehtävien joukkoon samaan tapaan kuin perinteisessäkin testissä eli yleensä testin aluksi ja lopuksi annettaviksi helpommiksi tehtäviksi.

Viiteluettelo

[Afuah, 2011] Allan Afuah. 2011. Crowdsourcing and open innovation. In: Allan Afuah, *Business Model Innovation: Concepts, Analysis, and Cases*, Taylor and Francis, 63-83.

[Albert *et al.*, 2010] Bill Albert, Tom Tullis and Donna Tedesco. 2010. *Beyond the Usability Lab: Conducting Large-scale Online User Experience Studies*. Morgan Kaufmann Publishers Inc.

[Alonso ja Baeza-Yates, 2011] Omar Alonso and Ricardo Baeza-Yates. 2011. Design and implementation of relevance assessments using crowdsourcing. In: *Proc. of the 33rd European conference on Advances in Information Retrieval (ECIR 2011), Lecture Notes in Computer Science 6611*. Springer, 153-164.

[Anttonen, 2005] Jenni Anttonen. 2005. Osallistujien valinta. Saila Ovaska *et al.* (toim.) *Käytettävyystutkimuksen menetelmät*, 283-298. Tampereen yliopisto, Tietojenkäsittelyn laitos B-2005-1.

[Bastien, 2010] J. M. Christian Bastien. 2010. Usability testing: a review of some methodological and technical aspects of the method. *International Journal of Medical Informatics*, 79, 18-23.

[Bruun ja Stage, 2015] Anders Bruun and Jan Stage. 2015. New approaches to usability evaluation: Barefoot and crowdsourcing. *The Journal of Systems and Software*, 105, 40-53.

[Carta *et al.*, 2011] Tonio Carta, Fabio Paterno and Vagner Figueredo de Santana. 2011. Web usability probe: a tool for supporting remote usability evaluation of web sites. In: *Proc. of Human-Computer Interaction – INTERACT 2011, Lecture Notes in Computer Science 6949*. Springer, 349-357.

[Chynal ja Sobecki, 2015] Piotr Chynal and Janusz Sobecki. 2015. Web-systems remote usability tests and their participant recruitment. In: *Proc. of Human-Computer Interaction: Design and Evaluation (HCI 2015), Lecture Notes in Computer Science 9169*. Springer, 175-183.

[Fitzgerald ja Stol, 2015] Brian Fitzgerald and Klaas-Jan Stol. 2015. The dos and don'ts of crowdsourcing software development. In: *Proc. of SOFSEM 2015: Theory and Practice of Computer Science, Lecture Notes in Computer Science 8939*. Springer, 58-64.

[Gomide *et al.*, 2014] Victor H. M. Gomide, Pedro A. Valle, José O. Ferreira, José R. G. Barbosa, Adson F. da Rocha and Talles M. G. de A. Barbosa. 2014. Affective crowdsourcing applied to usability testing. *International Journal of Computer Science and Information Technologies*, 5, 1, 575-579.

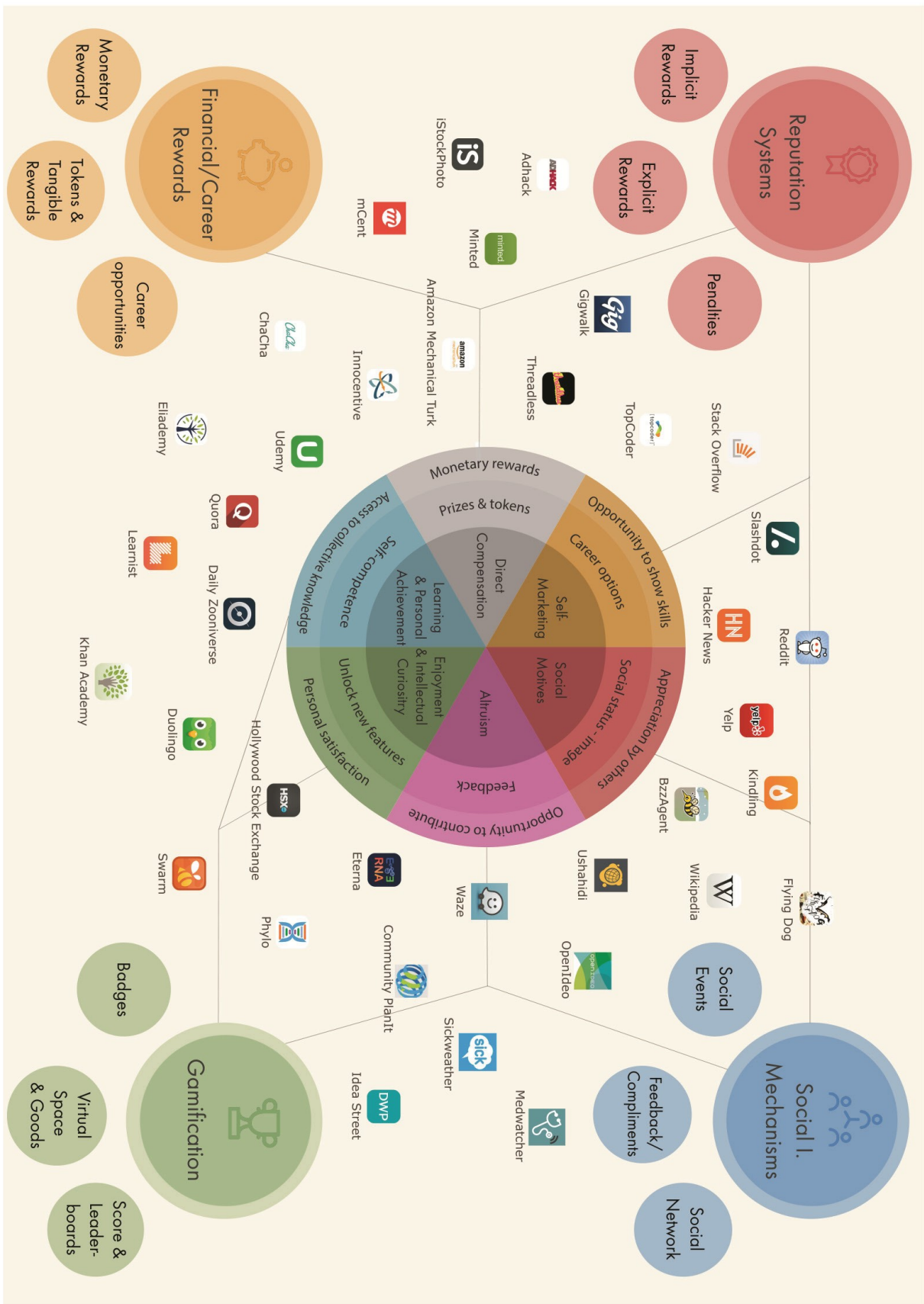
[Hammon ja Hippner, 2012] Larissa Hammon and Hajo Hippner. 2012. Crowdsourcing. *Business & Information Systems Engineering*, 4, 3, 163-166.

- [Heimonen, 2005] Tomi Heimonen. 2005. Käytettävyyden automaattinen arviointi. Saila Ovaska *et al.* (toim.) *Käytettävyyystutkimuksen menetelmät*, 169-186. Tampereen yliopisto, Tietojenkäsittelyn laitos B-2005-1.
- [Hertzum *et al.*, 2015] Morten Hertzum, Pia Borlum and Kristina B. Kristoffersen. 2015. What do think-aloud participants say? a comparison of moderated and unmoderated usability sessions. *International Journal of Human-Computer Interaction*, 31, 9, 557-570.
- [Hornbæk, 2006] Kasper Hornbæk. 2006. Current practice in measuring usability: challenges to usability studies and research. *Int. J. Human-Computer Studies*, 64, 79-102.
- [Hossain, 2012a] Mokter Hossain. 2012. Crowdsourcing: activities, incentives and users' motivation to participate. In: *Proc. of International Conference on Innovation, Management and Technology Research (ICIMTR2012)*, 501-506.
- [Hossain, 2012b] Mokter Hossain. 2012. Users' motivation to participate in online crowdsourcing platforms. In: *Proc. of the International Conference on Innovation, Management and Technology Research (ICIMTR2012)*, 310-315.
- [Hossain ja Kauranen, 2015] Mokter Hossain and Ilkka Kauranen. 2015. Crowdsourcing: a comprehensive literature review. *Strategic Outsourcing: An International Journal*, 8, 1, 2-22.
- [Howe, 2009] Jeff Howe. 2009. *Crowdsourcing: How the Power of the Crowd is Driving the Future*. Cornerstone.
- [Ilves, 2005] Mirja Ilves. 2005. Ääneenajattelu. Saila Ovaska *et al.* (toim.) *Käytettävyyystutkimuksen menetelmät*, 209-222. Tampereen yliopisto, Tietojenkäsittelyn laitos B-2005-1.
- [ISO, 1998] International Organization for Standardization. 1998. Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability. <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en>, Haettu: 15.05.2017
- [Jones, 2012] Gareth J.F. Jones. 2012. An introduction to crowdsourcing for language and multimedia technology research. In: Maristella Agosti *et al.* (eds.), *Information Retrieval Meets Information Visualization, Lecture Notes in Computer Science 7757*. Springer, 132-154.
- [Katmada *et al.*, 2016] Aikaterini Katmada, Anna Satsiou and Ioannis Kompatsiaris. 2016. Incentive mechanisms for crowdsourcing platforms. In: *Proc. of Internet Science (INSCI 2016), Lecture Notes in Computer Science 9934*. Springer, 3-18.
- [Kazai *et al.*, 2013] Gabriella Kazai, Jaap Kamps and Natasa Milic-Frayling. 2013. An analysis of human factors and label accuracy in crowdsourcing relevance judgements. *Inf. Retrieval*, 16, 138-178.

- [Kittur *et al.*, 2008] Aniket Kittur, Ed H. Chi and Bongwon Suh. 2008. Crowdsourcing user studies with mechanical turk. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, 453-456.
- [Koiranen *et al.*, 2016] Ilkka Koiranen, Pekka Räsänen ja Caj Södergård. 2016. Mitä digitalisaatio on tarkoittanut kansalaisen näkökulmasta? *Talous ja yhteiskunta*, 6, 24-29.
- [Korvenranta, 2005] Heta Korvenranta. 2005. Asiantuntija-arvioinnit. Saila Ovaska *et al.* (toim.) *Käytettävyystutkimuksen menetelmät*, 111-124. Tampereen yliopisto, Tietojenkäsittelyn laitos B-2005-1.
- [Koskinen, 2005] Joni Koskinen. 2005. Käytettävyystestaus. Saila Ovaska *et al.* (toim.) *Käytettävyystutkimuksen menetelmät*, 187-208. Tampereen yliopisto, Tietojenkäsittelyn laitos B-2005-1.
- [Lehtinen, 2005] Merja Lehtinen. 2005. Katseenseuranta. Saila Ovaska *et al.* (toim.) *Käytettävyystutkimuksen menetelmät*, 223-236. Tampereen yliopisto, Tietojenkäsittelyn laitos B-2005-1.
- [Liu *et al.*, 2012] Di Liu, Randolph G. Bias, Matthew Lease and Rebecca Kuipers. 2012. Crowdsourcing for usability testing. In: *Proc. of the American Society for Information Science and Technology*, 49, 1, 1-10.
- [Nielsen, 1993] Jacob Nielsen. 1993. *Usability Engineering*. Academic Press, Inc.
- [Nguyen *et al.*, 2013] Quoc Viet Hung Nguyen, Thanh Tam Nguyen, Ngoc Tran Lam and Karl Aberer. 2013. An evaluation of aggregation techniques in crowdsourcing. In: *Proc. of Web Information Systems Engineering (WISE2013), Lecture Notes in Computer Science 8181*. Springer, 1-15.
- [Ovaska *et al.*, 2005] Saila Ovaska, Anne Aula ja Päivi Majaranta. 2005. Johdatus käytettävyystutkimukseen. Saila Ovaska *et al.* (toim.) *Käytettävyystutkimuksen menetelmät*, 1-16. Tampereen yliopisto, Tietojenkäsittelyn laitos B-2005-1.
- [Rockwell, 2012] Geoffrey Rockwell. 2012. Crowdsourcing and humanities: social research and collaboration. In: Willard Mccarthy and Marilyn Deegan (eds.), *Digital Research in the Arts and Humanities: Collaborative Research in the Digital Humanities*. Routledge. 135-154.
- [Romanovs *et al.*, 2014] Andrejs Romanovs, Boris V. Sokolov, Arnis Lektauers, Semyon Potryasaev and Viacheslav Shkoryrev. 2014. Crowdsourcing interactive technology for natural-technical objects integrated monitoring. In: *Proc. of Speech and Computer (SPECOM 2014), Lecture Notes in Computer Science 8773*. Springer, 176-183.

- [Simula, 2012] Henri Simula. 2012. The rise and fall of crowdsourcing? In: *Proc. of 46th Hawaii International Conference on System Sciences (HICSS2013)*, 2783-2791.
- [Spool ja Schroeder, 2001] Jared Spool and Will Schroeder. 2012. Test web sites: five users is nowhere enough. In: *Proc. of CHI EA '01 Extended Abstracts on Human Factors in Computing Systems*. ACM, 285-286.
- [TaY, 2016] Tampereen yliopiston kirjasto. 2016. Andor-opas: Suomi, <http://libguides.uta.fi/andorguide/>, Haettu 21.10.2016.
- [Thuan *et al.*, 2013] Nguyen Hoang Thuan, Pedro Antunes and David Johnstone. 2013. Factors influencing the decision to crowdsource. *Information System Frontiers*, 18, issue 1, 47-68.
- [UIAH, 2007] University of Art and Design Helsinki. 2007. Kurssimateriaali. Tuote ja tieto: tuotteiden tutkimus ja kehittäminen, Tietojen rekisteröinti, <http://www.uiah.fi/projects/metodi/054.htm>, Haettu 07.04.2017.
- [Zhao ja Zhu, 2012] Yuxiang Zhao and Quighua Zhu. 2012. Evaluation on crowdsourcing research: current status and future directions. *Information System Frontiers*, 16, issue 3, 417-434.
- [Župič, 2013] Ivan Župič. 2013. Social media as enabler of crowdsourcing. In: Miguel R. Olivás-Lujan and Tanya Bondarouk (eds.), *Social Media in Human Resources Management*, Emerald Group Publishing Limited, 243-255.

Liite 1: Käyttäjien motiivit, kannustimet ja kannustinmekanismit [Katmada et al., 2016]



Liite 2: Jeff Howen [2009] ohjeet joukkoistamiseen

1. **Valitse oikea malli** (Pick the right model): Joukkoistaminen on sateenvarjokäsite useille eri malleille. Valitse näistä malleista sopivin päämääräsi mukaan.
2. **Valitse oikea joukko** (Pick the right crowd): Muotoile viestisi tarkoituksen mukaan ja jaa sitä oikeiden kanavien kautta löytääksesi oikeat osallistujat.
3. **Tarjoa oikeat kannustimet** (Offer the right incentives): Saadakseen ihmiset mukaan on tiedettävä, mikä heitä motivoi ja osallistujien on tunnettava, että heitä palkitaan käytetyn väkivallan mukaan, vaikka palkkio olisikin pieni.
4. **Pidä pinkit tohvelit laatikossa** (Keep the pink slippers in the drawer): Kuulostaa houkuttelevalle teettää työ ilmaiseksi joukolla, kuin maksaa työntekijälle sen tekemisestä. Kuitenkaan joukkoistaminen ei ole niin helppoa, vaan tarvitaan työntekijöitä esimerkiksi käymään läpi osallistujien vastauksia.
5. **Joukkojen tyhmyys vai hyväntahtoinen diktaattori -periaate** (The dumbness of crowds, or the benevolent dictator principle): Osallistujat tarvitsevat ohjeita ja ohjausta sekä jonkun vastaamaan heidän kysymyksiinsä. Usein yhteistyötä vaativassa joukkoistamisessa joku joukosta ottaa tämän tehtävän, koska yhteisö tarvitsee aina johtajan.
6. **Pidä yksinkertaisena ja hajota osiin** (Keep it simple and break it down): Joukkoistamisessa jokainen tekemisen arvoinen tehtävä on myös pienempiin osiin jakamisen arvoinen, jolloin tehtävät ovat selkeitä ja yksinkertaisia.
7. **Muista Sturgeonin laki** (Remember Sturgeon's law): Suurin osa joukkoistamisella saatavista tuloksista ei ole haluttua laatutasoa, eli Sturgeonin lain mukaan kaikesta yhdeksänkymmentä prosenttia on roskaa (90 percent of everything is crap).
8. **Muista kymmenen prosenttia, eli Sturgeonin lain vastalääke** (Remember the 10 percent, the antidote to Sturgeon's law): Joukko korjaa itse taipumustaan hukuttaa Internettiä lukuisilla huonolaatuisilla tuotoksilla, käytä sitä apuna löytääksesi vastauksista parhaimmat.
9. **Yhteisö on aina oikeassa** (The community's always right): Kohdassa 5 mainitun diktaattorin auktoriteetti on moraalinen ja suostutteleva, mutta ei ehdoton, eli voit yrittää ohjata yhteisöä, mutta lopulta päädyt kuitenkin seuraamaan sitä.
10. **Älä kysy mitä joukko voi tehdä sinun hyväksesi, mutta mitä sinä voit tehdä joukon hyväksi** (Ask not what the crowd can do for you, but what you can do for the crowd): Joukkoistaminen toimii parhaiten, kun joukkoistaja antaa joukolle jotain mitä se tarvitsee. Ihmisillä on tapana osallistua, kun heillä on joku tarve täytettävänä (psykologinen, sosiaalinen tai emotionaalinen) ja jos tarve ei täyty, he eivät osallistu.