

Käytettävyyden arviointi ja käytettävyydestauksen soveltaminen terveydenhuollon tietojärjestelmien valinnassa

Janne Pitkänen, DI^{1,2}, Matti Pitkäranta, DI¹

¹ Adusso Oy, Riihimäki, ² Aalto-yliopiston tietotekniikan laitos (käytettävyydestutkimus), Espoo

Janne Pitkänen, DI, Adusso Oy, 11130 Riihimäki. Sähköposti: janne.pitkanen@adusso.com, janne.p.pitkanen@aalto.fi

Abstract

Procurement of a Healthcare Information System is a significant investment and its successful implementation is of a great interest while good overall economic efficiency is the target. This principle of increasing the efficiency of the use of public funds is also stated in the Act on Public Contracts, which is to be followed when Healthcare Information Systems are bought for public healthcare. Successful information system procurement necessitates the alternative products to be compared according their realistic utility and this can be enhanced by measuring usability. We have trialed out the user-triggered usability testing for means of procurement and development of Healthcare Information Systems. The trials suggest that the traceability of usability related remarks is sufficient with the used testing arrangements. For exploiting the usability test results for comparison, we have created an optimized analysis model for more commensurate measures.

Keywords: information systems, hospital information systems (MeSH), usability, usability testing, benchmarking, public procurement

Tiivistelmä

Terveydenhuollon tietojärjestelmän hankinta on merkittävä investointi, jonka onnistuminen halutaan varmistaa ja tehdä kokonaistaloudellisesti hyvä ratkaisu. Julkisten varojen käytön tehostaminen ja laadukkaiden hankintojen edistäminen on keskeisenä tavoitteena myös hankintalaissa, joka koskee tietojärjestelmähankintoja julkisen terveydenhuollon tapauksessa. Järjestelmähankinnan onnistuminen edellyttää tuotteiden todellista käyttökelpoisuutta edustavaa vertailua, mitä voidaan edesauttaa käytettävyyden mittaamisella. Olemme kokeilleet käyttäjäherätteen käytettävyydestauksen järjestelyjä terveydenhuollon tietojärjestelmähankinnan ja kehityksen tarpeisiin. Pilottitestikokeilun tuloksena on testausjärjestelyn tuottamien havaintojen jäljitettävyyden osoitettu toimivaksi ja testien tuottaman vertailukelpoisuuden optimoimiseksi esitetään siihen soveltuva analysointimalli.

Avainsanat: tietojärjestelmät, sairaalan tietojärjestelmät (FinMeSH), käytettävyys, testaus, benchmarking, julkiset hankinnat (YSA)

Johdanto

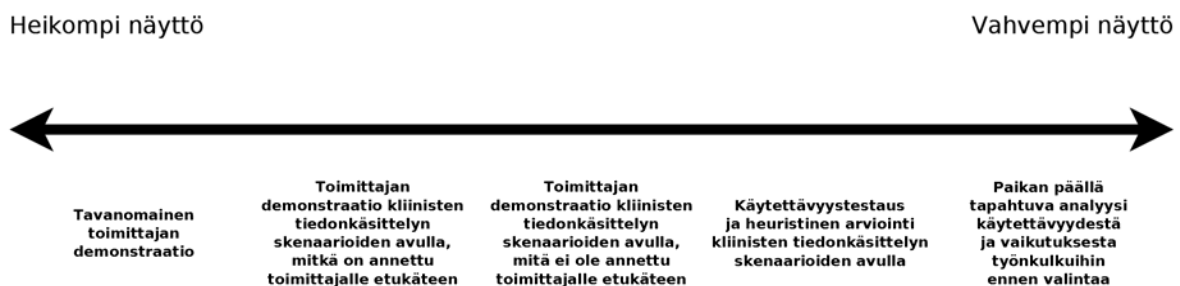
Terveydenhuollon tietojärjestelmän hankinnassa on kysymys investoinnista, joka kohdistuu yleisesti osittain valmiiseen ohjelmistotuotteeseen ja käyttäjäorganisaation tarpeisiin toteutettaviin hankintakohtaisesti määriteltäviin ominaisuuksiin. Haasteet suurten tietojärjestelmäkokonaisuuksien hankinnoissa liittyvät usein niiden käyttökelpoisuuden varmistamiseen suunniteltua tarkoitusta varten sekä hankinnan onnistumiseen investointina mukaan lukien kokonaiskustannusten, toteutusaikataulun ja teknisen toteutettavuuden epävarmuustekijät. Lisäksi suuria hankintoja julkisten organisaatioiden tarpeisiin säätelee hankintalaki [1].

Olemme kokeilleet käyttäjäherätteisen käytettävyydestä järjestelyjä terveydenhuollon tietojärjestelmä-hankinnan ja -kehityksen tarpeisiin. Testauksen toteuttamisella tavoitellaan tietojärjestelmän valintaa tukevaa arviointia osana muita mahdollisia arviointitapoja, joilla voidaan eritasoisesti vertailla järjestelmätuotteiden käyttökelpoisuutta niiden suunnitellussa tarkoituksessa. Kuvassa 1 on esitetty testauksen lisäksi muita käytettyjä arviointitapoja, joiden tuottama näytön vahvuus arvioinnin tueksi on kuvattu niiden keskinäisellä asemoinnilla heikommasta vahvempaan näyttöön järjestyksessä vasemmalta oikealle. Kliinisten tiedonkäsittelyn skenaarioiden avulla toteutettu käytettävyydestä asettuu heuristisen arvioinnin kanssa näistä toiseksi oikealta. Tämän mukaan vahvempaa näyttöä voidaan saada vain

kokeilemalla järjestelmää sen todellisessa työympäristössä, mikä ei välttämättä tule kysymykseen ennen todellista valintaa sen vaatimien työläiden järjestelyjen takia, mutta sitä on kuitenkin mahdollisuuksien mukaan hyvä harkita vahvana ja toimivana arviointitapana.

Käytettävyyden edistäminen ja arviointi

Tuotteen tai tietojärjestelmän kokonaislaatua ja käyttökelpoisuutta voidaan edistää nimeämällä haluttuja tuoteominaisuuksia, joiden toteuttaminen johtaa asiakkaalle mielekkääseen investointiin ja käyttäjille kelvolliseen lopputulokseen. On kuitenkin odotettavissa, että tällainen "rasti ruutuun" -tyyppinen hankintatapa ei välttämättä tuota asiakkaalle hyvin ennustettavaa investoinnin toteutumista, eikä käyttäjille luotettavaa lupausjärjestelmällä hoidettavien asioiden todellisesta helpottumisesta. Yleisesti ominaisuuksilla kauppaa käytävässä hankinnassa ostajan ja myyjän erisuuntaiset tavoitteet voivat johtaa joko siihen, että kauppaan tulee tingityksi mukaan tarpeettoman suuri määrä ominaisuuksia tai tarpeelliset ominaisuudet sisältävä kauppa lyödään lukkoon epärealistisen halvalla hinnalla. Molemmissa tapauksissa keskittyminen pääasiassa tekniikkiin ja näennäisesti täsmällisiin ominaisuusvaatimuksiin on omiaan vaarantamaan hankinnan todellisen käyttökelpoisuuden sekä usein myös aikatauluihin ja kokonaiskustannuksiin liittyvien arvioiden toteutumisen.



Kuva 1. Terveydenhuollon tietojärjestelmän valintaa tukevan näytön vahvuus erilaisilla järjestelmän käyttöä edustavilla arviointitavoilla [mukailten 2].

Mielekkään investoinnin ja käyttäjille hyvin kelvollisen lopputuloksen saavuttamiseksi on teknisten ominaisuuksien lisäksi perusteltua tarkastella myös tuotteen käytettävyyttä. Varsinkin, jos tuote on riittävän valmis sen kokeilemiseksi tavoiteltua käyttöä edustavalla tavalla, on sen käytettävyyden arvioinnilla mahdollisuus suoraviivaisempaan tapaan todeta hyvinkin monimutkaisten teknisten ratkaisujen toimintamahdollisuudet riippumatta siitä ovatko ne asiakkaan määrittelemiä tarpeita vai tuotteen tarjoamia mahdollisuuksia helpottaa asioiden tekemistä käytännössä. On myös syytä todeta, että tekemällä määrittelyt ja arvioimalla tuotteiden kyvykkyyttä oikeaa käyttöä tarkastelemalla voidaan usein jättää ottamatta kantaa moniin teknisiin yksityiskohtiin. Tällä tavalla toimimalla on mahdollista rohkaista selkeästi radikaalimpiin parannuksiin tuotekehityksessä ja välttää yksityiskohtaisiin teknisiin vaatimuksiin mahdollisesti jääneet aukot, jotka hankaloittavat todellista käyttöä.

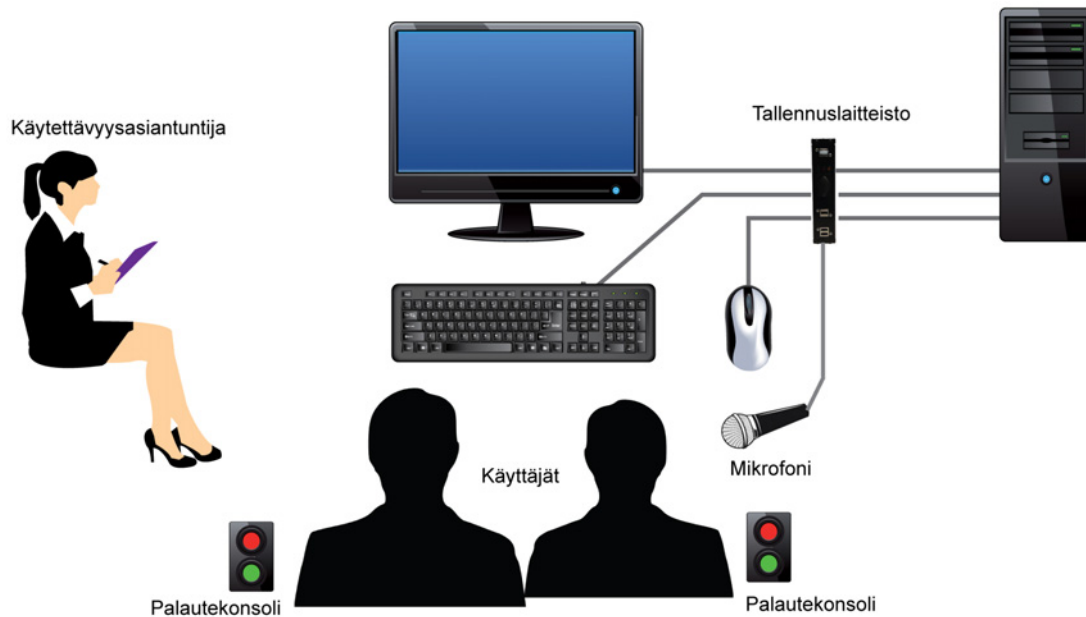
Käytettävyyden tarkastelua varten on monenlaisia arviointimenetelmiä, joiden kyvykkyys koostuu Hartsonin [3] mukaan ainakin luotettavuudesta, kattavuudesta, validiteetista, tulosten sovellettavuudesta sekä tehokkuudesta. Tehokkuudelle on edelleen esitetty yleistaloudellisuuden lisäksi tarkemmaksi kriteeriksi se, miten menetelmä löytää todellisia käytettävyysongelmia ja miten niiden todellisuus voidaan käytännössä vahvistaa. Erääksi ratkaisuksi tällaisten ns. standardiongelmiin määrittelemiseksi on esitetty useammalla eri käytettävyyesarviointimenetelmällä tapahtuvaa samojen ongelmien tunnistamista. Tällainen ratkaisu on kuitenkin asetettu kyseenalaiseksi, koska löydettyjen ongelmien on laatu on riippuvainen voimakkaasti käytetystä tulosaineistosta ja siten myös sovelletuista menetelmistä. Todellisten käytettävyyteen vaikuttavien asioiden tehokkaaksi löytämiseksi on ehdotettu oikeaan käyttöön perustuvien, erityisen hyvien ja huonojen ominaisuuksien tunnistamista, mitä on tutkittu kriittisten tapahtumien menetelmään perustuvilla tavoilla. Käytettävyysongelmien ja muiden havaintojen merkityksellisuuden arviointiin kuuluu myös niiden vakavuuden luokittelu. [3,4] Ongelmien todellinen esiintyminen voidaan siis havaita ensisijaisesti oikeiden käyttäjien avulla kun taas asiantuntija-arviolla on jossain määrin

mahdollista tarkastella valmiiksi tunnistettujen ongelmien samanlaisuutta, esiintymismahdollisuutta ja vakavuutta kohteena olevassa järjestelmässä.

Aineisto ja menetelmät

Käytettävyydestauksen soveltamiseksi terveydenhuollon tietojärjestelmähankkeissa kokeilimme käyttäjäherätteen käytettävyydestauksen järjestelyjä [5,6]. Testitavassa on tarkoituksena tunnistaa ne käyttötilanteet, jotka käyttäjät kokevat ongelmallisiksi tai erityisen hyviksi. Nämä kokeilutestit toteutettiin paritesteinä, jotka noudattivat muuten yleisiä käytettävyydestin ohjeita ja järjestelyjä [7]. Testitehtäviksi oli valittu joukko tyypillisiä hoitajien ja lääkärin työtehtäviä. Nämä käytettävyydestien kokeilutestit suoritettiin paritesteinä, joissa testikäyttäjäparin muodostivat kaksi hoitajaa tai kaksi lääkäriä. Testauksen kohteena oleva järjestelmä ei kuitenkaan ollut testeihin osallistuville käyttäjille välttämättä tuttu vaan heidän piti suoriutua testitehtävistä pääasiassa omaan ammattitaitoon ja kokemukseen perustuvien tiedoin. Kokeilu suoritettiin yhden järjestelmän parissa, joten siitä saadut kokemukset kertovat lähinnä testijärjestelyn soveltavuudesta terveydenhuollon tietojärjestelmän käytön testaamiseen sen varsinaiseen käyttäjäkuntaan kuuluvilla ammattilaisilla. Vastavaa menetelmää on tutkittu ammattikäyttöön tarkoitettujen isännöintialan tietojärjestelmien parissa [6] ja sovellettu kaupallisesti mm. toiminnanohjausjärjestelmien ja kuluttajaverkkopalveluiden testauksessa, missä on saatu nopeasti käytäntöön soveltuvia kehitysehdotuksia työn ja käytön helpottamiseksi.

Testijärjestelyn muodostivat ATK-koulutustilan nurkkaan sijoitettu tietokonetyöasema ja siihen liitetty käyttötilanteiden tallennukseen tarkoitettu laitteisto. Testijärjestely poikkesi perinteisestä käytettävyydestin järjestelystä siten, että käyttäjillä oli mahdollisuus merkitä ongelmallisia ja hyviä käyttötilanteita vapaasti testien aikana heille annetuilla palautekonsoleilla napinpainalluksin (kuva 2). Käyttötilanteiden tallennuksen kohteena olivat näyttökuvan videointi, näppäimistöpainallusten ja hiirenklikkausten kaappaaminen sekä äänentallennus mikrofoniin avulla puheen ja mahdollisten tietokoneen merkkiäntäen taltioimiseksi. Testejä



Kuva 2. Käytettävyysasiantuntija ohjaa testin kulkua ja testikäyttäjillä on langattomat kahden painikkeen palautekonsolit. Näytön, näppäimistön ja hiiren, sekä tietokoneen väliin on liitetty tallennuslaitteisto, joka tallentaa käyttötapahtumien lisäksi mikrofoniin ääntä videotallenteen yhteyteen.

järjestettiin yhteensä kolme, ja niissä oli eri tehtävät. Testien tarkoituksena oli paitsi pilotoida testiasetelmaa, myös erilaisia testitehtäväkokonaisuuksia.

Testeissä merkittyjä ja tunnistettuja käyttötilanteita oli yhteensä 58, joista kahdeksan liittyi kriittisiin ongelmiin kun havaintojen vakavuutta arvioitiin neliportaisella luokittelusteikolla (kriittinen, tärkeä, kohtalainen ja vähäinen) perustuen käyttötapahtumatalleiden läpikäyntiin asiantuntija-arviona. Positiivisiksi merkittyjä käyttötapahtumia oli yhtä paljon kuin negatiivisiksi merkittyjä ongelmallisia tilanteita (29), mikä on tavanomaisesta poikkeavaa, koska negatiivisia merkintöjen osuus on ollut muiden kokemusten perusteella selvästi suurempi lähes aina. Pilotoinnin tuloksena testausjärjestely todettiin toimivaksi tavaksi taltioida käytettävyydestien tapahtumat. Sen lisäksi katsottiin tarpeelliseksi, että käytettävyysasiantuntijalla olisi myös painikekonsoli tehtävien suoritusta koskevaa automatisoidumpaa jälkianalyysiä varten ja käytettävyyshavaintojen löytämisen varmistamista varten myös niitä tilanteita varten, missä selvästi ongelmalliset tai erityisenkin

hyvät käyttötapahtumat jäävät epähuomiosta merkittävää.

Miten käytettävyytestaus voi tukea hankintalain mukaista kilpailutusta

Hankintalain tavoitteina on mainittu julkisten varojen käytön tehostaminen ja laadukkaiden hankintojen tekemisen edistäminen sekä tasapuolisten mahdollisuuksien turvaaminen yrityksille julkisten hankintojen tarjouskilpailuissa [1]. Julkinen projekti nojaa etupäässä vaatimusten määrittelyyn, toimintamalleihin ja valintakriteereihin, joiden perusteella vaihtoehtoisia tarjouksia voidaan arvioida. Arviointi, erityisesti käytettävyydestaus, on kriittinen osa käyttökelpoisen järjestelmän toteutusta ja keskittyy siihen, miten hyvin järjestelmä tukee käyttäjien tavoitteiden saavuttamista. Lisäksi huomioitavaa on testikäyttäjien tarkoituksenmukainen määrä sekä haasteet toteuttaa hyviä testauskäytäntöjä ja pitää huolta toimintatapojen läpinäkyvyydestä julkisissa projekteissa [8].

Käytettävyydestä, jossa käyttäjävuorovaikutus ja käyttöliittymätapahtumat tallennetaan, mahdollistaa sen käyttämisen hankintalain mukaisissa kilpailutuksissa jäljitettävyyttä ansiosta. Tällaisen järjestelyn voi toteuttaa modernilla käytettävyydestestauslaboratoriokalustolla kohtuullisella vaivalla. Jäljitettävyyttä saavutetaan tallentamalla kaikki käytettävyydestestien tuloksiin olennaisesti vaikuttaneet käyttäjän, tuotteen ja käyttöympäristön tapahtumat. Vertailukelpoisuus vaatii lisäksi sitä, että samaa järjestelmän osakokonaisuuden testeihin osallistuvat käyttäjät ovat osaamiseltaan riittävän samantasoisia. Esimerkiksi kunkin järjestelmän ajanvarausominaisuutta testaa kokeneempi hoitajapari (5-10 v. työkokemus) ja toinen pari juuri valmistuneita hoitajia, joista kummallekaan pareista kyseinen järjestelmä ei ole entuudestaan tuttu. Henkilöiden osallistuminen useampaan samanlaiseen testiin ei saa myöskään vääristää testiasetelmaa vaan jokaisen järjestelmän testaukseen osallistuvat uudet henkilöt tai henkilöiden kiertämisyjärjestyksellä ratkaistaan tasapuolinen käyttäjien kokemus edeltävistä testeistä.

Käytettävyydestestien tulokset osana hankintakriteeristöä

Tuotteiden vertailua varten niiden arviointia varten on oltava kriteerit, joiden perusteella vaihtoehtojen asettaminen paremmuusjärjestykseen on tarkoituksenmukaista ja toistettavaa. Käytettävyyden ollessa osana arviointikriteeristöä, niiden on oltava mitattavissa ja toistettavissa sillä tavalla, että arvioitavana oleville vaihtoehdoille lopputuloksena saatava paremmuusjärjestys on objektiivisesti arvioitavissa. Tämän tulee olla vaatimuksena myös käytettävyydestestaukselle siitkin huolimatta, että merkittävät testaushavainnot on osin ennalta arvaamattomia ja tuovat esille sellaisia asioita, joiden listaaminen kattavasti etukäteen on käytännössä mahdotonta. On myös syytä muistuttaa, että samasta syystä käytettävyydestestauksen hyödyntäminen on osoitettu tehokkaaksi tavaksi arvioida todellista käyttökelpoisuutta. Vertailun vuoksi esimerkiksi pelkällä asiantuntija-arviolla on taipumus löytää tehokkaasti suuri määrä käytettävyyso ongelmia, mikä perustuu asiantuntijan heuristiseen kykyyn ja sen täydentämiseen pitkällä

muistilistoilla heuristisista säännöistä. Käytännössä heuristinen arviointi löytää suuren osan merkittävistä käytettävyyso ongelmista ja mittavan määrän pienempiä ongelmia. Käytettävyydestestauksen avulla tarkennetaan arviota ongelmien todellisesta merkittävyydestä ja erityisesti löydetään uusia merkityksellisiä käytettävyyteen vaikuttavia tekijöitä, joiden heuristinen havaitseminen ei ole ennustettavaa. [1,8,9]

Arvioitavien tietojärjestelmätuotteiden kyvykyys ja tuotekehitys tulevat kyseeseen hankintahetken lisäksi niiden varsinaisen käyttöelinkaaren aikana. Koska järjestelmien tarkoituksena on mahdollistaa ja helpottaa tiettyjen työnkulkujen tekemistä, on testauksen keinoin mahdollista todentaa niiden olevan käytännöllisiä. Suorituskyvyn mittaaminen esimerkiksi tehtäväkokonaisuuksiin kuluneilla ajoilla tai käyttötapahtumissa toteutuneiden vasteaikojen mittauksilla mahdollistaa laitteiston ja käyttöympäristön huomioimisen pelkkien teknisten suorituskykylupausten sijaan. Järjestelmän vertailua varten tehtyjen testien tallenteita voidaan siis käyttää sen varmistamiseen, että hankinnan kohteena olleen tuotteen ominaisuudet tulevat toteutetuksi myös tuotantokäytössä vastaavalla tavalla kuin mikä niiden eduksi on laskettu vertailuhetkellä. Mahdollisuuksien mukaan testaus soveltuu myös vertailukehittämisen eli ns. benchmarkingin tarpeisiin kun halutaan verrata käyttöön valittua järjestelmää kilpaileviin tuotteisiin ja ottaa niistä opittuja ideoita mukaan oman järjestelmän kehittämismahdollisuuksien ideointiin.

Optimoitu analysointimalli käytettävyydestestaus-tulosten vertailuun

Käytettävyydestestauksessa tarkoituksenmukaisena otosmääränä pidetään yleisesti viittä testikäyttäjää [10] silloin kun on kysymyksessä yksilötesti (ei paritesti). Jo viiden testin järjestäminen jokaista vertailtavaa tuotetta kohti tuottaisi mittavan määrän testikertoja esimerkiksi tapauksessa, jossa testattavia tuotteita olisi kuusi ja jokaiselle tuotteelle tehtäisiin neljä erilaista testiä käyttäjien ja käyttöskenaarioiden vaihdellessa. Yhteensä testikertoja koituisi tällaisessa järjestelyssä 20 per tuote ja kokonaisuudessaan 120 kaikille kuudelle vertailtavalle tuotteelle. Jos testit suoritettaisiin paritesteinä ja

noin tunnin kestäviä testejä tehtäisiin neljä päivässä, sitoutuisi testeihin noin kolme käyttäjäyökuukautta (2 käyttäjää x 120 testiä / 4 testiä päivässä = 60 htp.) sekä puolitoista kuukautta käytettävyyssiantuntijan työtä pelkästään testien seurantaan ja ohjaamiseen. Tämä ei ole ylitsepääsemätön uhraus riittävän suuressa hankinnassa verrattuna sillä odotettavasti saavutettaviin etuihin nähden, mutta kaippaa silti kriittistä tarkastelua.

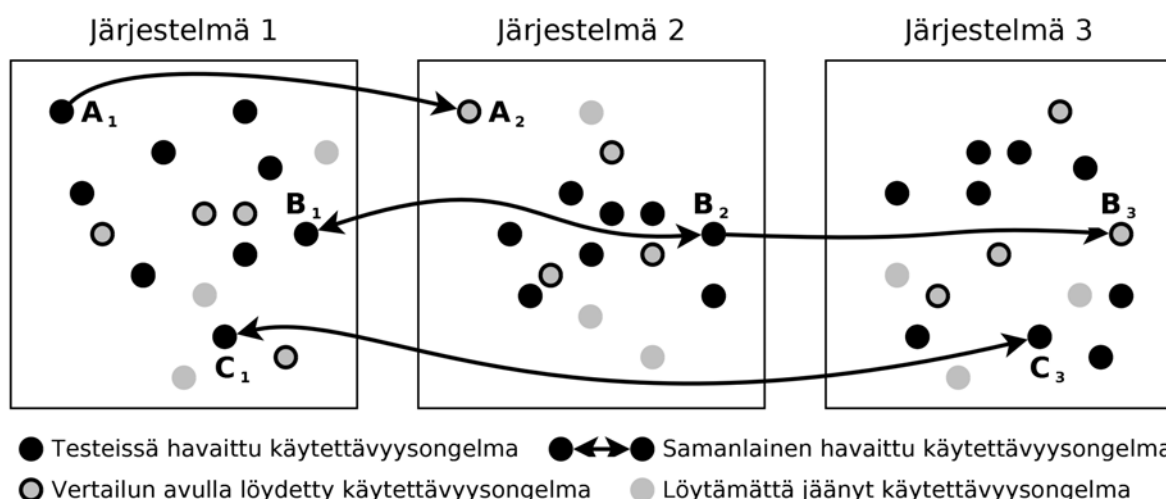
Kuten aikaisemmin todettiin, vaatii vertailukelpoisten testitulosten tuottaminen merkityksellisten havaintojen tunnistamista. Vain tällä tavalla havaintojen painoarvo voidaan luokitella ja toisaalta laskea yhteen samoista syistä johtuvien havaintojen määrällistä esiintymistä. Tarkastellaan seuraavassa hypoteesiä siitä, kuinka muista tuotteista löydettyjä käytettävyyshavaintoja voidaan hyödyntää testitulosten luotettavuuden ja vertailukelpoisuuden parantamiseen:

1. Kerätään kunkin tuotteen testauksessa ilmenneet käytettävyyshavainnot ja kootaan yhteen tunnistettavasti samanlaiset esiintymät.
2. Listataan kaikkien tuotteiden testauksessa esiintyneet havaintoryhmät ja tunnistetaan tuotteiden välillä keskenään samanlaiset esiintymät.

3. Arvioidaan kunkin tunnistettavasti erilaisen havaintoryhmän esiintymismahdollisuutta myös niissä tuotteissa, joiden kohdalla kyseistä havaintoa ei testeissä esiintynyt.

Tällä tavalla toimimalla oletetaan löydettävän sellaisia käytettävyyteen liittyviä ongelmia ja positiivisia asioita niistäkin testattavista tuotteista, joiden kohdalla ne eivät tulleet esille pienestä otannasta johtuen. Oletus perustuu siihen, että samaan käyttötarkoitukseen tehdyissä tuotteissa käytettävyyteen liittyvät tekijät ovat todennäköisesti suurelta osin samanlaisia.

Hypoteesiä voidaan kokeilla tekemällä varsinaiset käytettävyysetestit ja tarkastelemalla niiden tuloksia syntyneistä tallenteista. Kuvassa 3 on havainnollistettu mahdollisten käytettävyyshavaintojen tunnistamista hyödyntäen muissa testatuissa tuotteissa esiintulleita havaintoja. Tässä tapauksessa on yksinkertaisuuden vuoksi esitetty kolmessa tuotteessa havaitut käytettävyysongelmat (negatiiviset havainnot) ja kunkin tuotteen muiden mahdollisten käytettävyysongelmien tunnistamisen periaate vertailemalla niitä aina kahden muun tuotteen käytettävyysetestausten tuloksiin.



Kuva 3. Kolmen kuvitteellisen järjestelmän käytettävyysongelmien löytyminen testeillä ja niiden täydentäminen arvioimalla ovatko samat ongelmat mahdollisia muissa järjestelmissä; esim. järjestelmästä 1 löydetty ongelma A voi esiintyä järjestelmässä 2 ($A_1 \rightarrow A_2$).

Vertailun perustana käytetään järjestelmistä testien avulla löydettyjä käytettävyyso ongelmia ja käytettävyyso asiantuntijan arviota siitä, onko ongelma todennäköinen (tai edes mahdollinen) kussakin järjestelmässä, jossa sitä ei testien puitteissa havaittu. Tällaisia löydöksiä oletetaan saatavan merkittävässä määrin sellaisessa testausasetelmassa, jossa seuraavat ehdot täyttyvät:

- i. Käytettävyyso testeissä otanta on pienempi kuin viisi (5), jota pidetään yleisesti riittävänä määränä merkittävimpien käytettävyyso ongelmien löytämiseen.
- ii. Vertailtavat järjestelmät täyttävät toiminnallisuuksiltaan sellaiset minimivaatimukset, että niiden voidaan katsoa olevan soveltuvia samoihin suunniteltuihin käyttötarkoituksiin.

Jos käytettävyyso ongelmia tai -havaintoja löydetään tällä tavalla lisää joidenkin järjestelmien kohdalla vähintään siinä määrin, mikä tarvitaan pienimmän kahden järjestelmän välisen vertailusuureen alkuperäisen erotuksen kumoamiseksi, voidaan vertailun avulla saatuja tuloksia pitää merkityksellisinä. Tämä tekee mahdolliseksi sen, että vertailun tuloksena vaihtoehtoisten järjestelmien paremmuusjärjestys voi muuttua, edellyttäen että vertailun tuottaman luotettavuuden parantuminen voidaan tilastollisesti osoittaa.

Yhteenveto

Sovelsimme käyttäjäherätteen käytettävyyso testauksen menetelmiä terveydenhuollon tietojärjestelmä hankinnan ja -kehityksen tarpeisiin kokeilemalla testijärjestelyjä tyypillisillä hoitajien ja lääkärin työtehtäviä vastaavilla skenaarioilla erään tietojärjestelmän parissa. Testijärjestely poikkesi perinteisestä käytettävyyso testin järjestelystä siten, että käyttäjillä oli mahdollisuus merkitä ongelmallisia ja hyviä käyttötilanteita vapaasti testien aikana napinpainalluksin, samalla kun kaikki käyttäjän ja järjestelmän välinen vuorovaikutus tallennettiin suullisine kommentteineen. Tämän järjestelyn todettiin soveltuvan käytettäväksi hankintalain mukaisissa kilpailutuksissa jäljitettävyyso tentä ansiosta, mikä saavutettiin taltioimalla kaikki käytettävyyso testien tuloksissa huomi-

oitavat tekijät. Tämän lisäksi huomattiin, että järjestelyä voitaisiin parantaa käytettävyyso asiantuntijan ohjaamisessa testeissä antamalla hänelle mahdollisuus merkitä testitehtävien suoriutumiseen ja käytettävyyso havaintojen varmistamiseen liittyviä havaintoja myös napinpainalluksin.

Käytettävyyso testauksessa tuotettavien testitulosten luotettavuuden ja toistettavuuden parantamiseen esitettiin optimoitu analysointimalli ja hypoteesi sen testaamiseksi. Analysointimalli perustuu testeissä yksittäistä järjestelmistä tehtyjen havaintojen hyödyntämiseen jälkiarvioinnissa, jossa havainnon mukaisen löydöksen esiintymismahdollisuutta arvioidaan myös muissa järjestelmissä. Tällä oletetaan parannettavan tulosten luotettavuutta ja testaustavan toistettavuutta, minkä todentaminen vaatii täysmittaisten testien suorittamista sekä analyysiä tulosten luotettavuuden tilastollisesti merkittävästä paranemisesta, varsinkin siinä tapauksessa, että esitetyn analysointimallin soveltaminen johtaisi vertailtavien järjestelmien paremmuusjärjestyksen muuttumiseen.

Lähteet

- [1] Laki julkisista hankinnoista. 30.3.2007/348, Eduskunta.
- [2] Kushniruk A, Beuscart-Zéphir M-C, Grzes A, Borycki E, Watbled L, Kannry J. Increasing the Safety of Healthcare Information Systems through Improved Procurement: Toward a Framework for Selection of Safe Healthcare Systems. *Healthc Q* 2010 Sep;13 Spec No:53-8.
- [3] Hartson HR, Andre TS, Williges RC. Criteria for evaluating usability evaluation methods. *International Journal of Human-Computer Interaction* 2003;15(1):145-181.
- [4] Ovaska S, Aula A, Marjaranta P. Käytettävyyso tutkimuksen menetelmät. Tampere: Tampereen yliopisto, Tietojenkäsittelytieteen laitos; 2005. B-2005-1.
- [5] Pitkänen J, Pitkärananta M, Kaipio J. Uusi menetelmä terveydenhuollon tietojärjestelmien kehittämisen avuksi: käyttäjäherätteen käytettävyyso testaus aidossa

käyttöympäristössä. Finnish Journal of eHealth and eWelfare 2013;5(2-3):120-127.

[6] Pitkänen J, Pitkäranta M, Nieminen M. Usability testing in real context of use: the user-triggered usability testing. Proceedings of NordiCHI'12. ACM Press; 2012.

[7] Nielsen J. Usability Engineering. Morgan Kaufmann; 1993.

[8] Buie E, Murray D. Usability in Government Systems - User Experience Design for Citizens and Public Servants. Morgan Kaufmann; 2012.

[9] Albert W, Tullis T. Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics. Morgan Kaufmann; 2008.

[10] Nielsen J. How Many Test Users in a Usability Study? Fremont, CA: Nielsen Norman Group; 2012 [viitattu 5.12.2012]. Saatavissa: <http://www.nngroup.com/articles/how-many-test-users/>

[11] Pitkäranta M. Menetelmä ja laitteisto käytettävyyden tutkimiseen aidossa käyttötilanteessa. Diplomityö, Tieto- ja sähkötekniikan tiedekunta, Tampereen teknillinen yliopisto; 2010.